

ISSN: 2612-226X

Studi AISV 7

LA VARIAZIONE LINGUISTICA IN CONDIZIONI DI CONTATTO: CONTESTI ACQUISIZIONALI, LINGUE, DIALETTI E MINORANZE IN ITALIA E NEL MONDO

LANGUAGE CHANGE UNDER CONTACT
CONDITIONS: ACQUISITIONAL CONTEXTS,
LANGUAGES, DIALECTS AND MINORITIES
IN ITALY AND AROUND THE WORLD

a cura di
Luciano Romito

LA VARIAZIONE LINGUISTICA
IN CONDIZIONI
DI CONTATTO: CONTESTI
ACQUISIZIONALI, LINGUE,
DIALETTI E MINORANZE
IN ITALIA E NEL MONDO

LANGUAGE CHANGE UNDER CONTACT CONDITIONS:
ACQUISITIONAL CONTEXTS, LANGUAGES, DIALECTS
AND MINORITIES IN ITALY AND AROUND THE WORLD

a cura di
LUCIANO ROMITO

Studi AISV è una collana di volumi collettanei e monografie dedicati alla dimensione sonora del linguaggio e alle diverse interfacce con le altre componenti della grammatica e col discorso. La collana, programmaticamente interdisciplinare, è aperta a molteplici punti di vista e argomenti sul linguaggio: dall'attenzione per la struttura sonora alla variazione sociofonetica e al mutamento storico, dai disturbi della parola alle basi cognitive e neurobiologiche delle rappresentazioni fonologiche, fino alle applicazioni tecnologiche.

I testi sono sottoposti a processi di revisione anonima fra pari che ne assicurano la conformità ai più alti livelli qualitativi del settore.

I volumi sono pubblicati nel sito dell'Associazione Italiana di Scienze della Voce con accesso libero a tutti gli interessati.

Curatore/Editor

Cinzia Avesani (CNR-ISTC)

Curatori Associati/Associate Editors

Franco Cutugno (Università di Napoli), Barbara Gili Fivela (Università di Lecce), Daniel Recasens (Università di Barcellona), Antonio Romano (Università di Torino), Mario Vayra (Università di Bologna).

Comitato Scientifico/Scientific Committee

Giuliano Bocci (Università di Ginevra), Silvia Calamai (Università di Siena), Mariapaola D'Imperio (Rutgers University), Giovanna Marotta (Università di Pisa), Stephan Schmid (Università di Zurigo), Carlo Semenza (Università di Padova), Claudio Zmarich (CNR-ISTC).

© 2020 AISV - Associazione Italiana Scienze della Voce
c/o LUSI Lab - Dip. di Scienze Fisiche
Complesso Universitario di Monte S. Angelo
via Cynthia snc
80135 Napoli
email: presidente@aisv.it
sito: www.aisv.it



Edizione realizzata da
Officinaventuno
Via F.lli Bazzaro, 18
20128 Milano - Italy
email: info@officinaventuno.com
sito: www.officinaventuno.com

ISBN edizione digitale: 978-88-9765-749-1
ISSN: 2612-226X

Sommario

LUCIANO ROMITO	
Prefazione	7
Renata Savy	9

PARTE I

La variazione fonetica in contesti dialettali

PIA GRECA, JONATHAN HARRINGTON	
An acoustic analysis of metaphony in the dialects of the Lausberg Area (Southern Italy)	23
MARGHERITA DI SALVO	
Contatto interdialettale e cambiamento linguistico in un dialetto italiano all'estero	43
TERESA CARBUTTI	
L'italiano: fonte di riscatto per parlanti dialettofoni	59
MANUELA FRONTERA	
Calabrian migrants in Argentina: assessing first language phonetic attrition	73

PARTE II

Lingue minoritarie e varietà regionali

MICHELA RUSSO, FABIO APREA	
Identità linguistica di una lingua minoritaria alloglotta della Puglia (Italia meridionale): il francoprovenzale di Faeto	97
ALESSANDRO VIETTI, DANIELA MEREU	
Sistemi vocalici in contatto nell'italiano di Bolzano: un'analisi esplorativa <i>corpus-based</i>	125
ANDREA TARASI, LUCIANO ROMITO	
La consonante occlusiva labiovelare sorda nei dialetti calabresi: uno studio pilota	145

DALILA DIPINO Lunghezza vocalica e consonantica in due varietà acquisite di ligure: uno studio pilota	157
---	-----

PARTE III

La variazione interlinguistica

STEPHAN SCHMID, KRISTIJAN RAJIC La pronuncia delle vocali nasali nel francese L2 di apprendenti italofofoni: acquisizione e insegnamento esplicito	189
--	-----

FABIO ARDOLINO Stereotypical dimensions and Second Language development in a migrant community: A pilot study	211
---	-----

ANCUTA BUDEANU, ANNA DE MEO, MASSIMO PETTORINO Caratteristiche fonetiche dell'italiano di romeni in Calabria. Riflessioni su ritmo e lunghezza vocalica e consonantica	233
--	-----

MARINA CASTAGNETO, IRINA STAN <i>Code-switching</i> in contesto italiano L2: una famiglia croata e due famiglie romene a confronto	245
--	-----

FRANCESCA PEZZELLA La geminazione in un gruppo di apprendenti albanofoni: percezione e produzione di una nuova categoria fonologica	269
---	-----

PARTE IV

Variazione e prosodia

MAYARA DA SILVA NETO, BARBARA GILI FIVELA, ELISABETTA SANTORO, FLAVIANE ROMANI FERNANDES SVARTMAN Do mitigation strategies affect prosodic correlates? An investigation on orders and requests in Italian	295
--	-----

GLENDIA GURRADO Indici fonetici dell'esclamativa wh nell'italiano parlato a Bari	319
---	-----

PATRIZIA SORIANELLO Sul riconoscimento percettivo delle domande retoriche	341
--	-----

RICCARDO ORRICO, VIOLETTA CATALDO, MARIAPAOLA D'IMPERIO The effect of early exposure in the production of Salerno Italian question intonation	361
---	-----

SIMONA SBRANNA, FRANCESCO CANGEMI, MARTINE GRICE Quantifying L2 interactional competence	383
---	-----

PARTE V

Relazioni a tema libero

SONIA CENCESCHI, CHIARA MELUZZI, NICHOLAS NESE Speaker's identification across recording modalities: a preliminary phonetic experiment	409
BARBARA GILI FIVELA, SONIA I. D'APOLITO, GIORGIA DI PRIZIO Controllo motorio e disartria nella malattia di Parkinson: uno studio pilota sulla labializzazione	429
Autori	441

LUCIANO ROMITO

Prefazione

Il presente volume è una raccolta di contributi orientati allo studio e all'analisi della variazione linguistica in condizioni di contatto: contesti acquisizionali, lingue, dialetti e minoranze in Italia e nel mondo.

Nella struttura della raccolta, emergono diverse prospettive come la variazione fonetica in contesti dialettali, le lingue minoritarie e le varietà regionali, il contatto linguistico in ambito acquisizionale, la variazione interlinguistica e la variazione prosodica.

La manifestazione ha accolto alcune sessioni plenarie come quella della collega Elisa Fernández Rei dell'ILGA, Universidad de Santiago de Compostela dal titolo *Ibridazione intonativa di varietà galiziane e spagnole a contatto*, o quella della collega Esther de Leeuw della Queen Mary University of London dal titolo *Variation in Phonetic and Phonological L1 Attrition*.

I contributi provengono da studiosi di tutto il mondo pertanto la raccolta può essere considerata internazionale.

Il convegno si è concluso con un'interessante Tavola Rotonda sul *contatto linguistico nell'odierna società globale*. Il moderatore, Luciano Romito, ha invitato gli ospiti a dibattere sui problemi relativi alla presenza della lingua arbëreshë in Calabria (prof. Francesco Altimari), del valdese nel comune di Guardia Piemonte (prof. Antonio Romano), del Grieco in alcuni paesi della provincia di Reggio Calabria (prof.ssa M. Olimpia Squillaci). Gli argomenti trattati durante la Tavola Rotonda e le contaminazioni tra minoranze linguistiche, italiano e dialetti locali sono stati anche trattati musicalmente. All'interno del Museo Civico del Comune di Rende l'organizzazione del convegno ha programmato un evento culturale dal titolo – HYBRID TUNES – Mediterranean Traditional Music che ha previsto per l'appunto, musica e presentazione di strumenti musicali tipici delle minoranze linguistiche presenti in Calabria come quella grecanica (tamburo a cornice, zampogna a paro, “a moderna”/ organetto, lira e canto), o la musica Valdese (voce, flauto, fisarmonica, galoubet, chitarra e tamburo a cornice), la musica arbëreshë rappresentata dal Coro di Voci femminili VUXHË GRASH e la musica sefardita. Sono anche stati presentati i modelli musicali e gli strumenti della cultura calabrese (chitarra battente, tamburello, voce, zampogna a chiave lucana, surdulina, organetto, bottiglia a chiave, organetto 8 bassi, fisarmonica, mandolino, ballo).

Infine è stato presentato un Duo composto da un artista calabrese, Checco Pallone (chitarra e tamburo a cornice) e un artista africano Moussa Ndao (percussioni africane e voce). Il duo rappresenta una contaminazione musicale tra i nuovi

immigrati africani la loro musica i loro strumenti e la cultura della regione Calabria con i propri modelli e strumenti musicali.

Il convegno si è concluso con un evento satellite organizzato dall'Osservatorio sulla Linguistica Forense OLF dal titolo: *Il valore scientifico della Trascrizione*. Il coordinatore dell'OLF prof. Luciano Romito ha affrontato l'annoso problema della trascrizione forense in ambito giuridico, del suo valore scientifico e delle competenze richieste insieme ai suoi ospiti rappresentanti delle diverse cariche istituzionali come il rappresentante della Commissione Giustizia al Parlamento italiano, il presidente del Tribunale di Palmi (dott. Gianfranco Grillone), il rappresentante dell'Unione Camere Penali Italiane (avv. Marcello Manna), il rappresentante della Procura DDA di Reggio Calabria (dott. Stefano Musolino) e in rappresentanza delle Università la collega prof.ssa Franca Orletti (Università Roma Tre) per la Linguistica Generale e il collega prof. Giorgio Spangher (Università Sapienza di Roma) per la Giurisprudenza.

Renata Savy (30 novembre 1967 – 9 ottobre 2019)

L'Associazione Italiana di Scienze della Voce ha voluto dedicare il suo XVI Convegno Nazionale e questo volume a Renata Savy, professoressa associata presso l'Università di Salerno, socia dell' AISV sin dalla nascita all'Associazione e, prima ancora, socia del Gruppo di Fonetica Sperimentale. Queste pagine e la bibliografia aprono il volume con l'intento di onorare la memoria di Renata, una studiosa, un'amante della ricerca, un'instancabile osservatrice della lingua, del parlato e delle sue strutture. Ripercorrendo il suo percorso di studio e delineando i suoi interessi di ricerca, non è possibile non pensare alla sua forte personalità, alla sua prontezza, alla sua spontaneità e alla sua energia. Ricorderò con voi, quindi, solo alcune tappe, fondamentali, del suo percorso scientifico, nelle quali penso tutti noi potremo riconoscere l'essenza della personalità di Renata.



Il contributo che Renata Savy ha dato al nostro settore è ricchissimo di approfondimenti, di pubblicazioni e progetti, come è già stato messo in luce a Salerno, a inizio dicembre 2019, dalla prof.ssa Voghera nel primo incontro organizzato in ricordo di Renata (dal titolo "Renata Savy: un talento per la ricerca"). Un contributo, quello di Renata, che va ben oltre i singoli lavori pubblicati, perché è intriso di curiosità e forza propulsiva, di potere aggregante attorno a un tema di interesse comune.

Curiosità, sì, che si evince già dal fatto che Renata, con un percorso in Lettere classiche alla Federico II di Napoli, sotto la guida di Federico Albano Leoni, si avvicina alla linguistica e alla fonetica, in particolare all'analisi acustica. Nonostante la scelta di Lettere classiche, inizia a collaborare con il CIRASS, il Centro Interdipartimentale per l'Analisi e Sintesi dei Segnali che sorgeva proprio in quegli anni sotto la direzione di Albano Leoni, e si laurea nel 1992, con una tesi sull'*Analisi acustica della coarticolazione di vocale e consonante fricativa*. Con i due articoli estrapolati dalla tesi e pubblicati negli Atti dei Convegni Nazionali dell'Associazione Italiana di Acustica (AIA) del 1992 e 1993, inizia la sua produzione scientifica.

Una produzione basata sull'analisi rigorosa del dato linguistico, il più possibile spontaneo, che si dirama in varie direzioni a partire dallo studio della variabilità e della variazione linguistica, prendendo in esame sia aspetti segmentali che prosodici. In collaborazione con vari autori, tra cui Albano Leoni, Franco Cutugno, Pietro

Maturi, si occupa in diversi contributi tra il 1994 e il 1996 del sistema vocalico nel parlato connesso, considerando la variazione diafasica e diatopica in Italiano. Ed è nella sua tesi di dottorato, svolto a Roma Tre, sulle *Riduzioni foniche nel parlato. Il ruolo della morfologia flessiva nell'interpretazione del messaggio e nella comunicazione*, che si manifesta anche l'interesse per la prosodia. Renata la mette in evidenza, infatti, come dimensione rilevante addirittura nel preservare dai fenomeni di ipoarticolazione e riduzione che possono riguardare anche unità morfologicamente importanti come i morfi flessivi. Ne scaturiscono due articoli pubblicati negli Atti della Società di linguistica italiana del 1999 e 2001.

È proprio in quel periodo, alla fine degli anni Novanta, che Renata collabora in modo decisivo ai progetti di raccolta, trascrizione e annotazione di Corpora di parlato: AVIP (Archivio delle Varietà di Italiano Parlato), API (Archivio del Parlato Italiano) e, più tardi, CLIPS (Corpora e Lessici di Italiano Parlato e Scritto). Una collaborazione di fondamentale importanza perché, sin dal primo progetto, Renata è stata un punto di riferimento, ad esempio per i criteri di trascrizione. Ricordo ancora una delle prime riunioni per il progetto AVIP, riunione alla quale partecipavano le diverse unità. Penso che sia in quell'occasione, ad esempio, che la sottoscritta ha iniziato a conoscere Renata. Aveva il piglio di chi sa qual è la direzione da intraprendere e la mostra con determinazione! E infatti era un punto di riferimento per molti di noi che cominciarono ad avvicinarsi alla fonetica e alle varie fasi dell'analisi dei corpora di parlato. Oltre a tutta la sua produzione scientifica di quegli anni, penso che la maggior parte di noi ricorderà come preziosissimi i documenti relativi alle norme di trascrizione, di cui menziono qui la versione pubblicata in *Italiano Parlato. Analisi di un dialogo* (a cura di Albano Leoni F., Giordano R., Napoli, Liguori) del 2005, intitolata "Specifiche per la trascrizione ortografica annotata dei testi".

Ma era solo l'inizio, perché il suo ruolo era destinato a diventare sempre più importante. In CLIPS, come riconosce lo stesso Albano Leoni, Renata, insieme a Franco Cutugno, fornisce un apporto cruciale nella progettazione del lavoro, una raccolta di 100 ore di parlato, acquisito in 17 località italiane registrando varietà diafasiche diverse, come il parlato radiotelevisivo (notiziari, interviste, talk show), i dialoghi task-oriented (map task e 'gioco delle differenze'), il parlato letto e il parlato telefonico. Il corpus, in parte trascritto, segmentato ed annotato sul piano fonetico segmentale, rappresenta ancora oggi una preziosa risorsa per gli studiosi di fonti orali (si veda, ad esempio, l'articolo pubblicato con F. Cutugno *CLIPS. Diatopic, diamesic and diaphasic variations in spoken Italian* nei Proceedings of V Corpus Linguistics Conference – CL2009, University of Liverpool).

Come dimenticare, inoltre, l'apporto di Renata ai progetti PRIN che, coordinati da Miriam Voghera, hanno portato alla realizzazione del portale www.parlari-taliano.it tra il 2004 e 2008. Ma voglio ricordare anche tutto il lavoro che Renata ha svolto tra il 2009 e il 2012 per la digitalizzazione del corpus di italiano parlato LIP, oggi disponibile, infatti, nella versione audio VoLIP. Un lavoro che Renata ha portato avanti anche con Franco Cutugno, alla guida del suo gruppo di Salerno. Perché intanto insegnava a Salerno, appunto, dove affascinava e attraeva allo studio

della fonetica numerosi studenti, alcuni dei quali portano e porteranno avanti il suo lavoro, rappresentando una parte importante della sua eredità scientifica. Eredità che riguarda anche lo studio della pragmatica e delle lingue straniere, soprattutto lo spagnolo. Per farsi un'idea, basta comunque guardare il programma di questo nostro convegno, nel quale comunque Renata è ben presente.

A Salerno ha dato un grande contributo all'apertura del laboratorio PAROLE (nel 2006/07), e in seguito ha realizzato il suo Laboratorio di Linguistica applicata. Ma noi la conosciamo ancora di più per il suo contributo alla vita della nostra Associazione. Perché lei, che era già parte del Gruppo di Fonetica Sperimentale, ed è entrata a far parte anche del Gruppo di Studio della Comunicazione parlata (interno alla Società di Linguistica Italiana), è stata membro molto attivo anche dell' AISV, e sin dalle sue origini, nel 2004. Ha curato l'organizzazione di una Tavola rotonda molto interessante sulla prosodia, dal titolo "Modelli di analisi e sistemi di annotazione prosodica a confronto", proprio all'interno di un convegno AISV, nel 2005. Ha organizzato il convegno AISV del 2016, di cui ha curato anche la pubblicazione degli atti. Un punto di riferimento per l' AISV, quindi. Ma è con l'energia impiegata nel tenere vivo l'interesse per le scuole estive e nel garantire i finanziamenti agli studenti che Renata comunicava non solo le sue doti scientifiche e organizzative, ma anche le sue doti umane, il suo amore per la ricerca, da condividere e da far crescere.

Tutti noi abbiamo ben presente la sua lucidità di pensiero, la sua visione indipendente dalle logiche del sistema, quello universitario, al quale ha cercato sicuramente di dare una scossa. Sì, perché Renata è una dei pochi accademici, almeno che io conosca, che ha esplicitamente rifiutato di aderire al sistema, quando questo ha preso la deriva che ben conosciamo, verso una valutazione della ricerca che presenta ombre e lati oscuri.

È difficile racchiudere in poche righe ciò che Renata ci ha lasciato e, nel tentativo di farlo, ancora oggi le parole che mi vengono in mente sono quelle che vi ho già scritto il 9 ottobre 2019: Renata resterà nei nostri cuori con la sua grinta, la sua spontaneità e il suo approccio appassionato.

*Barbara Gili Fivela
Presidente AISV*

Renata Savy – Pubblicazioni

1992

RISPOLI, GIAMPIERO, SAVY RENATA (1992). Una analisi acustica della coarticolazione di vocale e consonante in relazione al tratto di sonorità. In Brambilla Giovanni (a cura di), *Atti del XX Convegno Nazionale della Associazione Italiana di Acustica*, 373-378.

1993

RISPOLI, GIAMPIERO, SAVY RENATA (1993). Alcune considerazioni spettroacustiche sulla [v] italiana. In Peretti Alessandro, Ferrero Franco (a cura di), *Atti del XXI Convegno Nazionale della Associazione Italiana di Acustica*, 91-94.

1995

CUTUGNO, FRANCESCO, SAVY, RENATA (1995). On phonetic boundary across categories for synthetic and natural vocalic speech sounds. In *ISCA Proceedings of Eurospeech95*, Madrid, ISCA, 2297-2300.

ALBANO LEONI, FEDERICO, CUTUGNO, FRANCESCO, SAVY, RENATA (1995). The vowel system of Italian connected speech. In *Proceedings of the 13th International Congress of Phonetic Sciences*, Stockholm KTH and Stockholm University, Stockholm, 396-399.

1996

CUTUGNO, FRANCESCO, SAVY, RENATA (1996). Percezione e categorizzazione di foni vocalici: adeguatezza delle procedure sperimentali. In *Atti del XXIV Convegno Nazionale AIA*, Associazione Italiana di Acustica, Trento, 179-182.

LANDI, ROSSELLA, SAVY, RENATA (1996). Durata vocalica, struttura sillabica e velocità di eloquio nel parlato connesso. In *Atti del XXIV Convegno Nazionale AIA*, Associazione Italiana di Acustica, Trento, 65-70.

1997

SAVY, RENATA, CUTUGNO, FRANCESCO (1997). Ipoarticolazione, riduzione vocalica, centralizzazione: come interagiscono nella variazione diafasica? In *Fonetica e fonologia degli stili dell'Italiano parlato*. Roma: Esagrafica, 177-194.

SAVY, RENATA, CUTUGNO, FRANCESCO (1997). Hypospeech, vowel reduction, centralization, how do they interact in diaphasic variations?. In *Actes du XVIe Congrès International des Linguistes*, Oxford: Pergamon-Elsevier, 1-13.

1998

SAVY, RENATA, CUTUGNO, FRANCESCO (1998). Analisi acustica di vocali in monosillabi cliticici in parlato spontaneo italiano. In *Atti del XXVI Convegno Nazionale AIA*, Istituto Elettrotecnico Nazionale G. Ferraris, Torino, 301-306.

1999

CUTUGNO, FRANCESCO, SAVY, RENATA (1999). Correlation between segmental reduction and prosodic features in spontaneous speech: the role of tempo. In *Proceedings of the 14th International Congress of Phonetic Sciences*, Berkeley Linguistics Department, University of California at Berkeley, 471-474.

SAVY, RENATA, CUTUGNO, FRANCESCO (1999). Su alcune correlazioni tra riduzioni segmentali e tratti prosodici nel parlato spontaneo: il ruolo del fattore tempo. In *Atti del XXVII Convegno Nazionale AIA*, Università di Genova, Genova, 183-187.

SAVY, RENATA (1999). Riduzioni foniche nella morfologia del Sintagma Nominale nel parlato spontaneo: indagine quantitativa e aspetti strutturali. In *Fonologia e Morfologia dell'italiano e dei dialetti d'Italia*, Roma: Bulzoni, 201-221.

SAVY, RENATA (1999). *Riduzioni foniche nel parlato spontaneo: il ruolo della morfologia nell'interpretazione del messaggio e nella comunicazione*. Tesi di dottorato, Università di Roma Tre.

2000

SAVY, RENATA, GAMAL, DALIA, D'ANNA, LEANDRO (2000). Metodi di analisi e codifica del piano suprasegmentale a confronto. In *Atti del XXVIII Convegno Nazionale AIA*, Bari, Politecnico di Bari - Università di Bari, 257-260.

GOMEZ PALOMA, FRANCESCA, SAVY, RENATA (2000). Problemi di analisi e codifica di alcuni fenomeni fonetici del parlato di bambini ipoacusici. In *Atti del XXVIII Convegno Nazionale AIA*, Bari, Politecnico di Bari - Università di Bari, 261-264.

2001

SAVY, RENATA (2001). L'interfaccia tra livelli di analisi del parlato: rapporti tra riduzioni segmentali e schemi prosodici. In *Dati Empirici e Teorie Linguistiche*. Roma, Bulzoni, 309-328.

2002

VALLONE, MARIANNA, CANIPAROLI, VALENTINA, SAVY, RENATA (2002). Una prima indagine su fenomeni di riduzione consonantica nella varietà napoletana nel corpus AVIP. In *La fonetica acustica come strumento di analisi della variazione linguistica in Italia*, Roma: Il Calamo, 83-90.

MANFRELOTTI, OLGA, SAVY, RENATA, GOMEZ PALOMA, FRANCESCA (2002). Un procedimento di codifica fonologica basato sull'analisi segmentale del parlato di bambini ipoacusici. In *La fonetica acustica come strumento di analisi della variazione linguistica in Italia*, Roma: Il Calamo, 211-218.

GIORDANO, ROSA, SAVY, RENATA, CROCCO, CLAUDIA, PAONE, ESTER (2002). La realizzazione dell'accento in due tipologie pragmatiche: relazioni tra codifica fonetica e fonologica. In *La fonetica acustica come strumento di analisi della variazione linguistica in Italia*, Roma: Il Calamo, 145-152.

CROCCO, CLAUDIA, SAVY, RENATA, MANFRELOTTI, OLGA, GIORDANO, ROSA, PAONE, ESTER, CANIPAROLI, VALENTINA (2002). La relazione tra sintassi e prosodia in due tipologie pragmatiche tratte dal corpus AVIP: alcuni dati preliminari sulla varietà napoletana. In *La fonetica acustica come strumento di analisi della variazione linguistica in Italia*. Roma: Il Calamo, 137-144.

2003

GIORDANO, ROSA, SAVY, RENATA (2003). The Intonation of instruct and explain in Neapolitan Italian. In *Proceedings of the 15th International Congress of Phonetic Sciences*, Barcelona, 603-606.

CROCCO, CLAUDIA, SAVY, RENATA, CUTUGNO, FRANCESCO (a cura di) (2003). *API: Archivio del Parlato Italiano*. Università degli Studi di Napoli Federico "II".

CROCCO, CLAUDIA, SAVY, RENATA (2003). Fenomeni di esitazione e dintorni: una rassegna bibliografica. In *API-Archivio del Parlato Italiano*, Università degli Studi di Napoli Federico "II".

2004

SAVY, RENATA, CROCCO, CLAUDIA (2004). Aspetti teorico-pratici di una codifica multilivello di corpora vocali. In *Costituzione, gestione e restauro di corpora vocali*, Roma: Esagrafica, 139-144.

LO PREJATO, MANUELA, CLEMENTE, GIULIANA, SAVY, RENATA (2004). Su alcuni aspetti della riduzione vocalica nella varietà napoletana. In *Costituzione, gestione e restauro di corpora vocali*, Roma: Esagrafica, 183-188.

ALBANO LEONI, FEDERICO, CUTUGNO, FRANCESCO, PETTORINO, MASSIMO, SAVY, RENATA (a cura di) (2004). *Il Parlato Italiano*. Atti del Convegno Nazionale Il Parlato Italiano. Napoli: D'Auria Editore.

2005

SAVY, RENATA, LO PREJATO, MANUELA, CLEMENTE, GIULIANA (2005). Per una caratterizzazione e una misura della riduzione vocalica in italiano. In *Misura dei parametri. Aspetti tecnologici ed implicazioni nei modelli linguistici*, Padova: EDK Editore, 135-160.

SAVY, RENATA, CROCCO, CLAUDIA, GIORDANO, ROSA (2005). Geminate e geminazioni tra codifica fonologica e codifica fonetica: esempi dal corpus AVIP. In *Tradizione & Innovazione. Linguistica e filologia italiana alle soglie di un nuovo millennio*, Firenze: Franco Cesati Editore, 179-197.

2006

SAVY, RENATA, CROCCO, CLAUDIA (a cura di) (2006). *Analisi prosodica: teorie, modelli e sistemi di annotazione*, Padova: EDK Editore.

SAVY, RENATA (2006). Fonetica. *Il linguaggio: strutture linguistiche e processi cognitivi*, Roma-Bari: Laterza, 3-26.

SAVY, RENATA, CUTUGNO, FRANCESCO, CROCCO, CLAUDIA (2006). Multilevel corpus analysis: generating and querying an AGset of spoken Italian (SpIt-MDb). In *Proceedings of 5th International Conference on Language Resources and Evaluation*. PARIS, European Language Resources Association (ELRA), 1654-1659.

SAVY, RENATA (2006). Specifiche per la trascrizione ortografica annotata dei testi raccolti. In Albano Leoni F., Giordano R. (a cura di), *Italiano Parlato. Analisi di un dialogo*, Napoli, Liguori, 1-37.

SAVY, RENATA (2006). Specifiche per l'etichettatura dei livelli segmentali. In Albano Leoni F., Giordano R. (a cura di), *Italiano Parlato. Analisi di un dialogo*, Napoli, Liguori, 38-73.

CLEMENTE, GIULIANA, SAVY, RENATA, CALAMAI, SILVIA (2006). Sistemi vocalici in diatopia. In *Analisi prosodica: teorie, modelli e sistemi di annotazione*, Padova-Rimini: EDK Editore, 439-460.

2006

SAVY, RENATA, DE LEO, SIMONA (2007). Specifiche per l'etichettatura pragmatica dei testi in Pra.Ti.D., 1-60.

CROCCO, CLAUDIA, SAVY, RENATA (2007). Topic in dialogue: prosodic and syntactic features. In *ISCA Proceedings of Interspeech 2007*, ANTWERP ISCA, 114-117.

ALFANO, IOLANDA, LLISTERRI, JOAQUIM, SAVY, RENATA (2007). The Perception of Italian and Spanish Lexical Stress: A first cross-linguistic study. In *Proceedings of the 16th International Congress of Phonetic Sciences*, 1793-1796.

2008

PETTORINO, MASSIMO, GIANNINI, ANTONELLA, VALLONE MARIANNA, SAVY, RENATA (a cura di) (2008). *La Comunicazione Parlata*. Atti del Convegno Internazionale, Vol. 3, Napoli, Liguori.

SAVY, RENATA, SOLÍS GARCÍA, INMACULADA (2008). Strategie pragmatiche in italiano e spagnolo a confronto: una prima analisi su corpus. *Testi e Linguaggi*, 2, 214-242.

ALFANO, IOLANDA, SAVY, RENATA, LLISTERRI, JOAQUIM (2008). Las características acústicas y perceptivas del acento léxico en español y en italiano: los patrones acentuales paroxítonos. *Language design*, 9, 23-30.

2009

SAVY, RENATA, CUTUGNO, FRANCESCO (2009). CLIPS. Diatopic, diamesic and diaphasic variations in spoken Italian. In Michaela Mahlberg, Victorina González-Díaz, Catherine Smith, *Proceedings of Vth Corpus Linguistic Conference (CL2009)*, Liverpool, University of Liverpool, 1-24.

SAVY, RENATA, CASTAGNETO, MARINA (2009). Funzioni comunicative e categorie d'analisi pragmatica: dal testo dialogico allo schema xml e viceversa. In *Linguistica e modelli tecnologici di ricerca*, Roma: Bulzoni, 569-579.

ALFANO, IOLANDA, SAVY, RENATA, LLISTERRI, JOAQUIM (2009). Sulla realtà acustica dell'accento lessicale in italiano ed in spagnolo: la durata vocalica in produzione e percezione. In *La fonetica sperimentale. Metodo e applicazioni*, Rimini: EDK Editore, 22-39.

2010

SAVY, RENATA, VOGHERA, MIRIAM (2010). A corpus-based study on syntactic and phonetic prosodic phrasing boundaries in spontaneous Italian speech. In *ISCA Speech Prosody 2010*, Chicago, SpeechProsody2010.Illinois.edu.

SAVY, RENATA, ALFANO, IOLANDA (2010). Requests in dialogic speech: a prosodic analysis on Italian and Spanish task-oriented dialogues. In *Parlare con le persone, parlare alle macchine: la dimensione interazionale della comunicazione verbale*, Rimini: EDK-Editore, 31-52.

SAVY, RENATA (2010). Pr.A.T.I.D: a coding scheme for pragmatic annotation of dialogues. In *ELRA/ELDA Proceedings of the 7th Conference on International Linguistic Resources and Evaluation*. Paris, ELRA/ELDA, 2141-2148.

CUTUGNO, FRANCESCO, MATURI, PIETRO, SAVY, RENATA, ABETE, GIOVANNI, ALFANO, IOLANDA (a cura di) (2010). *Parlare con le persone, parlare alle macchine: la dimensione interazionale della comunicazione verbale*, Rimini: EDK-Editore.

2011

ORIGLIA, ANTONIO, ABETE, GIOVANNI, CUTUGNO, FRANCESCO, ALFANO, IOLANDA, SAVY, RENATA, LUDUSAN, BOGDAN (2011). A divide et impera algorithm for optimal pitch stylization. In *International Speech Communication Association Speech Science and Technology for Real Life*, ISCA, Firenze, 1993-1996.

IACOPONI, LUCA, SAVY, RENATA (2011). Sylli: Automatic Phonological Syllabification for Italian. In *International Speech Communication Association Speech Science and Technology for Real Life*, ISCA, Firenze, 641-644.

ALFANO, IOLANDA, SCHWAB, SANDRA, SAVY, RENATA, LLISTERRI, JOAQUIM (2011). Cross-language speech perception: lexical stress in Spanish with Italian and Francophone subjects. In *La dimensione temporale del parlato*, Rimini: EDK-Editore, 455-473.

ALFANO, IOLANDA, SCHWAB, SANDRA, LLISTERRI, JOAQUIM, SAVY, RENATA (2011). La percepción del acento léxico en español como lengua extranjera por parte de hablantes nativos de italiano y de francés. In Antonio Hidalgo Navarro, Yolanda Congosto Martín, Mercedes Quilis Merín, *El estudio de la prosodia en España en el siglo XXI: perspectivas y ámbitos*, Valencia, Universitat de València, 249-266.

2012

SOLÍS GARCÍA, INMACULADA, SAVY, RENATA (2012). Diferentes estrategias comunicativas en diálogos Task-oriented españoles e italianos. In *Il dialogo. Lingue, letterature, linguaggi, culture*, Roma: AISPI Edizioni, 443-457.

SCHWAB, SANDRA, ALFANO, IOLANDA, SAVY, RENATA, LLISTERRI, JOAQUIM (2012). La percepción del acento léxico en una lengua extranjera. In *Perception phonique et parole. Percepción fónica, habla y hablar*. Mons, Centre International de Phonétique Appliquée, 279-296.

ROMITO, LUCIANO, SAVY, RENATA, TARASI, ANDREA, LIO, ROSITA (2012). Analisi percettiva, musicale e automatica dell'italiano L1 e L2. In *La voce nelle applicazioni*, Roma: Bulzoni Editore, 93-106.

GIORDANO, ROSA, SAVY, RENATA (2012). Sulla standardizzazione del consonantismo dell'italiano: consonanti geminate, rafforzate e fricative alveolari in contesto intervocalico. In *La variazione nell'italiano e nella sua storia Varietà e varianti linguistiche e testuali, Atti dell'XI Congresso SILFI Società Internazionale di Linguistica e Filologia Italiana*, Firenze: Franco Cesati, 2, 431-445.

ALFANO, IOLANDA, SAVY, RENATA (2012). Los estilos conversacionales en la interacción dialógica: un análisis de las peticiones en italiano y en español. *Oralia*, 15, 35-62.

ALFANO, IOLANDA, CROCCO, CLAUDIA, SAVY, RENATA (2012). Requesting in Italian as a foreign language: pragmatic and prosodic features. In *Prosodic and Rhythmic Aspects of L2 Acquisition: The Case of Italian*, Newcastle upon Tyne: Cambridge Scholars Publishing, 153-175.

2013

ALFANO, IOLANDA, SAVY, RENATA (2013). Los patrones entonativos como marcas de frontera: un análisis de algunos tipos de peticiones en italiano y en español. In *Frontiere: soglie e interazioni i linguaggi ispanici n ella tradizione e nella contemporaneità*, Trento, Università degli Studi di Trento, 7-26.

2014

VOGHERA, MIRIAM, IACOBINI, CLAUDIO, SAVY, RENATA, CUTUGNO, FRANCESCO, DE ROSA, AURELIO, ALFANO, IOLANDA (2014). VoLIP: a searchable Italian spoken corpus. In *Complex Visible Out There (Proceedings of the Olomouc Linguistics Colloquium 2014, Language Use and Linguistic Structure)*, 627-640.

PALMERINI, MARIA, SAVY, RENATA (2014). Gli errori di un sistema di riconoscimento automatico del parlato. Analisi linguistica e primi risultati di un progetto di ricerca interdisciplinare. In *Proceedings of the First Italian Conference on Computational Linguistics CLiC-it*, Pisa, Pisa: University Press, 1, 281-285.

ALFANO, IOLANDA, CUTUGNO, FRANCESCO, DE ROSA, AURELIO, IACOBINI, CLAUDIO, SAVY, RENATA, VOGHERA, MIRIAM (2014). VOLIP: a Corpus of Spoken Italian and a Virtuous Example of Reuse of Linguistic Resources. In *Proceedings of the Ninth International Conference on Language Resources and Evaluation (LREC'14) Paris European Language Resources Association (ELRA)*, 3897-3901.

ALFANO, IOLANDA, SAVY, RENATA (2014). La strategia del Questioning nell'interazione dialogica: verso una definizione di continua pragmatico-funzionali. *Lingue e Linguaggi*, 12, 7-22.

ALFANO, IOLANDA, SAVY, RENATA (2014). Pragmática y prosodia de las preguntas pronominales en italiano y en español. *DIÁLOGO DE LA LENGUA*, VI, 1-34.

2015

SAVY, RENATA, LUQUE MOYA, JOSE ANTONIO (2015). Aspectos prosódicos de las interrogativas en aprendientes italianos de ELE. *Normas*, 7, 283-296.

2016

VOGHERA, MARIA, IACOBINI, CLAUDIO, CUTUGNO, FRANCESCO, SAVY, RENATA, ALFANO, IOLANDA, DE ROSA, AURELIO (2016). Il VoLIP: una risorsa per lo studio della variazione nel parlato della lingua italiana. In *Actes du XXVIIe Congrès international de linguistique et de philologie romanes Strasbourg ELiPhi*, 2, 1651-1663.

SAVY, RENATA, ALFANO, IOLANDA (a cura di) (2016). *La fonetica nell'apprendimento delle lingue*, Milano: Officinaventuno, 2, 1-416.

SAVY, RENATA, ALFANO, IOLANDA (2016). La "richiesta di informazione" in prospettiva intra- e inter-linguistica: aspetti pragmatici e prosodici. In *Livelli di analisi e fenomeni di interfaccia. Atti del xlvii congresso internazionale di studi della società di linguistica italiana (SLI)*, Roma: Bulzoni, 205-230.

ORRICO, RICCARDO, CATALDO, VIOLETTA, SAVY, RENATA, BARONE, LINDA (2016). Transfer, Fossilization and Prosodic Drift in Foreign Language Learning. In *La fonetica nell'apprendimento delle lingue*, Milano: Officinaventuno, 117-132.

2017

SAVY, RENATA, ALFANO, IOLANDA, ORRICO, RICCARDO (2017). La prosodia del tópicó en italiano, español e inglés como lenguas extranjeras. In *Tendencias actuales en fonética experimental. Cruce de disciplinas en el centenario del Manual de Pronunciación Española (Tomás Navarro Tomás)*, Madrid: UNED-CSIC, 156-157.

LUQUE MOYA, JOSE ANTONIO, SAVY, RENATA (2017). Un método para la enseñanza asistida de la prosodia en aprendientes italoófonos de ELE. *Marco ELE*, 24, 192-217.

CATALDO, VIOLETTA, ORRICO, RICCARDO, SAVY, RENATA (2017). Phonetic variations of f0 range in L1 and L2: a comparison between Italian, English, and Spanish native and nonnative speakers. In *Fattori sociali e biologici nella variazione fonetica / Social and biological factors in speech variation*, Milano: Officinaventuno, 227-256.

ALFANO, IOLANDA, SAVY, RENATA (2017). La codificación prosódica: un análisis contrastivo de la 'request' en italiano y en español. *Marco ELE*, 24, 146-169.

2018

ORIGLIA, ANTONIO, SAVY, RENATA, POGGI, ISABELLA, CUTUGNO, FRANCESCO, ALFANO, IOLANDA, D'ERRICO, FRANCESCA, VINCZE, LAURA, CATALDO, VIOLETTA (2018). An Audiovisual Corpus of Guided Tours in Cultural Sites: Data Collection protocols in the CHROME Project. In *AVI-CH 2018 Advanced Visual Interfaces for Cultural Heritage in AVI '18 Proceedings of the 2018 International Conference on Advanced Visual Interfaces*, New York: ACM Publications, 1-4.

ALFANO, IOLANDA, SAVY, RENATA, SBRANNA, SIMONA, SCHETTINO, LOREDANA (2018). Strategie discorsive in spagnolo L1 ed L2 a confronto: un'indagine su corpora dialogici. *Chimera*, 5, 1, 27-57.

2019

SAVY, RENATA, ALFANO, IOLANDA (2019). Strategie pragmatiche di introduzione di un Topic Discorsivo in dialoghi *task-oriented*: lingue a confronto. In *Studi AIItLA 9: Lingua in contesto. La prospettiva pragmatica*, Milano: Officinaventuno, 109-127.

SAVY, RENATA, ALFANO, IOLANDA (2019). La estructura conversacional en los intercambios dialógicos: un análisis de las estrategias discursivas en el desarrollo textual. In *Pragmática del español hablado. Hacia nuevos horizontes*. Valencia, Universitat de València, 85-98.

ORRICO, RICCARDO, SAVY, RENATA, D'IMPERIO, MARIAPAOLA (2019). Salerno Italian: Intonational phonology and dimensions of variation. In Duccio Piccardi, Fabio Ardolino, Silvia Calamai (a cura di) *Gli archivi sonori al crocevia tra scienze fonetiche, informatica umanistica e patrimonio digitale || Audio archives at the crossroads of speech sciences, digital humanities and digital heritage*, Milano: Officinaventuno, 309-328.

ORRICO, RICCARDO, SAVY, RENATA, D'IMPERIO, MARIAPAOLA (2019). The perception of speaker certainty in Salerno Italian intonation. *Proceedings of the 19th International Congress of Phonetic Sciences*, Melbourne, Australia, 2946-2950.

ORIGLIA, ANTONIO, SAVY, RENATA, CATALDO, VIOLETTA, SCETTINO, LOREDANA, ANSANI, ALESSANDRO, SESSA, ISORA, CHIERA, ALESSANDRA, POGGI, ISABELLA (2019). Human, All Too Human: Towards a Disfluent Virtual Tourist Guide. In *ACM UMAP '19 - Adjunct Publication of the 27th Conference on User Modeling, Adaptation and Personalization*, New York, NY, USA ACM Order Department, 393-399.

CATALDO, VIOLETTA, SCETTINO, LOREDANA, SAVY, RENATA, POGGI, ISABELLA, ORIGLIA, ANTONIO, ANSANI, ALESSANDRO, SESSA, ISORA, CHIERA, ALESSANDRA (2019). Phonetic and functional features of pauses, and concurrent gestures, in tourist guides' speech. In Duccio Piccardi, Fabio Ardolino, Silvia Calamai (a cura di) *Gli archivi sonori al crocevia tra scienze fonetiche, informatica umanistica e patrimonio digitale / Audio archives at the crossroads of speech sciences, digital humanities and digital heritage*, Milano: Officinaventuno, 205-231.

2020

ALFANO, IOLANDA, SAVY, RENATA, SCETTINO, LOREDANA (2020). Strategie pragmatiche di gestione di un *Discourse Topic* in dialoghi *task-oriented*. Una prospettiva interlinguistica. In *La comunicazione parlata / Spoken Communication 2018*, Roma: Aracne, 43-62.

PARTE I

LA VARIAZIONE FONETICA
IN CONTESTI DIALETTALI

PIA GRECA, JONATHAN HARRINGTON

An acoustic analysis of metaphony in the dialects of the Lausberg Area (Southern Italy)

This acoustic analysis focuses on metaphony - a type of regressive vowel harmony - in a restricted dialectal area in Southern Italy, across the border between Basilicata and Calabria, the so-called “Lausberg Area” (Lausberg, 1939). The first aim is to relate synchronically metaphony to V_1CV_2 coarticulation, in which V_2 influences V_1 and can lead to associated sound changes. Secondly, we want to determine if metaphony varies between younger and older speakers and between males and females. Overall, the results showed an influence mainly of suffix vowel height on mid stem vowels in both age groups. The generally weaker metaphonic influence in younger speakers – especially in young men – suggests their tendency to converge to Standard Italian as regards stem vowel quality.

Keywords: dialects, metaphony, coarticulation, sociophonetics, sound change.

1. Introduction

The present study is a phonetic analysis of metaphony - a type of regressive vowel harmony typical of many Romance languages - in a restricted dialectal area in Southern Italy, across the border between Basilicata and Calabria, called the Lausberg area (Lausberg, 1939) after the author who first described it. In particular, the dialects in this area have undergone far fewer phonological changes from Latin compared to other Southern Italian dialects. The vowel systems in the Lausberg area are particularly interesting, since we find the coexistence, and very often also mixing, of more than one vowel system in a relatively restricted territory. Basically, most dialects in the Lausberg area share the so-called Sardinian vowel system, keeping the Latin vowel qualities but neutralising vowel length. In particular, the *Mittelzone* in the east is thought to be the sub-area characterized by the uniform spreading of the Sardinian vowel system (Lausberg, 1939; Martino, 1991). Along the transition area between the Lausberg area and the *Südzone* (as defined by Lausberg, 1939) we often find traces of the Sicilian vowel system, merging the Latin /i/, /i:/ and /e:/ into short /i/, and /u/ and /o:/ into short /u/ (Martino, 1991: 46). There is a more complex area in the north-west, the so-called *Zwischenzone*, where a mixed vocalic system occurs, taking elements of the Sardinian and the Sicilian vowel system. In this sub-area, we can distinguish two phonologically contrasting degrees of aperture for mid vowels: /e/-/ɛ/ and /o/-/ɔ/ (as in the Neapolitan vowel chart).

Figure 1 - *Geographical localisation of the Lausberg area (between the two borders highlighted in red) and of its sub-areas (adapted from Pellegrini, 1977 and Martino, 1991). The Zwischenzone corresponds to the striped area on the left, while the remaining area left blank represents the Mittelzone. The area just below the southern red border represents the transition territory to the Südzone (below the black southern border)*



The map of the Lausberg area has been reviewed more than once (Rensch, 1964; Falcone, 1976; Trumper, Maddalon, 1988; Martino, 1991), but we lack recent studies completely mapping the status of the dialects in this area today. Literature on the dialects of the Lausberg area is unfortunately limited. Some systematic studies, apart from the very first ones that first detected and described this linguistic area (Lausberg, 1939; Rensch, 1964), were carried out by Trumper (1979, 1997), Martino (1991), Romito (Romito et al., 2006), Conte (as concerns the Basilicata territory, 2014).

Metaphony is pervasive in most Romance and Italian dialects, especially southern ones, and is triggered by a still existing or previously existing high vowel in the unstressed suffix, influencing the quality of the stressed root vowel, or less typically also the pre-tonic vowel in trisyllabic words. These suffixes may indicate a variety of morpho-syntactic categories, mainly gender and number in nouns and person, and number and tense in verbs. The numerous studies on metaphony and associated sound changes for Italian dialects (Lausberg, 1939; Rensch, 1964; Rohlf, 1966; and many others) have been mostly auditorily based without any quantitative analysis (more recent studies are also Maiden 1991; Gaglia, 2011; Savoia, 2015; the volume on metaphony by Torres-Tamarit et al., 2016). Only very recently, some first acoustic analyses of metaphony in Southern Italian dialects have also appeared (see Grimaldi, 2003; Grimaldi et al., 2016; Grimaldi, Calabrese, 2018; Romito et al., 2006; Romito, Gagliardi, 2009; Romito et al., 2011).

Metaphony can result not only into vowel shifting (typically raising), but also diphthongization. In particular:

- Vowel raising usually affects mid-low vowels, less frequently mid-high ones and more rarely the low central vowel: $e \rightarrow e/_i$, u ; $o \rightarrow o/_i$, u ; $e \rightarrow i/_i$, u ; $o \rightarrow u/_i$, u ; $a \rightarrow \epsilon/_i$, u .

(1) Mormanno, Calabria: [ˈvɛcca] – [ˈvɛccu] (‘old woman, old man’) (Savoia, 2015: 209)

(2) Papisidero, Calabria: [ˈmɔru] – [ˈmurisi] (‘I die, you die’) (Trumper, 1997: 362)

- Diphthongisation usually involves mid-low and more rarely low vowels: /a, ɛ/ → /je, je, i:ə/; /ɔ/ → /wɔ, wo, u:o/. Diphthongs are typical of – but not exclusive to – the *Zwischenzone* (though not shared by all varieties) and in the northern part of the *Mittelzone*.

(3) Trebisacce, Calabria: [ˈbellə] – [ˈbiəllə] (‘beautiful’, fem. sg. vs. masc. sg.)

(IPA-adapted from Rensch, 1964: 25)

(4) Papisidero, Calabria: [ˈlɔrda] – [ˈluordu] (‘dirty’, fem.sg. vs masc.sg.)

(IPA-adapted from Rensch, 1964: 46)

Curiously, however, diphthongising areas seem not to be geographically compact (Martino, 1991: 45). As far as the southern part of the *Mittelzone* is concerned (broadly coinciding with the High Ionic Calabrian coast), monophthongisation follows diphthongisation historically (Martino, 1991: 14; Trumper, 1997: 361) and is typically also accompanied by compensatory lengthening: /je, ie/ → /i:/, /wo/ → /u:/. Consequently, the following changes are possible within the relatively geographically restricted Lausberg area.

(5) *[ˈbellu] > *[ˈbiellu] > [ˈbi:llə] (‘beautiful’, masc. sing.) (Romito et al., 2006: 4)

(6) *[ˈkɔktu] > *[ˈkuottu] > [ˈku:ttu] (‘cooked’, masc. sing.) (Trumper, 1997: 361)

Metaphonic effects have been shown to be confined principally to the influence of vowel height on the mid vowels /e, o/ (Lausberg, 1939; Rensch, 1964; Rohlfs, 1966), which in the Sardinian vowel system can be also phonetically realised as [ɛ, ɔ]. We expect most metaphonic effects in mid vowels not only because this is suggested by the literature (Lausberg, 1939; Rensch, 1964; Rohlfs, 1966), but also because in the Sardinian vowel system the opposition between tonic mid-high and mid-low vowels – in non-metaphonic contexts – is not phonologically relevant, but rather conditioned by the phonological environment. More specifically, a mid target stem vowel is probably more open in words where the trigger is /e, a/, but should get higher to a certain degree (thus undergoing metaphony) when the trigger is high. Fronting or backing effects are expected principally when both trigger (i.e. the suffix) and the target (i.e. the metaphonising stem vowel) are either both front or back (that is, we do not expect a categorical shift in frontness or backness, but basically only in vowel height, see also Maiden, 1991; Savoia, 2015).

It is currently unclear whether metaphony in the Lausberg area still exists or whether it is waning under the influence of Standard Italian in which there are no categorical metaphonic effects. For this purpose, an apparent-time analysis (Bailey, Wikle, Tillery & Sand, 1991; Weinreich, Labov, Herzog, 1968) was carried out, in which younger and older speakers were compared acoustically in order to assess the

metaphonic influence of suffix vowel on stem vowels. Given that certain kinds of sound changes in progress are known to be influenced by sex and sometimes led by women (e.g. Eckert, 1989; Labov, 1990; Maclagan et al., 1999), coupled with some suggestions that women might be less inclined to make use of broad dialect features than men in the Lausberg area (Trumper, 1979), a further test was made on whether any such weakening of metaphony towards the standard was different for women and for men.

In summary, the main aim of the study was to test whether metaphony occurs in the Lausberg area. Assuming metaphony does occur, the further questions that were considered were the following:

1. Is the metaphonic influence of the suffix vowel strongest for mid (as opposed to high or open) vowels?
2. Is metaphony principally due to a change in height rather than fronting?
3. Are metaphonic effects weaker for younger than older speaker and, if so, does this weakening interact with speaker sex?

2. Method

2.1 Speakers and villages

34 participants (18 females) from 10 villages in the Lausberg area were recorded in quiet conditions at their homes. The speakers were recruited either from personal contacts of the first author (who is herself a native speaker of this region), or through contacts given by previous participants, or by broadcasting the research activity on social media. All participants were paid a small amount of money for their participation. Before carrying out the recordings, all participants were invited to answer some questions related to their age, degree of education, and use of dialect in everyday life. Only participants who declared to be able to speak the local variety of the village they are from and to use the dialect at least relatively often were involved in the recordings. The speakers include 18 older (41-92 years) and 16 younger (13-32 years) speakers. Fig. 2 shows the villages and the number of speakers per village from which recordings were made. Unfortunately, Basilicata was underrepresented (Lauria), since most of the villages from which recordings were made are in Calabria, in the Province of Cosenza. Nevertheless, as Fig. 2 also shows, the whole Lausberg-Calabrian area has been broadly covered, taking both *Zwischenzone* and *Mittelzone* into account and also including some villages at the border between the Lausberg area and the transition zone to the *Südzone* (cf. § 1). The further details are as follows: Laino Borgo and Laino Castello, Mormanno and S. Domenica Talao belong to the *Zwischenzone*; Lauria is on the northern border between *Zwischenzone* and *Mittelzone*; Canna, Montegiordano and Cerchiara belong to the *Mittelzone*, while Castrovillari and Schiavonea are both at the southern border of the Lausberg area (Trumper, Maddalon, 1988). A further summary of some of the speaker attributes is shown in Table 1 in the *Appendix*, including village of origin, sex and age.

Figure 2 - The geographical distribution of the villages represented in the data. The border between the two regions Basilicata and Calabria is highlighted in dark blue. The villages belonging to the *Zwischenzone* are circled in red, those belonging to the *Mittelzone* are circled in dark grey, while those on the borders of the area are circled in blue. The number of speakers for each village is also indicated. URL: maps.google.com



While most of the dialects belonging to the Lausberg area should share the Sardinian vowel system (see references in § 1), Castrovillari and Schiavonea might indeed show a greater influence of the Sicilian vowel system (Martino, 1991: 12). However, the vowels of the *Zwischenzone* might also be influenced by the *Südzone*, characterized by the Sicilian vowel system, at least in some lexical items (Lausberg, 1939; Martino, 1991). Finally, the exact borders of the Lausberg area are not easy to trace, especially considering that these might have changed in the many decades since the first studies were undertaken. For instance, a closer contact between villages, possibly due to slightly improved infrastructure and mobility linked to study and work, might be contributing to dialect levelling (Trumper, 1979; Trumper, Maddalon, 1988). These are some of the reasons why recordings were made from a wide selection of villages in each zone.

2.2 Lexical items and vowel tokens

The lexical items in this study were elicited through a picture-naming task in order to avoid the use of Standard Italian and to encourage the participants to talk directly in the dialect. Also, all interactions between the investigator and each participant were carried out in the dialect, as far as this was possible. Each speaker produced a total of 102 inflected words (51 stem types) including inflected nouns, adjectives and verbs, in randomised order and in two repetitions. In order to elicit inflected lexical items just by using pictures, a slightly different strategy was adopted for each lexical category, as visible in Figs. 7 to 9 in the *Appendix*. In particular, while nouns could be elicited on their own, inflected adjectives had to be elicited in combination with a noun (e.g. [ˈpuma ˈrus:a], ‘red apple’, where [ˈrus:a] is the target word, see Fig. 8 in the *Appendix*) and inflected verbs had to be elicited within sentences which

were graphically “suggested” by the picture and thus different for each verb (see Fig. 9 in the *Appendix*). In order to make sure that the participants had understood the task, a training phase, consisting in observing the pictures and clarifying their meaning in case of ambiguities, preceded the recording phase. The speakers repeated each word twice, first in isolation and then embedded in a carrier sentence (“I say __ two times”, in the dialect [jɛ 'diku __ dui 'votə] as regards nouns and adjectives, while verbs could not be elicited in isolation, so that the same sentence containing the target verb had to be repeated twice).

The mostly bisyllabic words were composed of a stem and either front vowel suffixes (henceforth: V_{front} = /i, e/, e.g. /'kani/ - /'kane/, ‘dogs’ - ‘dog’), or back vowel suffixes (V_{back} = /u, a/, e.g. /'vec:u/ - /'vec:a/, ‘old man’ - ‘old woman’). 25 and 26 stem types preceded V_{front} and V_{back} suffixes respectively. The stem vowel (V_{stem}) varied over /i, e, a, o, u/ (e.g. in /kani/ - /kane/, ‘dogs’ - dog’, the V_{stem} is /a/). Following the removal of those words that had been misarticulated or produced in Standard Italian, 2444 stem vowels preceding V_{front} suffixes and 2535 stem vowels preceding V_{back} suffixes remained for analysis.

The lexical items considered in the analysis are listed in Italian in Table 2 in the *Appendix*. The expected pronunciation of the items in the dialect is also indicated. As anticipated, these words that vary in stem and suffix vowels include all lexical categories that could be affected by metaphony, i.e. nouns, adjectives and inflected verbs (first, second and third person singular of the present indicative). In particular, the high vowel suffixes /i, u/ are the ones expected to trigger metaphony (§ 1). Table 2 also shows that the lexical items are organised into pairs, where the stem is shared and there are two competing suffixes. From a morphological point of view, the /i/ marks either (a) the plural counterpart for nouns with the/e/suffix in the singular (e.g. /'verme, 'vermi/, ‘worms’ - ‘worm’), or (b) the second person singular for some verbs vs. the third person singular ending with /e/, e.g. /'tenisi, 'tene/, ‘(I) have’ - ‘(you) have’. The suffix /u/ marks in verbs the first person singular vs. the third person singular, e.g. /'trovu, 'trova/, or it represents in nouns and adjectives gender, e.g. /'bona, 'bonu/ (‘good’, feminine vs masculine).

2.3 Instrumentation and software

The speakers were recorded at their homes using a laptop and a headset with integrated microphone (*Sennheiser SC 60*). The words were elicited in a picture-naming task using *SpeechRecorder 3.28.0* (Draxler, Jänsch, 2004). The raw speech data were then semi-automatically segmented and labelled using MAuS (*Munich Automatic Segmentation System*) (Kisler et al., 2017), which is integrated in the *emuR* package (version 1.1.2) (Winkelmann et al., 2017) in the R software environment (version 3.5.3). Even though the segmentation process used for the data was based on the phonological set for Italian, a phonological transcription in the dialect for each word of the dataset was given to the system. This is why the segmentation of dialectal data could be possible even by using a language-dependent segmentation system. Any obviously misplaced segment boundaries were manually corrected.

2.4 Formant analysis

The first two formant frequencies were calculated with the *Praat* formant tracker (using the *PraatR* package in R) (Albin, 2014) between the acoustic onset and offset of the stem vowel using a 25 ms window and a 5 ms frame shift. Visibly mis-tracked formants were manually corrected. They were then linearly time normalised into 11 equidistant time points. The strength of metaphonic effects was assessed separately in the context of front vowel (V_{front}) and back vowel suffixes (V_{back}) and separately for F1 and F2.

2.5 Statistical analysis

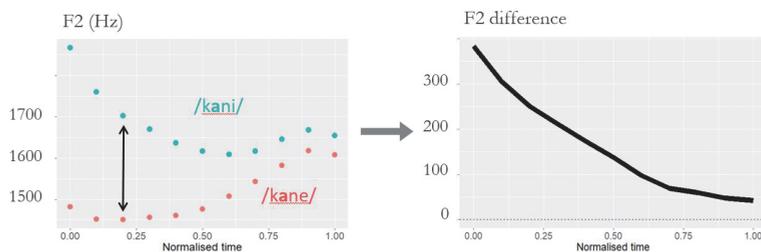
Linear mixed effects (LME) models were applied to the data by using the *lmerTest* package (version 3.1-0) on R. The dependent variable was the formant value extracted at time-normalised point 0.1, with fixed factors stem (five levels: one of the stem vowels /i, e, a, o, u/), suffix (two levels: one of the suffix vowels /i, u/), age group (two levels: young, old), sex, and with the speaker and stem (e.g. ‘mes’ for ‘mese/mesi’) as the random factors. The random factors included intercepts and all possible slopes to measure the interaction between the fixed and random factors; these were dropped if they were non-significant. Four mixed models were applied: one for each of the two formants separately, and one for each of the front, /i, e/, and back, /u, a/, suffix pairs. The motivation for basing the dependent variable on time-normalised point 0.1 was that this was the time point in which metaphonic effects were most marked (see § 3.1). After applying the LME models to the data, estimated marginal means (EMMs) post-hoc tests between different factor combinations were computed by using the *emmeans* package (version 1.4.5) in R.

3. Results

3.1 Does metaphony occur in the Lausberg area?

An initial analysis of metaphony was made by subtracting separately by speaker and stem the formant values at each time point in the non-metaphonic context from those in the metaphonic context. For example, the mean F2-trajectory in the stem vowel of /'kane/ (non-metaphonising) was subtracted from the mean F2 of the stem vowel /'kani/ (metaphonising) for a given speaker (Fig. 3). If the result of the subtraction is zero, then the suffix /i, e/ has no influence on the target, i.e. there is no metaphony.

Figure 3 - *An example of application of formant difference plotting of F2 for a particular stem and speaker*



Formant difference plots (Fig. 4 and Fig. 5) were obtained by subtracting formants in the stem vowel in the context of the two front or back vowel suffixes separately for each speaker, and then by grouping the differences according to target type and aggregating them across speakers and items. The plots for V_{front} (Fig. 4) suffixes were separated from those with V_{back} suffixes (Fig. 5). Thus, the upper plots in Fig. 4 are based on subtracting F2 of e.g. /a/ in /'kane/ from F2 of /a/ in /'kani/; and the upper plots in Fig. 5 subtracting F2 of e.g. /e/ in /'vec:a/ from F2 of /e/ in /'vec:u/. The plots in the lower rows are based on similar calculations but for F1. If the suffix influences the stem vowel, then the F2 difference plot should be positive regarding a stem in the context of a V_{front} suffix (given that $F2 /i/ > F2 /e/$) (Fig. 4) and negative in a V_{back} context (since $F2 /u/ < F2 /a/$) (Fig. 5). Similarly, the F1 difference plot should be negative for stems before V_{front} suffix ($F1 /i/ < F1 /e/$) (Fig. 4) and negative in a V_{back} context ($F1 /u/ < F1 /a/$) (Fig. 5). Trajectories further away from zero indicate a greater influence of the suffix on the stem vowel.

In summary, Figs. 4 and 5 provide clear evidence for metaphony. They also show that formant differences are more marked at the vowel onset.

Figure 4 - *F2 and F1 differences in V_{front} contexts for all stem types, aggregated across speakers and items, also including the confidence interval of the mean difference values plotted*

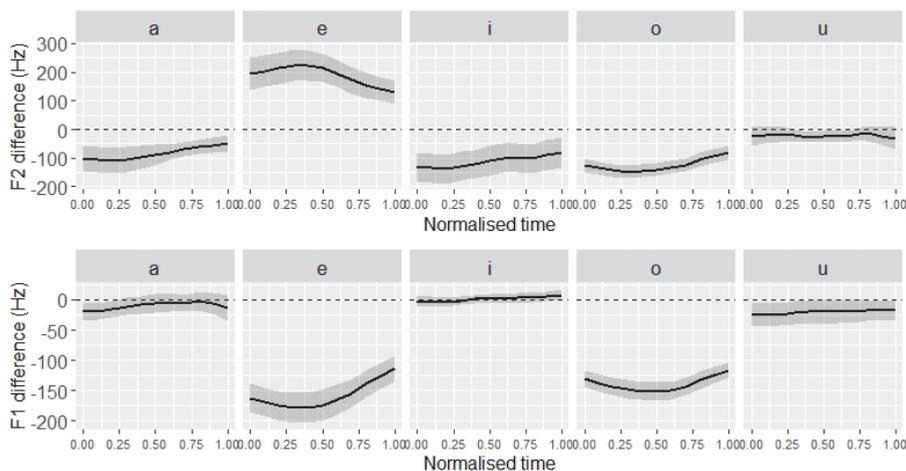
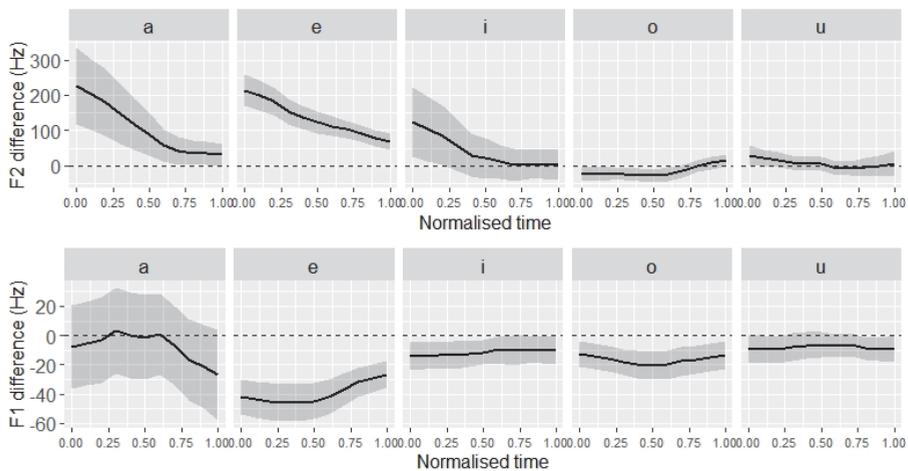


Figure 5 - F2 and F1 differences in V_{back} contexts for all stem types, aggregated across speakers and items, also including the confidence interval of the mean difference values plotted



3.2 Are metaphonic effects stronger in mid-vowels?

Figs. 4 and 5 clearly show that the metaphonic influences are greatest on mid vowel stems /e, o/. In addition, there is a weaker metaphonic influence on /a/ vowel stems, especially in F2. This suggests that /a/ might be slightly fronted due to the influence of the /i/ suffix. Also, a slight tensing of /i/ as regards V_{front} contexts (since F2 and F1 get more distant), and a small retraction as regards V_{back} (noticeable because of the F2 lowering), are visible from the figures.

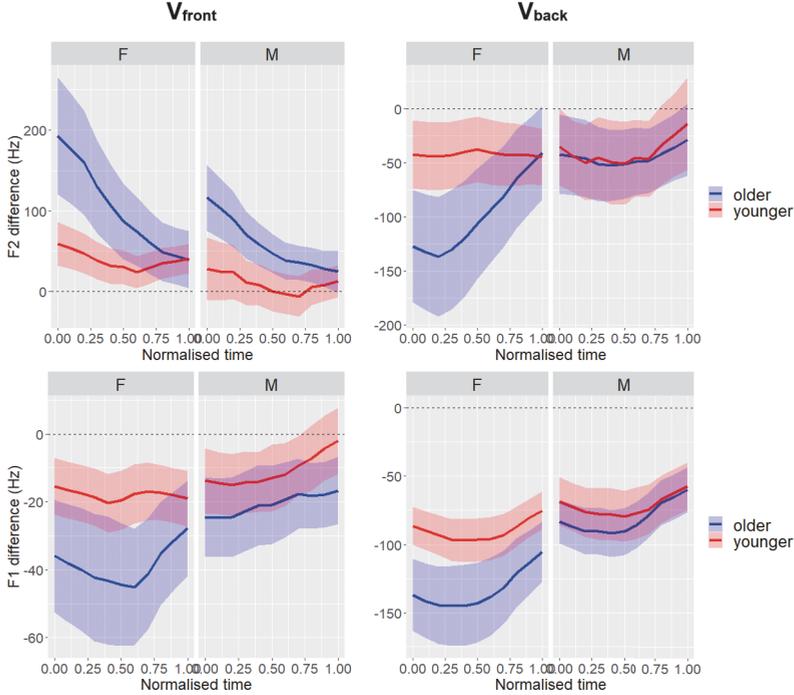
3.3 Does metaphony have a greater influence on height than fronting?

Figs. 4 and 5 show that the influence of the suffix vowel on the stem vowel was stronger in F1 than in F2, thereby also confirming that metaphony affects vowel height to a greater extent than vowel fronting. The most marked F2 effects were an F2 lowering of stem-/o/ in the context of suffix-/u/ and F2 raising of stem-/e/ in the context of suffix-/i/.

3.4 Are there age and sex influences on metaphony?

Following the reasoning in Fig. 3, if metaphony is present, then F2 difference plots (indicating fronting) should be positive for the $V_{front} = /i, e/$ suffixes and negative for the $V_{back} = /a, u/$ suffixes; and F1 difference plots (indicating raising) should be negative in both V_{front} and V_{back} contexts. Based on these difference plots, the results in Fig. 6 show that metaphony was stronger in older (blue lines) than in younger (red lines) speakers and that this age effect was more marked for females and especially for V_{front} suffixes.

Figure 6 - F2 and F1 differences for stems preceding $V_{front} = /i, e/$ and stems preceding $V_{back} = /u, a/$, grouped by age and sex and aggregated across items and speakers, also including the confidence interval of the mean difference values plotted. “F” stands for females, “M” for males, while age groups are distinguished by the blue (= old) and red (= young) colours respectively



3.5 Statistical analysis

Since there were in all cases four-way interactions between the fixed (independent) factors, the results are discussed in terms of post-hoc tests (cf. § 2.5).

Turning firstly to the mid-vowel targets /e, o/ where the major effects of metaphony were expected, the results – listed in detail in Table 3 in the *Appendix* – show the following. The /a, u/ suffix pair (V_{back}) had a significant influence on /e/ and /o/ stems in both F1 and F2 for all speaker (age, sex) groups with the exception of F2 in /o/ for younger men. The /e, i/ suffix pair (V_{front}) only influenced /e/ in both F1 and F2 in women and in older men.

The metaphonic effects for other vowel stems were generally not significant, with the following two exceptions: F2 in /i/ stems was significantly lower preceding an /u/ vs. /a/ suffix and to a greater extent in older ($p < .001$) than in younger women ($p < .05$); and F2 in /a/ stems was significantly ($p < .001$) raised preceding an /i/ vs. /e/ suffix in older women.

4. Discussion

The study has provided acoustic data from over 30 speakers to show that metaphony occurs in the dialects of the Lausberg area, as also attested in most Southern Italian varieties and to a certain extent also in Northern Italy (Rohlf, 1966; Savoia, Maiden, 1997; Grimaldi, 2003; Grimaldi, Calabrese, 2018; and Walker, 2005 and Delucchi, 2012 as regards some northern Italian dialects). Consistently with earlier studies (Martino, 1991; Romito et al., 2006; Savoia, 2015), the suffix vowel had the greatest influence on stem mid vowels /e, o/. There was also clear evidence that the influence of the suffix on the stem vowel was stronger in F1 than in F2, thereby also confirming that metaphony affects vowel height to a greater extent than vowel fronting. Nevertheless, F2 effects emerged visibly for /e/ stems in V_{front} contexts and /o/ stems in V_{back} contexts: /e/ was more front (as shown by a raised F2) in the metaphonising /i/suffix context, and /o/ was more back (as shown by a lowered F2) in the metaphonising /u/ suffix context.

This aspect makes metaphony different from other kinds of phonologised coarticulation processes, like e.g. Umlaut in German, where we systematically have vowel fronting (e.g. “grün” [grʏ:n] from Old High German “gruoni”). In the dialects of the Lausberg area, and consistently with most southern dialects having metaphony (Rohlf, 1966; Savoia, 2015), we have hardly any significant effect on frontness or backness of the stem vowel, when stem vowel and suffix are not both either back or front. At the same time, our data show some minor (but significant) fronting or backing effects, also when the stem and the suffix do not agree in [\pm back] or [\pm front]. This shows that the metaphonic phenomenon is actually coarticulatorily complex and cannot be reduced to a “simple” vowel raising.

The study also shows that there were some more minor influences of age on metaphony in men. More specifically, stem-/o/ was retracted due the following suffix-/u/ and stem-/e/ was raised and fronted due to the following suffix-/i/ in older but not younger men. By contrast, there were no such significant age differences in women. These results suggest a potentially greater influence of Standard Italian (and hence waning of the dialect) in men than in women. The reason why women continue to make greater use of dialect could be that they have less contact with speakers of Standard Italian. Trumper (1979) suggested that this difference in the use of dialect between men and women could come about because women tended not to have occupations and hence very limited contact with other villages or big social networks. Whether this remains the case today needs to be demonstrated; but irrespective of the cause, our results do point to an influence of Standard Italian on the dialects of the Lausberg area that seems to be more advanced in men than in women. Analysis is currently in progress on a larger number of speakers in order to investigate whether the findings demonstrated in this study vary by sub-region and by village.

Acknowledgments

We would like to thank our colleague Michele Gubian for his help in the statistical analysis. Also, a special thanks to all participants who offered their time and their voices to collect the data.

This study was funded by the European Research Council advanced grant N° 742289 ‘Human interaction and the evolution of spoken accent’ (2017–2022) and by the Graduate School of Language and Literature, Class of Language, of the Ludwig-Maximilians-University of Munich.

Bibliography

- ALBIN, A. (2014). PraatR: An architecture for controlling the phonetics software “Praat” with the R programming language. *Journal of the Acoustical Society of America*, 135(4), 2198.
- BAILEY, G., WIKLE, T., TILLERY, J., SAND, L. (1991). The apparent time construct. In *Language Variation and Change*, 3, 241–264.
- CONTE, M. G. (2014). *Dialetti in contatto nella Valle del Mercure: La variazione microdialettale e la sua percezione nell’Area Lausberg*. Frankfurt am Main: Peter Lang.
- DELUCCHI, R. (2012): *Fonetica e fonologia dell’armonia vocalica. Esiti di –A nei dialetti della Svizzera Italiana in prospettiva romanza*. Tübingen: A. Francke Verlag.
- DRAXLER, C., JÄNSCH, K. (2004). SpeechRecorder - a Universal Platform Independent Multi-Channel Audio Recording Software. In *Proc. of the IV. International Conference on Language Resources and Evaluation*, 559-562.
- ECKERT, P. (1989). The whole woman: sex and gender in variation. In *Language Variation and Change*, 1, 245-267.
- FALCONE, G. (1976): Calabria. In CNR, Centro di Studio per la dialettologia italiana, *Profilo dei dialetti italiani*. Pisa, Pacini.
- GRIMALDI, M, MIGLIETTA, S., SIGONA, S., CALABRESE, A. (2016). On integrating different methodologies in phonological research: acoustic, articulatory, behavioral and neurophysiological evidence in the study of a metaphonic system. In TORRES-TAMARIT, F., LINKE, K., VAN OOSTENDORP, M. (Eds.), *Approaches to Metaphony in the Languages of Italy*. Berlin/Boston: De Gruyter, 195-219.
- GAGLIA, S. (2011). Representational aspects of morphomic vowel variation in Southern Italy. In GOLDBACH, M., HINZELIN, M.C., MAIDEN, M., SMITH, J.C. (Eds.). *Morphological Autonomy: Perspectives from Romance Inflectional Morphology*. Oxford: Oxford University Press, 95-118
- GRIMALDI, M. (2003). *Nuove ricerche sul vocalismo tonico del Salento meridionale*. Alessandria: Ed. dell’Orso.
- GRIMALDI, M., CALABRESE, A. (2018). Metaphony in Southern Salento: New analysis and new data. In D’ALESSANDRO, R., PESCARINI, D. (Eds.), *Advances in Italian dialectology. Sketches of Italo-Romance grammars*. Leiden: Brill, 2018, 253-291.
- KISLER, T., REICHEL, U., SCHIEL, F. (2017). Multilingual processing of speech via web services. In *Computer Speech & Language* 45, 326-347.

- LABOV, W. (1990). The intersection of sex and social class in the course of linguistic change. In *Language Variation and Change*, 2, 205-254.
- LAUSBERG, H. (1939). *Die Mundarten Südlukaniens*. Halle: Niemeyer.
- MACLAGAN, M., GORDON, E., LEWIS, G. (1999). Women and sound change: conservative and innovative behavior by the same speakers. In *Language Variation and Change*, 11, 19-41.
- MAIDEN, M. (1991). *Interactive morphonology: Metaphony in Italy*. London/ New York: Routledge.
- SAVOIA, L., MAIDEN, M. (1997). Metaphony. In MAIDEN, M., PARRY, M. (Eds.): *The Dialects of Italy*. London/New York: Routledge, 15-25.
- MARTINO, P. (1991). *L' "area Lausberg". Isolamento e arcaicità*. Roma: Dipartimento di studi di glottoantropologici dell' Università di Roma "La Sapienza".
- PELEGRINI, G. B. (1977). *Carta dei dialetti d' Italia*. Pisa: Pacini.
- RENSCH, K. H. (1964). *Beitrage zur Kennmis der nordkalabrischer Mundarten*. Miinster: Aschendorff.
- ROHLFS, G. (1966.) *Grammatica storica della lingua italiana e dei suoi dialetti*, vol. I, *Fonetica*. Turin: Einaudi.
- ROMITO, L., TARASI, A., RENZELLI, M. (2011). Verso una ridefinizione dei confini linguistici dialettali: uno studio fonetico-fonologico in alcuni centri della Calabria. In GILI FIVELA, B., GARRAPA, S., GRIMALDI, M. (Eds.), *Contesto comunicativo e variabilità nella produzione e percezione della lingua. Atti del VII convegno nazionale AISV - Associazione Italiana Scienze della Voce (AISV)*, Università del Salento – Lecce, 26-28 gennaio 2011. Rome: Bulzoni, 37-47.
- ROMITO, L, GAGLIARDI, D. (2009). La metafonìa in alcuni centri del nord Calabria: verso una mappa regionale. In ROMITO, L., GALATÀ, V., LIO, R. (Eds.): *La fonetica sperimentale: metodo e applicazioni. Atti del IV convegno nazionale AISV - Associazione Italiana di Scienze della Voce (Università della Calabria, 3-5 dicembre 2007)*. Torriana: EDK, 423-437.
- ROMITO, L., GALATÀ, V., LIO, R., STILLO, F. (2006). La metafonìa dei dialetti dell'area Lausberg: un' introspezione sulla natura della sillaba. In SAVY, R., CROCCO, C. (Eds.): *Analisi prosodica - Teorie, modelli e sistemi di annotazione. Atti del II Convegno Nazionale AISV - Associazione Italiana di Scienze della Voce*, Fisciano (Salerno), 30 Novembre - 2 dicembre 2005. Torriana: EDK, 538-565.
- SAVOIA, L. (2015). *I dialetti italiani: sistemi e processi fonologici nelle varietà di area italiana e romancia*, Ospedaletto-Pisa, Pacini, 203-352.
- TORRES-TAMARIT, F., LINKE, K., VAN OOSTENDORP, M. (Eds.) (2016). *Approaches to Metaphony in the Languages of Italy*. Berlin/Boston: De Gruyter.
- TRUMPER, J, MADDALON, M. (1988). Converging Divergence and Diverging Convergence: The Dialect- Language Conflict and Contrasting Evolutionary Trends in Modern Italy. In DI LUZIO, A., AUER, P. (Eds.), *Variation and Convergence: Studies in Social Dialectology*. Berlin/Boston: De Gruyter, 1988, 242-252.
- TRUMPER, J. (1979). La Zona Lausberg ed il problema della frammentazione linguistica. In ALBANO LEONI, F. (Ed.), *I Dialetti e le minoranze di fronte all'italiano. Atti dell' XI Congresso internazionale di studi*. Cagliari, 27-30 maggio 1997. Roma: Bulzoni, 267-303.

TRUMPER, J. (1997). Calabria and Southern Basilicata, in MAIDEN, M., PARRY, M. (Eds.): *The Dialects of Italy*. London/New York: Routledge, 355-364.

WALKER, R. (2005). Weak triggers in Vowel Harmony. In *Natural Language and Linguistic Theory* 23(4), 917-989.

WEINREICH, U., LABOV, W., & HERZOG, M. (1968). Empirical foundations for a theory of language change. In W. LEHMANN, & Y. MALKIEL (Eds.), *Directions for historical linguistics* Austin: University of Texas Press, 97-195.

WINKELMANN, R., HARRINGTON, J., & JÄNSCH, K. (2017). EMU-SDMS: Advanced speech database management and analysis in R. *Computer Speech & Language*, 45, 392–410.

Appendix

Table 1 - *The recorded speakers grouped by age and sex. The first two letters of the speakers' code refer to the village they come from: CA = Canna, CC = Cerchiana, CV = Castrovillari, LA = Lauria, LI = Laino, MG = Montegiordano, MM = Mormanno, SD = S. Domenica Talao, SH = Schiavonea. The last letter of the speakers' code refers to their sex. The column "Age" indicates the biological age of each speaker when recorded*

Age group	Sex	Speaker	Age
older	F	CA01F	44
		CC01F	65
		CC03F	44
		CC04F	51
		CC05F	81
		CC08F	44
	M	CC02M	47
		CC03M	46
		LI01M	82
		LI02M	80
		LI03M	90
		LI04M	92
		LI05M	67
		LI06M	85
		MG01M	45
		MG02M	67
		MM09M	73
		SH01M	41
younger	F	CC02F	13
		CC06F	14
		CC07F	19
		CV01F	32
		CV02F	24
		CV03F	22
		LA01F	31
		MM02F	25
		MM03F	28
		MM04F	26
	MM05F	25	
	SD01F	27	
	M	CC01M	27
		MM03M	26
MM04M		25	
MM05M		22	

Table 2 - List of elicited lexical items used for this study, listed in alphabetical order. The expected metaphonic changes are highlighted in bold type. In bisyllabic words, the stressed syllable is always the first one

Target word	Expected pronunciation	Target word	Expected pronunciation	Target word	Expected pronunciation
NOUNS and ADJECTIVES					
bella	bella	lunga	lɔŋga	uova	ɔva
bello	bellu, billu	lungo	lɔŋgu, lun̄gu	uovo	ovu, uvu
braccia	vrattsa	mese	mɛsɛ	vecchia	vɛcca
braccio	vrættsu	mesi	mesɪ, misi	vecchio	veccu, viccu
buona	bɔna	morta	mɔrta	verde	vɛrde
buono	bɔnu, bunu	morto	mortu, murtu	verdi	verdi, vir̄di
cane	kane	nera	'ni(v)ura	verme	vɛrme
cani	kæni	nero	'ni(v)uru	vermi	vermi, vir̄mi
capretta	kra'petta	nipote	nipɔtɛ	volpe	vurpe
capretto	kra'pettu, krapittu	nipoti	nipoti, niputi	volpi	vurpi
corna	kɔrna	noce	nufɛ	zoppa	tsɔppa
corno	kɔrnu, kurnu	noci	nufi	zoppo	tsoppu, tsuppu
corta	kurta	nuova	nɔva	VERBS	
corto	kurtu	nuovo	novu, nuvu		
cotta	kɔtta	ossa	ɔssa	tieni	'tenisi
cotto	kottu, kuttu	osso	ossu, ussu	tiene	tene
croce	krufɛ	pettine	'pettine	dormi	'dormisi
croci	krufi	pettini	'pettini, pittini	dorme	dɔrme
cuore	kɔrɛ	pezza	pɛttsa	penso	penzu
cuori	kori, kwori	pezzo	pettsu, pittsu	pensa	pɛnza
dente	dɛntɛ	piede	pɛdɛ	trovo	trovu
denti	denti, dinti	pedi	pedi, pidi	trova	trova
dita	'jidita, jita	ponte	pɔntɛ	corri	'kurrisi
dito	'jiditu, jitu	ponti	ponti, punti		

Target word	Expected pronunciation	Target word	Expected pronunciation	Target word	Expected pronunciation
dolce	durtʃe	prete	ˈprevete	corre	kurre
dolci	durtʃi	preti	ˈpreviti, priviti	esci	ˈjessisi
fiore	fjɔre	rossa	russa	esce	jesse
fiori	fjɔri, fjuri	rosso	russu	apri	ˈ(j)aprisi
fredda	fridda	santa	santa	apre	(j)apre
freddo	friddu	santo	sæntu	bevi	ˈvivisi
ginocchia	jiˈnucca(i)	sposa	spɔsa	beve	vive
ginocchio	jiˈnuccu	sposo	sposu		
grossa	grɔssa	topi	ˈsɔritʃi, suritʃi		
grosso	grɔssu, grussu	topo	ˈsɔritʃe		
legna	linna	uomo	ˈommini		
legno	linnu	uomini	ˈɔmmine		

Table 3 - *Post-hoc tests that were significant in the stem vowels /e, o/. The columns show (from left to right) the suffix vowel pair, the stem vowel, age group, sex, the dependent variable (dV), t-ratio (t), degrees of freedom (df), and probability of significance (p-value). For example, row 1 means that the suffix vowel /a/ vs. /u/ had a significant influence on F2 of the stem vowel /e/ in older women*

dV	Suffix pair	Stem	Age group	Sex	t	df	p-value
F2	a, u	e	older	F	5.0	77.0	$p < .001$
			younger	F	5.1	35.0	$p < .001$
			older	M	3.5	56.0	$p < .001$
F1	a, u	e	older	F	7.7	34.5	$p < .001$
			younger	F	5.6	23.3	$p < .001$
			older	M	4.3	29.1	$p < .001$
F2	a, u	o	younger	M	3.2	45.4	$p < .01$
			older	F	6.4	75.8	$p < .001$
			younger	F	5.2	32.7	$p < .001$
			older	M	4.0	58.0	$p < .001$
F1	a, u	o	younger	M	2.1	110.3	$p < .05$
			older	F	10.9	34.5	$p < .001$
			younger	F	22.5	6.9	$p < .001$
			older	M	5.7	30.0	$p < .001$
			younger	M	4.4	42.6	$p < .001$

dV	Suffix pair	Stem	Age group	Sex	t	df	p-value
F2	e, i	e	older	F	6.6	50.5	$p < .001$
			younger	F	3.0	29.0	$p < .01$
			older	M	4.8	42.3	$p < .001$
F1	e, i	e	older	F	4.2	71.3	$p < .001$
			younger	F	2.4	34.0	$p < .05$
			older	M	4.1	56.1	$p < .001$

Figures 7, 8 and 9 - *The following figures show some examples of visual stimuli used for the picture-naming task described in the Method (§ 2.2). In particular, Fig. 7 is an example of how inflected nouns were elicited; Fig. 8 shows an example of how inflected adjectives were elicited; Fig. 9 shows an example of how conjugated verbs were elicited*

Figure 7 - *Picture stimulus used to elicit the word 'egg' (in the dialect ['ovu, 'uvu]) on the left, vs. picture stimulus to elicit the word 'eggs' (in the dialect ['ova]), on the right*

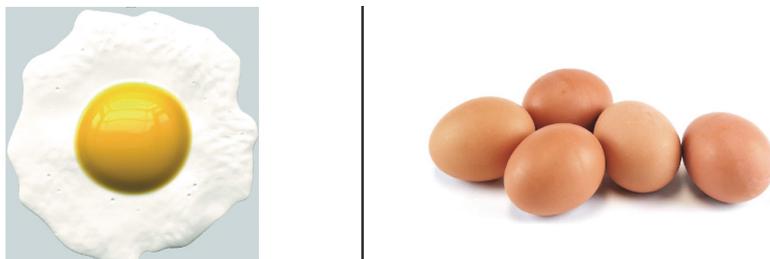
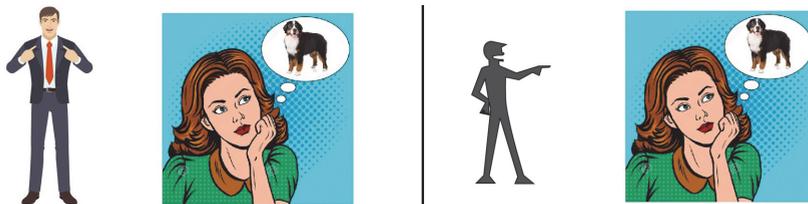


Figure 8 - *Picture stimulus used to elicit the word 'red', masc. sing. (in the dialect ['rus:u]), on the left, vs. picture stimulus to elicit the word 'red', fem. sg. (in the dialect ['rus:a]), on the right.*

In the picture on the right, the red apple is circled in order to lead the speaker to specify its colour, i.e. the participants tended to say 'red apple', in the dialect ['puma 'rus:a], and not just the isolated target word



Figure 9 - Picture stimulus used to elicit the word '(I) think' (in the dialect [ˈpenzu]), on the left, vs. picture stimulus used to elicit the word '(he/she) thinks' (in the dialect [ˈpenza]), on the right. In the picture on the left, the man pointing at himself graphically represents the subject of the sentence to be uttered by the speaker, in this case 1st person singular. In the picture on the right, the man pointing in the direction of the drawing graphically suggests to the speaker that the subject of the sentence to be pronounced is the 3rd person singular



MARGHERITA DI SALVO

Contatto interdialettale e cambiamento linguistico in un dialetto italiano all'estero

Dialect contact and language change in an Italian dialect abroad

This paper focuses on dialectal variation due to the contact with other dialects in a migratory context. This is carried out through the comparison of a group of return migrants in the Irpine town of Montefalcione and a group of migrants born and raised in the English town of Bedford. Both the forms of rhoticity in the Latin double lateral alveolar sound -LL- and the retention of the labiovelar approximant [w] in the secondary consonant cluster [kw] in the case of the Italian demonstrative *quello* 'that' and in the adverb *kwa* 'here' were analysed in a sociolinguistic perspective.

Keywords: livellamento dialettale, migrazione, sociolinguistica, variazione fonetica.

1. *Prospettive di analisi*¹

Nel recente volume di Vedovelli dedicato alla *Storia linguistica dell'emigrazione italiana*, l'Autore individua, sulla base di specifiche tendenze linguistiche, tre fasi della storia migratoria italiana che coincidono con *il parallelismo, lo slittamento, e la discontinuità*. Con Vedovelli (2011: 31), il *parallelismo* si riferisce alle "dinamiche linguistiche che hanno coinvolto da un lato gli emigrati italiani nel mondo e dall'altro la società italiana di origine", dinamiche che "si sono svolte secondo vie parallele e hanno mostrato, nonostante le distanze e le separazioni, esiti simili o assimilabili". Per l'emigrazione italiana successiva alla seconda guerra mondiale, composta prevalentemente da dialettofoni con competenza limitata dell'italiano, la migrazione ha comportato la creazione di modelli linguistici comuni: la condivisione degli stessi spazi sociali e fisici tra migranti di diversa origine regionale ha infatti generato processi di livellamento dialettale e convergenza. Tali processi sono stati raramente descritti per l'italiano come *heritage language*, nella misura in cui, nella bibliografia di impronta sociolinguistica sull'argomento, è prevalsa l'attenzione per gli esiti del contatto con la lingua del paese di immigrazione (Di Salvo 2012, Rubino 2014, Gorla 2015) e per i processi di erosione dovuta ad una scarsa esposizione alla lin-

¹ Il presente contributo rientra nella ricerca dal titolo "Transnational migrations: the case of the Italian communities in the UK", finanziata, nell'ambito del programma STAR (Sostegno territoriale alle Attività di Ricerca), dalla Compagnia di San Paolo in collaborazione con l'Università Federico II, dove è stata svolta sotto la supervisione della sottoscritta.

gua di origine (Scaglione 2000, 2003, Caruso 2010, Celata e Cancila 2010, Nodari, Celata, Nagy 2018).

È stato infatti soprattutto a partire da altre lingue di eredità che la prospettiva della convergenza e del livellamento dialettale è stata elaborata e formalizzata². Mi riferisco, in modo particolare, ai lavori sullo spagnolo a New York condotti sull'Otheguy Zentella Corpus of Spanish in New York City (OZC), formato da oltre 140 interviste raccolte tra il 2000 e il 2002 con ispanofoni immigrati da vari Paesi e regioni del Sud America con caratteristiche sociobiografiche differenziate. Questi studi hanno mostrato una diminuzione delle caratteristiche regionali delle diverse varietà di spagnolo (Erker 2012, Erker e Otheguy 2016) e una convergenza dovuta al contatto con l'inglese. Erker e Otheguy distinguono tra *dialectal levelling* per designare "the intergenerational reduction of regionally differentiated linguistic behaviour" e *linguistic convergence* per indicare, invece, "the enhancement of inherent structural similarities found between two linguistic systems", lo spagnolo e l'inglese (Erker e Otheguy (2016: 132). Tale distinzione ha come ripercussione sul piano dell'analisi che "*dialectal levelling* e *dialectal contact* will be used in reference to changes taking place in the Spanish of Latinos in New York due to the influence of other Latinos' Spanish, while *linguistic convergence and linguistic contact* will be reserved for changes due to the influence of their own and other people's English" (Erker e Otheguy 2016: 132). Tuttavia, soprattutto per l'uso dei pronomi tonici e la -s in fine di parola si assiste ad una riduzione della variazione tra diverse varietà dello spagnolo ma ciò accade in maniera statisticamente più significativa nel gruppo di residenti a New York da un maggiore numero di anni. Ciò induce gli Autori a formulare l'ipotesi che il numero di anni trascorso all'estero possa incidere sui processi di livellamento e contatto interdialeale: il numero di anni trascorso all'estero e il livellamento dialettale sono quindi in relazione di corrispondenza biunivoca nella misura in cui all'aumentare del primo corrisponde un aumento del secondo.

Tale studio costituisce un punto di riferimento per il caso di studio che qui si presenta in quanto la prospettiva del contatto interdialeale e del livellamento sono generalmente trascurati dalla bibliografia di stampo sociolinguistico sulle comunità italiane all'estero.

2. Il contesto della ricerca

2.1 Gli italiani a Bedford

La storia della comunità italiana di Bedford risale alla fine della Seconda Guerra Mondiale, quando, a seguito di accordi bilaterali tra il governo italiano e il governo britannico, dal giugno del 1951 iniziarono ad arrivare i primi nuclei di migranti che erano stati arruolati, mediante un sistema ufficiale di reclutamento, per la locale in-

² Non prendiamo in esame in questa sede i lavori precedenti condotti sul livellamento dialettale e i processi di convergenza che sono stati condotti su lingue parlate non nel contesto dell'emigrazione.

dustria di mattoni. L'emigrazione, qui come in altri contesti, coinvolse inizialmente i giovani adulti con titolo di studio basso, spinti dalla necessità di trovare un'occupazione stabile. A Bedford però sin dalle prime fasi, l'emigrazione riguardò anche le donne, che furono impiegate in una fabbrica di cioccolata e caramelle. Con il progressivo esaurirsi del sistema ufficiale di reclutamento, l'emigrazione fu alimentata dalle catene migratorie che favorirono l'arrivo nella città di migranti provenienti da aree ristrette dell'Italia meridionale: la provincia di Avellino, all'interno della quale, stando ai dati di Colpi (1991), i migranti nati nel comune di Montefalcione costituiscono il primo gruppo per presenze nella città inglese; la provincia di Campobasso e il comune di Busso; la provincia di Agrigento con il comune Sant'Angelo Muxaro in primis. Migranti da altre province e da altre regioni (Puglia, Basilicata e Calabria) sono presenti nella città inglese ma in minore proporzione.

I diversi gruppi regionali hanno adottato, soprattutto nei primissimi anni trascorsi in Inghilterra, strategie di chiusura all'interno della propria rete paesana: entro questa rete si sceglieva il coniuge e si svolgeva il tempo libero. In questi anni i migranti furono però costretti a vivere insieme, prima in ostelli e successivamente in case condivise (Colucci 2009), tutte concentrate nella zona della stazione in quello che per molti versi era un vero e proprio ghetto: ciò ha comportato la creazione di reti a base italiana che superavano il limite della regione di origine.

Le peculiari condizioni di vicinanza fisica e di condivisione degli spazi lavorativi e del tempo libero con parlanti di altre varietà italo-romanze hanno reso fortemente pluridialeltale il contesto in cui la prima generazione e la successiva hanno vissuto. I nati in Inghilterra, pur avendo avuto come lingua materna il dialetto d'origine dei genitori, sono cresciuti a contatto con vicini, parenti e amici di altre varietà dialettali e anche se, in accordo con rilevamenti precedenti (Di Salvo 2012, 2019), era l'inglese la varietà adoperata tra i pari, il dialetto ha continuato ad essere usato per rivolgersi alle generazioni più anziane. La socializzazione esterna alla famiglia è avvenuta in un ambiente anglofono e ciò rende questi parlanti, che, di fatto, non hanno esperito la migrazione come i propri genitori, completamente proiettati sul Paese di nascita piuttosto che su quello di origine dei genitori. Proprio per tali ragioni, essi possono essere considerati, sulla stregua del modello interpretativo elaborato da Turchetta (2018, 2019), come generazione 0:

La generazione 1 è così caratterizzata da individui di età adolescenziale o età adulta, già parzialmente o totalmente scolarizzati nel Paese di provenienza (tranne analfabeti), culturalmente e socialmente radicati nel Paese 1 (P1) con una o più lingue 1 (L1) di partenza. Nel novero di queste ultime rientrano le lingue di minoranza e tutte le variazioni dialettali caratterizzanti la dialettologia diffusa nei migranti italiani fino a pochi decenni fa. [...] La successiva generazione nata e cresciuta nel nuovo Paese rappresenta virtualmente una generazione 0 rispetto al Paese delle origini e riguarda individui in età preadolescenziale, anche nati nel Paese delle origini o nel nuovo Paese, ma comunque scolarizzati, culturalmente e socialmente radicati nel nuovo Paese e con una lingua materna diversa da quella dei propri genitori, perché coincidente con quella della socializzazione e della scolarizzazione nel Paese di effettiva residenza.

Le successive generazioni 0 sono caratterizzate da spazio linguistico italiano trasmesso da una prima generazione (1) in una condizione di tendenziale erosione, nell'ambito di repertori linguistici in cui le nuove lingue prime sono dominanti³.

2.2 I rientrati a Montefalcione

A partire dalla fine degli anni Settanta risale l'inizio del fenomeno dei rientri che interessò alcune decine, forse centinaia, di migranti. Non si hanno informazioni sul numero effettivo di coloro che, spinti soprattutto da motivazioni familiari come indicato in altra sede (Di Salvo 2015, 2019), decisero di tornare nel comune di nascita, ma rilevamenti precedenti hanno mostrato la consistenza del fenomeno, soprattutto in comuni come Montefalcione (AV). I protagonisti dei ritorni furono tanto i migranti di I generazione quanto i loro discendenti, mediante percorsi e modalità divergenti, tanto sul piano biografico che sul piano sociolinguistico. Questa diversità di storie ha portato la sottoscritta a ipotizzare che nelle migrazioni di ritorno sia possibile distinguere, sulla base di specifici comportamenti linguistici, una prima generazione rientrata, una prima generazione migrata, e una generazione 0⁴. Con "prima generazione rientrata" mi riferisco a coloro che sono migrati, per un numero variabile di anni, e successivamente sono ritornati a vivere nel comune di nascita: hanno quindi la conoscenza diretta di entrambi i Paesi. Il loro repertorio linguistico è contraddistinto dal dialetto, lingua materna e dominante; per loro l'emigrazione non ha significato uno *slittamento* verso i modelli linguistici del Paese di immigrazione. Proprio questo parametro permette di distinguere questi parlanti da coloro che sono stati inclusi nella categoria di "prima generazione migrata", categoria che include coloro che si sono trasferiti in Irpinia senza avere una conoscenza diretta ed effettiva del paese natio dei genitori: confluisce in questa casistica parte dei membri della generazione 0 individuata da Turchetta (2018, 2019). Sono considerati come "migranti" e non come "rientrati" in quanto, contrariamente ai propri genitori, non avevano una conoscenza diretta dell'Italia prima di sceglierla come residenza. Sul piano sociolinguistico, questi parlanti hanno avuto come lingua materna il dialetto ma la loro intera socializzazione è avvenuta in inglese, lingua che ha acquisito una posizione dominante all'interno del proprio repertorio individuale, per lo meno fino al momento del rientro.

La categoria di "generazione 0", infine, è stata assegnata a quei parlanti nati all'estero e che non hanno potuto scegliere di rientrare ma che sono stati portati in Italia da bambini o in età prepuberale dai propri genitori; hanno avuto come lingua materna il dialetto e, per via dei pochissimi anni vissuti in Gran Bretagna, non hanno avuto molti contatti con la società britannica e non hanno oggi alcuna competenza dell'inglese.

³ Turchetta (2019, 117-18).

⁴ Tale proposta è formulata in Di Salvo (in stampa).

3. *Ipotesi e obiettivi della ricerca*

Il caso di studio offerto dalla comunità di Bedford consente di formulare l'ipotesi di un livellamento linguistico indotto dalla compresenza, nel medesimo spazio urbano, di parlanti di diverse varietà dialettali. In questa città inglese, la prima generazione è stata a contatto prevalentemente con migranti originari di altre regioni. La condivisione delle medesime occupazioni lavorative e la socialità mediata dalla Chiesa cattolica (italiana) e dai pochi bar italiani presenti sul territorio permettono di ipotizzare che il comportamento linguistico dei migranti possa essere caratterizzato da un'innovazione dovuta al contatto interdialettale con varietà dialettali altre presenti nel medesimo spazio linguistico.

L'ipotesi di una convergenza verso modelli condivisi è formulabile in quanto nel caso di studio che qui si presenta sussistono tutte le condizioni che, in letteratura, sono considerate diagnostiche per la tendenza verso la genesi di modelli condivisi: in primo luogo, un contatto stabile, prolungato e costante di parlanti di varietà diverse, come peraltro ampiamente dimostrato anche a partire dagli studi già ricordati sullo spagnolo di emigrazione; una forte chiusura etnica che tuttavia prevede incontri al di fuori della rete regionale (Di Salvo 2012); la presenza di un modello linguistico, il napoletano, verso cui i parlanti potrebbero muoversi (Di Salvo, Matrisciano 2020). Su questo ultimo aspetto, studi precedenti hanno mostrato come, nel contesto della migrazione internazionale, i migranti originari delle regioni meridionali riconoscano nel dialetto napoletano un "modello" linguistico cui conformarsi. Questo sembra valere sia per la I generazione sia per la generazione 0 (Di Salvo, Matrisciano 2020) che, indipendentemente dalla varietà attivata, assumono il napoletano a *varietà di riferimento extralinguistico* (Di Salvo, Matrisciano, Maulella 2018), ossia a varietà che si crede/ si vorrebbe parlare. Non a caso, analizzando gli atteggiamenti linguistici, Di Salvo e Matrisciano (2020) hanno evidenziato che parlanti meridionali (ma non siciliani) residenti a Bedford dichiarano di parlare e di identificarsi, sul piano culturale e linguistico, con Napoli e i napoletani: questo atteggiamento è stato considerato anche un marcatore identitario nella misura in cui consente a tutti i migranti di I generazione nati nel Meridione continentale di distinguersi dai siciliani, considerati al contrario come modello culturale negativo. In tale prospettiva, si potrebbe ipotizzare che il processo di graduale riduzione della variazione interdialettale possa coincidere con l'adozione di varianti prossime al napoletano, percepite dai parlanti come più prestigiose⁵.

Su queste basi, è lecito chiedersi che cosa accade a quei migranti che dopo un contatto prolungato con varietà dialettali altre ritornano a vivere nel paese natio, ri-immersedosi nuovamente in un ambiente in cui il dialetto di origine è vitale: si potrebbe supporre, da un lato, un riappropriarsi del dialetto locale, anche per la volontà di reinserirsi, sul piano dell'identità, nel paese, ma, dall'altro, non si può

⁵ Un processo analogo, del resto, è stato descritto anche per le varietà irpine a ridosso dell'area napoletana (Abete e Vecchia 2018).

escludere che, proprio per rimarcare la loro non-appartenenza al comune di nascita, i rientrati continuino ad adottare varianti innovative apprese durante l'emigrazione.

Lo studio si è quindi proposto di capire se l'emigrazione e il contatto con varietà dialettali altre abbiano innescato un livellamento linguistico e se il successivo ritorno nel comune natio abbia invece determinato un nuovo adattamento al modello linguistico montefalcionese. Esso si è quindi proposto due obiettivi. Il primo consiste nell'appurare se, con il contatto prolungato con varietà dialettali altre, vi sia stato un allargamento delle varianti linguistiche in relazione a precisi tratti fonetici o, se, al contrario, l'emigrazione abbia avviato un processo di attrito linguistico, anche per le peculiari condizioni di (ridotta) esposizione al dialetto natio che caratterizzano qualsiasi esperienza migratoria. Il secondo obiettivo consiste nel verificare che cosa accade con il rientro e il reinserimento nell'ambiente nativo, con particolare riferimento alla variazione e alle varianti apprese durante gli anni trascorsi all'estero: il confronto con i rientrati ha mirato a capire se e quanto fosse ampio lo spettro di variazione tra emigrati ancora all'estero e rientrati, argomento che, per quanto i rientri siano stati un carattere stabile della storia italiana, non è stato affrontato in maniera sistematica da linguisti e dialettologi. I pochi studi precedenti hanno misurato solamente il tasso di adesione dei rientrati alla parlata locale senza prendere in esame la variabilità e l'erosione indotte dalla più o meno lunga permanenza all'estero.

4. Metodologia della ricerca

La prospettiva adottata in questo contributo si fonda su due premesse: la prima è che per descrivere la variabilità dialettale presente nel contesto dell'emigrazione è indispensabile ricostruire cosa accade nel dialetto di partenza; la seconda è che, per capire il comportamento dei rientrati, è necessario descrivere preliminarmente la variabilità acquisita/determinata dall'esperienza migratoria. Pertanto, in una fase iniziale del lavoro è stata descritta la variabilità interna al comune di partenza in relazione ad alcune variabili opportunamente selezionate; tale variabilità è stata successivamente comparata con quella rilevata nel contesto dell'emigrazione prima, e del rientro poi.

La ricerca ha previsto una fase iniziale di osservazione partecipante a cui è seguita la raccolta delle interviste, condotte dalla sottoscritta con la collaborazione di Cesarina Vecchia che ha partecipato in qualità di assegnista di ricerca al progetto da me coordinato e in seno al quale si colloca il presente contributo (Di Salvo 2019). Le interviste sono state condotte in dialetto e in italiano regionale, con microfono a vista. Hanno tutte una durata superiore ai 40 minuti.

Si è optato per un'intervista libera che potesse ricostruire la variabilità del parlato, ma sono stati proposti ai parlanti temi comuni, quali la storia migratoria e del rientro a Montefalcione nel caso dei rientrati, la vita nei due Paesi, con particolare riferimento all'ambito lavorativo e a quello privato.

5. *Il campione e il corpus*

L'analisi ha perso in esame un corpus raccolto con migranti e rientrati, di prima generazione e dei loro discendenti, equamente distribuiti per genere, come sintetizzato in tabella:

Tabella 1 - *Prospetto dei parlanti per contesto di residenza, generazione e genere*

	Nome	Generazione	Genere	Fascia di età
Emigrati	Giuseppe	I	M	>80
	Luigi	I	M	>80
	Raffaella	I	F	>80
	Caterina	I	F	>80
	Tonino	0	M	Tra 45 e 60
	Pasquale	0	M	Tra 45 e 60
	Carmela	0	F	Tra 45 e 60
	Saveria	0	F	Tra 45 e 60
Rientrati	Antonio	I	M	>80
	Gaetano	I	M	>80
	Antonio F.	I	M	>80
	Antonietta	I	F	>80
	Celeste	I	F	>80
	Anna	I	F	>80
	Teresa	0	F	Tra 45 e 60
	Gerarda	0	F	Tra 45 e 60
Olimpia	0	F	Tra 60 e 80	

Da un punto di vista socio-biografico, i membri della prima generazione hanno tutti un livello di studi medio-basso; gli uomini hanno lavorato nella fabbricazione dei mattoni, mentre le donne sono state impiegate nel settore industriale o in quello delle pulizie. Nati o arrivati in Inghilterra prima della pubertà, i loro discendenti che, come indicato in precedenza sono qui considerati, sulla scia di Turchetta, come generazione 0, hanno invece raggiunto un livello di istruzione maggiore grazie al quale hanno avuto accesso a un maggior numero di professioni.

Dal punto di vista linguistico, se, prima dell'emigrazione, i membri della prima generazione non solo hanno avuto come lingua materna il dialetto ma sono anche cresciuti in un ambiente in cui il dialetto montefalcionese era la sola varietà utilizzata, a Bedford hanno vissuto in un contesto anglofono all'interno di una rete pluridialettale, contraddistinta da frequentazioni, ora occasionali ora più assidue, con migranti di altra origine regionale; i membri della generazione nata in Inghilterra, pur avendo avuto come L1 il dialetto montefalcionese dei genitori, sono cresciuti in questo ambiente diversificato dal punto di vista dialettale, sebbene ben presto si siano gradualmente inseriti all'interno di reti sempre più anglofone.

Una nota riguarda l'assenza di informatori della generazione 0/migrata di sesso maschile tra i rientrati: tale mancanza è dovuta all'impossibilità di rintracciare nel comune indagato soggetti che potevano rientrare in questa tipologia.

I dati raccolti secondo questa metodologia sono stati comparati con quelli realizzati da un gruppo di confronto formato da quattro parlanti, due uomini e due donne, residenti stabilmente nel comune di Montefalcione e senza alcuna pregressa esperienza migratoria.

6. Variabili interne e variabili esterne, strategie di analisi

Al fine di verificare l'impatto del livellamento dialettale così come descritto per lo spagnolo *heritage language* da Erket e Otheguy (2016) e di descrivere l'adozione di varianti napoletane, ho scelto di partire da quelle variabili che più di altre permettessero di distinguere la varietà dialettale montefalcionese dal napoletano. Esse coincidono con il mantenimento della labiovelare /w/ nel nesso secondario /kw/ e la resa rotacizzata della geminata laterale latina -LL-.

6.1 Mantenimento della labiovelare nell'avverbio *kkwa*

Il mantenimento dell'approssimante labiale [w] nel nesso [kw], secondo studi precedenti sul dialetto di Montefalcione (Di Salvo 2019), si realizza in questa varietà sia davanti alle vocali anteriori [e] e [i], come nei dimostrativi *questo* e *quello*, sia davanti ad [a], ad esempio nel tipo avverbiale *qua* (ECCU HAC > [kwa])⁶.

Sulla base di tale variabile, il dialetto di Montefalcione si distingue da molte varietà meridionali e campane che, al contrario, presentano la riduzione a velare [k]: così, ad esempio, in napoletano la forma dell'avverbio è *cà* [ka] (Rohlf, 1966, § 163; Ledgeway, 2009: 118; Loporcaro, 1988: 149)⁷.

La problematicità dell'analisi del tipo *kwa* è correlata alla situazione di contatto tra italiano e dialetto in quanto l'avverbio montefalcionese e quello italiano sono omofoni: da un punto di vista metodologico, è stato pertanto deciso di limitare l'analisi alle porzioni di testo in dialetto, in cui pare improbabile, anche se non impossibile, una commutazione intrafrastica limitata al solo avverbio. Sono state quindi considerate per l'analisi tutte le occorrenze dell'avverbio *kwa* presenti in frammenti dialettali, al fine di analizzare l'eventuale variazione con la forma *kà*, priva della labiovelare e propria delle altre varietà dialettali presenti nel contesto di Bedford e soprattutto nel napoletano.

⁶ Per l'Irpinia cfr. Abete (2017, p. 53) e Vecchia (2018: 219).

⁷ In alcuni dialetti della Puglia inoltre il nesso /kw-/ secondario davanti a vocale anteriore in alcuni casi si palatalizza dando esito [tʃ], in altri ha sviluppo [ku] come nei dimostrativi [kuddu] e [kuru] 'quello' e [kustu] 'questo' (Rohlf, 1966, § 163).

6.2 Esiti del dimostrativo

Poiché, come accennato in precedenza, la labiovelare è mantenuta anche davanti a vocale anteriore, per avere un numero più consistente di tokens da analizzare, l'analisi è stata estesa anche alle occorrenze del dimostrativo (maschile, femminile, neutro), in cui, accanto alla presenza/riduzione della labiovelare, è altresì osservabile la rotacizzazione della laterale alveolare geminata⁸.

Nei dialetti meridionali al di sotto della Linea Salerno-Lucera (Avolio, 1989), il nesso latino -LL- dà esiti occlusivi retroflessi [dɖ] o alveolari [dd]⁹, al contrario di quanto accade nel napoletano, nel Sannio beneventano e dell'area irpina occidentale prossima alla provincia di Napoli (Avolio, 1989; Maturi, 2002: 102-103; Abete, Vecchia, 2018: 453)¹⁰ in cui al contrario si ha l'esito laterale alveolare geminato¹¹. Rispetto a questo quadro generale, le varietà irpine più interne divergono in quanto presentano lo sviluppo rotacizzato negli articoli determinativi, nei pronomi personali e negli aggettivi e pronomi dimostrativi (Abete, 2017: 52-53; Vecchia, 2017)¹².

Nel dialetto di Montefalcione l'esito rotacizzato è presente solo nel dimostrativo *quello*, sia in funzione aggettivale sia pronominale, in tutte le categorie di genere e numero (Como, 2007, p. 161; Vecchia, 2017, pp. 87-90¹³), con la conseguenza che la forma patrimoniale del dimostrativo del dialetto di Montefalcione è *kwiro*, contraddistinta sia dall'esito rotacizzato che dal mantenimento della labiovelare.

Negli studi precedenti sulla comunità montefalcionese di Bedford (Di Salvo e Guzzo 2019), al contrario, è stato rilevato un forte polimorfismo nella realizzazione del dimostrativo. Accanto alla forma patrimoniale, sono attestati altri esiti che sembrano risentire tanto del modello italiano quanto del modello napoletano: *kwello* è il tipo italiano, mentre *killo* e *kiro*, per quanto dialettali, sono da considerare innovativi in quanto frutto del processo di napoletanizzazione per ciò che riguarda l'evoluzione della laterale e per l'assenza del nesso labiovelare. Il prospetto delle forme

⁸ Per una rassegna degli esiti di -LL- in area romanza si vedano Celata (2006) e Müller (2012).

⁹ Per una indicazione del limite settentrionale dei dialetti con esiti alternativi alla laterale alveolare in area alto-meridionale si veda la mappa riportata in Abete, Vecchia (2018: 453).

¹⁰ A nord della linea Salerno-Lucera-Gargano, oltre all'area con mantenimento della laterale alveolare, si attesta l'esito palatale laterale [ʎʎ] condizionato da /i, u/ finali in un'area che comprende la parte montana interna dell'Abruzzo e del Molise, il Lazio meridionale e la Campania settentrionale (alto casertano, area del Matese e parte del Sannio beneventano, cfr. Abete, Vecchia, 2018, p. 450).

¹¹ Sono presenti tuttavia esiti alternativi alla laterale intensa nel contesto del dimostrativo in alcuni dialetti dell'area flegrea, area situata a nord del capoluogo campano (Como, 2006: 121-127).

¹² Si tratta in tutti i casi di forme etimologicamente connesse al pronome latino ILLE. Nel caso delle parole lessicali piene, invece, le varietà irpine centro-orientali hanno esiti di tipo retroflesso, occlusivi [dɖ] o affricati [dʒ], palatali laterali [ʎʎ] o occlusivi [ʝʝ] e occlusivi alveolari [dd] (Abete, Vecchia, 2018: 553-556). Tali sviluppi non sono tuttavia attestati a Montefalcione.

¹³ A riguardo si vedano anche le carte AIS 1587 – 'andiamo a scegliere un cappello nella bottega. Perché hai scelto questo e non quello?'; 1589 – 'pagatemi quello che mi dovete'; 828 – 'quei cavalli non valgono niente eppure costano il doppio dei miei'; 1045 – 'se potessero comprerebbero quella vacca'.

attestate in questi studi precedenti condotti a Bedford è quindi quello che si ritrova in tabella:

Tabella 2 - *Polimorfismo del dimostrativo nel dialetto di Montefalcione secondo Di Salvo e Guzzo (2019)*

	<i>kwello</i>	<i>kwillo</i>	<i>killo</i>	<i>kiro</i>	<i>kwiro</i>
/i/	-	+	+	+	+
/-r-/	-	-	-	+	+
/-ll-/	+	+	+	-	-
/kw-/	+	+	-	-	+
/k-/	-	-	+	+	-

6.3 Variabili esterne: genere, generazione, contesto di residenza (emigranti vs rientrati)

Per quanto riguarda le variabili esterne, è in primo luogo stata valutata la variazione tra residenti ancora in Inghilterra e rientrati in modo da capire se l'emigrazione abbia determinato un ampliamento delle varianti o piuttosto una loro perdita e se, nella successiva fase del rientro, il reinserimento in un ambiente italofono abbia spinto i parlanti ad adottare nuovamente le forme del dialetto arcaico o quelle dell'italiano (regionale). Tali dinamiche sono state valutate tenendo conto anche della generazione e del genere dei parlanti.

7. Risultati dell'analisi

7.1 Il mantenimento della labiovelare nell'avverbio qua

I dati provenienti da un gruppo di confronto composto da 4 parlanti nati e cresciuti a Montefalcione senza mai emigrare all'estero presentano solo il tipo *kwa*: in tutte le occorrenze, al di là di una variazione nella resa dell'occlusiva e della vocale finale, talvolta dipendente dal contesto fonologico, la labiovelare è mantenuta in tutte le forme dell'avverbio (38 occorrenze). Qualora, quindi, nei corpora raccolti con i migranti e con i rientrati fosse presente la variante priva della labiovelare, la sua comparsa potrebbe essere quindi considerata come un'innovazione indotta dal contatto interdialeale proprio del contesto migratorio. Nelle porzioni dialettali di testo di montefalcionesi ancora residenti all'estero, compare, unicamente in alcune tipologie di parlanti, la variante *kà*, non presente nel dialetto montefalcionese di origine. Tale comparsa sembra essere la risultante del contatto con varietà dialettali altre che, al contrario di quella montefalcionese, non presentano il mantenimento della labiovelare.

Tabella 3 - *Distribuzione degli esiti relativi al mantenimento della labiovelare nell'avverbio (valori percentuali e assoluti)*

	Emigrati				Rientrati		
	M		F		M	F	
	I	0	I	0	I	I	0
/kwa/	72,22% (31 occorrenze)	100% (17 occorrenze)	100% (26 occorrenze)	100% (16 occorrenze)	96,30% (25 occorrenze)	100% (35 occorrenze)	100% (18 occorrenze)
/ka/	27,78% (12 occorrenze)	0	0	0	3,70% (1 occorrenza)	0	0

Con il rientro le rare forme *napoletanizzanti* si contraggono e continuano a comparire, molto sporadicamente, solo negli informatori uomini. Sulla base di questi dati, sembra configurarsi un livellamento. Non si ha una riduzione delle varianti, bensì un ampliamento per effetto della convergenza verso altre varietà dialettali. Tra queste si può ipotizzare che, vista la coincidenza di esiti con le forme *napoletane*, vi sia influenza del modello napoletano che, come esposto in altra sede (Di Salvo, Matrisciano 2020), funge da varietà di prestigio.

Questo quadro è però relativo solo ai migranti ancora residenti all'estero, ma non ai rientrati, il cui comportamento è invece caratterizzato da una riduzione degli esiti di tipo napoletano e una maggiore incidenza di forme patrimoniali, come se il reinserimento nel comune natio abbia indotto un livellamento sulla varietà di origine, con conseguente riduzione della polimorfia innestata dalla migrazione.

7.2 L'analisi del dimostrativo: il caso del dimostrativo maschile

Per quanto riguarda le forme del dimostrativo, è stato già segnalato altrove come esse consentano di analizzare sia il mantenimento della labiovelare seguita da vocale anteriore, sia gli sviluppi del nesso latino -LL-.

Nel gruppo di confronto, la variante patrimoniale *kwiro* convive con la forma italiana *kwello*, secondo le percentuali riportate in tabella:

Tabella 4 - *Distribuzione delle forme del dimostrativo nel gruppo di confronto*

	<i>kwello</i>	<i>kwiro</i>
M	90 (38 occorrenze)	10 (4 occorrenze)
F	100 (24 occorrenze)	0

I dati provenienti dal gruppo di confronto mostrano come le forme patrimoniali siano presenti unicamente negli informatori di sesso maschile, in cui tuttavia costituiscono una piccola percentuale del totale delle occorrenze. Sia gli uomini che le donne prediligono nella quasi totalità dei casi la variante italiana. I dati sono in parziale controtendenza con quelli riportati da Vecchia (2019: 86) secondo la quale nel gruppo di confronto da lei esaminato (ma formato unicamente da due parlanti) la forma *kwello* risultava essere assente, mentre è attestata la forma *kwillo*. I dati di Vecchia, inoltre, non fornendo percentuali relative a questo gruppo di

confronto non permettono pienamente di valutare né la dinamica interna al punto tra conservazione e innovazione dialettale per il processo di italianizzazione né quanto sia significativo, sul piano statistico, l'ampliamento delle varianti da parte di chi è emigrato.

A Bedford invece la distribuzione delle varianti evidenzia (i) un'incidenza statisticamente significativa di varianti di tipo napoletano, ma limitatamente agli uomini di I generazione; (ii) la riduzione delle forme innovative successiva al rientro.

Tabella 5 - *Distribuzione delle varianti del dimostrativo percentuali e assoluti*

			<i>kwello</i>	<i>kwillo</i>	<i>killo</i>	<i>kiro</i>	<i>kiwo</i>
Emigrati	1	M	1,33	18,67	20,67	30,67	28,67
		F	0	0	0,60 (1 occorrenza)	0,60 (1 occorrenza)	98,80 (168 occorrenze)
Emigrati	0	M	11,76 (11 occorrenze)	47,06 (45 occorrenze)	8,24 (8 occorrenze)	1,18 (1 occorrenza)	31,76 (30 occorrenze)
		F	18,75 (18 occorrenze)	30,21 (29 occorrenze)	5,21 (5 occorrenze)	3,13 (3 occorrenze)	42,71 (41 occorrenze)
Rientrati	1	M	40,38 (42 occorrenze)	9,62 (10 occorrenze)	5,77 (6 occorrenze)	8,65 (9 occorrenze)	35,58 (37 occorrenze)
		F	22,06 (16 occorrenze)	27,94 (21 occorrenze)	0	0	50,00 (35 occorrenze)
	0	F	100 (45 occorrenze)	0	0	0	0

I dati evidenziano un comportamento sbilanciato a favore delle varianti dialettali nei migranti ancora residenti all'estero, soprattutto nelle donne, indipendentemente dall'appartenenza generazionale, e a favore dell'italiano nei rientrati. In questo secondo caso, la variazione di genere non è pienamente valutabile in quanto, per le motivazioni esposte in precedenza, non sono riuscita ad arruolare informatori maschi appartenenti alla generazione più giovane. Limitando le considerazioni alla sola prima generazione, la scelta in favore della variante italiana è maggiore negli uomini che nelle donne. Tuttavia, considerando come anche l'esito *kwillo* possa essere influenzato dal modello italiano *kwello*, si ha nelle donne una percentuale maggiore di esiti italiani/italianizzanti. Capovolgendo la prospettiva in modo da guardare al polo del dialetto e non a quello dell'italiano, le forme dialettali sono preponderanti nelle donne, in maniera indipendente dal contesto di residenza, e lo scarto tra rientrati e migranti ancora in Inghilterra non è particolarmente significativo. Quello che conta, però, ai fini della prospettiva adottata, è la presenza, soprattutto negli uomini ancora all'estero ma in misura significativa anche nei rientrati, di varianti dialettali non patrimoniali che sono, quindi, la risultante di un contatto interdialettale con varietà che non presentano il mantenimento della labiovelare.

8. *Discussione dei risultati e considerazioni conclusive*

L'analisi linguistica ha mostrato l'adozione di esiti privi della labiovelare e dell'assenza di rotacizzazione della -LL- geminata, esiti che sono favoriti dal contatto con varietà dialettali di altri gruppi di parlanti: è lecito chiedersi quali siano queste varietà e perché l'effetto del contatto interdialettale sia osservabile soprattutto negli uomini ancora residenti all'estero. Vista la composizione per gruppi regionali e provinciali della comunità di Bedford, si potrebbe escludere un influsso del napoletano in quanto non sono presenti numerosi napoletani residenti nella città inglese. Tuttavia, sulla base di studi di impronta sociolinguistica e di dialettologia percettiva (Di Salvo, Matrisciano 2019, 2020), è ipotizzabile che il napoletano eserciti, come avviene per alcune aree del Meridione continentale italiano, il ruolo di modello linguistico: ciò potrebbe incoraggiare i parlanti nati a Montefalcione ad adottare esiti napoletanizzanti indipendentemente dal numero di emigrati napoletani residenti nel contesto inglese. Se tale ipotesi fosse verificata a partire da un numero maggiore di variabili linguistiche, si potrebbe dimostrare la presenza di un processo di convergenza linguistica verso il modello linguistico più prestigioso, il napoletano.

Nei rientrati, invece, il numero di esiti non patrimoniali del dialetto montefalcionese si riducono, non scomparendo del tutto: con il rientro, infatti, si ha una contrazione delle forme dialettali e degli esiti non patrimoniali a favore di quelli italiani/italianizzanti. Se, quindi, per quanto riguarda il gruppo di parlanti ancora residenti all'estero è emersa una convergenza verso il modello più prestigioso, il napoletano, nei rientrati si osserva un livellamento, da intendere come riduzione del polimorfismo dialettale, a favore dell'adozione di varianti o patrimoniali o italiane/italianizzanti.

Tali processi sono tuttavia soggetti a variazione sociolinguistica. La variabile genere infatti sembra essere cruciale: nelle donne, infatti, si ha una minore significatività statistica del contatto interdialettale e della convergenza verso forme napoletane/napoletanizzanti e ciò sembra essere ricondotto alla conformazione della loro rete sociale. Le donne hanno relazioni quasi prevalentemente all'interno di una rete familiare, composta da sorelle, cognate, vicine di casa, spesso nate nel proprio stesso paese. Gli uomini, al contrario, trascorrendo il tempo libero al bar italiano, a fare una partita a carte con gli amici o nei comitati organizzatori di eventi di carattere religioso, hanno intessuto, anche al di fuori del lavoro, relazioni sociali con italiani di altre parti d'Italia, con una maggiore esposizione a modelli linguistici diversi. Al netto di tali differenze, la prospettiva del livellamento *vs* convergenza così come formulata per lo spagnolo negli Stati Uniti sembra essere una chiave di lettura interessante, per quanto ancora da approfondire, anche per l'italiano come lingua di eredità, tradizionalmente studiato a partire da paradigmi incentrati sul contatto con la lingua di immigrazione o sul concetto di attrito linguistico. Questo ribaltamento di prospettiva, al contrario, consente di guardare alle varietà migrate anche in termini di innovazione e non, secondo i paradigmi teorici tradizionali, come varietà erose o semplificate.

Riferimenti bibliografici

- ABETE G. (2017). *Parole e cose della pastorizia in Alta Irpinia*. Napoli: Giannini Editore.
- ABETE G., VECCHIA C. (2018). *Variabilità degli esiti di -LL- in Irpinia: dettagli fonetici e implicazioni diacroniche*. In ANTONELLI, R., VIDESOTT, P., GLESSGEN, M. (eds.), *Atti del XXVIII Congresso Internazionale di Linguistica e Filologia Romanza, Roma 18-23 luglio 2016*, Strasbourg: ELiPhi, pp. 448-459.
- AIS: *Atlante Linguistico ed Etnografico dell'Italia e della Svizzera Meridionale*.
- AVOLIO F. (1989). *Il limite occidentale dei dialetti lucani nel quadro del gruppo «alto meridionale»: considerazioni a proposito della linea Salerno-Lucera*. In *L'Italia dialettale*, 52, 1-22.
- CARUSO M. (2010). *Italian language attrition in Australia. The verb system*. Milano: Franco Angeli.
- CELATA (2006). *Analisi dei processi di retroflessione delle liquide in area romanza con dati sperimentali dal corso e dal siciliano*, Tesi di Dottorato, Università Normale Superiore di Pisa.
- CELATA, C., CANCELILA, J. (2010). Phonological attrition and the perception of geminate consonants in the Lucchese community of San Francisco (CA). In *International Journal of Bilingualism*, 14(2), 185-209.
- COLPI, T. (1991). *The Italian Factor. The Italian Community in Great Britain*, Edimburgo e Londra: Mainstream Publishing.
- COLUCCI, M. (2009). *Lavoro in movimento*, Roma: Donzelli.
- COMO, P. (2007). *La variabilità del dialetto. Uno studio su Monte di Procida*. Napoli: Liguori.
- DI SALVO, M. (2012). *“Le mani parlavano inglese”: percorsi linguistici e culturali tra gli italiani d’Inghilterra*. Roma: Il Calamo.
- DI SALVO, M. (2015). Italiani in Inghilterra. Inglese in Irpinia. In *Studi Italiana di Linguistica Teorica e Applicata*, 2, 295-316.
- DI SALVO, M. (2019). *Repertori degli italiani all'estero*. Pisa: Pacini.
- DI SALVO, M. (in stampa). Modello generazionale dell'emigrazione di ritorno: generazione rientrata, generazione migrata, generazione zero. In *L'Italia dialettale*.
- DI SALVO, M., MATRISCIANO, S., MAULELLA, M. (2018). Il dialetto napoletano in prospettiva lucana. Varietà di prestigio o inesistente nell'immaginario collettivo dei lucani? In *Lingua e Stile*, 1, 75-104.
- DI SALVO, M., GUZZO, S. (2019). Aspetti della variazione morfosintattica nell'emigrazione di ritorno, in *Lo spazio linguistico globale dell'italiano: scenari a confronto*. In DI SALVO M. (Eds.), *Lo spazio linguistico globale dell'italiano: scenari a confronto*, Alessandria: Dell'Orso, 69-92.
- DI SALVO, M., MATRISCIANO, S., (2019). Il siciliano come varietà di distanziamento extralinguistico nel contesto migratorio. In DI SALVO M. (Ed.), *Lo spazio linguistico globale dell'italiano: scenari a confronto*, Alessandria: Dell'Orso, 45-68.
- DI SALVO, M., MATRISCIANO, S., (2020). Il dialetto di Campobasso è su per giù come il napoletano. Neapolitan as a linguistic identity marker in transnational migrations. In *Studi italiani di Linguistica Teorica e Applicata*, XLIX, 2, 139-162.

- ERKER, D. (2012), *An acoustically based sociolinguistic analysis of variable coda /s/ production in the Spanish of New York City (Ph.D. dissertation)*. New York University.
- ERKER, D., OTHÉGUY, R. (2016). Contact and coherence: Dialectal leveling and structural convergence in NYC Spanish. In *Lingua*, 172-173, 131-146.
- GORIA, E. (2015), Il piemontese in Argentina. Considerazioni generali e analisi di un caso di studio. In *Rivista italiana di dialettologia*, 39, 12-158.
- LEDGEWAY, A. (2009), *Grammatica diacronica del napoletano*. Tübingen: Max Niemeyer Verlag.
- LOPORCARO, M. (1988), *Grammatica storica del dialetto di Altamura*. Pisa: Giardini Editori.
- MATURI, P. (2002). *Dialetti e substandardizzazione nel Sannio Beneventano*. Frankfurt am Main: Peter Lang.
- MÜLLER, D. (2012), *Developments of the lateral in Occitan dialects and their Romance and cross-linguistic contexts*, PhD Thesis, Università di Tolosa.
- NODARI, R., CELATA, C., NAGY, N. (2019). Socio-indexical phonetic features in the heritage language context: VOT in the Calabrian community in Toronto. In *Journal of Phonetics* 73, 91-112.
- ROHLFS, G. (1966). *Grammatica storica dell'italiano e dei suoi dialetti*. Torino: Einaudi.
- RUBINO, A. (2014). *Trilingual talk*. Palgrave: McMillan.
- SCAGLIONE, S. (2000), *Attrition. Mutamenti sociolinguistici nel lucchese di San Francisco*. Milano: Franco Angeli.
- SCAGLIONE, S. (2003), *Segnali discorsivi allogeni nelle varietà di emigrazione*. In DE FINA, A., BIZZONI, A. (Eds.), *Italiano e italiani fuori d'Italia*. Guerra: Perugia, 105-109.
- TURCHETTA, B. (2018). Modelli linguistici interpretativi della migrazione italiana. In TURCHETTA, B., VEDOVELLI, M. (Eds.), *Lo spazio linguistico italiano globale: il caso dell'Ontario*. Pisa: Pacini, 73-104.
- TURCHETTA, B. (2019). Proiezione simbolica e innovazione nelle identità linguistiche migrate, in ORIOLES, V., BOMBI, R. (Eds.), *Valori identitari e imprenditorialità*. Udine: Forum ed., 113-122.
- VECCHIA, C. (2017). *La variazione fonetica degli esiti di -LL- in Irpinia. Processi di rotacizzazione e di retroflessione nelle varietà dell'alta valle del Calore*, Tesi di dottorato, Università degli Studi di Napoli "Federico II".
- VECCHIA, C. (2018). Dialetti e comunità dell'alta valle del Calore. Profilo linguistico e socio-culturale dei centri dell'Irpinia centro-meridionale. In *Rendiconti della Accademia di archeologia, lettere e belle arti*, 78, 213-224
- VECCHIA, C. (2019). *Analisi della variazione dialettale nel parlato di migranti e rientrati*, in DI SALVO, M. (Ed.), *Repertori linguistici degli italiani all'estero*. Pisa: Pacini Editore, 83-124.
- VEDOVELLI, M. (2011). *Storia linguistica dell'emigrazione italiana*, Roma: Carocci.

TERESA CARBUTTI

L'italiano: fonte di riscatto per parlanti dialettofoni

The italian language: redemption for dialect speakers

Lucanian dialects, as other Italo-Romance dialects, are conditioned by the socio-cultural pressure of Italian.

The analysis of linguistic data from some lucanian dialects shows that native speakers are replacing some of their original features with those of Italian. This situation can be observed at all levels, especially at the lexical level and at the phonetic and phonological level.

The paper aims to demonstrate that the loss of some original phonetic features of dialects in favour of phonetic features of standard language does not occur casually. Some original phonetic features are replaced by phonetic features of Italian because they are considered as “vulgar” by native speakers.

Thus, in the eyes of dialect speakers, standard language appears to be a social advancement instrument and a form of emancipation from their past poverty condition as symbolized by dialects.

Keywords: dialects, Italian, phonetic features, Italianization.

1. *Italiano e dialetto in Basilicata: alcuni presupposti*

Sul rapporto fra italiano e dialetto si è soffermato recentemente Gaetano Berruto, sottolineando come questo si risolve nei termini della dilalia (Berruto, 1987) più che della diglossia. Il concetto di dilalia tende a inglobare quei casi nei quali

«a clear functional differentiation exists (with a High variety and a Low variety), as in (classic) diglossia, but unlike diglossia, there is a functional overlap in spoken domains, with both L and H varieties used in ordinary conversation and primary socialization» (Berruto, 2018: 497).

In effetti, sempre più spesso, e la Basilicata non fa eccezione, l'italiano affianca il dialetto nell'ambito della conversazione informale, come confermano anche i dati ISTAT (2015), secondo i quali il 69,4% della popolazione lucana sopra i 6 anni parla dialetto, ma affiancandolo all'italiano, anche in famiglia o nei contesti informali.

La perdita di terreno del dialetto a favore dell'italiano è un processo che si avvia già nei primi decenni del XX secolo, quando la scuola iniziò la sua opera di diffusione della lingua nazionale, e, tuttavia, già all'indomani dell'Unità d'Italia l'utilizzo del dialetto o dell'italiano era indicatore di una data estrazione sociale: all'uso dell'italiano (prevalentemente scritto) erano associati i “galantuomini”, ricchi e possidenti, mentre all'utilizzo del dialetto (esclusivamente parlato) erano associati i “cafoni”, poveri e analfabeti (De Blasi, 1994: 103-109). All'attribuzione di prestigio

alla lingua italiana ha concorso, in epoca contemporanea, anche il ruolo dei professionisti che «hanno collegato il piccolo mondo di ogni paesino alla vita culturale nazionale» (De Blasi, 1994: 159).

Dunque, la naturale conseguenza è stata la svalutazione del dialetto e la “sopravalutazione” della lingua italiana, idea rafforzata dal sistema d’istruzione, dalla diffusione della stampa e dall’azione dei nuovi media (De Mauro, 2017: 57-147).

Attualmente, sebbene si stia registrando una rivalutazione del dialetto, persiste, nelle generazioni più anziane, ancora dialettofone, la scelta di non trasmettere la propria lingua ai più giovani:

«uno dei problemi della Basilicata che [...] incide fortemente sul problema della trasmissione e della conservazione del dialetto, è la migrazione dei giovani lucani verso altre Università, nella maggior parte dei casi, tra l’altro, non verso Università meridionali, ma, prevalentemente, toscane, lombarde o emiliane. Ciò causa timori nei genitori riguardo l’integrazione dei figli e porta spesso i ragazzi ad assumere in breve tempo cadenza locale, se non proprio il lessico, abbandonando il proprio dialetto e mostrando così un incomprensibile senso di inferiorità culturale nei confronti del luogo dove si recano per gli studi universitari» (Del Puente, 2018: 131).

Dunque, nonostante l’alto numero di dialettofoni, soprattutto nella generazione più anziana, la pressione dell’italiano sui dialetti lucani e il conseguente contatto tra la varietà alta e quelle basse potrebbero condurre alla perdita dei tratti più peculiari di ciascuna parlata o, nella peggiore delle ipotesi, all’italianizzazione delle varietà locali, il processo «per cui i dialetti tendono ad evolversi verso una riduzione della loro distanza strutturale con l’italiano, a tutti i livelli di analisi» (Ricca, 2010).

Il presente studio si propone in primo luogo di verificare l’impatto dell’italiano su alcune varietà linguistiche lucane, soprattutto a livello fonetico e fonologico, con particolare attenzione alla lingua dei parlanti più anziani tradizionalmente custodi delle forme più conservative e, in secondo luogo, di dimostrare che l’acquisizione di alcuni modelli fonetici dell’italiano avviene spesso a scapito di caratteristiche delle varietà locali percepite come marcatamente dialettali proprio perché non presenti nel repertorio fonologico dell’italiano e quindi percepite dai parlanti come “volgari”.

2. L’inchiesta

2.1 Metodologia

La metodologia adoperata rispecchia quella adottata nell’ambito delle ricerche per l’Atlante Linguistico della Basilicata (d’ora in poi indicato come A.L.Ba.).

I comuni lucani considerati nella ricerca sono 7: Atella, Avigliano, Cancellara, Oppido Lucano, Ruvo del Monte, San Fele, Venosa. Per ogni comune sono stati scelti 10 informanti di entrambi i sessi, di età compresa fra i 65 e i 100 anni sempre vissuti nel paese di nascita, figli di genitori anch’essi nati e vissuti *in loco*, e di diversi gradi d’istruzione per verificare eventuali differenze legate a questo parametro. Tutti i parlanti hanno il dialetto come L1. Di seguito i dati quantitativi relativi agli interlocutori coinvolti:

Tabella 1 - *Dati informanti per ciascun comune*

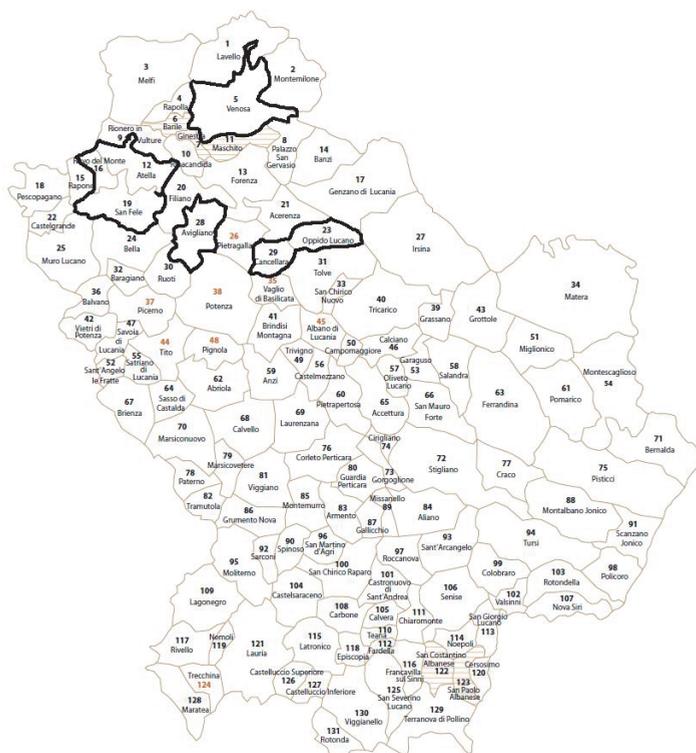
Paesi indagati	Numero	Informanti	
		Età	Sesso
Atella	10	70 – 85 anni	5 donne 5 uomini
Avigliano	10	65 – 96 anni	5 donne 5 uomini
Cancellara	10	65 – 90 anni	5 donne 5 uomini
Oppido Lucano	10	65 – 90 anni	5 donne 5 uomini
Ruvo del Monte	10	65 – 100 anni	5 donne 5 uomini
San Fele	10	65 – 100 anni	5 donne 5 uomini
Venosa	10	65 – 100 anni	5 donne 5 uomini

Una volta selezionati gli informanti, si è proceduto all'analisi di conversazioni libere registrate durante le interviste, per un totale di circa trenta ore, e di dati ottenuti attraverso la somministrazione di un questionario mirato per il nostro scopo. Le interviste sono state registrate grazie a un supporto professionale, *sony IC recorder*, e poi trasferite su pc. Sono state poi da me riascoltate, sbobinate e trascritte in grafia IPA per l'analisi dei dati. Si è tenuto conto, inoltre, delle valutazioni che i parlanti hanno espresso circa alcune caratteristiche fonetiche proprie della loro lingua madre, emerse durante le interviste. Lo studio si avvale di inchieste condotte nell'arco di cinque anni e utilizza parole-spia appartenenti a diversi campi semantici per un totale di 401 lessemi:

Tabella 2 - *Dati sui campi semantici indagati*

Campi semantici indagati	Numero lessemi
Nomi di parentela	34
Parti del corpo	70
Difetti e pregi fisici	36
Tempo non meteorologico	49
Tempo umano (età dell'uomo)	57
Tempo meteorologico	23
Numerali	33
Parti della casa	19
Mobili e arredi	35
Utensili domestici	45

Figura 1 - *Cartina della Basilicata riportante la numerazione progressiva di A.L.Ba. con indicazione dei comuni indagati*



2.2 Analisi dei dati

2.2.1 Il livello fonetico e fonologico: il caso studio

L'impatto della lingua tetto sulle varietà lucane è evidente a livello fonetico e fonologico, ma anche a livello lessicale, come si vedrà più avanti. L'impressione generale è che i modelli fonetici dell'italiano stiano sostituendo quelli dei singoli dialetti. Ad esempio, in alcuni dialetti, secondo i dati dell'A.L.Ba., si registra la graduale sostituzione di (-)[ʏ]- < lat. (-)G + A, O, U- con (-)[g]-, come è possibile osservare a Venosa dove si registra [ˈgran:ə]/[ˈʎran:ə] 'grande (d'età)', [ˈgrɔs:ə]/[ˈʎrɔs:ə] 'grande (di corporatura) (f. sg./pl.)', [ˈsɔgrə]/[ˈsɔʎrə] 'suocera', [ˈgardzə]/[ˈʎardzə] 'guancia', [aˈgostə]/[aˈʎostə] 'agosto'. Un altro esempio è fornito dalla progressiva sostituzione di (-)[ð]- < lt.(-)D- con (-)[d]-, come si può osservare dai seguenti esempi nel dialetto di San Fele: [ˈðotə]/[ˈdotə] 'dote', [ˈwɛðəvə]/[ˈwɛdəvə] 'vedova', [ˈsudə]/[ˈsudə] '(tu) sudi', [ˈðif:ətə]/[ˈdif:ətə] 'dito'.

Il mutamento più significativo che si può registrare nei dialetti oggetto del presente studio riguarda, però, gli esiti di lat. (-)DJ-, (-)G + E, I-, (-)J-. Si tratta di un

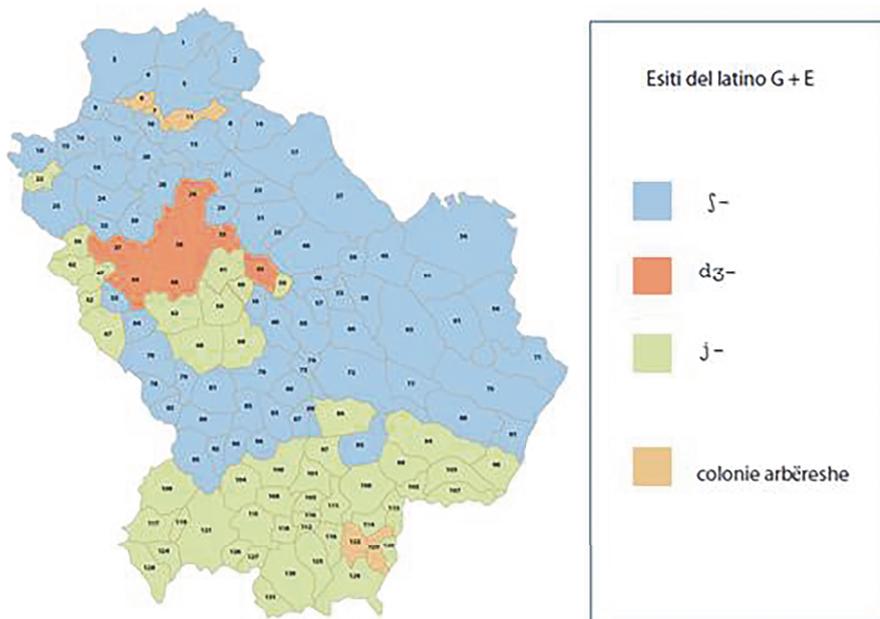
mutamento interessante soprattutto perché, come si vedrà, sembra essere in una fase iniziale e sembra seguire uno schema ben preciso.

In Lucania l'esito (-)[dʒ]- < (-)DJ-, (-)G + E, I-, (-)J- è proprio dei dialetti galloitalici (Figura 2 e 3), qui infatti si registrano, ad esempio, [ˈdʒenərə] 'genero' [dʒərˈnarə] 'giornata' [ˈwordʒə] 'orzo' (i dati si riferiscono al dialetto di Vaglio Basilicata). Nei dialetti indagati, invece, tutti situati nella parte settentrionale e nord-occidentale della regione, l'esito comune è (-)[ʃ]-, pertanto avremo [ˈʃjenərə] 'genero', [ʃərˈnatə] 'giornata' e [ˈworʃə] 'orzo' (i dati si riferiscono al dialetto di Avigliano, ma riguardo all'esito in questione, esso è condiviso da tutte le parlate investigate). In questi dialetti è possibile registrare per poche parole anche l'esito (-)[j]- (si veda figura 2), che però sembra essere "importato" dal napoletano, dal momento che «le voci con esito [j] [...] risultano [...] accomunate da connotazioni latamente "cittadine" o, quanto meno, "sovralocali"» (Fanciullo, 2013: 34).

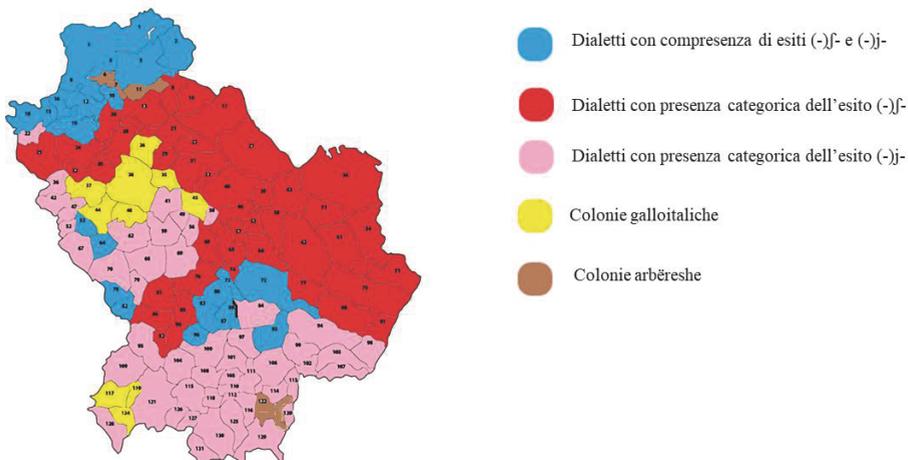
In questi idiomi (-)[dʒ]- è tuttavia presente:

1. come variante forte di /ʃ/ (Fanciullo, 1989: 671-673), come ad esempio in [rə d:ʒəˈnək:jə] 'le ginocchia' ~ [u ʃəˈnuk:jə] 'il ginocchio' o [ndʒəˈnuk:jə] 'in ginocchio';
2. come sonorizzazione post-nasale di /tʃ/, come ad esempio in [ˈndʒində] 'incinta';
3. in prestiti della lingua letteraria, come ad esempio in [ˈd:ʒendə] 'gente'¹.

Figura 2 - *A.L.Ba., vol. I, "Genero" sez. I - c. 26 bis*



¹ In questi casi [dʒ] si registra sempre intensa.

Figura 3 - Carta mutuata dai dati presenti in *A.L.Ba., I, II, III, IV voll*

Si riportano di seguito alcuni dati emersi dalle ricerche effettuate:

Atella

- [sar'taʃənə] 'padella';
- ['tʃel:ə] 'pentola a due manici';
- ['riʃətə] 'dito';
- [ʃə'nuk:jə] 'ginocchio';
- ['ʃjenərə]/['ʃinərə] 'genero';
- [n:at:sə'keʃə] '(lei/lui) culla';
- [ʃə'lamə] 'brina';
- [ʃəla'turə] 'gelo';
- [tsup:sə'keʃə] '(lei/lui) zoppica';
- ['oʃə] 'oggi';
- [d:ʒə'n:ajə] 'gennaio';
- ['mad:ʒə] 'maggio';
- ['d:ʒuŋ:ə] 'giugno';
- [sta'd:ʒonə] 'estate';
- [d:ʒuvə'rijə] 'giovedì'.

Avigliano

- [sar'taʃənə] 'padella';
- [ti'eɖ:ə] 'pentola a due manici';
- ['riʃtə] 'dito';
- [ʃə'nuk:jə] 'ginocchio';
- ['ʃjen:ərə] 'genero';
- [ʃə'lama] 'brina'
- ['ɔʃə] 'oggi';
- [ʃə'n:arə] 'gennaio';
- ['maʃə] 'maggio';
- ['d:ʒuŋ:ə] 'giugno';

[sta'd:ʒonə] 'estate';
[ʒuvə'rijə] 'giovedì'.

Cancellara

[sar'taʃənə] 'padella';
[ti'ed:a] 'pentola a due manici';
[riʃtə] 'dito';
[ʒə'nuk:ʒə] 'ginocchio';
[ʃinərə] 'genero';
[ʃilə] 'gelo'/'brina';
[tsup:ə'kiʃə] '(lui) zoppica';
[ʎuʃə] 'oggi';
[ʒə'n:arə] 'gennaio';
[maʃə] 'maggio';
[d:ʒuŋ:ə]/[ʃuŋ:ə] 'giugno';
[sta'd:ʒunə] 'estate';
[ʒuvə'rejə] 'giovedì'.

Oppido Lucano

[sar'taʃənə] 'padella';
[ti'el:ə] 'pentola a due manici';
[ðiʃtə] 'dito';
[ʒə'nuk:ʒə] 'ginocchio';
[ʃjen:ərə] 'genero';
[oʃə] 'oggi';
[ʒə'lata] 'brina';
[ʒə'n:arə] 'gennaio';
[maʃə] 'maggio';
[d:ʒuŋ:ə] 'giugno';
[sta'ʃonə] 'estate';
[ʃwovə'ði] 'giovedì'.

Ruvo del Monte

[sar'taʃənə] 'padella';
[ti'el:ə] 'pentola a due manici';
[riʃ:ətə] 'dito';
[ʒə'nuk:ʒə] 'ginocchio';
[ʃjen:ərə] 'genero';
[ʒə'lamə] 'brina';
[ʒlə'turə] 'gelo';
[tsup:ə'kejə] 'zoppica';
[oʃə] 'oggi';
[d:ʒə'n:ajə] 'gennaio';
[mad:ʒə] 'maggio';
[d:ʒuŋ:ə] 'giugno';
[sta'd:ʒonə] 'estate';
[d:ʒuvə'ri] 'giovedì'.

San Fele

[sar'taʃənə] 'padella';
 [ti'el:ə] 'pentola a due manici';
 ['diʃ:ətə] 'dito';
 [ʃə'nuk:jə] 'ginocchio';
 ['ʃjenərə] 'genero';
 [ʃə'ləmə] 'brina';
 [tsup:ə'kejə] 'zoppica';
 ['oʃə] 'oggi';
 [d:ʒə'n:əjə]/[ʃə'n:arə] 'gennaio';
 ['mad:ʒə]/['maʃə] 'maggio';
 ['d:ʒuŋ:ə] 'giugno';
 [sta'd:ʒonə] 'estate';
 [d:ʒuvə'di] 'giovedì'.

Venosa

[sar'taʃənə] 'padella';
 [ti'ed:ə] 'pentola a due manici';
 ['rɪʃtə] 'dito';
 [ʃə'nək:jə] 'ginocchio/-a';
 ['ʃinərə] 'genero';
 [nat:sə'kejə] '(lei/lui) culla';
 [ʃə'ləmə] 'brina';
 ['ʃilə] 'gelo';
 [tsup:ə'kejə] 'zoppica';
 ['oʃə] 'oggi';
 [d:ʒə'n:əjə] 'gennaio';
 ['mad:ʒə] 'maggio';
 ['d:ʒuŋ:ə] 'giugno';
 [sta'd:ʒonə] 'estate';
 [d:ʒuvə'dejə] 'giovedì'.

Come si può osservare dagli esempi qui sopra riportati, nel lessico inerente al tempo cronologico (-)[dʒ]- si registra in luogo di (-)[ʃ]- < (-)DJ-, (-)G + E, I-, (-)J-². Tuttavia, alcuni lessici e dizionari compilati da autori locali registrano per questi dialetti, negli stessi lessemi, l'esito atteso [ʃ] (come per esempio D'Annucci, 1996). Si tratta senza dubbio di una conseguenza dell'influenza della lingua italiana, veicolata dai mass media e dal sistema d'istruzione, come si è accennato al paragrafo 1, e parlata dalle nuove generazioni; è da escludere, invece, un'influenza dei dialetti galloitalici, in quanto i comuni nei quali questi sono parlati, tra cui il capoluogo di regione, non costituiscono centri di irradiazione di modelli linguistici (Lüdtke,

² L'unico lessema del tempo non meteorologico nel quale continua a registrarsi sempre l'esito autotono è 'oggi'.

1979: 8) e, inoltre, è in atto un processo di perdita dei tratti marcatamente settentrionali a favore di tratti meridionali o di tratti della lingua italiana³.

Ci sembra importante evidenziare un altro dato: nelle parole in cui si registra (-)[d₃]-, la soluzione (-)[ʃ]-, evidentemente più antica e peculiare dei dialetti analizzati, da me suggerita agli interlocutori per verificarne la conoscenza passiva, è ritenuta “volgare” dagli informanti, indipendentemente dal sesso e dal grado d’istruzione, ed è associata all’uso di altre comunità ritenute più arretrate. Queste ultime, per gli interlocutori, sono rappresentate da quei coloni arrivati *in loco* per la riforma agraria da comuni lucani più o meno lontani, come ad esempio la comunità di aviglianesi insediata nel comune di Atella o quella dei sanfelesi nel comune di Ruvo del Monte (effettivamente, dai dati risulta che nel dialetto di Avigliano e in quello di San Fele vi è una maggiore conservatività dell’esito [ʃ] anche nel lessico del tempo). Di contro, gli interlocutori sostengono che l’esito (-)[d₃]-, che definiscono più “civile”, sia proprio del loro dialetto.

2.2.2 Fonetica e lessico: una possibile spiegazione

Non si può tuttavia analizzare il livello fonetico e fonologico senza contestualmente aver dato uno sguardo alla situazione generale del lessico.

Come noto, il lessico è la parte del sistema linguistico che maggiormente è esposta al contatto con altre lingue ed è la porta d’ingresso delle innovazioni dell’ambito fonetico e fonologico, e spesso anche dell’ambito morfologico (si veda Gusmani, 1987).

Dall’analisi dei dati, si registra una massiccia acquisizione di lessemi dall’italiano evidentemente dialettizzati. Si tratta molto spesso di prestiti che affiancano i corrispondenti dialettali, i quali, per conseguenza, vanno incontro a un restringimento di significato⁴ come accade, ad esempio, per i lessemi che esprimono il significato di ‘incinta’ e ‘pelle’. Per la prima delle due parole citate sono stati registrati i termini [ˈprenə] e [ˈndʒində]⁵. Di questi il primo, chiaramente dialettale, si usa solo in riferimento agli animali, mentre il secondo, prestito italiano, si usa solo per gli esseri umani. Nel caso di ‘pelle’, sono stati registrati [ˈpɛd:ə] e [ˈpɛl:ə]. La prima forma, con esito fonetico dialettale, indica la pelle degli animali, mentre la seconda, prestito italiano, indica la pelle umana.

Ma molti sono anche i casi di perdita di lessemi dialettali sostituiti da prestiti italiani. Il campo semantico nel quale si registra il maggior numero di casi del genere è, ancora una volta, quello del tempo cronologico. Fino a qualche tempo fa per i nomi dei mesi dell’anno o per le parti della giornata, nei dialetti indagati, erano ancora conosciuti lessemi collegati a una società di tipo agricolo e pastorale o collegati alla

³ Tra i tratti galloitalici più caratteristici, attualmente in regressione, vi è la lenizione delle occlusive sorde intervocaliche: nel dialetto potentino, accanto a forme lenite come [aˈtʃirə] (a Potenza -[d]- < -T- subisce rotacismo come -D- primaria), si registrano e sono assai più diffuse forme non lenite, come ad esempio [aˈtʃitə] (Fanciullo, 1989: 675), e le prime sono considerate arcaiche; altrettanto regressiva è la caduta della sillaba finale composta da nasale [n] + vocale: forme come [pa] ‘pane’, registrate da Rohfs negli anni ’30, sono ormai state sostituite da forme come [ˈpanə] (Del Puente, 2007: 78).

⁴ Processo descritto come “conservazione specializzata” in Massariello Merzagora, 1985: 431.

⁵ Si prescinde in questa sede dal riportare tutte le diverse varianti fonetiche del singolo lessema.

religiosità. Di seguito si riportano alcuni esempi tratti dal II volume dell'A.L.Ba. e riferiti al dialetto di San Fele: [tʃərə'salə] 'giugno' così chiamato perché le ciliegie venivano raccolte soprattutto in questo mese; [mətə'torə] 'luglio' così chiamato per l'attività di mietitura tipica del periodo; ['sandə mar'tinə] 'novembre' dalla ricorrenza dell'11 novembre, giorno nel quale si aprivano le botti del vino novello; [na'talə] 'dicembre' dal nome della festa religiosa più importante del mese⁶.

Dalle inchieste condotte risulta che i parlanti, oggi, hanno solo conoscenza passiva di questi termini e che, per designare i mesi dell'anno, adoperano forme quali [d:ʒuŋ:ə], [l'uɫ:ə], [nə'vembərə] e [di'tʃembərə].

Il motivo per il quale il lessico del tempo non meteorologico risulta così innovativo e maggiormente esposto all'influenza della lingua tetto sta probabilmente nel fatto che è stato adoperato dai parlanti dialettofoni in molti contesti "istituzionali", quali la burocrazia e la scuola, settori di maggior contatto fra le due varietà (Cortelazzo, 1972: 44). D'altro canto, venuta meno la società agricola e pastorale di un tempo, sono evidentemente venuti meno anche tutti i riferimenti ad essa nell'ambito delle denominazioni del tempo.

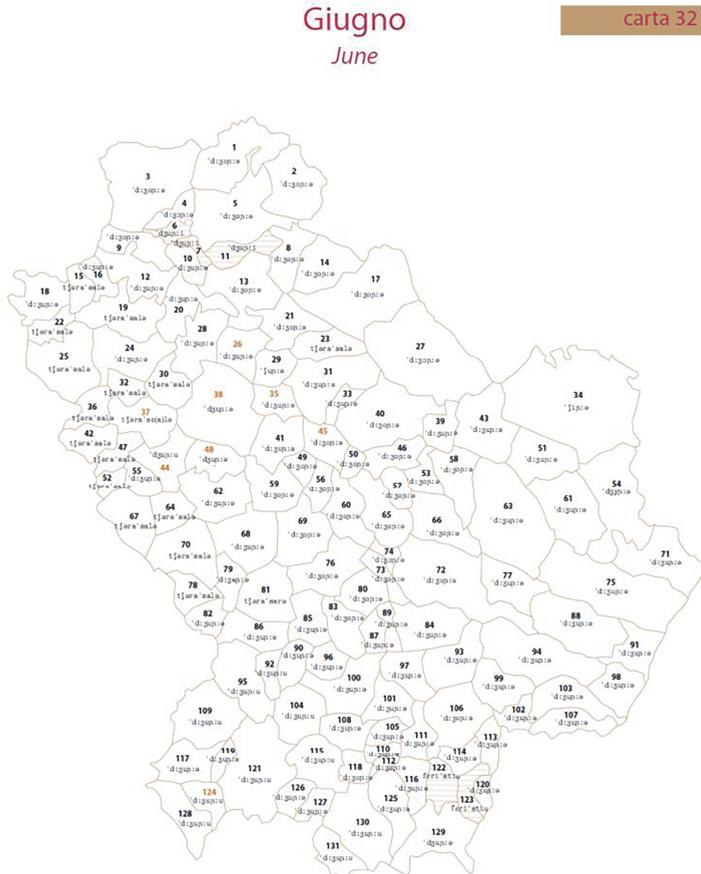
Tornando alla nostra ricerca, si cercherà di dare una spiegazione alla diffusione di [dʒ], alla luce di quanto visto fin qui. La sua diffusione all'interno del lessico del tempo non meteorologico si potrebbe spiegare proprio attraverso la dinamica della diffusione lessicale.

L'innovazione dei modelli fonetici può essere partita nello specifico proprio dal lessema utilizzato per indicare il mese di giugno. Se si osserva la c. 32, sez. II del II vol. dell'A.L.Ba. (Figura 4) si potrà osservare che il tipo [d:ʒuŋ:ə]⁷, con [dʒ], è stato registrato pressoché in tutti i dialetti della regione, anche in quelli che dovrebbero presentare l'evoluzione (-)J- > (-)[j]-. Si sarà trattato, in questi casi⁸, di un prestito importato dall'italiano (un prestito necessario, in quanto l'avvento della società contemporanea, non più agricola e pastorale, rendeva incomprensibile forme come [tʃərə'salə]), che potrebbe aver fatto da apripista all'affermazione di [dʒ] in altre parole appartenenti alla serie dei lessemi relativi al tempo non meteorologico, *in primis* nei nomi indicanti i mesi dell'anno, in questo facilitato anche dalla comune etimologia dei lessemi dialettali e di quelli italiani (si pensi alla conseguente vicinanza strutturale delle forme italiane e dialettali per 'gennaio', 'maggio', etc.).

⁶ A questi nomi si aggiungono, in altri comuni, le denominazioni [urtu'lanə] 'maggio', così chiamato perché in questo periodo venivano messe a dimora le piantine per l'orto, e ['sandə 'lu:kə] 'ottobre', dalla ricorrenza del 18 ottobre, giorno nel quale, in tempi antichi, iniziava la transumanza del bestiame dai paesi interni, montuosi e collinari, verso le zone pianeggianti della costa.

⁷ Si prescinde in questa sede dal riportare tutte le diverse varianti fonetiche del singolo lessema.

⁸ Per il dialetto di Cancellara, la variante [ʃuŋ:ə] (che si registra in co-occorrenza con [d:ʒuŋ:ə]) potrebbe essere un prestito proveniente dalla vicina Puglia, penetrato in qualche modo attraverso lo stretto contatto con l'area pugliese.

Figura 4 - *A.L.Ba., vol. II, "Giugno" sez. II – c. 32*

Di qui, successivamente, si sarebbe diffuso anche in altri termini relativi al lessico del tempo e così si può immaginare che gradualmente si diffonderà anche in altri campi semantici.

3. *Conclusione*

Come si è potuto osservare, anche i dialetti lucani, come altre varietà italo-romanze (si veda Scivoletto, 2014; Cerruti, 2016), sono interessati da processi di italianizzazione.

Dall'analisi effettuata, pare di poter dire che le innovazioni provenienti dalla lingua tetto si affermano grazie al prestigio che tale varietà detiene all'interno della società lucana, di contro alla svalutazione cui è stato soggetto il dialetto.

La dimostrazione risiede nel fatto che alle forme italiane viene sempre assegnata una valenza positiva, mentre alle forme dialettali una valenza negativa, "bassa".

Prova ne è l'affermazione nel lessico del tempo di [dʒ], considerata, dagli stessi parlanti, forma "civile", e la contestuale stigmatizzazione di [ʃ], connotata, invece, come forma "volgare" propria di comunità più arretrate, legate ancora a un tipo di economia agricola e pastorale; anche gli esempi riportati per il livello lessicale mostrano o la perdita di interi lessemi legati al mondo contadino (e la loro sostituzione con italianismi) o la loro relegazione al mondo animale (conseguentemente all'acquisizione di italianismi che ricoprono significati relativi al mondo umano). D'altro canto, come si è detto nel paragrafo 1, al dialetto sono stati da sempre associati gli analfabeti e i contadini, mentre all'italiano i galantuomini e i professionisti; si spiega così la scelta da parte delle generazioni più anziane di non trasmettere la lingua madre ai giovani (si veda paragrafo 1 e Del Puente, 2018: 131).

Dunque, l'adozione di elementi dell'italiano deve aver rappresentato una forma di promozione sociale⁹ ed emancipazione dalla povertà e dall'analfabetismo che hanno caratterizzato in passato le comunità lucane.

Riferimenti bibliografici

- BERRUTO, G. (1987), Lingua, dialetto, diglossia, dilalia, in Holtus, G. (Eds.), Kramer, J., *Romania et Slavia Adriatica. Festschrift für Žarco Muljačić*, Hamburg: Buske, 57-81.
- BERRUTO, G. (2018), The languages and dialects of Italy, in AYRES-BENNET, W., CARRUTHERS, J., *Manual of Romance Sociolinguistics*, Berlin/Boston: De Gruyter, 494-525.
- CERRUTI, M. (2016), L'italianizzazione dei dialetti: una rassegna, *Quaderns d'Italià*, 21, 63-74.
- CORTELAZZO, M. (1972), *Avviamento critico allo studio della dialettologia italiana*. Lineamenti di italiano popolare, vol. III, Pisa: Pacini.
- D'ANNUCCI, T. (1996), *Atella. Del villaggio pre-globale*, Atella (PZ): Basiliskos editrice.
- DE BLASI, N. (1994), *L'italiano in Basilicata*, Potenza: Il Salice.
- DEL PUENTE P. (2007), Il dialetto di Potenza: un'analisi in tempo reale, in *L'Italia Dialettale. Rivista di dialettologia italiana*, LXVIII.
- DEL PUENTE P., (2010-2011), *Atlante Linguistico della Basilicata (A.L.Ba.)*, I e II voll., Rionero in Vulture (PZ), CalicEditori.
- DEL PUENTE P., (2015-2018), *Atlante Linguistico della Basilicata (A.L.Ba.)*, III e IV voll., Lagonegro (PZ), Zaccara Editore.
- DEL PUENTE, P. (2018), Varianti sintopiche e diastratiche in Basilicata, in MARCATO, G. (Ed.), *Dialetto e società*, Padova: Cleup, 125-134.
- DE MAURO, T. (2017), *Storia linguistica dell'Italia unita*, Bari: Laterza.
- FANCIULLO F. (1989), Lukanien, in HOLTUS, G., METZELTIN, M., SCHMITT C. (Eds.), *Lexicon der Romanistischen Linguistik 4*, Tübingen: Niemeyer, 669-688.

⁹Questo aspetto è stato osservato anche da Pautasso, 1969: 105 a proposito degli immigrati a Pettinengo.

- FANCIULLO, F. (2013), Il napoletano, gli altri dialetti meridionali e i relativi, reciproci rapporti. Qualche considerazione, in *Andirivieni linguistici nell'Italo-romània*, Alessandria: Edizioni dell'Orso, 29-50.
- GUSMANI, R. (1987), Interlinguistica, in LAZZERONI, R. (Ed.), *Linguistica storica*, Roma: Nuova Italia Scientifica, 87-146.
- Istat, L'uso della lingua italiana, dei dialetti e delle lingue straniere (2015), https://www.istat.it/it/files//2017/12/Report_Usoitaliano_dialetti_altrelingue_2015.pdf
- LÜDTKE, H. (1979), *Lucania*, Pisa: Pacini editore.
- MASSARIELLO MERZAGORA, G. (1985), Fenomeni di italianizzazione del dialetto nell'area lombarda, in AGOSTINIANI, L., BELLUCCI MAFFEI, P., PAOLI, M. (Eds.), *Linguistica storica e cambiamento linguistico*, Roma: Bulzoni, 427-438.
- PAUTASSO, M. (1969), *Dialetto, lingua e integrazione linguistica a Pettinengo*, Torino: Giappichelli.
- RICCA, D. (2010), Italianizzazione dei dialetti, in Simone, R. (Ed.), *Enciclopedia dell'italiano*, Roma: Istituto dell'Enciclopedia Italiana Treccani.
- SCIVOLETTO, G. (2014), L'italianizzazione dei dialetti come caso di convergenza linguistica, in *Rivista Italiana di Dialettologia. Lingue dialetti società*, a. XXXVIII.

MANUELA FRONTERA

Calabrian migrants in Argentina: assessing first language phonetic attrition

This paper discusses the research plan and first results of the CMA project (*Calabrian Migrants in Argentina*), aimed at exploring linguistic interaction processes active in first generation Italian-Argentinian migrants from Central-South Calabria and living in the provinces of Buenos Aires, Córdoba and Santa Fe from the 1950s onwards. These are the last living people who offer evidence of linguistic maintenance/attrition processes affecting both L1 Calabrian dialects¹ and L2 Calabrian regional Italian (both homeland linguistic resources) after decades of permanence in their L3-dominated (Spanish) host country.

Keywords: language contact, Italian migrants in Argentina, heritage languages, L1 attrition, VOT.

1. Introduction

The massive migratory waves, which affected the global population between the 19th and 20th centuries, lasting until almost thirty years after the end of the Second World War (1876-1978), saw a huge number of Italians leaving their native country to reach the coasts of Argentina. Among these, more than 450,000 came from Calabria, which has been the region of origin for the majority of Italian-Argentines for generations (in 1973, 60 Italian immigrants out of 100 were from Calabria, and 42 ten years later). They settled mainly in Buenos Aires, in the territories of Pampa Gringa, spreading up to the littoral zone of the Paraná River (Rosario) and the cities of Córdoba, Mendoza and Tucumán. Many people moving in their early adulthood with the very last wave (1946-1960) still live there with their families experiencing a long-lasting language contact process, started at the time of their settlement and continuing over the years.

These speakers can be referred to as first generation migrants (Nagy, 2015), speakers who had lived at least until the age of puberty in their homeland, subsequently moved to Argentina, where they still live. These immigrants can be identified as trilingual speakers, since their linguistic repertoire involves at least: a heritage language² (Calabrian dialects of Area 3 or 4, according to Trumper's (1997) parti-

¹ Coeval geographical varieties of the dialect from which the standard language descends, *primary dialects* in the sense expressed by Coseriu (1980).

² Language acquired spontaneously and available to young children, which is spoken at home but which is not a dominant language within the social community. See Bayram, Pascual y Cabo & Rothman (2019) for broader definitions of heritage language and heritage speakers.

tion) acquired as mother tongue; Calabrian regional variety of Italian – acquired and or learned in Italy – and Argentinian Spanish, being the language of their host country. Despite the minority-language status and the unofficial role generally ascribed to heritage languages (as commented by Rothman, 2009 or Nagy, 2015), Calabrian dialects for these migrants represent the first language acquired at home while they accessed to regional Italian at school or in formal situations, and only some among them still maintain a passive competence mainly through exposure to media (such as Tv programs).

Thus, from a chronological perspective, Calabrian dialects are informants' L1³; nevertheless, Spanish has turned to be the dominant code across their life span both in private and public domains. Consequently, various kinds of linguistic change may have occurred. These processes have been scarcely investigated for late Italian-Argentinian speakers so far. As a matter of fact, greater attention has been addressed to language contact phenomena as far as *cocoliche* or *lunfardo* are concerned (see Cancellier, 1994). Therefore, it could deserve attention the extent to which heritage languages have experienced change: namely in what direction and under what specific conditions Spanish has affected migrants' linguistic resources (dialects and Calabrian Italian) mainly as far as interlanguage phonology is concerned.

In Section 2 of this paper, the research plan for the CMA project will be outlined. In Section 3, the materials and methodological framework of the entire project will be presented. In Sections 4 and 5 the very first experimental analysis on phonetic-acoustic data (VOT measurements) will be described and results will be discussed within the scope of L1 phonetic attrition.

2. *An outline of the CMA project*

The *Calabrian Migrants in Argentina* (CMA) project focuses on the linguistic interaction processes active in Italian migrants from Central-South Calabria who have been living in Argentina from the second half of the 19th century onwards. CMA is the acronym used along this paper to refer to the research project “Effects of intragenerational linguistic contacts on Italian migrants in Argentina”, carried out by the author of this paper as Principal Investigator.

2.1 Scientific national and international background

The starting point for this research is a study by Romito, Graziano, Frontera, Tarasi, Ciardullo, Piemonti and Strano (2016) which investigates the linguistic contaminations of Calabrian dialects in contact in Argentina. This study outlined both cases

³ A language acquisition chronological criterion is adopted along the whole article to identify informants' L1, L2 and L3. This classification is functional in order to analyse and interpret all possible outcomes of language interference. On this regard, it is worth mentioning that the distinction between local dialects and regional Italian is not always clear to speakers, who do not properly identify definite boundaries within their regional linguistic space (see Romito et al., 2016).

of maintenance and apparent change of typical morpho-syntactical and phonological Calabrian dialect features, not only due to interference from other dialects of the same area, but also to attrition triggered by Spanish. This evidence has given rise to other research questions related to long-lasting plurilingual contact. Since L1 phonetics and phonology are the most resistant levels to be affected when acquiring other languages in adulthood (see Romaine, 1984), attrition in this area has been considered worth-noting. Though much research has been conducted in this field, only few studies have focused on phonetic/phonological attrition concerning Italian varieties and Italian dialects. Moreover, studies carried out so far have considered productions by Italian migrants in Australia and in North America, therefore focusing on the linguistic couple Italian-English, while the Italian-Argentinian community has not been investigated from this phonetic-phonological perspective. Some inspiring works, such as Avesani, Galatà, Vayra, Best, Di Biase, Tordini & Tisato (2015) and Avesani, Galatà, Best, Di Biase, Vayra & Ardolino (2017), observe phonetic attrition in Italian-Australian immigrants' production of coronal obstruents, taking into account possible effects of English as L3 on both dialect L1 (varieties of Veneto) and L2 Italian, as well as convergence in terms of standardization/dialectisation phenomena between immigrants' heritage and homeland languages. Thus, they show the importance of evaluating every possible evidence of phonetic drifts from/to each variety used in the first-generation community's repertoire. Nagy, Nodari, Celata (2018), Nodari, Celata, Nagy (2019) and Nagy (2015) examine voice onset time values across generations of various immigrants, including Italians (from Calabria), in Toronto. The latter includes a comparison as far as the null subject structure use is concerned, so presenting a multi-faceted approach to the study. More recently, De Iacovo, Mairano and Hajek (2020) explore levels of gemination produced by Italian migrants in Australia. Attrition has also been explored on the front of perception and foreign accent detection. In a pioneering work by Major and Baptista (2009)⁴, levels of attrition in terms of change of native-like pronunciation in the L1 of bilingual speakers were measured. Any study on Italian attrition has been based on such an approach that it will be applied in this work to contribute to the problem of the dearth of evidence in this field.

2.2 Theoretical framework: cross-language interference processes

The informants engaged in the present research are all trilingual speakers, first generation migrants, who moved to a new country and had to acquire another language in their adulthood. This gives rise to some brief, but necessary, considerations on the interference dynamics involved during the acquisition process of other languages⁵. Moreover, typological markedness (Maddieson 1984; Eckman 1977, 2008) may affect the acquisition and L1-L2 distance between the systems may slow down,

⁴ Followed by de Leeuw et al. (2010), Hopp & Schmid (2013) and Bergmann et al. (2016), for example.

⁵ References to L2 acquisition process throughout the paper are to be intended to refer to the natural acquisition process of languages (L2, L3, L4, etc.) other than the L1.

and even hinder, the acquisition process (Major 2014; Derwing and Munro 2015, among others).

Theories on maturational constraints and critical or sensitive period/s hypotheses state that interference can be stronger and more prolonged if associated with adult learners (see Lenneberg 1967; Patkowski 1994; Ellis 2003; Long 2008; de Leeuw, Mennen & Scobbie, 2013).

Nevertheless, these claims are counter-balanced by extensive research, which demonstrate the fundamental role of psycho-social factors, such as speakers' attitudes, motivation, or sense of identification/assimilation (see Weinreich 1957; Bongaerts et al. 1997; Moyer, 2007; Major 2014, among others), as well as extra-linguistic variables relevant to second language acquisition: age of acquisition, length of residence, sex, usage of L1 and L2 (Flege et al. 1997; Piske et al. 2001; Flege 2009).

A native-like phonological competence represents the most challenging achievement for non-native speakers, whose perceptual and productive (both at phonological and phonetic level) are guided and deeply filtered by their native language and languages previously acquired (Best, 1995; Best, Tyler 2007; Mori 2007; Flege 1987, 1995; Brown 2000; Kuhl 2000). It has been demonstrated, however, that late bilinguals, create new phonological spaces through merging and equivalence classification processes (*Speech Learning Model* by Flege, 1987) and, in cases of constant and prolonged contact with both languages, they may suffer a bidirectional interference: from L1 to L2 and vice versa (Flege 1987; Major 1992). Moreover, the lack of L1 use in a non-native environment, counter-balanced by a largely widespread use of the L2 – especially within a multilingual context – may provoke functional and structural changes in the L1. This condition may lead to the construction of mid-point systems (see Mori, Barkat, 2005) intermediate realisations between L1 and L2 and complete or partial modifications of L1, in terms of shift (inter-generational) or attrition (intra-generational) (see Bartoli 1945; Cook 2003; Pavlenko 2004; de Leeuw et al. 2010; Schmid and Köpke 2013; Schmid et al. 2013, among others).

Language change is a wide-ranging phenomenon, which can act at different linguistic levels, from phonetics/phonology and morphology to syntax, semantics and lexical units (better defined as restructuring or borrowing). According to van Els' taxonomy (de Bot, Weltens, 1985), evidence can be observed both on L1 affecting L2 and viceversa. When this latter takes place, first language attrition is reported in literature: this is generally the consequence of a progressive overuse of, and increasing proficiency in L2, stronger in casual than in formal speech (see Major, 1992), but somehow reversible by training (Schmid, Köpke, 2013).

Therefore, L1 attrition can be considered as a complete or partial modification of L1 features, as a result of the interaction with another language, in post-adolescence (de Leeuw, 2019). Again, it is important to note that, in phonology, if it is so difficult to control L1 interference on L2, the same L1 mature phonological system will be much more stable and inaccessible to L2/L3 influence. That is why there exist some essential ideal conditions for attrition to occur: emigration, extensive use of the FL/reduced use of L1 in daily life and a long-time span (Schmid, Köpke, 2013;

Major, 1992). Moreover, any linguistic domain is differently affected by attrition (see Schmid, Köpke, 2019, for an exhaustive review), even if recent literature proposes the existence of a trickle-down effect of phonetic changes on other linguistic domains (*ibidem*).

2.3 Research scope

The CMA project aims at:

- a. detecting contact phenomena in the heritage dialects, in order to find evidence of maintenance/substitution/alternation of peculiar phonological, morphological and morpho-syntactical traits (see Frontera, in preparation);
- b. assessing L1 phonetic attrition, in order to understand to what extent Argentinian Spanish influences migrants' heritage code (local dialect) as far as a salient phonetic feature, such as voiceless stop aspiration, is concerned (see Section 4);
- c. evaluating foreign accent perception, in order to check the degree of perceived nativeness of heritage speakers' production in local dialects and regional Italian by native-Italian listeners.

The project is multidimensional, aiming at focusing both on production and perception. As far as the former is concerned, the analysis has been divided into two steps: the first explores the effects of intra-group language contact between migrants' Calabrian dialects and/or between these and regional Italian. Namely, by focusing on possible convergence or interference phenomena, occurring in their local dialects, at morpho-syntactical and phonological levels. The second one concerns a specific phonetic feature, the cross-linguistic VOT values of L1/L2/L3, and it will be dealt with in the preliminary analysis presented in this paper (Section 4 and 5). The aim is to verify whether attrition has occurred, so that the aspirated voiceless stops /p t k/ of dialect (realised as such in specific phonological contexts) are reduced to short-lag productions under the effect of Argentinian Spanish.

As for the perception analysis of L1 attrition, this will be tested in the near future by asking native Italian subjects (from the same areas of Calabria of Italian-Argentinian migrants) to judge levels of foreign accent detected in L1/L2 productions by the same migrants.

3. Methodology

The CMA project comprises 29 hours of recorded speech: spontaneous, semi-spontaneous and read in Calabrian dialects, regional Italian and Argentinian Spanish⁶, as well as oral translations from Italian/Spanish to dialects. The recordings were con-

⁶All subjects gave their informed consent for inclusion before they participated in the study. The study was conducted in accordance with the Declaration of Helsinki, and the protocol was approved by the Ethics Committee of the University of Calabria, Italy (project identification code CALA171337, n°4087).

ducted by means of a Shure WH20 headset microphone, a Tascam DR-100mkIII recorder, with a sampling frequency of 44.100 Hz, 24-bit, mono-channel.

3.1 Tasks and materials

Spontaneous and semi-spontaneous speech was collected through the accomplishment of two tasks: a questionnaire and an interview. The former was conceived in order to extract information related to age, sex, place of origin, education, occupation, age of immigration, LOR (length of residence in Argentina), L1 dialect, age of first contact with Italian, age of first contact with Spanish, formal instruction in Italian/Spanish. The latter was designed to derive information on migrants' psycho-social profile (L1/L2/L3 use, attitudes towards dialects and degree of integration in the host country culture).

Read speech was elicited via two tasks consisting of a 54-item wordlist and 12 sentences (containing the same target words) both in Italian and Spanish, respectively. These materials were chosen to perform the analyses on VOT durations, so the selected words were balanced in order to include /p/, /t/ or /k/ in any phonetic context of interest, as detailed in §4.3.

Translated speech was elicited through two translation tasks from Italian/Spanish, as source language, to informants' native dialects, as target language. The first task consisted of a word list of 36 items and 13 sentences containing the same words. These words subsumed the same target sounds and phonetic contexts as the ones used for the above-mentioned reading tasks. The second translation task was conceived to elicitate 30 isolated real words and 20 sentences. These words were partially selected from the Italian Linguistic Atlas (ALI, Aa.Vv., 1995) referring to villages and towns of the third corridor of Calabria, in order to control and check the glottalization/aspiration of /f/ (see Falcone, 1976; Rohlf, 1966; Trumper, 1997) and Latin LL outcomes (see Romito, Milelli, 1999; Romito, Soriano, 1998). The sentences were conceived in order to get other target features such as the enclitic possessive for singular forms (like in *sorma*, 'my sister'; see Krefeld, 2007; Manzini, Savoia, 2005), and the periphrastic infinitive constructions using the modal particles *ma/mu/u* + V (e.g. *hai ma/mu/u chiudi a porta*, 'you have to close the door'; see Damonte, 2009; De Angelis, 2014).

Trilingual speakers were asked to face three different recording sessions, one for each language involved. All sessions were preceded by a phase of attunement to the language under investigation.

As for the study of the perception side, sentences in dialect and Italian (the same used for the elicitation of VOT occurrences) by our informants and by a group of monolingual Italians (control stimuli, see §) will be used to build our future test on perceptual attrition. This test will be performed by Calabrian monolingual listeners, with the aim of verifying whether they still perceive them as Italians or, conversely, identify their voices as belonging to foreign speakers.

3.2 Participants

In order to lead our production analyses (i.e. (a) and (b) in § 2.3), 37 late trilingual first-generation Italian immigrants (23 females and 14 males, aged 70-88) were selected. All these speakers acquired Spanish at the moment of migration, which took place during or after their puberty. The trilingual participants were recruited among the members of five Italian-Argentinian associations, two in the province of Córdoba, two in the province of Santa Fe, one in Buenos Aires.

Two other groups were involved: the first is a control group of 6 monolingual Italian adults, males and females from the geographical areas of origin of the Italian-Argentinian migrants involved, and of similar age and level of education; the second control group comprises 6 Argentinian Spanish monolingual speakers, adult males and females living in Argentina. Control group participants were selected by means of a questionnaire (see 3.1) including questions about speakers' sociolinguistic profile (age, sex, place of origin, education, occupation, L1, other languages acquired, age of first contact with Italian and Spanish respectively). The control group of Italians was recruited in two villages in Calabria, one in Area 3, the other in Area 4 while the informants for the Spanish control group were from the capital city of Santa Fe.

As for the fulfilment of c) (see § 2.3) a group of 30 native Italian adult listeners from the same target dialectological areas of Calabria will be further selected to perform the perceptual test on foreign accent detection.

4. *Assessing L1 phonetic attrition: a preliminary analysis*

The aim of this first experiment is to verify the hypothesis of phonetic attrition (as explained in §2.3) in the production of late trilingual Italian-Argentinian migrants. Due to a massive and prolonged use of Argentinian Spanish, Voice Onset Time values produced by elderly Italian-Argentinian speakers may have suffered a structural reduction within their heritage language speech (see Flege, 1987; Major, 1992; de Leeuw, 2019).

In the present research, the first parameter used to “assess” a possible drift of Calabrian dialect L1 towards Spanish L3 values is voiceless stop consonants' aspiration. The acoustic correlates of VOT have shown themselves to be a precious index of specificity at a cross-linguistic level, being language-specific (see Cho, Ladefoged, 1999), thus they play a role for detecting degrees of interference or shift in cross-linguistic studies, above all when bilingual or plurilingual speakers are involved.

4.1 Evidences of stop aspiration in Calabrian varieties

Recent studies have confirmed acoustically the presence of stop aspiration within the Central-Calabrian language varieties, when the voiceless stops /p t k/ are geminated or in post-liquid (-lC-, -rC-) and post-nasal (-nC-, -mC-) position, in regional Italian (see Stevens, Hajek, 2010; Nodari, 2015; Frontera, 2018) as well as in dialect-

tal speech (Frontera, Tarasi, Graziano, 2019). According to the above-mentioned studies, post-aspiration seems to be a peculiar feature of two main dialectological areas of Calabria, namely Area 3 and 4 (see Trumper, 1997, 2016), even though this trait is sometimes associated with Calabrian regional Italian as well (Telmon, 1993; Fanciullo and Librandi, 2002; De Blasi, 2014). Nevertheless, the dialectal varieties across Areas 3 and 4 are the ones in which aspiration remains stable and stronger in every given phonetic context, even if long lag VOT (see Lisker & Abramson, 1964)⁷ still appears in Area 5 and Area 2, in post-nasal context (see respectively Frontera et al. 2019; Romito, Ciardullo, Tarasi, 2015). In addition, aspiration is found in Area 2 when stops are geminated or follow a rhotic sound (Soriano, 1996). Moreover, elderly speakers seem to retain this acoustic trait more than adolescents (Frontera, 2018; Nodari, 2015) and to preserve it even in cases of long-term immersion in another linguistic context, namely in the case of Italian speakers of L2 English living in Canada (see Nagy, Nodari, Celata, 2018; Nodari, Celata, Nagy, 2019).

On the other hand, long lag VOT is not present for unvoiced stops in the Argentinian Spanish variety, where only short lag VOT realisations are attested (Borzzone, 1980). Long VOT realisations in initial, non-initial or intervocalic position are not reported neither for the American varieties of Spanish (see Lisker and Abramson, 1964; Abramson and Lisker, 1973; Borzzone and Guerlekian, 1980; Flege and Eefting, 1987; Roldán and Soto-Barba, 1997; Soto-Barba and Valdivieso, 1999) nor for Peninsular Spanish (see Castañeda Vicente, 1986; Asensi et al., 1998; Amengual, 2012; Martínez Celdrán, 2013). According to previous research, Spanish VOT durations fluctuate between 8 and 28 ms on average, remaining stable under the limit of aspiration and in compliance with the variability related to the stop constriction point (VOT durations are very low in /p/, slightly higher in /t/ and significantly longer in /k/).

4.2 Research hypotheses

For first generation Calabrian migrants in Argentina, from the central area of the Italian region, the presence of unaspirated voiceless stops in Spanish L3 has already been reported (Frontera, 2019), even if their VOT values are, in some cases, significantly higher than those produced by native Argentinians. These VOT values are extremely lower if compared to those evidenced, in the same phonetic contexts, within the Italian speech of Calabrian peers who had always been living in Italy (Frontera, 2018; Frontera et al., 2019). However, this former reference work lacks primary acoustic data in Italian language and dialect. So, in order to confirm whether attrition has occurred or not, the ongoing research will improve and implement data from the aforementioned study, by adding a comparison between VOT values produced in dialectal voiceless stops by the same heritage speakers in Argentina. Thus,

⁷ According to Lisker and Abramson's taxonomy, the world languages all belong to three main VOT categories, where the long lag category is the one entailing post-aspiration, with VOT durations equal or superior to 60 milliseconds.

the hypothesis for the present research is that, due to a massive and prolonged use of Spanish, Voice Onset Time values in the production of elderly Italian-Argentinian speakers may have suffered a structural reduction in their heritage language speech. As far as Spanish as L3 is concerned, effects of this heritage phonetic correlate could not be evident, though it is possible that VOT values do not reach the low degree of aspiration featuring Argentinian Spanish native productions⁸. The attested variability at intra-personal and inter-personal levels in attrition studies, however, may lead to unexpected results or exclude a unidirectionality. In order to take in due consideration all this, Argentinian Spanish and dialectal realisations of /p t k/ will be investigated in relation to VOT durations and a comparison among the two varieties will be outlined.

4.3 Selected materials

The sample for this preliminary research comprises 10 (out of 37) trilingual migrants, 5 females + 5 male speakers aged 80 (on average) and living in Córdoba, Santa Fe, Rosario and Buenos Aires. They all belong to the dialectological area 3 of Calabria (see Table 1). Their age of arrival (AoA) in Argentina varies from 9 to 20 years, while their LoR (length of residence) amounts to 65 years averagely. Most of these speakers can be also referred to as *late child L2 learners*, “children monolingual in the HL⁹, who received some elementary schooling in their home country” before migration (Montrul, 2011: 157). Indeed, they all learnt formally Italian at least for three years, only one among them attended secondary school in Italy (length of education in the country of origin).

Table 1 - *Informants' sociolinguistic information*

ID ¹⁰	Age	Dialect	AoA	LoR	LoE
TRFIII2	77	Badolato (CZ)	13 Y	54 Y	4 Y
TRMIII1	73	Gasperina (CZ)	9 Y	63 Y	3 Y
TRFIII4	74	Petronà (CZ)	9 Y	63 Y	4 Y
TRFIII5	75	San Sostene (CZ)	10 Y	64 Y	3 Y
TRMIII4	80	San Sostene (CZ)	16 Y	65 Y	4 Y
TRFIII6	85	Davoli Marina (CZ)	20 Y	65 Y	7 Y
TRMIII6	88	Satriano (CZ)	17 Y	71 Y	5 Y
TRMIII7	85	Falerna (CZ)	16 Y	70 Y	3 Y
TRMIII8	87	Nicastro (CZ)	16 Y	71 Y	9 Y
TRFIII11	77	Conflenti (CZ)	11 Y	65 Y	5 Y

⁸ For a review of the assimilation effects between L1 and L2 VOT of plosives in phonetic attrition, see de Leeuw (2019).

⁹ Heritage Language.

¹⁰ Every ID acronym indicates: informant's group (TR=trilingual speakers); sex (M/F); Calabrian dialectological area (III or IV); sequential number.

VOT values are extracted from productions in Argentinian Spanish and dialect elicited through read and translated speech, respectively (see §).

Spanish VOT productions are elicited via two reading tasks, consisting of a wordlist of 36 words and 12 sentences containing the same words.

Dialectal speech was elicited by translating to dialect a word list covering 36 items and 13 sentences containing the same words. The offered words encompassed a targeted sound, /p/, /t/ or /k/, in post-liquid or post-nasal contexts (-rC-, -nC-, -mC-), equally alternating point of articulation, stressed/unstressed syllables and adjacent vowels (/i, a, u/), as illustrated in (a), since each of these contexts intrinsically determines higher or lower VOT values:

e.g.	DIALECT	SPANISH
	<i>mancava</i>	<i>roncado</i>
	<i>urtimu</i>	<i>ártico</i>
	<i>'mpurnare</i>	<i>computar</i>

Spanish post-liquid occurrences here only included post-rhotic sounds. This choice was motivated by the observation that in the Calabrian dialects of interest, pre-consonant /l/ can undergo assimilation, velarization or rhotacism processes, namely if followed by dental or velar sounds (Rohlf's, 1966; Falcone, 1976). For this reason, and in order to ensure the comparability of our data, only those occurrences resulting from rhotacism were analysed.

A total amount of 1,210 target sounds were selected for the acoustic analysis, detailed as follows (see Table 2.):

Table 2 - Number of occurrences used for acoustic measurements

	/k/	/t/	/p/	N
Argentinian Spanish (L3)	165	156	145	466
Heritage Dialect (L1)	193	376	175	744
TOT.	358	532	320	1210

4.4 First analyses and results

All target productions were labelled and phonetically annotated using Praat (Boersma, Weenink, 2018). Voice Onset Times were measured starting from the first release burst of the stop consonant under investigation up to the acoustic periodic onset of the following vowel (Harrington, 2013). Thus, measures are automatically extracted¹¹, summing up REL (release) and ASP (aspiration) durations in each annotated sound, then checked manually.

¹¹Through the *get_vot.praat* script, created by J. Kang and D.H. Whalen.

Global mean durations and standard deviations associated to different consonants and contexts are extracted and compared between the groups. Thereafter, statistical analyses are implemented by means of the SPSS software (IBM, v. 25) in order to:

1. show significant drifts in VOT as regards the two language varieties (Argentinian Spanish and dialects), for the whole group as well as for each speaker;
2. examine variability related to linguistic factors (phonetic segment and context, task).

The descriptive analyses performed and presented in Table 3 as regards point 1) show how similar are VOT duration ranges for each language considered.

Table 3 - *Descriptive statistics*

Phone	N°	Mean	St. Deviation
/k/_HD	193	45,0187	19,36019
/k/_AS	165	44,9382	19,00226
/t/_HD	376	28,9226	13,48924
/t/_AS	156	24,1378	10,43352
/p/_HD	175	27,0091	18,47986
/p/_SPA	145	23,7903	13,98316

Global VOT duration means, calculated for each voiceless stop in heritage dialect(s), are included in a 27-45 ms range, where /p/ and /t/ are featured by very similar values (/p/_HD¹² = 27 ms, st. d. 18,5 *vs.* /t/_HD = 29 ms, st. d. 13,5), while /k/ productions are consistently longer (/k/_HD VOT = 45 ms, st. d. 19,4). If we compare data produced by Calabrian elderly speakers who have never leaved Italy (see Frontera, 2018), it is worth-noting how these informants' VOT productions consistently outdistance values extracted by trilingual migrants' speech as being much longer¹³.

Global VOT values, calculated for each voiceless stop in Argentinian Spanish, range between 23-44 ms. As for /p/ and /t/ almost equal results are reported (/p/_AS¹⁴ = 24 ms, st. d. 14 *vs.* /t/_AS = 24 ms, st. d. 10,4). VOT values for /k/ in AS amount to 45 ms (st. d. 19), being equal to those in HD. It is self-evident how similar global VOT mean values are between the two languages involved (HD and AS), even in terms of internal variability, as standard deviations clearly show.

After examining general trends related to the whole group, global /p t k/ VOT values are contrasted for each trilingual speaker. In particular, five case-studies were carried out considering those speakers (five out of ten) who completed the elicitation tasks entirely. Therefore, cross-linguistic comparisons at the intra-speaker level were performed for TRFIII2, TRMIII1, TRMIII4, TRFIII6 and TRFIII11 (see Table 1). By observing Table 4, we can state that VOT values in HD are almost constantly higher than the AS counterparts, except for /k/ elicited by speakers TRFIII2,

¹² Heritage Dialect.

¹³ Mean global values extracted from the above-mentioned study (Frontera, 2018): /k/ = 87ms; /t/ = 61ms; /p/ = 59ms.

¹⁴ Argentinian Spanish.

TRMIII4 and TRFIII6, who tend to produce longer VOTs in Argentinian Spanish rather than in their local dialect. Moreover, standard deviations show a remarkable VOT variability. It is worth-noting that VOT standard deviations for /t/ show a trend opposite to /p/ and /k/, being less dispersed and more consistent.

Table 4 - VOT durations and standard deviations for /p t k/ produced by five trilingual speakers in their heritage dialect (HD) and Argentinian Spanish (AS).

* Values are significantly higher (unpaired t-tests).

	HD			AS		
	/p/	/t/	/k/	/p/	/t/	/k/
TRFIII2	17,04 (13)	24 (7,3)	39 (15,9)	15,7 (10,1)	23 (8)	45,8 (24,2)
TRMIII1	29 (11,5)	31 (12,6)*	51,5 (18,6)*	23,9 (13)	22 (7)	36,8 (11,7)
TRMIII4	34,7 (29,1)	27 (11,3)	46 (22,8)	26,6 (14,7)	23,3 (10,4)	52,5 (20,2)
TRFIII6	42,3 (18,2)	29 (13)	42 (21,5)	32 (18,2)	25 (13,3)	51 (17)
TRFIII11	34,4 (15,4)	20,7 (6,8)	36,2 (11,5)	23,8 (11,7)	19,9 (6)	34,4 (10,5)

Unpaired t-tests assuming unequal variance between /p t k/ global VOT durations featuring in heritage dialect(s) and Argentinian Spanish for each informant were run. Only one speaker produced stop consonants with significant cross-linguistic differences in VOT lengths while, in spite of the high degree of internal variation, the other 4 informants did produce voiceless stops with similar VOT durations for L1 and L3 productions. In particular, speaker TRMIII1 still seems to retain long lag /k/ in his heritage language speech (51,51 ms – st. d. 18,6 versus 36,78 ms – st. d. 11,7) and, consequently, to discriminate them from aspirated dialectal sounds from less aspirated velar stops of Argentinian Spanish. This occurs in a significant way both for velar stops ($t(32)=2,03$, $p<.01$) and alveolar ones (HD L1=30,94 ms – st. d. 12,6; AS L3=22 ms, st. d. 7, so $t(41)=2,01$, $p<.01$).

As for point 2) above, the total means reported in Table 5 highlights a clear-cut distribution between VOT durations associated with the task: word-list (WL) vs. sentences (SNT).

Table 5 - VOT duration for /p t k/ elicited through wordlist (WL) and from words in sentences (SNT)

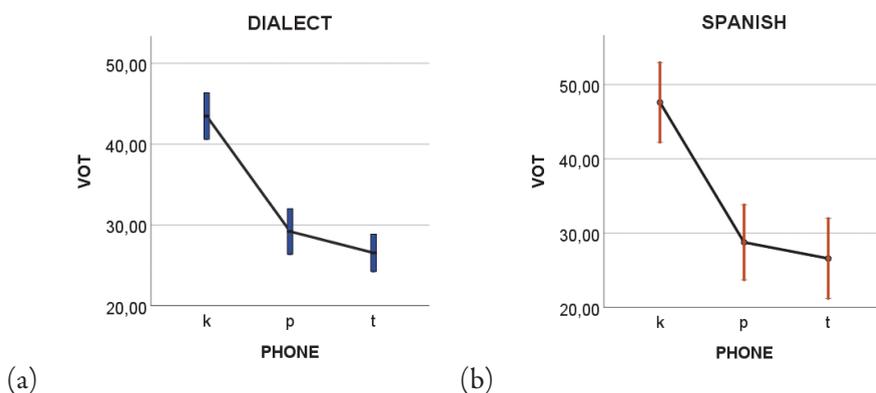
	Heritage Dialect		Argentinian Spanish	
	SNT	WL	SNT	WL
/k/	42,26126	48,75122	43,91744	46,04937
/t/	29,85751	27,3993	23,15926	25,19467
/p/	26,37857	27,81169	23,00694	24,56301

As it can be noted, the two tasks seem to trigger very slight divergences: apart from /t/ (HD), voiceless stops elicited through the wordlist entail longer release phases than those in words elicited in sentences, both in HD and in AS. Other linguistic factors are then examined as far as point 2) is concerned. Within this scope, VOT values are used as dependent variable in an UNIANOVA test, where *language*, *phone*, *context* and *task* are the factors selected to explore VOT variability, while the variable *subjects* represents the random effect.

A significant effect on the main variable is reported only for the following factors: *phone*, $F(2, 1422) = 229,3$, $p = .00$, and *context*, $F(8, 1422) = 19.4$, $p = .00$. This result could be easily associable to the intrinsic articulatory and acoustic features of the three stops. What strikes the attention more is the fact that *language*, the key factor of this first investigation, does not seem to entail any significant variability (as well as the factor *task*): HD mean value = 32,6 ms (st. d. 18) vs. AS mean value = 31,1 ms (st. d. 17,7).

The marginal means extracted through the Anova and displayed in Graph 1. show that estimated /p/ values, on average, have longer VOT durations than /t/. This seems to contradict the common trends reported by literature (see, for example, Lisker, Abramson, 1964 or Cho, Ladefoged, 1999). On the other hand, this trend is in line with what found by Frontera et al. (2019) in the dialectal speech by native Calabrian speakers: /k/ > /p/ > /t/. This evidence is contemplated both in DIA L1 and SPA L3. Graph 1 shows how the highest estimated value of aspiration can be ascribed overall to Spanish productions (1b), whose post-sonorant /k/ reaches the maximum peak, >45 ms duration.

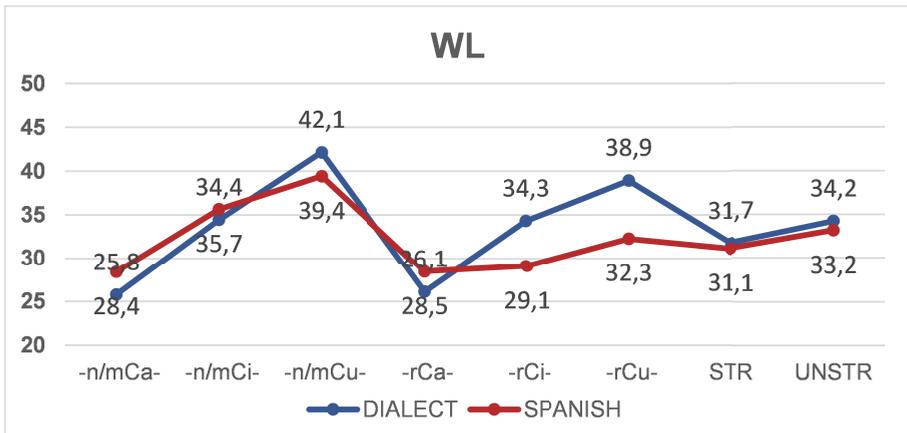
Graph 1 - Estimated marginal means and standard deviations for /p t k/ VOTs in HD and in AS



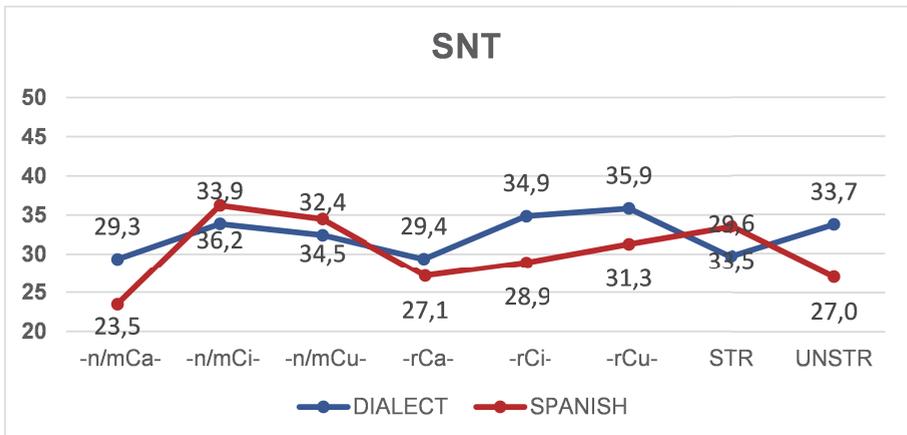
Given the significant effect of *context*, further t-tests are performed, in order to explore differences implied between Heritage Dialect (HD) and Argentinian Spanish (AS). Despite the clear distinctions that every context brings about at the intra-linguistic level (Graph 2.), the phonetic contexts here examined do not trigger any

significant variation in VOT values between L1 and L3. Nonetheless, it is necessary to observe that when high vowels (/i u/) follow the target segment they normally trigger longer VOT durations if compared to what occurs when with /a/. Moreover, as evidenced by acoustic data on Italian and Calabrian speech, stops in unstressed context are generally longer than in pretonic position. This was confirmed by our results related to WL and SNT tasks, both for HD and AS (see Graph 2x, 2y).

Graph 2 - Mean VOT durations of /p t k/ in all post-sonorant contexts, produced in HD and AS in isolated words (x) and words extracted from sentences (y)



(a)



(b)

5. Discussion and conclusions

This preliminary study allowed us to explore phonetic properties of VOT in the production of first-generation Italian-Argentinian migrants, as far as their heritage Calabrian dialect and L3 Argentinian Spanish are concerned.

First results suggest that migrants produced very similar VOT durations in their dialectal speech and in Argentinian Spanish. Apparently, VOT values are lower than those ascertained in monolingual Calabrian elderly speakers (see Frontera, 2018). Even if this trend needs to be verified through direct acoustic comparisons to data elicited by a monolingual Italian control group, it seems as if the acoustic correlate of VOT for voiceless stops for our trilingual speakers has reached a sort of convergence point, by constantly negotiating between the long lag realisations of HD and short lag releases of AS.

This last consideration is supported by evidences of /p/ release phases which are slightly higher than those reported for /t/, which is a peculiar trend acoustically confirmed by previous research on VOT aspiration (see Frontera et al., 2019) on native Calabrian speakers. This cue features VOT of Argentinian Spanish produced by our Italian-Argentinian migrants. Thus, in compliance with Flege's *merging hypothesis*, both first and third language phonetic traits of our informants seem to be deviant from the respective monolingual norms. Further acoustic VOT data on L2 Italian could provide a broader interpretation of these first results. In fact, it could be interesting to see to what extent the presence of a dominant code (AS) within migrants' linguistic repertoire has undermined the aspiration heritage trait as far as Calabrian dialects and regional Italian are concerned.

As reported by Cho and Ladefoged (1999), after examining a set of 18 different languages, the boundary between unaspirated (short lag) and aspirated stops can be set at 50 ms. As regards Spanish varieties, this threshold is placed at 40 ms by Martínez Celdrán (2013) while, more recently, in a study by Nodari, Celata and Nagy (2019) Calabrian stops are included in a VOT investigation of different languages by proposing to consider voiceless stops as aspirated when VOT ranges from 30 to 100 ms. By assuming their criterion, we can set 30 ms as a benchmark to classify our HD and AS results as slightly aspirated consonants. The assumed change in the phonetic realisations of voiceless stops could be interpreted as a cue of phonetic attrition. However, VOT undergoes a great variation across our speakers, both at the inter-personal and the intra-personal level (see standard deviation values reported in Graph 1 and Table 4). This could be due to the task (see Major's, 1992): the translation task, here used to elicit dialectal speech, may have triggered longer VOT realisations compared to reading, assuming a higher degree of accuracy for the latter. As a matter of fact, the highest VOT values are reported when translating into their heritage dialect. Nonetheless, mean global values confirm the vulnerability of the phonetic trait here investigated. Within the scope of language attrition studies, it is known that "L1 is malleable upon competition from a new language or dialect in adulthood" (de Leeuw, 2019: 215). In this framework, these very preliminary results suggest that attrition may have occurred. This may be seen in the conver-

gence of VOT values reported for languages comprised within the first-generation speakers' repertoires (see Flege, 1987).

The attested variability at intra-personal and inter-personal levels in attrition studies, however, may lead to unexpected results or exclude a univocal direction for drift. Therefore, these data need to be enriched and strengthened through deeper analyses, using normalized VOT durations (see Mori, Barkat, 2005) and further comparisons with the monolingual control groups. Moreover, further analyses on the whole group of 37 speakers may reveal a stronger effect of extralinguistic factors on language attrition, such as AoA, LoR, also considering the potential effect of education in Italy (number of years and education degree): a probable use and/or identification of Italian as the 'prestigious' code may lead to even lower aspiration values in L2 productions.

It will be interesting, furthermore, to explore possible correlations between acoustic data and other parameters such as in-groupness, use of L1-L2-L3 and attitudes toward dialects indicators, elaborated through responses to our interviews (see §). In addition, considered that acoustic realisations of VOT itself may also depend on variables of a sociolinguistic nature (thus conveying information about the potential presence of variation dynamics and/or about the speakers' status, see Nodari, 2015), future research could entail additional analyses, dealing with other segmental and/or suprasegmental elements.

In line with the future developments of the reference projects, if this first results of attrition in VOT values will be confirmed, aspiration may turn to be a relevant cue for accent detection by native listeners (see task (c in § 3.2): the trilingual speakers are then likely to be recognised as foreign speakers of their own heritage languages (dialects) and/or regional Italian.

To conclude, our whole research aims to provide new original data on outcomes of long-lasting contact between three close languages within migrants' trilingual repertoires in order to feed the scientific discussion on Italian and its varieties world-wide as far as L2/L3 language acquisition, L1 attrition and sociolinguistic dynamics are concerned.

Acknowledgment

I would like to express my deepest gratitude to Patricia Scarponetti (CEA, UNC) for her precious guide during my stay in Argentina. I wish to thank Antonio Antonuccio (La Falda, Córdoba) and the associations *Familia Calabresa* (Rosario), *Hermanos Calabreses al Exterior* (Santa Fe) and *Calabresa* (Buenos Aires), for their support in recruiting informants for this research. Thanks to all participants for their collaboration and helpfulness, without whom this project would not have been possible.

I am indebted to Laura Mori and to anonymous reviewers for their insightful suggestions which allowed me to improve the present paper.

Funding

The research field for the CMA project has been led in the context of INCASI Network, a European project that has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie GA No 691004 and coordinated by Dr. Pedro López-Roldán.

This paper was further developed in the context of the PRIN 2017 project «Italian language, global language market, Italian companies in the world: new linguistic, sociocultural, institutional, economic-productive dynamics» (Prot. 2017K79S7T; national coordinator: Prof. Massimo Vedovelli, Università per Stranieri di Siena), as a member of the local research unit of Università degli Studi Internazionali di Roma – UNINT (scientific responsible: Prof. Laura Mori).

Bibliography

ABRAMSON, A.S., LISKER, L. (1973). Voice-timing perception in Spanish word-initial stops. In *Journal of Phonetics*, 1, 1-8.

AMENGUAL, M. (2012). Interlingual influence in bilingual speech: Cognate status effect in a continuum of bilingualism. In *Bilingualism: Language and Cognition*, 15 (3), 517–530.

ASENSI, L., PORTOLÉS, S. AND DEL RÍO, A. (1998). Barra de explosión, VOT y frecuencia de las oclusivas sordas del Castellano. In *Estudios de Fonética Experimental IX*, Universitat de Barcelona, Laboratori de Fonètica, 221-242.

AVESANI, C., GALATÀ, V., VAYRA, M., BEST, C., DI BIASE, B., TORDINI, O., TISATO, G. (2015). Italian roots in Australian soil: dialect as heritage language in first generation bilinguals. In AVESANI, C., VAYRA, M. E TAMBURINI, F. (Eds.), *Il farsi e disfarsi del linguaggio. Acquisizione, mutamento e destrutturazione della struttura sonora del linguaggio/Language acquisition and language loss. Acquisition change and disorders of the language sound structure*, Milano: Officinaventuno, 75-98.

AVESANI, C.; GALATÀ, V.; BEST, C.; VAYRA, M.; DI BIASE, B.; ARDOLINO, F. (2017). Phonetic details of coronal consonants in the Italian spoken by Italian-Australians from two areas of Veneto. In BERTINI, C., CELATA, C., LENOCI, G., MELUZZI, C. E RICCI, I. (Eds.), *Fattori sociali e biologici nella variazione fonetica/Social and biological factors in speech variation*, Milano: Officinaventuno, 283–308.

BARTOLI, M. (1945). *Saggi di linguistica spaziale*. Torino: Rosenberg & Sellier.

BAYRAM, F., PASCUAL Y CABO, D. AND ROTHMAN, J. (2019). Intra-generational attrition: Contributions to heritage speaker competence. In SCHMID, M.S. AND KÖPKE, B. (Eds.), *The Oxford Handbook of Language Attrition*, Oxford: Oxford University Press, 446-459.

BERGMANN, C., NOTA, A., SPRENGER, S.A., & SCHMID, M.S. (2016). L2 immersion causes non-native-like L1 pronunciation in German attriters. In *Journal of Phonetics*, 58(Supplement C), 71–86.

BEST, C.T. (1995). A direct realist perspective on cross-language speech perception. In Strange, W. (Ed.) *Speech perception and linguistic experience: Theoretical and methodological issues in cross-language speech research*, York Press: Timonium, MD, 167–200.

- BEST, C.T. AND TYLER, M.D. (2007). Nonnative and second-language speech perception: Commonalities and complementarities, in Language Experience in Second Language Speech Learning. In BOHN, O.-S. AND MUNRO, M. (Eds.), *In honor of Jim Flege*, Amsterdam: John Benjamins, 13-34.
- BOERSMA, P., WEENINK, D. (2018). *Praat: doing phonetics by computer* [Computer program]. Version 6.0.43, retrieved 8 September 2018 from <http://www.praat.org/>
- BONGAERTS, T., VAN SUMMEREN, C., PLANCKEN, B. AND SCHILS, E. (1997). Age and Ultimate Attainment in the Pronunciation of a Foreign Language. In *Studies in Second Language Acquisition*, 19, 447-465.
- BORZONE, A.M. (1980). *Manual de fonética acústica*, Buenos Aires: Hachette.
- BORZONE, A.M., GUERLEKIAN, J. (1980). Rasgos acústicos de las consonantes oclusivas españolas. In *Fonoaudiología*, 26 (3), 326-330.
- BOT, K. DE AND WELTENS, B. (1985). Taalverlies: Beschrijven versus verklaren. In *Handelingen van het 38^e Nederlands Filologencongres*, 51-61.
- BROWN, C. (2000). The interrelation between speech perception and phonological acquisition from infant to adult. In ARCHIBALD, J. (Ed.), *Second Language Acquisition and Linguistic Theory*, Oxford: Blackwell, 4-63.
- CANCELLIER, A. (1996). *Lenguas en contacto. Italiano y español en el Río de la Plata*. Padova: Unipress.
- CANEPARI, L. (1986). *Italiano standard e pronunce regionali*. Padova: CLEUP.
- CASTAÑEDA VICENTE, M.L. (1986). El V.O.T. de las oclusivas sordas y sonoras españolas. In *Estudios de Fonética Experimental*, 2, 92-110.
- CHO, T., LADEFOGED, P. (1999). Variation and universals in VOT: evidence from 18 languages. In *Journal of Phonetics*, 27, 207-229.
- COOK, V. (2003). The changing L1 in the L2 user's mind. In COOK, V. (Ed.), *Effects of the second language on the first*, Clevedon: Multilingual Matters, 1-18.
- COSERIU, E. (1980). Historische Sprache und Dialekt. In J. GOSCHEL, P. IVIC, & K. KEHR (Eds.), *Dialekt und Dialektologie*, Wiensbaiden: Steiner, 106-122.
- DAMONTE, F. (2009). La particella mu nei dialetti calabresi meridionali. In *Quaderni di lavoro ASIT*, 9, 101 - 117.
- DE ANGELIS, A. (2016). Origini formali e funzionali della particella (m)i, (m)u, ma nell'area messinese e calabrese centro-meridionale. In P. DEL PUENTE (Ed.), *Dialetti per parlare e parlarne. Atti del quarto Convegno Internazionale di Dialettologia*, Venosa: Osanna edizioni, 75-95.
- DE BLASI, N. (2014). *Geografia e storia dell'italiano regionale*. Bologna: Il Mulino.
- DE IACOVO, V., MAIRANO, P. & HAJEK, J. (2020). Studio sul grado di geminazione in un corpus di emigrati italiani in Australia, lavoro presentato a *XVI Convegno Nazionale AISV, La variazione linguistica in condizioni di contatto: contesti acquisizionali, lingue, dialetti e minoranze in Italia e nel mondo*, Rende (CS), Italia, 29-31 gennaio 2020.
- DE LEEUW, E. (2019). Phonetic attrition. In SCHMID, M.S. AND KÖPKE, B. (Eds.), *The Oxford Handbook of Language Attrition*, Oxford: Oxford University Press, 202-217.
- DE LEEUW, E., MENNEN, I., & SCOBIE, J.M. (2013). Dynamic systems, maturational constraints and L1 phonetic attrition. In *International Journal of Bilingualism*, 17(6), 683-700.

- DE LEEUW, E., SCHMID, M.S., & MENNEN, I. (2010). The effects of contact on native language pronunciation in an L2 migrant setting. In *Bilingualism: Language and Cognition*, 13, 33–40.
- DERWING, T.M. AND MUNRO, M.J. (2015). *Pronunciation Fundamentals. Evidence-based Perspectives for L2 Teaching and Research*. Amsterdam-Philadelphia: John Benjamins Publishing Company.
- ECKMAN, F.R. (1977). Markedness and the contrastive analysis hypothesis. In *Language Learning*, 27, 315–330.
- ECKMAN, F.R. (2008). Typological markedness and second language phonology. In HANSEN EDWARDS, J.G. AND ZAMPINI, M.L. (Eds.), *Phonology and Second Language Acquisition*, Amsterdam-Philadelphia: John Benjamins Publishing, 95–115.
- ELLIS, R. (2003). *The Study of Second Language Acquisition* (10th ed.). Oxford: Oxford University Press.
- FALCONE, G. (1976). Calabria. In CORTELAZZO, M. (Ed.), *Profilo dei dialetti italiani*, Pisa: Pacini.
- FANCIULLO, F., LIBRANDI, R. (2002). La Calabria. In CORTELAZZO, M., MARCATO, C., DE BLASI, N. & CLIVIO, G. (Eds.), *I dialetti italiani. Storia, struttura, uso*, Torino: Utet, 793–833.
- FLEGE, J.E. (1987). The production of “new” and “similar” phones in a foreign language: Evidence for the effect of equivalence classification. In *Journal of Phonetics*, 15, 47–65.
- FLEGE, J.E. (1995). Second Language Speech Learning: Theories, Findings and Problems. In STRANGE, W. (Ed.), *Speech perception and linguistic experience: issues in cross-language research*, Timonium, MD: York press, 233–277.
- FLEGE, J.E. AND EEFING, W. (1987). Cross-language switching in stop consonant perception and production by Dutch speakers of English. In *Speech Communication*, 6 (3), 185–202.
- FLEGE, J.E., FRIEDA, E.M. AND NOZAWA, T. (1997). Amount of native-language (L1) use affects the pronunciation of an L2. In *Journal of Phonetics*, 25, 169–186.
- FLEGE, J.E. (2009). Give input a chance!. In PISKE, T. AND YOUNG-SCHOLTEN, M. (Eds.), *Input Matters in SLA*, Bristol: Multilingual Matters, 175–190.
- FRONTERA, M. (in preparation). *Cues of dialect attrition in Italian-Argentinian trilingual speakers*.
- FRONTERA, M. (2018). Aspirated voiceless stops in elderly speakers from Calabria: a pilot study. In BOTINIS, A. (Ed.), *Proceedings of the 9th Tutorial and Research Workshop on Experimental Linguistics*, Paris, France, 28–30 August 2018, 33–36.
- FRONTERA, M. (2019). Spanish voiceless stop production by Italian-Argentinian bilinguals from Calabria, poster presentato al convegno *Bilingualism in the Hispanic and Lusophone world (BHL)*, Leiden University, Netherlands, 9–11 January 2019.
- FRONTERA, M., TARASI, A. AND GRAZIANO, E. (2019). Le consonanti occlusive sorde aspirate in Calabria: un confronto tra aree dialettali. In CALAMAI, S., PICCARDI, D. E ARDOLINO, F. (Eds.), *Gli archivi sonori al crocevia tra scienze fonetiche. Informatica umanistica e patrimonio digitale*, Milano: Officinaventuno, 293–307.

- HARRINGTON, J. (2013). Acoustic Phonetics. In HARDCASTLE, W.J., LAVER, J. & GIBBON, F.E. (Eds.), *The Handbook of Phonetic Sciences*, (2nd ed.), Chichester: Blackwell Publishing, 81-129.
- HOPP, H., & SCHMID, M.S. (2013). Perceived foreign accent in first language attrition and second language acquisition: The impact of age of acquisition and bilingualism. In *Applied Psycholinguistics*, 34(2), 361-394.
- KREFELD, T. (2007). Variazione morfosintattica in calabrese: i possessivi al microscopio, tratto da <http://www.asica.gwi.uni-muenchen.de/downloads/poss-calab.pdf>.
- KUHL, P.K. (2000). A new view of language acquisition. In *Proceedings of the National Academy of Sciences USA*, 97, 11850-11857.
- LENNEBERG, E.H. (1967). *Biological foundations of language*. New York: Wiley.
- LISKER, L., ABRAMSON, A.S. (1964). A cross-language study of voicing in initial stops: acoustical measurements. In *Word*, 20, 3, 384-422.
- LONG, M.H. (2008). Stabilization and Fossilization in Interlanguage Development. In DOUGHTY, C. AND LONG, M. (Eds.), *The Handbook of Second Language Acquisition*, Oxford: Blackwell, 487-535.
- MADDIESON, I. (1984). *Patterns of sounds*, Paperback reprint 2009 ed., Cambridge: Cambridge University Press.
- MANZINI, M.R., SAVOIA, L.M. (2005). *I dialetti italiani e romanci. Morfosintassi generativa*, 3 vols. Alessandria: Edizioni dell'Orso.
- MARTINEZ CELDRÁN, E. (2013). Los sonidos obstruyentes en la cadena hablada. In *Panorama de la fonética española actual*, Madrid: Arcos Libros, 253-289.
- MAJOR, R.C. (1992). Losing English as a First Language. In *The Modern Language Journal*, 76 (2), 190-208.
- MAJOR, R.C., BAPTISTA, B.O. (2009). First language attrition in foreign accent detection. In WATKINS, M.A., RAUBER, A.S., BAPTISTA, B.O. (Eds.), *Recent research in Second Language Phonetics/ Phonology: Perception and production*, Newcastle Upon Tyne: Cambridge Scholars Publishing, 256-267.
- MAJOR, R.C. (2014). *Foreign accent: The ontogeny and phylogeny of second language phonology*. New York and London: Routledge.
- MONTRUL, S. (2011). Morphological errors in Spanish second language learners and heritage speakers. In *Studies in Second Language Acquisition*, 33, 155-161.
- MOYER, A. (2007). Do Language Attitudes Determine Accent? A Study of Bilinguals in USA. In *Journal of Multilingual and Multicultural Development*, 28, 502-518.
- MORI, L. (2007). *Fonetica dell'italiano L2. Un'indagine sperimentale sulla variazione nell'interlingua dei marocchini*. Roma: Carocci.
- MORI L., BARKAT-DEFRADAS M. (2005). VOT e variazione interlinguistica: il caso delle occlusive sorde aspirate in italiano L2. In P. Cosi (Ed.), *Misura dei parametri. Aspetti tecnologici ed implicazioni nei modelli linguistici*, Atti del I Convegno Nazionale AISV 2004, Padova: E.D.K., 511-524.
- NAGY, N. (2015). A sociolinguistic view of null subjects and VOT in Toronto heritage languages. In *Lingua*, 164 (2), 309-327.

- NAGY, N., NODARI, R. & CELATA C. (2018). A variationist analysis of Heritage Italian VOT: Phonetic but not phonological fidelity. In *NWAV 47*, New York, 18-21 October 2018.
- NODARI, R. (2015). Descrizione acustica delle occlusive sorde aspirate: analisi sociofonetica dell'italiano regionale di adolescenti calabresi. In VAYRA, M., AVESANI, C. & TAMBURINI, F. (Eds.), *Il farsi e disfarsi del linguaggio. Acquisizione, mutamento e destrutturazione della struttura sonora del linguaggio*, Milano: Officinaventuno, 139-153.
- NODARI, R., CELATA, C. & NAGY, N. (2019). Socio-indexical phonetic features in the heritage language context: Voiceless stop aspiration in the Calabrian community in Toronto. In *Journal of Phonetics*, 73, 91-112.
- PATKOWSKI, M.S. (1994). The critical age hypothesis and interlanguage phonology. In YAVAS, M. (Ed.), *First and second language phonology*, San Diego, CA: Singular Publishing Group, 205-221.
- PAVLENKO, A. (2004). L2 influence and L1 attrition in adult bilingualism. In SCHMID, M., KÖPKE, B., KEIJZER, M. & WEILEMAR, L. (Eds.), *First language attrition: Interdisciplinary perspectives on methodological issues*, Amsterdam: John Benjamins, 47-59.
- PISKE, T., MACKAY, I.R.A. AND FLEGE, J.E. (2001). Factors affecting degree of foreign accent in an L2: a review. In *Journal of Phonetics*, 29(2), 191-215.
- ROHLFS, G. (1966), *Grammatica storica della lingua italiana e dei suoi dialetti*. Torino: Einaudi.
- ROLDÁN, Y. AND SOTO-BARBA, J. (1997). El V.O.T. de /p-t-k/ y /b-d-g/ en el español de Valdivia: un análisis acústico. In *Estudios Filológicos*, 32, 27-33.
- ROMAINE, S. (1984). The status of sociological models and categories in explaining language variation. *Linguistische Berichte*, 90, 25-38.
- ROMITO, L., CIARDULLO, M.A. & TARASI A., (2015), Analisi acustica delle occlusive sorde aspirate del dialetto di San Giovanni in Fiore (CS). In VAYRA, M., AVESANI, C. & TAMBURINI, F. (Eds.), *Il farsi e il disfarsi del linguaggio. L'emergere, il mutamento e la patologia della struttura sonora del linguaggio*, Milano: Officinaventuno, 169-186.
- ROMITO, L., GRAZIANO, E., FRONTERA, M., TARASI, A., CIARDULLO, M.A., PIEMONTE, M.G., STRANO, M. (2016). Definizione delle radici (pan)calabresi in terra argentina: uno studio pilota. In CAPPELLI, V. E SERGI, P. (Eds.), *Traiettorie Culturali tra il Mediterraneo e l'America Latina*, Cosenza: Pellegrini, 281-296.
- ROMITO, L., MILELLI, R. (1999), Analisi elettroacustica degli esiti di -LL- in alcuni dialetti della Calabria. In *Quaderni del Dipartimento di Linguistica dell'Università della Calabria*, Vol. 7(16), 11-26.
- ROMITO, L., SORIANELLO, P. (1998), Ridefinizione delle consonanti retroflesse nei dialetti calabresi, presentato al V Convegno Internazionale *SILFI*, Catania, 15-17 ottobre 1998.
- ROTHMAN, J. (2009). Understanding the nature and outcomes of early bilingualism: Romance languages as heritage languages, In *International Journal of Bilingualism*, 13(2), 155- 163.
- SCHMID, M.S., KÖPKE, B. (2013). *First Language Attrition*, Amsterdam: John Benjamins Publishing.
- SCHMID, M.S., KÖPKE, B. (Eds.) (2019). *The Oxford Handbook of Language Attrition*, Oxford: Oxford University Press.

- SCHMID, M.S., KÖPKE, B., & BOT, K. DE. (2013). Language attrition as a complex, nonlinear development. In *International Journal of Bilingualism*, 17(6), 675–682.
- SORIANELLO, P. (1996), Indici fonetici delle occlusive sorde nel cosentino. In *Rivista Italiana di Dialettologia*, 20, 123-159.
- SOTO-BARBA, J. AND VALDIVIESO, H. (1999). Caracterización fonético-acústica de la serie de consonantes /p-t-k/ vs. /b-d-g/. In *Onomazein*, 4, 125-133.
- STEVENS, M., HAJEK, J. (2010). Post-aspiration in standard Italian: some first cross-regional acoustic evidence. In *Proceedings of Interspeech*, Makuhari, Japan, 1557-1560.
- TELMON, T. (1993). Varietà regionali. In SOBRERO, A.A. (Ed.), *Introduzione all'italiano contemporaneo, la variazione e gli usi*, Roma-Bari: Laterza, 93-149.
- TRUMPER, J., MADDALON, M. (1988). Converging Divergence and Diverging Convergence: The Dialect-Language Conflict and Contrasting Evolutionary Trends in Modern Italy. In AUER, P. e DI LUZIO, A. (Eds.), *Variation and Convergence: Studies in Social Dialectology*, Berlin-NewYork: Mouton-de Gruyter, 217-259.
- TRUMPER, J. (1997). Calabria and southern Basilicata. In MAIDEN M. AND PARRY M. (Eds.), *The dialects of Italy*, London: Routledge, 355-364.
- TRUMPER, J. (2016), *Geostoria linguistica della Calabria*. Roma: Aracne.
- WEINREICH, U. (1957). On the description of phonic interference. In *Word*, 13(1), 1-11.

PARTE II

LINGUE MINORITARIE
E VARIETÀ REGIONALI

MICHELA RUSSO, FABIO APREA

Identità linguistica di una lingua minoritaria alloglotta della Puglia (Italia meridionale): il francoprovenzale di Faeto

“Linguistic identity of an alloglot minority language in Apulia
(southern Italy): Faeto’s Francoprovençal”

Faetar is a Franco-Provençal language spoken in a small community near Foggia, in southern Italy (in Apulia region), which owes its origin to Franco-Provençal settlements dating back to the 13th century. Franco-Provençal, although it remains quite vital in Italy, is considered today an endangered language, which UNESCO has included in its Atlas of the *World’s Languages in Danger*. Through an auditory and spectrographic analysis of different data, we will interpret many phonetic features of the Franco-Provençal spoken in Faeto, and we will compare them with Franco-Provençal acoustic data collected in France (which are close to the Faetar geographical area), as well as with French acoustic data.

The aim is to assess the synchronic variety of southern Italian Faetar, in relation to the Franco-Provençal variety of France, and to identify the phonetic features belonging to the two varieties in contact: the southern Italo-Romance varieties and the Franco-Provençal variety. This is to build a diachronic and synchronic grammar of this unique blend dialect shown by language contact.

Keywords: Faetar; Franco-Provençal; Italo-Romance; vowel system; nasalization.

1. *Introduzione*

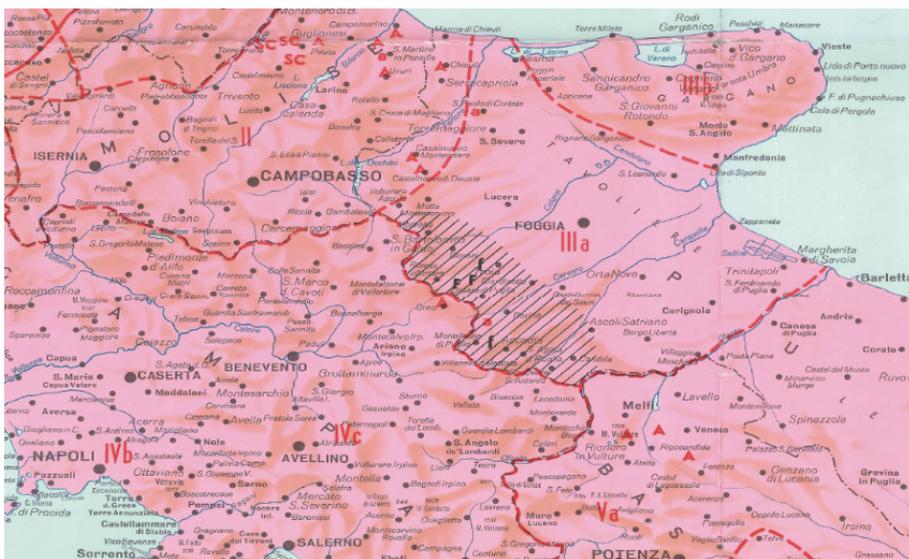
Il nostro lavoro si concentra sul dialetto di Faeto, comune italiano della provincia di Foggia, nella Puglia nord-occidentale, dove si parla il francoprovenzale dal XIII secolo (Morosi, 1890-92; Melillo, 1959; Nagy, 1993; 1996; 2000a/b). Insieme a Celle di San Vito si tratta di una situazione linguistica unica nella regione, riconosciuta nel 1999 dalla legge nazionale italiana e da una legge regionale nel 2012 che protegge le minoranze linguistiche storiche (cfr. Agresti, 2016)¹.

¹ Queste leggi si basano sulla Costituzione italiana, in particolare sull’art. 6: «La Repubblica tutela le minoranze linguistiche con norme speciali». Si tratta della legge dello Stato n. 482 del 15 dicembre 1999: «Norme per la tutela delle minoranze linguistiche storiche», e della legge regionale n. 5 del 22 marzo 2012: «Norme per la promozione e la tutela delle lingue minoritarie in Puglia». A livello europeo l’attenzione normativa verso le lingue minoritarie era stata riaccesa dalla *Convenzione-quadro per la protezione delle minoranze nazionali*, adottata dal Comitato dei Ministri del Consiglio d’Europa a Strasburgo il 1° febbraio 1995 ed entrata in vigore il 1° febbraio 1998; sarebbe stata sottoscritta e ratificata dall’Italia il 27 maggio 2000.

La Puglia settentrionale conosce anche altre isole linguistiche; è esistita una colonia occitana a Volturara (in provincia di Foggia), ormai estinta. Oltre alle colonie di lingua romanza, non molto lontano, in Campania (nella provincia di Avellino), esiste una colonia albanese, Greci.

L'insieme di queste colonie ha lasciato tracce sufficienti perché Pellegrini (1977) consideri quest'area geografica complicata da classificare linguisticamente (Figura 1):

Figura 1 - // Aree miste – di complessa classificazione; A Colonie albanesi, a Colonie albanesi estinte; F Colonie francoprovenzali, f Colonie francoprovenzali estinte (Pellegrini, 1977)



Il francoprovenzale, anche se rimane abbastanza vitale in Italia (soprattutto in Valle d'Aosta, Diemoz, 2012), è oggi una lingua considerata in via di estinzione, che l'UNESCO ha inserito nel suo *Atlas of the World's Languages in Danger* (2010)².

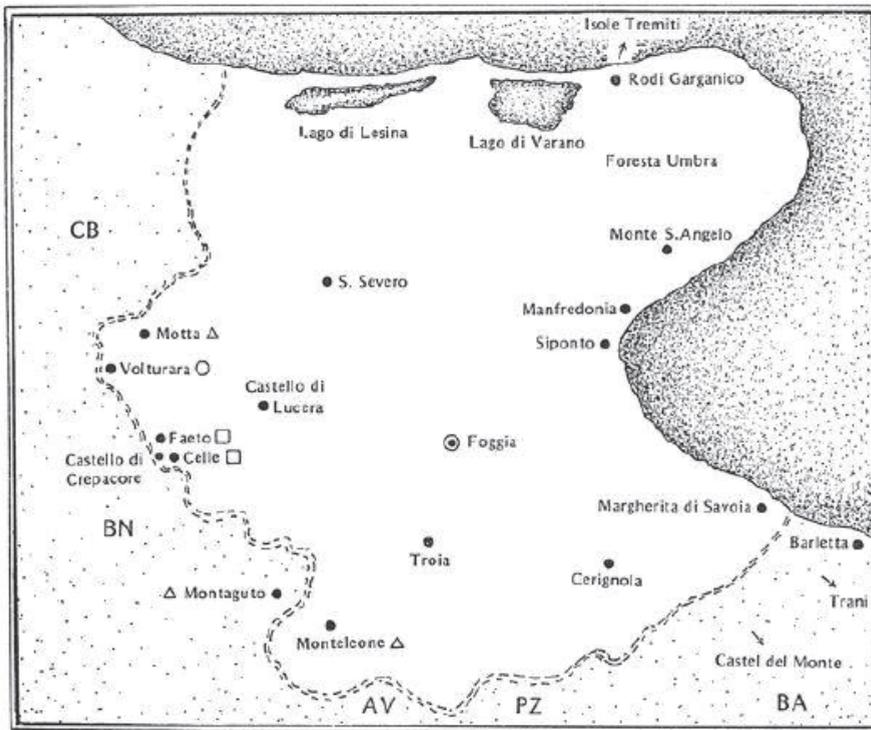
Gli abitanti di Faeto e Celle di San Vito provengono storicamente da una zona gallo-romanza compresa tra Grenoble, Le Grand-Serre, Coligny e Replonge, cioè da una zona ampia dell'Isère e dell'Ain, al limite orientale del Rodano (Melillo, 1959; Perrin, 2018).

Faeto ha conservato, grazie al suo isolamento geografico, la lingua originaria che viene trasmessa ai bambini in ambito familiare e presenta una situazione di contatto linguistico.

² In Francia il francoprovenzale è riconosciuto dal 1999 come lingua regionale e patrimonio immateriale della Francia (Cerquiglini, 1999), ma non è una lingua regionale protetta dalla legge, e non è entrata nel sistema educativo. Cerquiglini era stato incaricato dal ministero della Cultura (in qualità di vice-presidente del *Conseil Supérieur de la langue Française*) di redigere un inventario delle lingue regionali che avrebbero potuto beneficiare di una protezione della *Carta europea delle lingue regionali e minoritarie* (*European Charter for Regional or Minority Languages*, ECRML), ma la Francia, pur avendo firmato in un primo momento la Convenzione, non ha mai ratificato la *Carta europea*. Al contrario, in Valle d'Aosta il francoprovenzale gode di uno statuto autonomo, è protetto dalla legge e viene insegnato nelle scuole (cfr. Kasstan, 2015; Russo, Stich, 2020).

stico molto particolare: il dialetto di Faeto è circondato da varietà pugliesi e campane, in competizione con l'italiano (Stehl, 1980; Fanciullo, 2002). Il francoprovenzale di Faeto è in contatto con i dialetti italo-romanzi della Puglia settentrionale, dove vige un sistema vocalico di tipo napoletano, con le modificazioni osservate da Stehl (1980; 1988) in relazione agli sviluppi metafonetici di Ę e Ö latine e alla differenziazione vocalico-sillabica (Valente, 1975; Stehl, 1988; Coluccia, 1992; Russo, 2002; Fanciullo, 2002)³. Gli abitanti di Faeto sono bilingui francoprovenzale / pugliese, e trilingui francoprovenzale / pugliese / italiano (Nagy, 1996; 2000a/b; Puolato, 2011; 2016; Favre 2010; Perrin, 2018).

Figura 2 - □ colonie francoprovenzali del XIII secolo, ancora esistenti; ○ colonie occitane del XVI secolo, estinte nella prima metà del XVIII secolo;
 △ Probabili colonie francoprovenzali o occitane di epoca incerta, comunque poco fortunate dal punto di vista linguistico. Melillo (1981: 40)



³ “Vokaldifferenzierung nach freier und gedeckter Stellung” (Weinrich, 1958: 175). In particolare, il tipo metafonetico delle varietà dialettali limitrofe (assente a Faeto) è quello tipico dei dialetti foggiano-baresi e baresi settentrionali, con il dittongo accentato sul primo elemento vocalico o confluito con le vocali alte primarie /i, u/ (Pellegrini, 1975; Stehl, 1988; Russo, 2002; Loporcaro, 1997; 2013; 2016). Per le consonanti, si ricorda per le varietà italo-romanze pugliesi la presenza delle geminate lessicali (e di tipo fonosintattico), assenti a Faeto, e la conservazione delle occlusive sorde intervocaliche, o l'indebolimento non fonologico delle stesse. Faeto mostra, invece, in modo sistematico la lenizione fonologica delle occlusive, tipica del gallo-romanzo.

2. *Materiali e metodo di analisi*

In questa sede presentiamo un'analisi fonetica e dialettologica di diversi tratti linguistici francoprovenzali identificati a Faeto, propri alla situazione di contatto.

I materiali usati nella nostra indagine sono di due tipi: 1) liste di parole in contesto isolato, o inserite in un contesto di frasi cornice, pronunciate da locutori di diverse fasce d'età aventi il francoprovenzale di Faeto quale L1⁴; 2) parole elicitate a partire da un *Archivio Audio di Faeto* (materiale acustico estratto da video-registrazioni) che abbiamo collocato e creato presso il laboratorio di linguistica dell'università di Parigi 8, creato grazie alle donazioni di un insieme di locutori⁵. Tale *Archivio* è composto di due tipologie di parlato: parlato prodotto in esecuzione spontanea da locutori di diverse fasce d'età; parlato 'seminaturale' recitato da locutori nativi, monologhi, poesie, etnotesti (brani, racconti popolari filastrocche). Da tutto questo materiale, registrato a Faeto, abbiamo estratto un quantitativo sufficiente di parole bersaglio.

I materiali acustici sono stati inclusi come file WAV o MP3 alla frequenza di campionamento originale di 44,1 KHz e quantizzazione 16 bit. Il risultato dell'analisi fonetica e acustica manuale, effettuata mediante ascolto e visualizzazione del segnale, è stato riportato sotto forma di trascrizione fonetica stretta (utilizzando l'alfabeto IPA – *International Phonetic Alphabet revised*, 2020). L'analisi spettro-acustica è stata effettuata mediante l'utilizzo del software PRAAT (versione 6.1.08, Boersma, Weenink, 2019).

Attraverso l'analisi uditiva e spettro-acustica abbiamo interpretato i materiali sperimentali relativi alle vocali toniche e atone (sezioni 3 e 4), che riflettono tratti fonetici, fonologici e morfologici tipici del francoprovenzale parlato a Faeto. Ci siamo interessati in particolare all'analisi formantica di alcuni processi vocalici (cfr. Amélot, 2004; Martin, 2008; Cosi, Ferrero, Vagges, 1995; Ferrero, 1997) che distinguono il francoprovenzale dal gallo-romanzo settentrionale, e dall'italo-romanzo meridionale.

Faeto esibisce una sua specificità e degli ibridismi dovuti alla situazione di contatto e di interferenza tra la parlata alloglotta, i dialetti italo-romanzi e l'italiano regionale.

Abbiamo inoltre confrontato i dati acustici di questa lingua di minoranza, con altri dati acustici provenienti dal francoprovenzale di Francia (della Savoia), e dal francoprovenzale della Valle d'Aosta e della Svizzera Romanda. È noto infatti che il francoprovenzale comprende un gruppo di dialetti frammentato (Martin, 1990; Telmon, 2001; Kristol, 2016), che si estendono, con soluzioni di continuità dal sud-est della Francia alle zone limitrofe della Svizzera e dell'Italia (la Valle d'Aosta)⁶.

⁴ Materiale acustico disponibile sul sito web Forvo.com. L'analisi è stata effettuata su 4 locutori di Faeto (di età compresa tra 45 e gli 80 anni, 2 uomini e 2 donne).

⁵ Cinque locutori di Faeto (di età compresa tra 43 e 75 anni, 3 uomini e due donne) registrati durante conversazioni informali nelle quali si è discusso di *items* ed espressioni attinenti al lavoro e alla vita quotidiana della comunità locale.

⁶ Sono state inoltre segnalate alcune comunità in Germania, in Canada e negli Stati Uniti dove il francoprovenzale sarebbe in uso (Nagy, 2011; Kasstan, 2015).

I dati acustici che abbiamo analizzato, relativi alla Valle d'Aosta e alla Svizzera Romanda, provengono dal *Sound Comparisons Exploring Diversity in Phonetics across Language Families*, database online (dir. Paul Heggarty), frutto di un progetto elaborato al *Max Planck Institute for the Science of Human History*, a partire dal 2002 (cfr. Heggarty et al., 2019)⁷. *Sound Comparisons* ospita registrazioni di singole parole in centinaia di varietà linguistiche in tutto il mondo, ed è un progetto che concentra la sua attenzione sulle minoranze linguistiche. Nonostante il corpus non sia quantitativamente molto ampio, qualitativamente gli elenchi di parole che lo compongono sono stati concepiti per massimizzare la rappresentatività fonetica dei dati per ogni varietà.

Abbiamo fatto un confronto del materiale acustico relativo al francoprovenzale anche con dati acustici provenienti dal francese, il cui sistema vocalico di riferimento prevede dieci vocali orali e tre nasali /i e ε a o ɔ u y œ ø ē ā ã/ (Amélot, 2004; Georgeton, Paillereau, Landron, Gao, Kamiyama, 2012; Meynadier, 2013). I materiali acustici presi in considerazione nel nostro studio per il francese si basano invece su registrazioni personalmente effettuate a Lione e a Parigi nel 2019 (7 uomini e 4 donne tra i 21 e i 24 anni).

L'indagine effettuata è strutturata nel modo seguente. Nella sezione 3. ci siamo concentrati sull'analisi spettro-acustica (con annotazione in *textgrid*) di dati riscontrati sincronicamente a Faeto, che riflettono le evoluzioni del vocalismo tonico e atono del faetaro, il trattamento francoprovenzale specifico di /a/ tonica latina e atona finale da cui ci si aspetta quale risultato francoprovenzale (d'ora in avanti FP) prototipico [i] in contesto palatale e [a] in contesto non palatale (laddove il francese ha viceversa sempre [e/ε] in posizione tonica indipendentemente dal contesto, e elisione di /-a/ atona finale). Questi due tratti linguistici sono particolarmente significativi, in quanto distinguono il francoprovenzale da tutte le altre lingue romanze (Gardette, 1941; Bec, 1970-71; Martin, 1990; Tuailon, 2007).

Abbiamo effettuato inoltre in questa sezione l'analisi acustica di alcuni tratti della dittongazione romanza riscontrati a Faeto (per la serie velare dal latino Ő), quali il trattamento dell'approssimante *onglide* del dittongo in contesto palatale armonizzante (cfr. *infra* fr. *huit* 'otto' [ɥ] vs. FP [w]), e abbiamo valutato per Faeto la presenza vs. assenza, e il tipo di realizzazione delle vocali anteriori arrotondate, tipiche del francese /y œ ø/.

Nella sezione 4. abbiamo effettuato un'analisi con ispezione acustica delle vocali nasali riscontrate nel francoprovenzale di Faeto. Tali vocali presentano tratti ben distinti, come vedremo, rispetto al francese, che distingue unicamente in sincronia /ē ā ã/ (Amélot, 2004). Abbiamo interpretato le correlazioni tra il segnale acustico delle vocali nasalizzate e le tracce di occlusive nasali postvocaliche in posizione di coda sillabica. Abbiamo osservato l'andamento spettrale del segnale acustico, i livelli di energia (*amplitude*), le modifiche della sezione spettrale corrispondenti all'innalzamento e all'abbassamento del velo palatino.

⁷ https://soundcomparisons.com/#/en/Romance/map/eight/Lgs_Sln

Il nostro obiettivo in questa sede è quello di identificare la varietà sincronica del dialetto di Faeto in relazione alla varietà francoprovenzale di origine, ma anche rispetto alla continuità centro-meridionale italo-romanza, allo scopo di contribuire alla costruzione di una grammatica diacronica, sincronica e intergenerazionale di questa 'isola' linguistica francoprovenzale della Puglia.

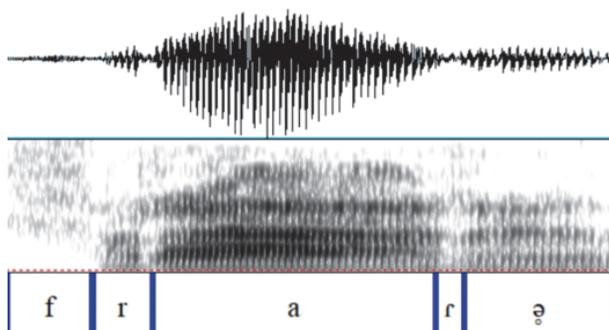
3. *Analisi dialettologica, acustica e fonetica del dialetto di Faeto*

3.1 Trattamento di /a/ tonica

Un tratto tipico del francoprovenzale, condiviso con l'occitano, assente nel gallo-romanzo settentrionale e quindi nel francese, è la conservazione di /a/ tonica latina in sillaba aperta (Gardette, 1941; Bec, 1970-71; Tuailon, 2007: XVII-XVIII, carte n° 20 e 21; ib.: 120-125; Russo, Stich 2020). Questo tratto in un *item* lessicale quale il lat. FRATRE viene realizzato nel francese con il *fronting* incondizionato di /a/ tonica [fɛʁ], mentre in dominio FP viene realizzato tramite la conservazione di /a/. Fa eccezione solo la realizzazione di /a/ tonica in contesto palatale come vedremo *infra* (Tuailon, 2007: XV-XVI, carte n° 18 et 19; ib. 115-120; Russo, in stampa a).

Tale conservazione tipicamente francoprovenzale è evidente a Faeto in *item* lessicali che presentano anche altri tratti francoprovenzali, si vedano gli spettrogrammi 1 e 2.

Spettrogramma 1 - FP Faeto ['frarɔ] 'fratello' vs. fr: [fɛʁ] – Trattamento FP di A tonica latina e lenizione gallo-romanza di /t/ nel cluster /tr/ (lat. FRATRE)



Il lat. FRATRE è alla base del francese (con *fronting* di A > [-posteriore]), del francoprovenzale (A = [+posteriore]), ma anche dell'italo-romanzo centro-meridionale (a nord della Roma-Ancona, com'è noto, la forma è suffissata con -ĒLLU).

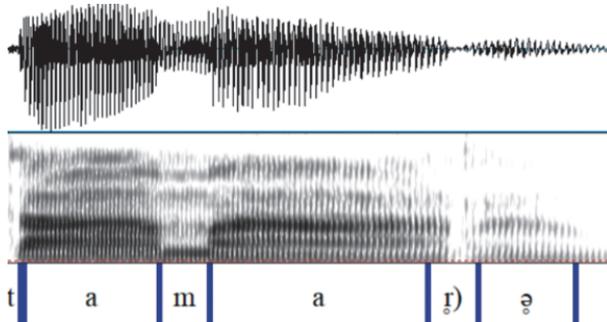
Se l'assenza di suffisso differenzia il faetaro dall'italiano, ma non dal contesto dialettale circostante, un altro tratto gallo-romanzo, la lenizione fonologica di /t/ intersonorica (qui nell'attacco biconsonantico latino TR a sonorità crescente) dif-

ferenza il dialetto di Faeto e l'italo-romanzo centro-meridionale⁸. La cancellazione gallo-romanza di T in posizione iniziale dell'attacco sillabico complesso /TR/ è visibile a Faeto nello spettrogramma 1, da cui la forma con la vibrante intervocalica in posizione di attacco della seconda sillaba [¹frarə]. Questa lenizione, fonologica, è condivisa dal francese, mentre l'italo-romanzo meridionale conserva le occlusive sorde intervocaliche. Lo si evince anche dagli esempi in (1) tratti dall' AIS 13 'tuo fratello, i tuoi fratelli' (cfr. NavigAIS online):

- (1) Faeto e Lucera a confronto in provincia di Foggia:
 [_{DP} [D° [def tum]]][_{NumP} [Num° SG] [_{NP} [N frar]]]
- a. [tum frar] (FP) (Faeto, p. 715) POSS + kin N
tuo-fratello
- b. [¹fratətə] kin N + POSS ENCL 2° Pers SG
fratello-tuo (Lucera/Foggia [*Dauno-Appenn.*], p. 707)

In (1) e nello spettrogramma 2 va osservata la linearizzazione diversa del possessivo rispetto all'italo-romanzo meridionale: il gallo-romanzo di Faeto presenta una linearizzazione prenominale del clitico nella posizione D (= determinante), identica a quella del francese (Zribi-Hertz, 1999) e dell'area francoprovenzale. Nelle altre località pugliesi e meridionali, ad eccezione delle varietà gallo-italiche del Meridione, invece, la posizione del clitico è post-nominale nei nomi di parentela (= *kinship names*, cfr. Delsing, Verner, 2002; Russo, 2021).

Spettrogramma 2 - FP [ta 'marə] 'tua madre' vs. fr. [ta mɛʁ] (*Fronting di A tonica*) MATRE



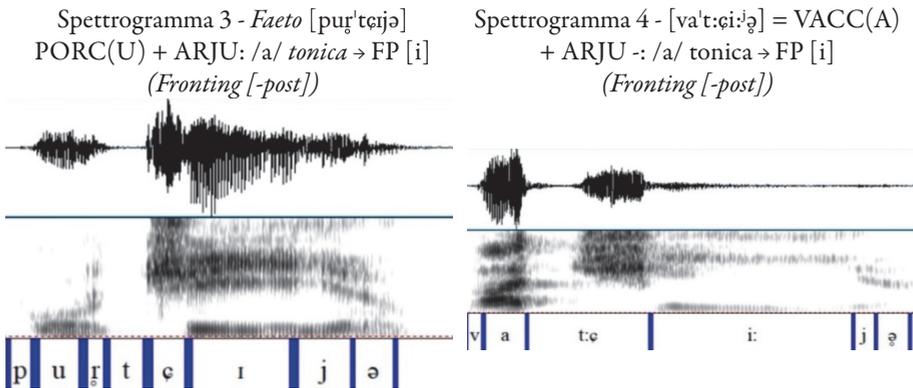
In generale, dopo i gruppi consonantici ostruente + liquida del latino (anche se ridotti), il FP presenta una vocale atona finale di appoggio (vedi spettrogramma 2). Negli spettrogrammi 1 e 2, in tale particolare contesto sillabico finale costituito da gruppo consonantico diacronicamente complesso (TR), la vocale finale centralizzata è propria sia dell'italo-romanzo, che del FP. Tale tendenza non è netta, in quanto nel francoprovenzale la cancellazione della vocale finale (diversa da /-a/) è anche possibile in contesti identici.

⁸ Tale tratto trova riscontro anche nei dialetti gallo-italici come quello di Picerno.

3.2 Trattamento francoprovenzale di /a/ tonica in contesto palatale a Faeto

La /a/ tonica diventa invece [i] se preceduta da segmento palatale, trattamento propriamente francoprovenzale (Tuailon, 2007: XV-XVI, carte n° 18 e 19; Russo, 2020a; Russo, Stich 2020). Generalmente nel dominio francoprovenzale tale trattamento palatale (il *fronting* condizionato di A tonica in [-i-]) è riservato principalmente alle classi verbali (cfr. spettrogramma 7 ‘accompagnare’)⁹. Tuttavia, abbiamo osservato che a Faeto la palatalizzazione di /a/ tonica si estende anche alla morfologia nominale, a lessemi autoctoni, quali quelli rappresentati dalla serie produttiva dei suffissi in -ĀRJU, indicanti nomi di mestiere, PORCĀRJU, VACCĀRJU ecc., si vedano gli spettrogrammi 4 e 5.

In questo processo, diacronia e sincronia s’incrociano, dato che il segmento palatale è fornito dalla palatalizzazione diacronica gallo-romanza di /ka/ in ambito francoprovenzale (Escoffier, 1958: 51-59 e carta n° 3 p. 55; Vurpas, 2004; Hinzelin, 2018; Russo, in stampa a; Russo, Stich 2020)¹⁰. La sillaba /ka/ subisce a Faeto palatalizzazione come accade nel gallo-romanzo, la vocale tonica /a/ viene a trovarsi in contesto palatale (con un segmento palatale precedente) e l’esito francoprovenzale di /a/ tonica in tale contesto è [i], come si vede negli spettrogrammi 3 e 4:



Va osservato che nel gallo-romanzo poi, non solo francoprovenzale, anche l’aprossimante /j/ seguente di -ĀRJU provoca un effetto di *Umlaut* sulla sillaba tonica precedente in alcuni *item* lessicali suffissati in -ĀRJU (fin dal latino medievale gallo-romanzo, cfr. Russo, 2014a/b: BUXARĪA *Boniras nel 1079 in Vallonia*, PŌMARĪUS > villa Pumirs supra Mosellam a. 1135)¹¹.

⁹ Cfr. le carte dell’*Atlas Linguistique et ethnographique du Lyonnais* (Ally) e dell’*Atlas linguistique de la France* (ALF) seguenti: Ally 592 e ALF 109 *Balayer*; Ally 918 *S’essuyer*; Ally 938 e ALF 329 *Coucher*.

¹⁰ Il risultato prototipico di questo processo di assibilazione in francoprovenzale è /ts θ tʃ/. Per la distribuzione diatopica e morfo-fonologica di questi esiti, cfr. Vurpas, 2004; Russo, Stich 2020.

¹¹ Questo processo è molto produttivo nel latino medievale gallo-romanzo, e investe principalmente le radici nominali suffissate in -ĀRJU, ma non solo. Per esempio dal lat. BĒSTJA in francese si sviluppano sia [biʃ] ‘fr. biche/it. cervia’ che [bɛʁ] ‘bestia’; l’effetto armonizzante di chiusura è da osservare solo

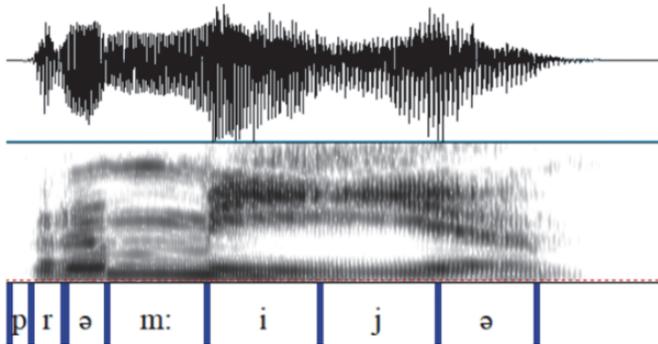
Nell'esito di PRIMĀRJU 'primo', base gallo-romanza che a Faeto prende il luogo di PRIMU, generalizzato in area italo-romanza, osserviamo la tonica armonizzata in [i], anche se la tonica non è preceduta da segmento palatale (si veda spettrogramma 5: [prəm'mijə] 'primo').

Il latino medievale PRIMĀRJU è anche alla base del francese *premier* [pʁemje], ma il *fronting* della vocale tonica in francese (/e/) non è condizionato da un *trigger* palatale e coincide con il *fronting* degli items [fʁɛʁ] 'fratello' e [mɛʁ] 'madre'. Il *fronting* del francese produce diacronicamente in sillaba aperta /e/, sia in [pʁemje], che in [mɛʁ], indipendentemente dal contesto palatale.

L'esito di /a/ armonizzato in [i] è prototipico nel francoprovenzale e si oppone all'esito francese realizzato con un *fronting* non condizionato da trigger [+alto]. L'esito francoprovenzale è quindi diverso dal francese. Nel francoprovenzale il *fronting* della tonica in [i] esibisce l'assimilazione dei tratti [+alto] e [-posteriore] integrati nella vocale tonica tramite un *trigger* metafonetico (/j/), con /a/ che diventa [i] (fenomeno di *Umlaut* più esteso geograficamente nel latino medievale gallo-romanzo).

Colpisce quindi a Faeto il caso di PRIMĀRJU con passaggio di -A tonica a [i] per *Umlaut* innescato dall'approssimante palatale seguente /j/, elemento [+alto] (presente nella sillaba seguente del suffisso -ARJU), che armonizza la vocale tonica precedente.

Spettrogramma 5 - Lat. PRIMĀRJU FP [prəm'mijə] 'primo' vs. Fr. [pʁemje] –
Fronting di /a/ condizionata da /j/ [+alto] nella sillaba finale

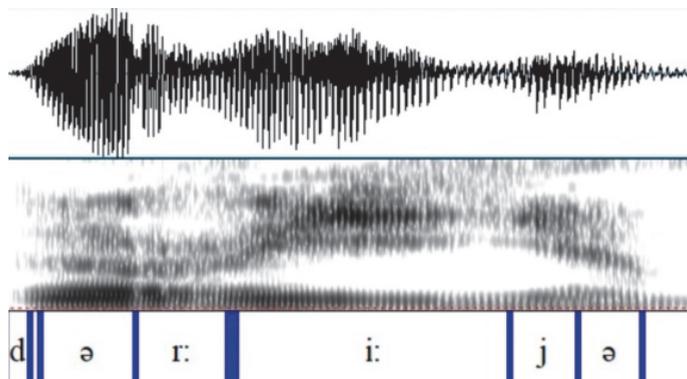


Questa forma gallo-romanza appare comunque adattata se si osserva la geminazione della nasale intervocalica, tipica dell'italo-romanzo meridionale e la presenza dello schwa finale nel suffisso -ĀRJU. Abbiamo verificato tale processo in altri lessemi riscontrati a Faeto, con suffisso in -ĀRJU. La nostra indagine dimostra che PRIMĀRJU non è l'unico lessema dove è osservabile la chiusura in [i] della vocale

sul primo *item*, [biʃ], nel quale l'approssimante palatale dell'ultima sillaba ha trasmesso per assimilazione il tratto alto alla vocale tonica precedente (cfr. Russo, 2014a/b).

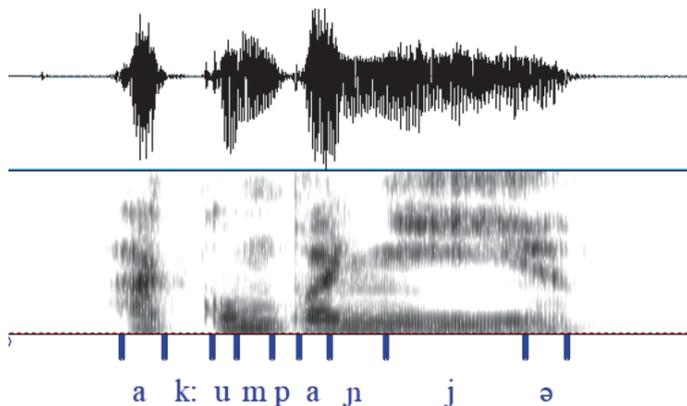
tonica; l'avverbio DE RETRO 'dietro', in francese 'derrière' è analogicamente ricomposto a Faeto sulla serie suffissale -ĀRJU, si veda lo spettrogramma 6:

Spettrogramma 6 - *Faeto Lat. DE RETRO FP [dər'rijə] 'dietro' – fr. [dɛʁjɛʁ] 'derrière' trattato analogicamente su -ĀRJU*



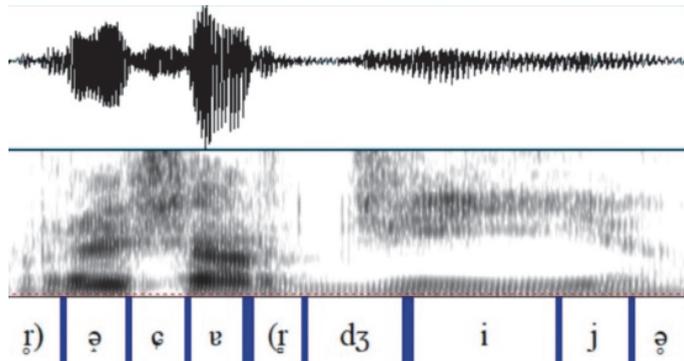
Il cambiamento fonologico di /a/ tonica in [i], tipico del francoprovenzale coinvolge a Faeto comunque anche le forme verbali aventi l'infinito della I^a classe in -ĀRE (e le forme coniugate dello stesso paradigma verbale), secondo quanto atteso in FP. Ciò accade quando una consonante palatale precede la /a/ tonica (cfr. lat. CA supra e spettrogramma 8), quest'ultima è armonizzata in [i]. Tale processo divide a Faeto, così come nel resto del dominio francoprovenzale, i verbi di prima classe in due categorie: una classe di verbi in -ĀRE che conserva la /a/ tonica quando non è preceduta da un segmento palatale e una classe derivata, caratterizzata dall'infinito con la tonica in [i]. Nell'esempio che segue il verbo *accompagnare* a Faeto mostra l'esito francoprovenzale palatalizzato all'infinito [akkumpa'ɲiə] nello spettrogramma 7; la forma francoprovenzale in [-i-] è seguita da uno schwa finale italo-romanzo:

Spettrogramma 7 - *Faeto FP [akkumpa'ɲiə] 'accompagnare' – ARE [-i-] -innescata da palatale precedente*



Dalla nostra ispezione spettro-acustica risulta che le forme in [i] precedute da palatale provenienti dall'infinito in -ĀRE sono sistematiche a Faeto, come si vede anche dallo spettrogramma 8, lat. CARRICĀRE (da confrontare con la carta dell'ALF 239 *charger du fumier*):

Spettrogramma 8 - Faeto Lat. CARRICĀRE -ĀRE → FP [i] [r̥]əçç(r̥dʒijə)
vs. fr. [ʃaʁʒe] 'charger'



Tuttavia, quello che ci è sembrato rilevante nell'ambito del trattamento di A tonica latina nel francoprovenzale di Faeto, è l'applicazione generalizzata dell'*Umlaut* di tipo gallo-romanzo innescato da /j/ seguente, che ha prodotto /i/ tonica a partire da A latina in tutti i contesti della serie suffissale -ĀRJU (anche in assenza di palatale nella consonante precedente), come in PRIMĀRJU. Le realizzazioni in [i] da -ĀRJU o -ĀRJA si giustificano per un effetto di chiusura innescato dall'approssimante palatale /j/ del suffisso.

Abbiamo effettuato un ulteriore controllo per accertare la natura del fenomeno, in quanto esso non sembra a Faeto legato al genere grammaticale. La nostra indagine fa emergere che, al di là del latino medievale gallo-romanzo, nell'area FP il suffisso -ĀRJU /-ĀRJA può quindi presentare risultati specifici, non solo per l'interazione con il femminile. In FP uno sviluppo atteso da PRIMĀRJA al femminile sarebbe FP *premeri* per influsso dell'approssimante seguente, con A atona finale (di -ĀRJA) che diventa [i] in FP quale desinenza di femminile preceduta da segmento palatale, in questo caso dal segmento /-j-/ di -ĀRJA.

A Faeto invece la A tonica dei suffissati in -ĀRJU diventa sistematicamente [-i-] anche al maschile, con trattamento identico a quello che nel francoprovenzale è riservato principalmente alle classi verbali del primo gruppo in -ĀRE, quando una consonante palatale precede il segmento vocalico tonico. Abbiamo potuto verificare che questo sviluppo particolare di -ĀRJU a Faeto è molto arcaico, dato che esso si ritrova in Francia nel Medioevo (Gonon, 1974), nell'area francoprovenzale del Forez (nella regione *Auvergne* Rhône-Alpes), e in altre regioni del dominio FP in Francia che hanno generalizzato -i- come a Faeto (cfr. Tabella 1):

Tabella 1 - Lat. A → [i] generalizzata per -ARJU a Faeto e nel Forez (Rhône-Alpes)

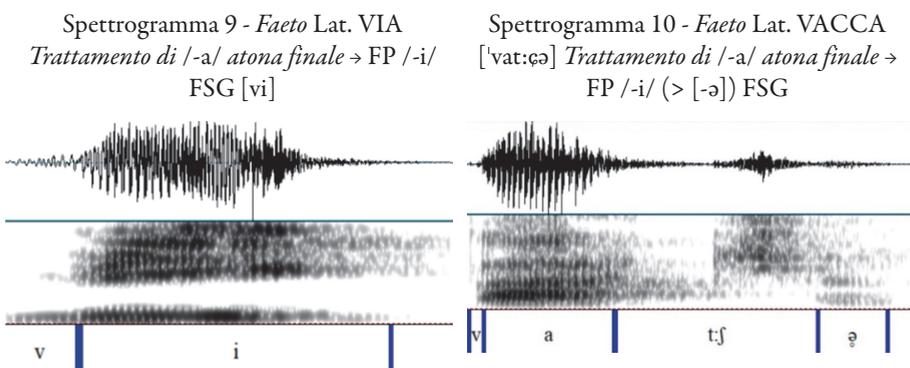
Forez (Loira)	Regione: Auvergne Rhône-Alpes	
Testi medievali	SG (M/F) -i/- -iri	PL (M/F) -ire
PRIMARJU	premi	A → [i]
POMARJU	pomi	A → [i]
*PRECARJA	prayiri	A → [i]
*COCHLEARJAS	quelires	A → [i]

Questo processo di *fronting* specifico nella serie suffissale -ĀRJU, evoluzione di -ĀRJU, nettamente distinta da quella dei dialetti italo-romanzi della stessa area, aumenta il carattere francoprovenzale e gallo-romanzo del dialetto di Faeto.

3.3 Trattamento di -A atona finale a Faeto

Il trattamento di -A atona finale a Faeto conferma, come il trattamento precedente, l'alto indice di francoprovenzalità presente in questa varietà alloglotta della Puglia nord-occidentale e le connessioni tra questa colonia e la sua area geografica di provenienza in Francia.

Il trattamento di -A latina femminile singolare (nella classe nominale e aggettivale afferente alla I^a declinazione latina), secondo cui -A atona finale latina preceduta da palatale muta in [-i], è un altro tratto tipico del dominio francoprovenzale (Tuaille, 2007: XVII-XVIII, carte n° 20 e 21; ib.: 120-125; Escoffier, 1958: 97-103 e carte n°14 p. 98, carte n° 16 p. 107; Russo, in stampa a). Esso è ben presente a Faeto, come si vede dagli spettrogrammi 9, 10 e 11.



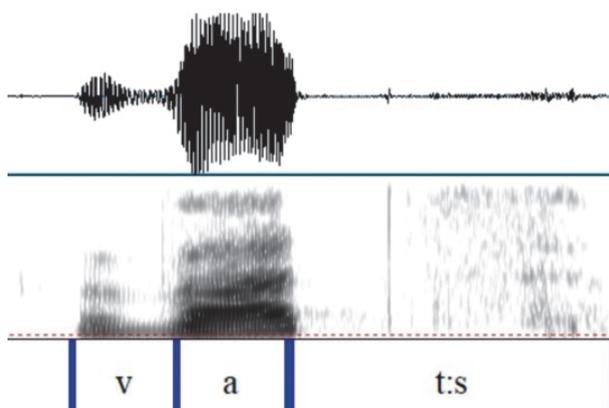
Nel lessema *la vacca* (spettrogramma 11) la sillaba /ka/ atona subisce la palatalizzazione diacronica e in tale contesto il francoprovenzale palatalizza la /a/ atona finale in /-i/ (risultato prototipico FP [vatsi]¹²). Questa /-i/ è realizzata come uno schwa a

¹² Cfr. Ally: 94 *Une torche de paille*, 112 (ALF 340) *Un coussinet*, 274 *La fane de rave*, (277* *Une vache*), 293 (ALF 348) *La creche*, 532 (ALF 1100) *Une puce*, 570 (ALF 267) *Une chenille*, 574 (ALF 770) *Une limace*, 580 (ALF 322) *Une coquille* 759 (ALF 664) *La grange*, 767 (ALF 133), *La bise*, 907 (ALF 453) *L'Église*, 882 *Le dimanche*, 1119 (ALF 264) *La chemise*.

Faeto, ma la consonante palatale al singolare femminile non lascia dubbi sulla ricostruzione della parola fonologica con /-i/ atona finale al FSG.¹³

Come si può osservare dallo spettrogramma 11 le forme di femminile singolare VACCA con elisione della vocale atona finale sono anch'esse riconducibili allo spazio francoprovenzale. Ad Evolène, baluardo della trasmissione del francoprovenzale oggi nella Svizzera romanda, la forma di FSG è sincronicamente [ˈvat:s], dove la palatalizzazione di /ka/ ha prodotto un'affricata alveolare, esito prototipico francoprovenzale (spettrogramma 11):

Spettrogramma 11 - Lat. VACCA [ˈvat:s] – *Trattamento di /-a/ atona finale* → FP /-i/ [-Ø]
FSG- Romandia (Evolène)



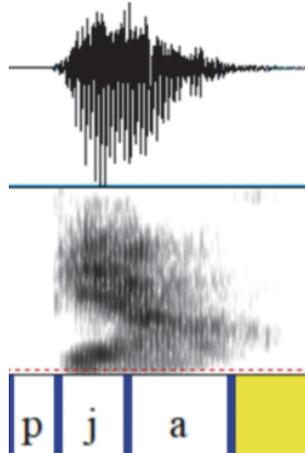
3.4. La dittongazione spontanea e le realizzazioni FP arcaiche del dittongo velare a Faeto

Abbiamo riscontrato a Faeto la dittongazione spontanea (non condizionata) delle vocali toniche medie semiaperte /e ɔ/ in sillaba aperta da È e Õ brevi latine. Questa dittongazione accomuna il francese e il francoprovenzale, tuttavia, il francoprovenzale non monotonga il dittongo velare, come invece accade in francese: CÔRE ‘cœur’ [kœʁ] (cfr. *infra*).

In questa comunità alloglotta della Puglia si riscontra il tipo di dittongo francoprovenzale più arcaico, sia per la serie palatale che per la serie velare.

Per la serie palatale, in FP la realizzazione prototipica da /e/ tonica semiaperta è [ja], per esempio da PËDE ‘piede’ abbiamo a Faeto [ˈpja]. L’apertura del nucleo vocalico sillabico è confermata dall’analisi formantica (spettrogramma 12):

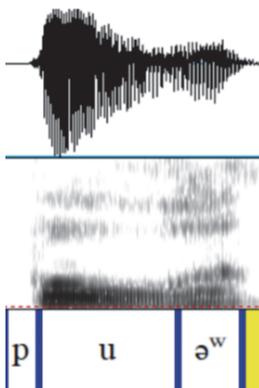
¹³ Nel FP medievale il singolare e il plurale della prima declinazione erano mantenuti distinti (-i < -A; -es < -AS). Nel francoprovenzale di Faeto, ma anche in larghe parti del dominio francoprovenzale, la distinzione di numero è ormai affidata al determinante e all’accordo sintattico di genere, in quanto [ˈvat:ɕə] è sia FSG sia FPL.

Spettrogramma 12 - *Faeto* – *Dittongazione spontanea* PĒDE ‘piede’: FP [‘pja]

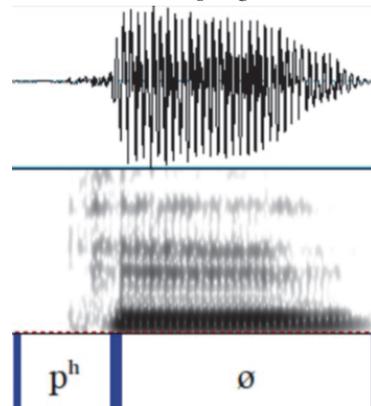
Nello spettrogramma 12 il nucleo del dittongo spontaneo è una vocale aperta, con F1 intorno ai 750 Hz e F2 1300 Hz.

Per la dittongazione della serie velare da Ö breve latina, quindi da /ɔ/ tonica semiaperta, il francoprovenzale conserva lo stato dell’antico francese, senza monotongazione successiva del dittongo, processo che nel sistema fonologico del francese in sincronia è alla base delle vocali anteriori arrotondate medie del francese /ø œ/, si vedano gli Spettrogrammi 13, 14 e 15. Nello spettrogramma 13 il dittongo velare dal lat. PAUCU è accentato sul primo elemento che a Faeto è velare e non labiale, come si osserva nello spettrogramma 14, in cui l’*item* lessicale ‘poco’ è pronunciato da un locutore parigino col monotongo labio-palatale. Lo si evince dall’altezza delle formanti F1, F2 e F3 (cfr. Georgeton et al., 2012), rispettivamente 458 Hz, 1700 Hz, 2600 Hz:

Spettrogramma 13 - *Faeto Lat.* PAUCU
FP *Dittongo velare* [‘puə] ‘poco’ /ɔ/ → /wɔ/
vs. fr. [pø] ‘peu’

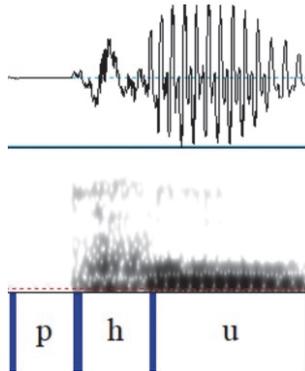


Spettrogramma 14 - *Monotongo labiale*
[pʰø] ‘poco’/fr. ‘peu’ – pronunciato
da locutore parigino



La pronuncia rappresentata nello spettrogramma 14 mette in relazione il faetaro con il resto delle varietà francoprovenzali, che presentano sia il dittongo velare composto da due segmenti vocalici (generalmente accentato sul primo elemento, che perde quindi il suo statuto di approssimante labio-velare per diventare elemento sillabico), sia il dittongo ridotto al primo elemento velare. Ciò è evidente dallo spettrogramma 15, che corrisponde all'*item* lessicale 'poco' pronunciato da un locutore francoprovenzale proveniente da Aosta:

Spettrogramma 15 - *Aosta Lat. PAUCI 'pochi' – Dittongo velare in FP – rappresentato solo dal primo elemento [phu]*



Sembra quindi che ad Aosta e a Faeto i dialetti FP moderni attestino il dittongo velare, realizzazione arcaica rispetto al francese. Infatti, nei primi documenti medievali dell'antico francese le vocali medie anteriori labiali [ø œ] non si erano ancora sviluppate in questo contesto.

3.5. Approssimanti a riprova del carattere arcaico e francoprovenzale dei dittonghi

È noto che il francese conosce tre approssimanti (anche se lo statuto fonologico di esse è discusso): [j w ɥ]. In francese, alcuni *item* lessicali con approssimante labio-palatale [ɥ] mostrano uno sviluppo palatale del dittongo in origine velare, originatosi in un contesto fonetico e fonologico palatale che favorisce la presenza dell'approssimante [ɥ], per esempio nel francese 'huit' [ɥit].

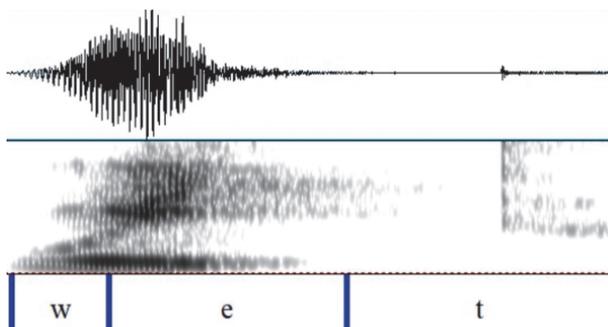
Nel lat. ÒCTO 'fr. huit/it. otto' oppure nel lat. HÖDJE 'fr. hui/it. oggi' (che sopravvive ormai solo agglutinato nel composto *aujourd'hui*), la /ɔ/ tonica (da lat. Ō) in francese diventa [ɥ] (= /y/), in quanto subisce un innalzamento (*raising*) in [u] e un *fronting* in [ɥ/y] per influsso di /j/ presente diacronicamente nel contesto sillabico¹⁴. Nel caso di ÒCTO /j/ si sviluppa a partire dalla consonante dorsale in posizione di coda sillabica, per lenizione consonantica in posizione debole, da cui si

¹⁴ Lo statuto sillabico [ɥ/y] (di nucleo vocalico o glide) dipende dalla posizione del segmento vocalico all'interno della sillaba.

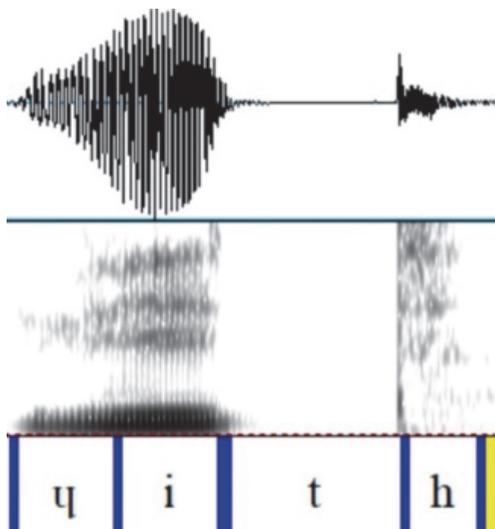
ottiene la vocalizzazione dell'ostruente /k/ in /j/). Nel caso di HÖDJE, invece, /j/ è già presente nella sillaba seguente¹⁵.

Lo sviluppo francoprovenzale della vocale tonica medioaperta /ɔ/ da Ö nello stesso contesto palatalizzante (tipo ÖCTO) si differenzia da quello del francese, dato che nel FP non si ha generalmente la riduzione del dittongo che precede l'approssimante /j/. In tali contesti in FP non appare l'approssimante labio-palatale /ɥ/, come si può osservare confrontando gli spettrogrammi 16, 17, 18 e 19, dove osserviamo l'*item* lessicale 'otto' pronunciato a Faeto, a Parigi, e nel francoprovenzale di Aosta e della Savoia (Chablais):

Spettrogramma 16 - *Faeto* [wɛt^h] 'otto' – assenza del glide [ɥ] vs. fr. [ɥit] 'huit'



Spettrogramma 17 - *Parigi Lat.* /kt/ [ɥit^h] 'otto'

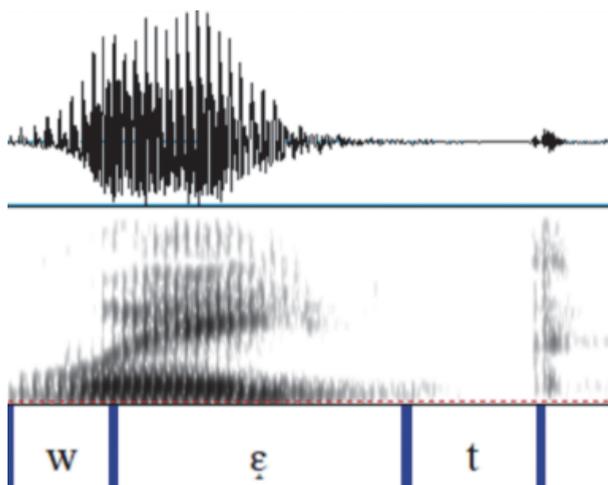


¹⁵ In francese, casi del tipo ÖCTO, in cui l'elemento palatale *trigger* del *fronting* e innalzamento della vocale tonica risulta dalla vocalizzazione della dorsale a partire da un *cluster* eterosillabico latino (/kt/ da cui [jt]), non sono isolati.

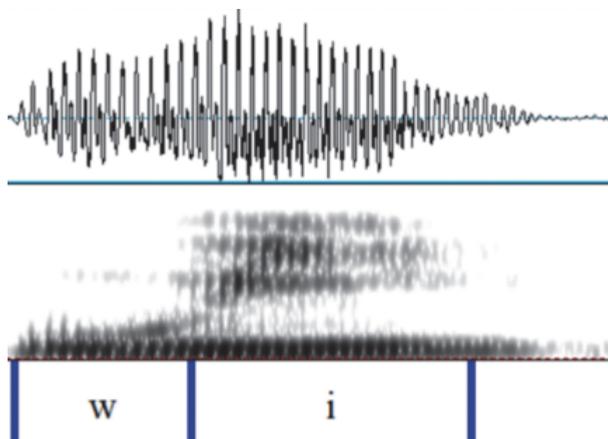
Negli spettrogrammi 16, 18 e 19, le transizioni formantiche indicano una modificazione della frequenza da un luogo di articolazione labio-velare per l'approssimante a un luogo palatale (della [i e] del nucleo vocalico successivo).

Nessuna delle varietà francoprovenzali esaminate mostra l'approssimante labio-palatale in questo contesto, le forme con l'approssimante labio-velare di Faeto sono coerenti con quelle della Valle d'Aosta e della Savoia, si vedano gli spettrogrammi 18 e 19:

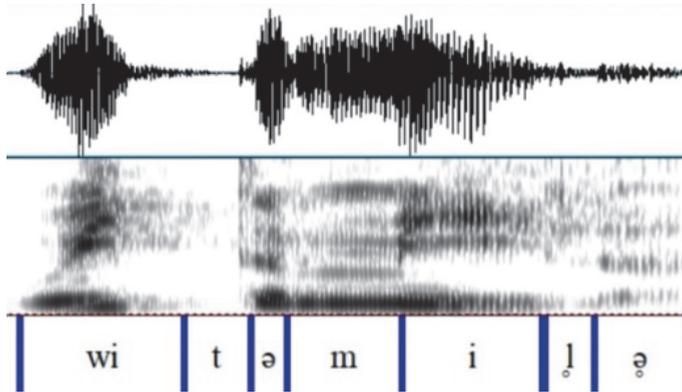
Spettrogramma 18 - *Aosta FP* ['wɛt] vs. *fr.* [ɥitʰ] 'otto'



Spettrogramma 19 - *Savoia – Chablais* ['wi] vs. *fr.* [ɥitʰ] 'otto'



A Faeto l'assenza di approssimante labio-palatale da questo tipo di contesto in cui è presente diacronicamente un *glide* palatale seguente è osservabile anche nei numerali composti (spettrogramma 20):

Spettrogramma 20 - *Faeto* FP [ˌwitə'miḷə] vs. *fr.* [qit mil] 'huit-mille' (ottomila)

La vocale tonica \bar{U} >/u/ mostra tuttavia tracce di palatalizzazione recenti in [y] nel francoprovenzale (Martin, 1983: 35; Kristol, 2016: 354; Kasstan, 2015; Hinzelin, 2018 con la bibliografia ivi indicata)¹⁶. A Faeto tale palatalizzazione risulta presente in alcuni lessemi, sebbene la vocale anteriore alta labiale [y] si presenti sotto forma di vocale delabializzata, cioè di [i]. Questo fenomeno è confermato da alcune carte dell' AIS (Faeto p. 715): 474 dove PŪLĪCE diventa [ˈppi:tʃ] 'pulce'; carta 858 dove MŪRUS diventa [ˈmmi:j] 'muro' e carta 993 dove CRŪDA diventa [ˈkritt] 'cru-da' (Perrin, 2018). Il fatto che questa labializzazione sia presente a Faeto potrebbe essere un indizio per datare la palatalizzazione di /u/ (\bar{U}) in [y] più precocemente in area francoprovenzale di quanto non sia stato fatto finora (Kattenbusch, 1982).

Anche nell'ALF, questa delabializzazione trova conferma al punto 920 (Le Grand-Serre nella regione di lingua francoprovenzale Auvergne-Rhône-Alpes) delle carte 364 'cru, crue', 890 'mur' e 1100 'puce', un punto appartenente alla presunta area di origine dei coloni di Faeto delimitata da Melillo (1959: 23-26; Perrin, 2018).

4. *Le vocali nasali a Faeto*

Il francoprovenzale conosce le vocali nasali (Kattenbusch, 1982; Kristol, 2016; Kasstan, 2015), e ha la particolarità di possedere nel suo sistema vocalico anche le vocali nasali alte [ī ū] in posizione sia tonica, che atona. Le vocali nasali del francoprovenzale risultano in generale più chiuse rispetto a quelle del francese (Gardette, 1941; Tuailon, 2007; Kristol, 2016); a Faeto '(tu) tieni' è pronunciato [tī] vs. francese [tjē]. Il contesto sillabico per la nasalizzazione delle vocali è storicamente VN con la nasale in posizione di coda sillabica.

Dalla nostra analisi emerge che Faeto conserva le vocali nasali del gallo-romanzo, integrate anche al lessico italo-romanzo.

¹⁶ Anche se alcuni documenti medievali scritti in lionese suggeriscono che questa palatalizzazione sarebbe stata introdotta tardivamente.

Le vocali nasali presentano caratteristiche più complesse delle vocali orali corrispondenti per quanto riguarda la descrizione acustica; interagiscono con vari parametri relativi ai tratti laringali (tra i quali *la breathy voice*). La complessità acustica delle nasali è data dalla presenza di extra-formanti e di antiformanti, gli ‘zeroes’, la cui presenza è dovuta all’uso congiunto della cavità nasale. In generale, i contrasti di altezza vocalica (come quelli tra una vocale alta [u] verso la medioaperta [ɔ]) appaiono acusticamente ridotti nelle vocali nasalizzate, in quanto la risonanza vocalica più bassa è nascosta dalla risonanza nasale e dall’antirisonanza. Tuttavia, la specificità delle vocali alte del francoprovenzale di Faeto ci è apparsa chiaramente.

Nella presente analisi abbiamo osservato diversi parametri per l’analisi delle vocali nasali pronunciate a Faeto, tra i quali la modificazione dei valori formantici rispetto alle vocali orali corrispondenti e l’abbassamento dell’energia (*amplitude*).

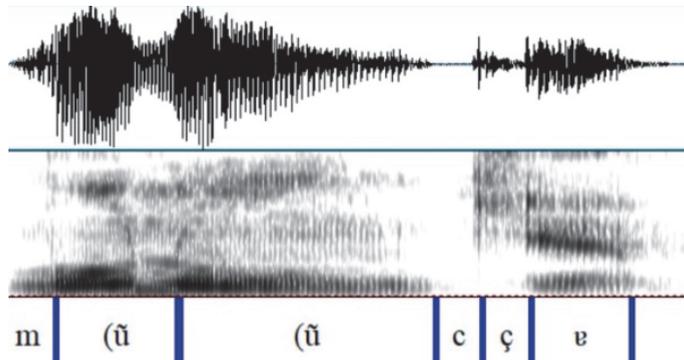
Abbiamo analizzato il *timing* e gli indici acustici delle vocali nasali del francoprovenzale di Faeto attraverso i diversi contesti sillabici, in posizione tonica e atona.

In particolare, nei nostri dati la nasalizzazione vocalica è espressa da un’elevazione dei valori di F1, un abbassamento di F2, e quale indice specifico di nasalizzazione, l’elevazione dei valori di F3.

A Faeto la classe funzionale dei possessivi è di tipo gallo-romanzo, con per esempio [mũ] ‘(il) mio’ che corrisponde al francese [mɔ̃] (spettrogramma 21).

Il possessivo con vocale alta nasale [mũ] (da confrontare col francese [mɔ̃]) è inserito nel sintagma determinativo, in posizione prenominali sintattica di determinante, illustrato dallo spettrogramma 21, con sequenza di due vocali nasali alte:

Spettrogramma 21 - Faeto FP [m(ũ)n(ũcʰə] ‘(la) mia unghia’ vs. fr. [mɔ̃nɔ̃gl]

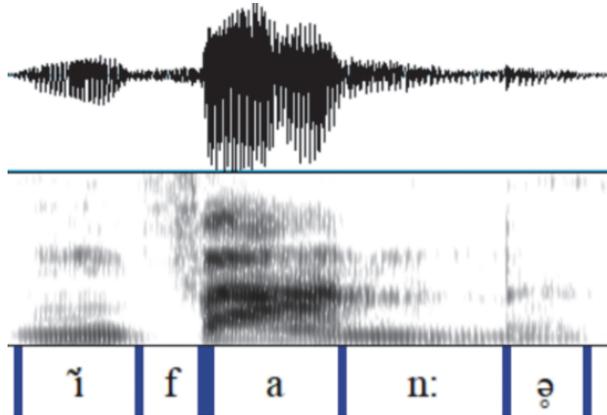


Il valore di F1 di una /u/ orale per una voce maschile si aggira intorno ai 300 Hz e di F3 ai 2000 Hz (Meunier, 2007; Martin, 2008; Meynadier, 2013). Si osserva nello spettrogramma 21 un’elevazione di F3 intorno ai 2500 Hz, e di F1 a 700 Hz (per il possessivo atono). Per quanto riguarda la vocale tonica nasalizzata della parola *unghia*, si osserva un netto abbassamento di F2 e un’elevazione di F3, oltre a un effetto *breathy* sulla seconda metà della durata vocalica.

La nasalizzazione riguarda i contesti di rima complessa VN in posizione atona e tonica, come si vede nello spettrogramma 22 con l’item [ĩ’fan:ə] ‘bambino piccolo’

che corrisponde al francese [ãfã], con la differenza che le vocali nasali medie del francese corrispondono alle vocali nasali alte in francoprovenzale:

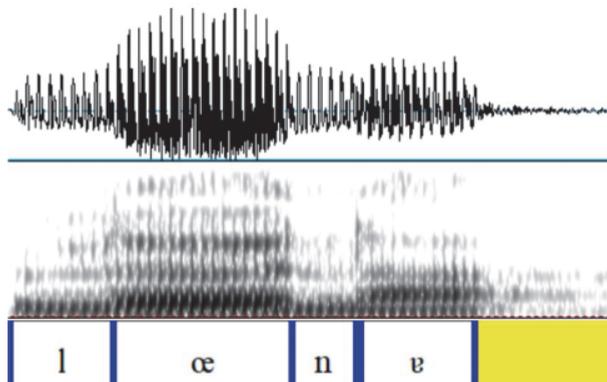
Spettrogramma 22 - FP [ĩ'fan:ə] 'bambino piccolo' vs. fr. [ãfã]



Nello spettrogramma 22 [ĩ'fan:ə], la vocale nasale è chiusa rispetto alla vocale francese corrispondente (così come nello spettrogramma 23). Inoltre F1 è intorno ai 430 Hz, mentre F2 si è abbassata (1600 Hz). È osservabile anche una notevole riduzione dell'energia.

Il processo di nasalizzazione delle vocali alte potrebbe essere dovuto a una sorta di *vowel shift* nel sistema vocalico del francoprovenzale, dato che la vocale Ū lunga del latino quando segue una consonante nasale tende ad abbassarsi [o ɔ], e ad essere realizzata nei dialetti odierni con timbri palatali arrotondati medi [ø œ] (cfr. spettrogramma 23 LŪNA [ˈlœna]):

Spettrogramma 23 - Aosta FP LŪNA [ˈlœna] 'luna'

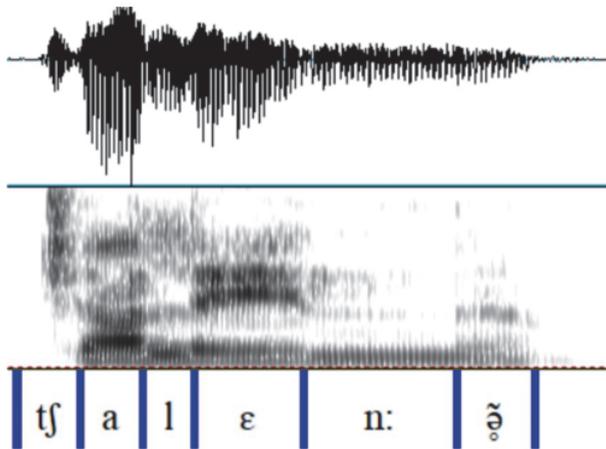


Nello spettrogramma 23, il carattere medio anteriore labiale di [œ] è evidente dalle misure di F1, F2 e F3, rispettivamente di 600 Hz (F1), 1580 (F2), e 2600 (F3).

Le vocali alte sono sistematicamente nasalizzate a Faeto se in contesto VN di sillaba chiusa; nel caso in cui invece la coda sillabica è parte di una nasale geminata risillabificata con l'attacco iniziale della sillaba successiva le vocali non sono nasalizzate (o lo sono solo parzialmente), come succede nel francese in cui per esempio l'aggettivo MSG [bɔ̃] 'bon/buono' perde la nasalizzazione della vocale al femminile FSG [bɔ̃n:ə] 'bonne/buona'.

Si veda qui la forma CALENDÀ 'Natale' nello spettrogramma 24 dove la nasalizzazione è debole su una durata ampia della vocale tonica in presenza di una nasale geminata:

Spettrogramma 24 - *Faeto* Lat. CALENDÀ FP [tʃa'ɛ̃n:ə] 'Natale'



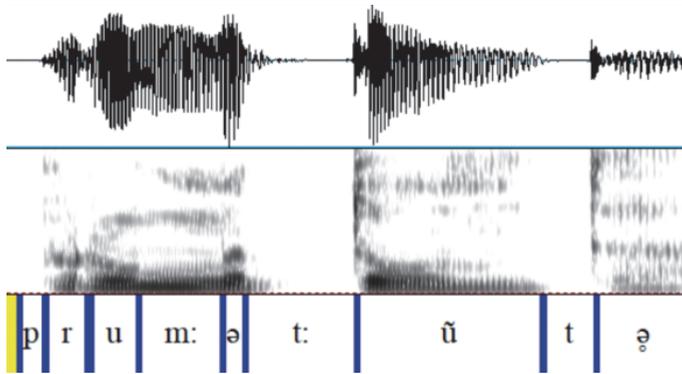
È da osservare inoltre in questa forma (CALENDÀ), usata a Faeto, e diffusa in tutto lo spazio francoprovenzale per indicare il Natale, la palatalizzazione gallo-romanza di CA latino iniziale in [tʃa] tipica del francoprovenzale e assente invece nell'italo-romanzo.

La vocale tonica risulta comunque parzialmente nasalizzata, ma l'energia nella parte bassa dello spettrogramma, intorno ai 300 Hz, segnala ancora la presenza della consonante nasale, con antiformali di cui si vede traccia a partire dai 1600 Hz (quindi con indicazione di luogo dentale)¹⁷.

La nasalizzazione delle vocali alte è sistematica anche nell'uso della 3PL *-unt* (Lat. *-UNT*) del presente indicativo, che abbiamo riscontrato a Faeto, desinenza tipica del francoprovenzale, si veda lo spettrogramma 25:

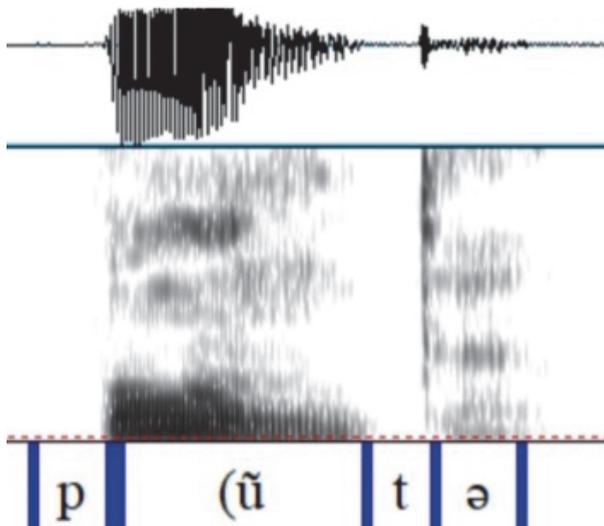
¹⁷ Per una descrizione dettagliata delle antiformali nelle consonanti e vocali nasali del francese, si veda Amélot (2004).

Spettrogramma 25 - *Faeto* [prum:ə't:ũtə] 'promettono' 3PL *presente indicativo*
Lat. -UNT [-ũtə]

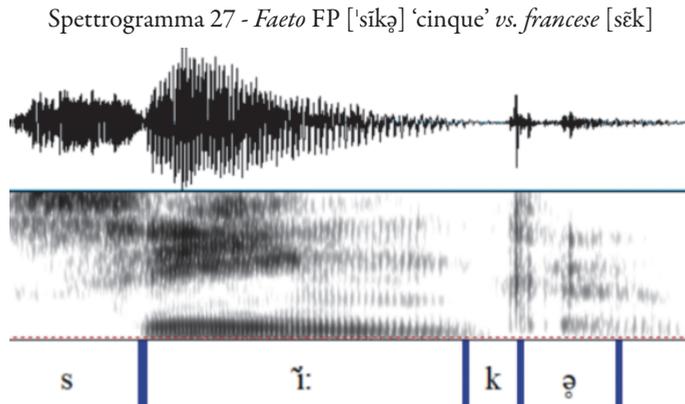


Nello spettrogramma 25 l'indice di nasalizzazione è dato da un elevato valore di F3 rispetto alla vocale orale corrispondente (3075Hz), l'elevazione di F1 e la diminuzione dell'energia sono visibili su tutta la durata vocalica.

Spettrogramma 26 - *Faeto* ['p(ũtə] 'ponte' vs. fr. [pɔ̃]



Nello spettrogramma 26, la nasalizzazione è visibile attraverso un abbassamento dell'energia su tutta la durata vocalica e un innalzamento di F3 (2700 Hz).



Nello spettrogramma 27, la nasalizzazione è tracciabile in base all’innalzamento di F1 e di F3 (3156 Hz).

L’associazione acustica delle cavità orale e nasale, mostra in genere formanti nasali supplementari, ma anche antiformalti, cioè una zona dove si osserva una caduta dell’intensità delle armoniche. Quest’antirisonanza o antiformalte è visibile (spettrogramma 27) nella seconda parte della vocale nasale (a 1538 Hz). Di conseguenza è visibile anche la riduzione dell’energia (*amplitude*) sulla F1 della seconda parte della [ĩ], fatto che indica un indice percettivo importante della nasalizzazione, anche se una diminuzione dell’energia (DB) appare comunque su tutta la durata della vocale tonica.

5. Conclusioni

Nonostante il carattere in fieri delle nostre analisi, svolte sul piano qualitativo e non quantitativo, riteniamo di aver aggiunto un tassello di riconoscimento al processo di definizione dell’identità linguistica di questa minoranza alloglotta della Puglia. Dalla nostra indagine è emerso chiaramente che, nonostante una situazione linguistica minoritaria che dura da diversi secoli, la lingua parlata a Faeto è ancora il francoprovenzale.

Per quanto riguarda l’influenza che i dialetti italiani, l’italiano standard o l’italiano regionale possono aver avuto sulla grammatica della lingua parlata a Faeto, tale influenza ci è apparsa ridotta. Faeto ha resistito ad alcuni cambiamenti morfologici emblematici dei dialetti centro-meridionali, in particolare alla metaforia italo-romanza e al raddoppiamento fonosintattico (Perrin, 2018). Nella grammatica francoprovenzale di Faeto i tratti francoprovenzali, fonetici, fonologici e morfologici, si sono combinati durante la trasmissione intergenerazionale al lessico italo-romanzo. Abbiamo mostrato, attraverso un campione di registrazioni rappresentativo, come le evoluzioni diacroniche discusse nelle sezioni 2. e 3. non risultano lessicalizzate, ma appaiono attive in sincronia, così come sembravano esserlo per Faeto all’epoca delle inchieste condotte per l’AIS tra il 1919 e il 1925.

Riferimenti bibliografici

- AGRESTI, G. (2016). L'enjeu de l'identité linguistique dans l'île francoprovençale des Pouilles. *Lengas* 79. <https://journals.openedition.org/lengas/1011>
- AIS = Jaberg, K., Jud, J. (1928–1940). *Sprach- und Sachatlas Italiens und der Südschweiz*, 8 vol. Zofingen: Ringier. <https://www3.pd.istc.cnr.it/navigais-web/> (NavigAIS Tisato, G. 2009-).
- ALF = GILLIÉRON, J., Edmont, E. (1902-1910). *Atlas linguistique de la France (ALF)*. Paris: Champion. <http://lig-tdcge.imag.fr/cartodialect5/#/>
- ALLy = Gardette, Pierre. 1950-1976. *Atlas linguistique et ethnographique du Lyonnais*. 5 vol. Lyon/Paris: Institut de linguistique romane des facultés catholiques/Éditions du CNRS.
- AMELOT, A. (2004). Etude aérodynamique, fibroscopique, acoustique et perceptive des voyelles nasales du français. Thèse de doctorat (en phonétique). Université Paris III – Sorbonne Nouvelle (dir. Jacqueline Vaissière).
- BEC, P. (1970-71). *Manuel pratique de philologie romane*, tome II, *Français, roumain, sarde, rhéto-frioulan, francoprovençal, dalmate*, Paris: A. & J. Picard.
- BOERSMA, P., WEENINK, D. (2019). *Praat: doing phonetics by computer*. University of Amsterdam. <http://www.fon.hum.uva.nl/praat/>
- CERQUIGLINI B. (1999). *Les Langues de la France*. Rapport au ministre de l'éducation nationale, de la recherche et de la technologie et à la ministre de la culture et de la communication.
- COLUCCIA, R. (1992). La Puglia, In Bruni, F. (Ed.), *L'italiano nelle Regioni. Lingua nazionale e identità regionali*, Torino: UTET, 685-719.
- COSI, P., FERRERO, F. & VAGGES K. (1995). Rappresentazioni acustiche e uditive delle vocali italiane. In *Atti del XXIII Congresso Nazionale AIA 1995*, Bologna, 12-14 Settembre, 151-156.
- DELSING, L.-O., EGERLAND, V. (2002). Kinship nouns in possessive constructions in Italian and Scandinavian. *Sprachtypol. Univ. Forsch. (STUF)*, Berlin 55, 103-120.
- DIÉMOZ, F. (Ed.) (2012). *Ernest Schüle: écrits sur la Vallée d'Aoste*, Quart. Vallée d'Aoste: Musumeci.
- ESCOFFIER, S. (1958a). *La rencontre de la langue d'oïl, de la langue d'oc et du francoprovençal entre Loire et Allier. Limites phonétiques et morphologiques*. Publication de l'Institut de linguistique romane de Lyon, vol. 11. Paris: Belles-Lettres.
- FANCIULLO, F. (2002). Puglia. In Cortelazzo, M., De Blasi, N. & Marcato, G. (Eds.), *Il dialetto nelle regioni*. Torino: UTET.
- FAVRE, S. (2010). Francoprovençale, comunità, http://www.treccani.it/enciclopedia/comunitafrancoprovencale_%28Enciclopedia_dell%27Italiano%29/.
- FERRERO, F. (1997). Problemi spettroacustici di misurazione delle vocali: un contributo, in *Atti delle VII Giornate di studio del Gruppo di Fonetica Sperimentale (A.I.A.)*, vol. XXIV, 235-264.
- GARDETTE, P. (1941). *Geographie phonétique du Forez*. Macon : Impr. Protat.
- GENDROT, C., ADDA-DECKER, M. (2005). Impact of duration on F1/F2 formant values of oral vowels: an automatic analysis of large broadcast news corpora in French and German. *Interspeech 2005*, Lisbon, Portugal. 2453-2456.

- GEORGETON, L., PAILLIEREAU, N., LANDRON, S., GAO, J. & TAKEKI K. (2012). Analyse formantique des voyelles orales du français en contexte isolé: à la recherche d'une référence pour les apprenants de FLE. Conférence JEP-TALN-RECITAL 2012, Grenoble, Francia. 145-152.
- GONON, M. (Ed.) (1974). *Documents linguistiques du Forez* (1260 -1498). Pierre Gardette, P., Monfrin, J. (Dir.) Paris: CNRS.
- HEGGARTY, P. (2019). Sound Comparisons: A New Resource for Exploring Phonetic Diversity across Language Families of the World. In Calhoun, S., Escudero, P., Tabain, M. & Warren, P. (Eds.), *Proceedings of the 19th International Congress of Phonetic Sciences (ICPhS 2019)*, 5-9 August, Melbourne, Australia. <https://assta.org/proceedings/ICPhS2019/>
- HINZELIN, M. (2018). Contact-induced change in Francoprovençal phonological systems caused by standard French. *International Journal of the Sociology of Language* 249, 49-70.
- KASSTAN, J. (2015). Illustrations of the IPA: Lyonnais (Francoprovençal). *Journal of the International Phonetic Association* 45.3, 340-355.
- KATTENBUSCH, D. (1982). Das Frankoprovenzalische in Südtalien. Studien zur synchronischen und diachronischen Dialektologie. Tübingen: Narr.
- KRISTOL, A. (2016). Francoprovençal. In Ledgeway, A., Maiden, M. (Eds.). *The Oxford Guide to the Romance Languages*. Oxford: Oxford University Press, 350-362.
- LOPORCARO, M. (2016). L'Italia dialettale. In Lubello, S. (Ed.), *Manuale di linguistica italiana*, Berlin: De Gruyter, 275-300.
- LOPORCARO, M. (2013). *Profilo linguistico dei dialetti italiani*. Roma/Bari: Laterza.
- LOPORCARO, M. (1997). Puglia and Salento. In Maiden, M., Parry, M. (Eds.), *The Dialects of Italy*. London-New York: Routledge, 338-348.
- MARTIN, J.-B. (1983). Le francoprovençal. *Nouvelles du Centre d'Études Francoprovençales René Willien* (Saint-Nicolas / Val d'Aoste), 8, 28-45.
- MARTIN, J.-B. (1990). Frankoprovenzalisch / Franco-Provençal. In Holtus, G., Metzeltin M., Schmitt C. (Eds.), *Lexikon der Romanistischen Linguistik* (LRL). Vol. 5.1. Tübingen: Niemeyer, 671-685.
- MARTIN, P. (2008). *Phonétique acoustique – Introduction à l'analyse acoustique de la parole*. Cursus. Paris: Armand Colin.
- MELILLO, M. (1959). Intorno alle probabili sedi originarie delle colonie franco-provenzali di Celle e Faeto. *Revue de linguistique romane*, 23, 1-34.
- MELILLO, M. (1981). Briciole francoprovenzali nell'Italia meridionale. *Vox Romanica*, 40, 39-47.
- MOSELEY, C. (Ed). 2010. *Atlas of the World's Languages in Danger*, 3th edition. Paris: UNESCO Publishing. www.unesco.org/culture/en/endangeredlanguages/atlas
- MOROSI, G. (1890-92). Il dialetto franco-provenzale di Faeto e Celle. *Archivio glottologico italiano*, 12, 33-75.
- MEUNIER, C. (2007). Phonétique acoustique. In Auzou, P., Rolland, V., Pinto, S. & Ozsancak, C. (Eds.), *Les dysarthries*. Marseille: Solal, 164-173.
- MEYNADIER, Y. (2013). Éléments de phonétique acoustique. In Adda-Decker, M., Nguyen, N. (Eds.), *Méthodes et outils pour l'analyse phonétique des grands corpus oraux*. Paris: Hermès, 25-83.

- NAGY, N. (1993). A geographic analysis of the origins of Faetar. *The Penn Review of Linguistics*, 17, 177-188.
- NAGY, N. (1996). *Language contact and language change in the faetar speech community*. University of Pennsylvania PhD dissertation, Philadelphia: IRCS.
- NAGY, N. (2000a). *Faetar*. Munich: Lincom Europa.
- NAGY, N. (2000b). Field work in Faeto, an endangered language community. *Southern Journal of Linguistics*, 1, 121-136.
- NAGY, N. (2011). Lexical change and language contact: Faetar in Italy and Canada. *Journal of Sociolinguistics* 15.3, 366-382.
- PELLEGRINI, G.B. (1975). I cinque sistemi dell'italo-romanzo. In Giovan Battista Pellegrini, *Saggi di linguistica: storia, struttura e società*, Torino: Boringhieri, 55-87.
- PELLEGRINI, G.B. (1977). *Carta dei Dialetti d'Italia*. Pisa: Pacini.
- PERRIN M. (2018). *Le francoprovençal des colonies du sud de l'Italie: Évolutions présentes à Faeto dans les cartes de l'AIS, l'Atlas linguistique et ethnographique de l'Italie et de la Suisse méridionale*. Sous la direction de Michela Russo. U. Lyon 3.
- PUOLATO, D. (2010). Lingua, dialetto e identità: percezioni e rappresentazioni della *enclave* francoprovenzale di Faeto. *Bollettino Linguistico Campano*, 18, 43-79.
- PUOLATO, D. (2011). Intrecci storici, linguistici ed identitari nella minoranza francoprovenzale di Celle di San Vito. *Bollettino Linguistico Campano*, 19/20, 91-149.
- PUOLATO, D. (2016). Langues en contact, langues en danger: une étude de cas autour de la négation verbale et autres mots-N en francoprovençal des Pouilles. *Congrès Mondial de Linguistique Française (CMLF)* 27, 13006. <https://doi.org/10.1051/shsconf/20162713006>
- RUSSO, M. (in stampa a). Les limites du Croissant dans l'Est francoprovençal (Forez) et dans le Nord occitan (Auvergne). In Esher, L., Guérin M., Quint, N., Russo, M. (eds.). Paris: L'Harmattan.
- RUSSO, M. (2021). The possessive enclitics with kinship nouns in Italo-Romance and the possessive determiners in the Franco-Provençal of Faeto. Two syntactic linearizations? *Zeitschrift für romanische Philologie* 137.1, 1-35.
- RUSSO, M. (2002). Metafonesi opaca e differenziazione vocalica nei dialetti della Campania. *Zeitschrift für romanische Philologie* 118.2, 195-223.
- RUSSO, M., VAN DER HULST, H. (2014). Proto-French Vowel Harmony: A Raising Metaphony in Merovingian Documents? Conference presented at the 12th Conference of the *French Phonology Network RFP*, Lille.
- RUSSO, M. (2014). Un processus métaphonique en latin mérovingien. Conférence présentée au Congrès international Latin Vulgaire–Latin Tardif 11, Oviedo.
- RUSSO, M., STICH, D. (2019). Les systèmes graphiques du francoprovençal: état de lieux et perspectives. Quels rapports diasystème entre graphie supra-dialectale et phonologie? *Lengas* 86. Url: <https://journals.openedition.org/lengas/>.
- STEHL, Th. (1988). Apulien und Salento / Puglia e Salento. In Holtus, G. – Metzeltin, M. e Schmitt, C., *Lexicon der Romanistischen Linguistik*, vol. IV, Tübingen, Niemeyer, 695-716.
- STEHL, Th. (1980). *Die Mundarten Apuliens. Historische und strukturelle Beiträge*. Münster: Aschendorffsche Verlagsbuchhandlung.

- TELMON, T. (2001). Piemonte e Valle d'Aosta. In Sobrero, A. (Ed.), *Profili linguistici delle regioni*. Bari: Laterza.
- TUAILLON, G. (2007). *Le Francoprovençal. Tome Premier*. Aosta: Musumeci.
- VALENTE, V. (1975), *Puglia, Salento* (Profilo dei dialetti italiani 15), Pisa, Pacini.
- VURPAS, A.-M. (2004). Les alternances de palatales en francoprovençal d'après le patois de Chambost-Allières (Rhône). *Travaux de dialectologie francoprovençale*. Lyon: IPG, UCLY, 13-22.
- WEINRICH, H. (1958), *Phonologische Studien zur romanischen Sprachgeschichte*, Münster, Aschendorff.
- ZRIBI-HERTZ, A. (1999). Le système des possessifs en français standard moderne. *Langue française* 122. *Le groupe nominal: contraintes distributionnelles et hypothèses de descriptions*. 7-29.

ALESSANDRO VIETTI, DANIELA MEREU¹

Sistemi vocalici in contatto nell'italiano di Bolzano: un'analisi esplorativa *corpus-based*

Vowel systems in contact in Bolzano Italian:
a *corpus-based* exploratory analysis

The aim of our paper is to illustrate an exploratory analysis on the vowel system in Italian spoken in Bolzano by both native and non-native speakers (i.e. Tyrolean German speakers). Data for the analysis come from DIA (Dialogic ItAlian) corpus, a speech corpus of spontaneous Italian spoken in Bolzano and based on dialogic interactions. After a short illustration of the studies on Italian vowels and a phonetic description of the Italian speech variety of Bolzano, we provide some information on the construction of DIA corpus. The central part of the contribution is devoted to the discussion of the specific methodological problems that arise in the acoustic analysis of vowels extracted from speech corpora. To address these problems, we carry out a spectral analysis based on DCT coefficients and we test the hypothesis that there are two vowel sub-systems using linear mixed models.

Keywords: Italian vowels; spontaneous speech; spectral analysis (DCT coefficients).

1. Introduzione

L'obiettivo principale di questo contributo è di presentare un'indagine esplorativa sulle vocali dell'italiano di Bolzano. Il suo carattere esplorativo si spiega con la necessità di definire una procedura di analisi che possa essere valida per lo studio delle vocali estratte da corpora di parlato spontaneo. Il lavoro che qui si presenta si basa infatti sul DIA-Dialogical ItAlian, un corpus di parlato spontaneo della varietà di italiano bolzanino, costruito su un campione di parlanti nativi e non nativi di italiano, ovvero tedescofoni. Prima di entrare nel vivo delle questioni metodologiche e procedurali dell'analisi, forniremo una descrizione della varietà oggetto di indagine, sulla base di una revisione della bibliografia esistente sull'argomento.

Inizieremo, quindi, con alcune note di sintesi sugli studi vocalici per l'italiano (par. 2), in seguito tratteremo una breve rassegna sull'italiano di Bolzano, dando una particolare attenzione alle vocali (par. 3), per poi concentrarci sulla metodologia adottata per la costruzione del corpus di riferimento e per la preparazione dei dati da analizzare (par. 4); la sezione centrale del lavoro si focalizzerà invece sull'a-

¹ Nel suo complesso, il lavoro è il frutto di una collaborazione tra i due autori; tuttavia, per fini accademici, segnaliamo che Alessandro Vietti è responsabile delle sezioni 5, 5.1, 5.2, 5.3, 6, mentre Daniela Mereu è responsabile delle sezioni 1, 2, 3, 4 (4.1, 4.2).

nalisi esplorativa delle vocali (par. 5), chiarendo dapprima la natura dei problemi metodologici legati all'analisi acustica di grandi quantità di dati spontanei (par. 5.1) e, successivamente, illustrando i risultati di un'analisi spettrale condotta con l'impiego di modelli lineari misti (par. 5.2). All'analisi seguiranno alcune osservazioni conclusive (par. 6).

2. *Le vocali dell'italiano*

Gli studi di fonetica sperimentale sulle vocali in italiano vantano una lunga tradizione, a partire dai primi lavori di Ferrero (1968; 1972), che si sono concentrati sulle vocali dell'italiano, senza fare distinzione tra varietà regionali diverse². Da questi lavori pionieristici, le ricerche sulle vocali sono state in seguito affinate secondo diverse prospettive. Una prima linea di indagine ha riguardato lo studio degli spazi vocalici in alcuni dialetti e nelle rispettive varietà locali di italiano (es. Mendicino & Romito, 1991; Romito & Lorenzi, 1997). Parallelamente, sono state sviluppate anche delle riflessioni di carattere metodologico sulla misurazione dei segmenti vocalici, come dimostrano i contributi scientifici presentati al primo convegno italiano su questo tema, sfociato nel volume a cura di Trumper e Romito (1993).

Progressivamente, la ricerca sulle vocali è stata ampliata sia in prospettiva diatopica, attraverso il confronto tra diversi punti linguistici di una stessa varietà e in relazione a diverse varietà di italiano regionale, sia in prospettiva sociolinguistica, con una particolare predilezione verso la dimensione diafasica di variabilità. Questo incremento di studi è stato reso possibile grazie anche alla costruzione di basi di dati di parlato, diatopicamente e stilisticamente diversificate, come il corpus AVIP (Bertinetto, 2001), API (Crocco, Savy, Cutugno, 2002), il DIVA (Albano Leoni, Cutugno & Savy, 1995; 1998) e il CLIPS (es. Albano Leoni, 2006; 2007), sui quali si sono basati molti dei lavori che hanno preso in esame le vocali.

Le principali direttrici di studio e di ricerca seguite hanno riguardato: a) la descrizione di singoli sistemi vocalici, con il confronto tra il sistema tonico e quello atono, es. Nocchi & Calamai (2009), per il parlato toscano in tre punti linguistici; b) l'analisi contrastiva dei sistemi vocalici in chiave diatopica – per es. Clemente, Savy & Calamai (2003), in cui vengono analizzati i sistemi vocalici di sette varietà diverse, riferibili alle città di Palermo, Napoli, Lecce, Pisa, Roma, Milano, Torino; c) l'analisi stilistica delle vocali, in funzione di diversi stili di parlato (lettura di singole parole, lettura di frasi, parlato semi-spontaneo), con una particolare attenzione verso i fenomeni di riduzione e centralizzazione – es. Savy & Cutugno (1996), che si focalizza su diversi fenomeni di riduzione, strutturale e non strutturale, in differenti varietà regionali (Lombardia, Toscana, Lazio e Campania), e in diversi stili di parlato, anche se solo per la varietà campana; Lo Prejato, Clemente, Savy (2003),

² In Calamai (2003a) è presente una dettagliata rassegna dei lavori condotti fino a quel momento sulle caratteristiche acustiche dei sistemi vocalici dell'italiano.

sulla riduzione vocalica nella varietà napoletana; Calamai & Sorianello (2003), sul sistema vocalico romano; Calamai (2003b), sul sistema vocalico pisano.

Prima di entrare nel vivo dell'italiano parlato a Bolzano e, in particolar modo, del suo sistema vocalico, sarà opportuno fornire un breve quadro del sistema vocalico dell'italiano standard, in relazione ai diversi italiani regionali.

Il sistema linguistico dell'italiano standard è composto da sette fonemi vocalici (/i, e, ε, a, ɔ, o, u/), ma non tutte le varietà di italiano mostrano opposizione fonologica tra vocali medio-alte e medio-basse in sillaba tonica. Inoltre, anche laddove esiste tale distinzione, questa si caratterizza per un peso funzionale molto basso.

Sintetizzando³, i diversi sistemi vocalici dell'italiano possono essere ricondotti a tre diversi tipi (Mioni, 2001b; Vietti, 2019): a) sistemi con sette fonemi vocalici, in cui la distribuzione delle vocali medio-alte e medio-basse può anche non coincidere con quella dello standard (oltre alla gran parte dell'Italia centrale, della Campania, della Basilicata e del Molise, in questo primo tipo rientrano anche Veneto e Trentino, di particolare interesse per il nostro discorso); b) sistemi in cui sono presenti sette foni, ma solo cinque fonemi vocalici, perché la distribuzione delle vocali medie è condizionata dal contesto fonetico e quindi non crea opposizioni fonologiche (Lombardia, Piemonte, Emilia-Romagna, Friuli e, probabilmente, Liguria; anche l'italiano regionale di Sardegna mostra una situazione analoga, per effetto della metaforia); c) sistemi vocalici con cinque foni e cinque fonemi vocalici, in cui le vocali medie mostrano valori intermedi tra il punto medio-alto e quello medio-basso (Italia meridionale estrema e Venezia Giulia).

3. *L'italiano di Bolzano con particolare riguardo per le vocali*

Sebbene l'analisi che verrà presentata al par. 5 abbia carattere esplorativo e sia finalizzata quindi più all'elaborazione e alla definizione di una procedura adatta allo studio sulle vocali nei corpora di parlato spontaneo che non alla ricerca di risultati sulla varietà (o sulle varietà) oggetto di indagine, ci sembra possa essere utile al lettore fornire dapprima una descrizione delle caratteristiche fonetiche proprie delle varietà di italiano presenti a Bolzano, ovvero la varietà dei parlanti nativi e la varietà dei parlanti tedescofoni. Il profilo linguistico qui presentato metterà in luce la posizione rilevante che le vocali occupano ai fini della caratterizzazione di queste due varietà.

Il repertorio linguistico della città di Bolzano viene tipicamente definito come un bilinguismo sociale di tipo bicomunitario, in quanto sono presenti due comunità linguistiche prevalentemente monolingui (una italiana e una tirolese) (cfr. Mioni, 1990b). Tuttavia, nel corso degli ultimi decenni la proporzione di parlanti bilingui, in particolare nella comunità tedesca tirolese, è in costante aumento nel contesto ur-

³ Per un'illustrazione esaustiva dei diversi sistemi vocalici dell'italiano, si rimanda a Schmid (1999), Mioni (2001b), Vietti (2019).

bano, al punto che la pratica bilingue e la presenza di varietà di contatto rappresenta la norma (Vietti 2011).

L'italiano parlato a Bolzano è da considerare un caso piuttosto singolare nel panorama sociolinguistico italiano, per il fatto che il suo sviluppo non si configura come l'esito di un processo storico di convergenza dei dialetti verso l'italiano, analogamente a quanto è avvenuto nel caso delle altre varietà regionali di italiano, ma piuttosto come il risultato di una trasformazione repentina del contesto sociolinguistico, che ha comportato un processo di livellamento e koinizzazione⁴.

Per la sua singolarità, l'italiano bolzanino è stato, a più riprese, oggetto di attenzione da parte soprattutto di Mioni (1990a, 2001a), che ha analizzato questa varietà a partire da dati provenienti da informatori riconducibili alla seconda generazione di italofoeni nati a Bolzano (o nella provincia). Sulla base di questi dati, e grazie anche al confronto con un campione di informanti padovani, Mioni (1990a: 201) ha definito l'italiano di Bolzano come una varietà di italiano di Nord-Est più avanzato di quello veneto verso la standardizzazione, attribuendo questa maggiore standardizzazione alla particolare storia dell'immigrazione di italofoeni, avvenuta nel periodo successivo alla Prima guerra mondiale verso la città di Bolzano. Nei primi anni, infatti, nonostante la presenza di un'importante componente di dirigenti lombardi ed emiliano-romagnoli, la parte più consistente della popolazione era costituita per lo più da trentini e veneti, mentre l'arrivo degli italiani provenienti dalle regioni centro-meridionali è da ricondurre al periodo del secondo dopoguerra, ovvero solo quando ormai il modello di riferimento locale per la pronuncia dell'italiano si era già formato e consolidato.

La pronuncia prevalente sarebbe quindi quella veneto-trentina ma con un grado maggiore di standardizzazione (Mioni, 1990a; 2001a). Dal confronto tra il campione padovano e i due campioni bolzanini, uno italofono e l'altro germanofono, Mioni (1990a) individua alcuni fenomeni che contraddistinguono queste tre varietà. Considerato il fine del presente lavoro, menzioneremo solamente i tratti caratteristici delle due varietà di italiano parlate a Bolzano.

L'italiano bolzanino del gruppo di italofoeni si caratterizzerebbe per:

- l'opposizione fonologica tra le vocali medie /e/ ed /ɛ/, in riferimento però a parole diverse rispetto a quelle previste dallo standard (questo tratto accomuna l'italiano di Bolzano a quello di Padova);
- la realizzazione più standardizzata rispetto a Padova delle vocali medie posteriori /o/ e /ɔ/;
- la realizzazione dei fonemi /ʃ/ e /ʎ/ con una durata non lunga analogamente a quanto si registra per Padova, ma con una frequenza maggiore di varianti standard rispetto alla varietà di italiano padovano;

⁴ Per una dimostrazione dell'applicazione al caso bolzanino dei principi di koinizzazione e dei fattori macro-sociolinguistici così come formulati da Kerswill, Williams (2000), si veda Vietti (2017).

- una distribuzione simile a quella padovana dei fonemi /ts/ e /dz/⁵, con tracce sporadiche di realizzazioni fricative dentali [s, z], da ricondurre all'influenza di famiglie di origine trentina o emiliana.

Verso un quadro dell'italiano bolzanino come varietà più standardizzata rispetto all'italiano regionale veneto e trentino, soprattutto in riferimento alla produzione delle vocali medie, sembra convergere anche la descrizione di Canepari (1999). L'autore sostiene infatti che negli item lessicali in cui l'italiano di Bolzano diverge rispetto alle altre due varietà nord-orientali, la differenza risiederebbe in una pronuncia maggiormente standardizzata per l'italiano altoatesino: es. la parola *nebbia* realizzata come [ˈnebbja], *bosco* come [ˈbɔsko], *nome* come [ˈnɔme] (Canepari, 1999). Tuttavia, anche per l'Alto Adige, secondo Canepari (1999), i fonemi /e, ɛ/ e /o, ɔ/ sarebbero dei semifonemi, ovvero delle “acquisizioni precarie e capricciose” che mostrano grandi oscillazioni, motivo per cui possono essere solamente descritte – con riferimento al sistema fonologico standard – le differenze nella loro distribuzione lessicale dei diversi timbri (Canepari, 1999).

Passando all'italiano parlato dai tedescofoni, Mioni (1990a) individua un elenco di tratti, riconducibili quasi totalmente al contatto con il tirolese e alle differenze presenti nei due inventari, per i quali questa varietà si contraddistingue. I fenomeni in questione sono:

- la differente modalità di distinzione tra consonanti sorde e sonore;
- la durata consonantica, inesistente in tedesco;
- la durata vocalica, con la tendenza a pronunciare come lunghe tutte le vocali in posizione finale di sillaba;
- la realizzazione di /r/, che nel dialetto tirolese è prodotta come un'uvulare⁶;
- la realizzazione delle affricate/tʃ/ e /dʒ/, in cui il componente occlusivo mostra un distacco rispetto a quello fricativo;
- la velarizzazione di /l/ in fine di sillaba;
- la frequente realizzazione delle sibilanti /s, z/ come dentali o pre-dentali;
- una maggiore standardizzazione delle vocali medie in posizione tonica, sia in sillaba aperta che chiusa.

La realizzazione delle vocali medie toniche come medio-alte [e, o] in sillaba aperta e come medio-basse in sillaba chiusa [ɛ, ɔ], per effetto del contatto col tirolese, spiega il trattamento molto standardizzato di /e/ in sillaba aperta, ma non fornisce una spiegazione del rendimento altrettanto buono di queste vocali in sillaba chiusa. Dal confronto con i dati padovani, Mioni (1990a) infatti, oltre all'esito standardizzato in sillaba aperta, osserva anche una maggiore standardizzazione in sillaba chiusa tra i parlanti tedescofoni di Bolzano rispetto ai parlanti padovani.

A conclusione del suo studio, l'autore sostiene che il modello di riferimento per l'apprendimento della fonologia italiana da parte dei tedescofoni è costituito dalla fonologia degli italo-foni altoatesini.

⁵ Per un'analisi sociofonetica di questi fenomeni si rimanda a Meluzzi (2020).

⁶ Un'analisi acustica di /r/ in tirolese è illustrata in Galatà, Spreafico, Vietti, Kaland (2016).

4. Metodologia

4.1 Il campione dei parlanti e i dati

I dati presi in considerazione in questa sede derivano dal corpus DIA-Dialogic ItAlian, un corpus di parlato spontaneo sulla varietà di italiano di Bolzano, proveniente da parlanti italofofoni e tedesconi, che rappresenta la base di dati per il progetto di ricerca NEB – *Networks, Exemplars and Bilinguals. The potential of quantitative social network analysis in sociophonetic research*. Considerato che lo scopo del progetto è l'esplorazione degli effetti delle strutture delle reti sociali personali sul comportamento linguistico di parlanti bilingui, la creazione di un corpus di parlato spontaneo è risultata essere la scelta metodologica più appropriata.

Il campione consta di 40 soggetti (età: 18-65 anni; 14 M, 26 F): 23 parlanti nativi di italiano, 12 parlanti nativi di tirolese (bilingui sequenziali) e 5 bilingui simultanei. Dal punto di vista del grado di istruzione e dell'occupazione lavorativa dei parlanti, il corpus si caratterizza per essere molto eterogeneo⁷.

Per la raccolta dei dati, è stato usato un registratore Zoom H4, con il campionamento a 44.100 Hz e la digitalizzazione a 16-bit; al fine di incrementare ulteriormente la qualità, sono stati usati anche due microfoni ad archetto (Shure SM35), uno per ogni informante. Ogni sessione di registrazione ha previsto tre fasi distinte: 1) un'interazione dialogica in coppia tra persone che si conoscevano bene tra loro (coppie di amici, marito e moglie, parenti, etc.); 2) la somministrazione di un questionario volto all'elicitazione di informazioni sulle reti sociali personali, oltre che ai dati di ordine socio-demografico; 3) la lettura di una lista di frasi. Complessivamente i dati ammontano a 30h 43' 25" di parlato, suddivise nel modo seguente: 9h 49' 32" di parlato spontaneo, 19h 03' 26" di parlato proveniente dalle interviste sulle reti sociali personali e 1h 50' 27" di parlato letto.

Per l'analisi che presenteremo nei prossimi paragrafi, sono stati usati i dati delle interazioni dialogiche, che si sono basate su una lista di argomenti comuni (es. problemi sociali a Bolzano, feste tradizionali locali, situazione politica della città, etc.), somministrata ai parlanti, i quali potevano decidere di prendere spunto o meno da questa lista. In totale, i dialoghi registrati sono 20, della durata di circa 30 minuti ciascuno.

La quantità di parlato spontaneo presa in esame in questa sede è di poco più di 30 minuti. Il campione di riferimento è costituito da 8 parlanti donne, di cui 4 native di italiano e 4 native di tirolese (bilingui sequenziali), di età compresa tra i 47 e i 59 anni. Per ogni parlante sono stati analizzati 4 minuti di parlato spontaneo.

4.2 Trascrizione ortografica e annotazione fonologica

I dati sono stati dapprima trascritti ortograficamente mediante lo strumento dei sottotitoli offerto da YouTube (Cangemi, Fründt, Hanekamp & Grice, 2019) e, in

⁷ Per una descrizione dettagliata della costruzione del corpus, ci permettiamo di rinviare a Mereu, Vietti (2021).

seguito, ricontrollati manualmente e corretti dalla seconda autrice. La trascrizione ortografica ha riguardato non solo il parlato, ma anche tutti i “rumori umani” (es. risate, tosse) e le pause piene. Al momento la trascrizione è stata completata per tutti i dati di parlato spontaneo provenienti dalle interazioni dialogiche (10 ore circa, per un totale di 100.966 tokens).

I file trascritti ortograficamente, accompagnati dai rispettivi file audio, sono stati processati su WebMAUS (cfr. Kisler, Reichel & Schiel, 2017), sfruttando lo strumento dell'allineamento forzato. Dopo una serie di passaggi tramite diversi servizi offerti dalla piattaforma, il risultato ottenuto è stato, per ogni file audio, un file Text. Grid contenente tre *tier*: un *tier* ORT con la trascrizione ortografica parola per parola, un *tier* KAN, indicante la trascrizione fonologica per parola, un *tier* MAU, che riporta la trascrizione fonico per fonico. Anche le segmentazioni e le annotazioni sono state completamente controllate e corrette manualmente dalla seconda autrice. Oltre a questi tre *tier*, al fine di poter svolgere l'analisi sulle vocali, è stato aggiunto manualmente un ulteriore *tier* (STR), contenente l'indicazione delle vocali toniche e atone.

Per quanto riguarda la correzione apportata alla segmentazione e all'annotazione prodotta in maniera automatica mediante WebMAUS, l'intervento manuale è avvenuto in modo uniforme per tutti i dati, tenendo conto di alcuni criteri. Per la correzione delle segmentazioni, tutte le vocali sono state segmentate nel punto di ampiezza zero, facendo corrispondere l'*onset* al punto di inizio del primo ciclo completo e l'*offset* al punto coincidente con l'ultimo ciclo completo incluso. Quanto all'annotazione, le vocali sono state etichettate seguendo un'annotazione fonologica e quindi facendo riferimento a quanto previsto dallo standard italiano, che ha funzionato come punto di riferimento per lo studio della distribuzione delle vocali. Le occorrenze segmentate ammontano a 7784 (2027 toniche, 5757 atone); tuttavia, l'analisi esplorativa dei dati di cui daremo conto prende in esame solamente le vocali toniche.

I dati trascritti ortograficamente e segmentati e annotati fonologicamente sono stati trasformati in un database interrogabile, mediante il sistema EMU (Winkelmann, Harrington & Jansch, 2017), che, oltre a offrire un'interfaccia online, consente di condurre l'analisi sul corpus attraverso il linguaggio R (pacchetto EmuR).

5. *Analisi esplorativa delle vocali: problemi*

Questa sezione dell'articolo è strutturata in questo modo: prima vengono enunciati i principali aspetti problematici che si presentano quando si debba condurre un'analisi acustica di vocali su un corpus di parlato spontaneo caratterizzato da alta variazione sociale; in seguito, sono avanzate alcune delle possibili soluzioni a questi problemi sotto la forma di un'analisi statistica esplorativa di un sottoinsieme di vocali estratte dal corpus DIA.

5.1 Problemi nell'analisi acustica di corpora

L'analisi acustica di vocali (ma non soltanto) condotta su dati elicitati con procedure di tipo naturalistico e poi raccolti in un corpus pone problemi specifici che sono riconducibili solo in parte alle procedure consuete di analisi di dati acquisiti in contesti con maggiore controllo sperimentale. Volendo riassumere le caratteristiche in pochi tratti, possiamo dire che nei corpora di parlato spontaneo, rispetto a quello di laboratorio, ci si trova di fronte a dati caratterizzati da un forte sbilanciamento nella frequenza relativa dei singoli foni e a una maggiore variazione contestuale, prosodica e individuale.

Nella nostra esperienza, tre ambiti metodologici hanno richiesto una particolare riflessione critica:

- a. il grado di affidabilità della misurazione automatica dei parametri acustici;
- b. l'alta variabilità del dato spontaneo;
- c. la complessità della modellizzazione statistica.

In questo contributo, nessuno di questi tre aspetti può essere trattato se non cursoriamente; la finalità è soltanto quella di sollevare il problema, osservarne alcune delle possibilità di soluzione e tentare una prima proposta di analisi.

Il primo problema riguarda le modalità di parametrizzazione acustica delle vocali quando ci si trovi ad analizzare ampi corpora di dati. La soluzione di default, nella ricerca fonetica acustica, è quella di basarsi su algoritmi di estrazione automatica delle formanti (*formant tracking*). Questi algoritmi sono basati in varia misura sulla tecnica di predizione lineare (*Linear Predictive Coding*). In quest'ottica è necessario scegliere l'algoritmo ottimale, ovvero quello più robusto, o per lo meno, quello sul quale i risultati della ricerca in ambito computazionale sembrano convergere. Due tra le alternative più comuni tra i *formant trackers* sono il metodo Burg (Andersen, 1974) e l'algoritmo Split Levinson (Delsarte & Genin, 1986). Il primo è quello implementato di norma anche in Praat, mentre il secondo è utilizzato, per esempio, all'interno del pacchetto *wrassp* nel sistema di database emuR.

Tuttavia, quando si dispone di grandi basi di dati, che presentano grande variabilità in termini di numero di parlanti e di grado elevato di spontaneità del parlato, la strategia di appoggiarsi ad algoritmi di *formant tracking* non sembra essere la più adatta, a meno di essere disposti ad affrontare una fase di controllo manuale intensivo. Il maggiore punto di debolezza del *formant tracking* è la sua rigidità di fronte alla variabilità delle caratteristiche del parlante e del parlato. È necessario infatti operare un continuo adattamento dei parametri non soltanto ai parlanti (in ragione di età e genere), ma anche in relazione a specifiche vocali e/o a interazioni tra vocali e caratteristiche dei parlanti. Anche in questo senso si possono evidentemente ottimizzare algoritmicamente (con opportune procedure iterative) il numero dei coefficienti dei polinomi di predizione lineare così come il *formant ceiling*⁸, tuttavia questa ope-

⁸ Il valore massimo al di sotto del quale cercare di identificare le formanti (Maximum formant in Praat); questo valore è legato alla lunghezza del tratto vocale e perciò può variare da individuo a individuo, ma in media per voci maschili è di 5000 Hz, mentre per voci femminili è di 5500 Hz.

razione risulta onerosa al crescere dei parlanti e in genere della variabilità del parlato (soltanto a titolo di esempio si possono segnalare le osservazioni metodologiche contenute in Escudero, Boersma, Rauber & Bion, 2009).

L'alta proporzione di errore operata dal *formant tracking* automatico eseguito su corpora si manifesta principalmente nella presenza consistente di *outlier*, ovvero misurazioni di F1 e/o F2 non solo distanti dalle tendenze centrali, ma anche chiaramente poco realistici rispetto alla natura del dato. In un contesto sperimentale, è consuetudine la revisione manuale dei valori formantici eseguiti automaticamente poiché la numerosità del dataset lo consente, mentre quando si opera in un contesto di *corpus phonetics* è necessario riflettere su possibili procedure automatiche alternative all'estrazione di formanti, che svincolino il ricercatore dall'attività di correzione di *outliers* (Harrington, 2010: 215).

La seconda caratteristica di questo tipo di analisi è connessa con la variabilità nella realizzazione delle diverse categorie fonetiche. Un corpus di parlato spontaneo di tipo dialogico comporta infatti un elevato grado di variabilità acustica che pone seri problemi all'individuazione di categorie definite e separabili. Volendo schematizzare, le possibili fonti di variazione presenti nel dato spontaneo (segmentale) si possono articolare secondo queste macro-categorie (Harrington 2010):

- i contesti fonetici e fonologici, e più in generale la variabilità posizionale e strutturale;
- le posizioni prosodiche a cui possono corrispondere le diverse funzioni sintattico-pragmatiche;
- gli stili di elocuzione, che a loro volta determinano i gradi di connessione del parlato e i fenomeni di riduzione;
- le caratteristiche sociali degli individui, per esempio in questo caso specifico la distinzione tra parlanti nativi e non nativi;
- le differenze di tipo biologico tra i parlanti, in primo luogo genere ed età.

Non c'è evidentemente qui spazio per riassumere neppure sinteticamente le diverse fonti di variabilità presenti nel dato acustico (per una schematizzazione v. Vietti, in stampa).

Il nocciolo della questione può essere riassunto così: se in un contesto sperimentale le fonti di variabilità possono essere controllate e manipolate in modo selettivo, nel parlato elicitato in contesti naturalistici queste si presentano tutte congiuntamente. Il compito è perciò quello di trovare il modo di organizzare questa variazione attraverso la scomposizione dello spazio di variazione in dimensioni, con lo scopo di meglio definire gli oggetti-categorie. Più in generale, l'opportunità che lo studio di corpora offre all'analisi acustica è quella di poter osservare il modo in cui la variazione contribuisce alla strutturazione di categorie fonetiche e fonologiche.

La necessità di rendere conto dell'alta variazione dei dati di corpora si traduce sul piano metodologico nella ricerca di modelli statistici adeguati. Un caso particolare in questa direzione è rappresentato dall'introduzione (e impiego sempre più esteso) di modelli lineari a effetti misti, che ha comportato un significativo avanzamento delle nostre possibilità di controllare, quasi in una logica sperimentale, quegli effetti

casuali legati alla ripetizione delle misurazioni su medesimi soggetti e alle possibili dipendenze interne tra osservazioni. In questo modo è ora possibile analizzare le relazioni tra variabili che presentano ripetizioni e in generale una struttura gerarchica o multilivello, tenendo sotto controllo alcune fonti di variabilità che altrimenti eserciterebbero un'indebita influenza sul modello finale.

Oltre a questa specifica innovazione, nel corso degli ultimi decenni le possibilità di analisi quantitativa delle vocali si è espansa enormemente, di pari passo con la necessità di formulare ipotesi via via più complesse sulla natura dinamica delle vocali e sull'interdipendenza di più variabili. Volendo di nuovo un po' semplificare il problema in dicotomie, potremmo distinguere tra (a) analisi statiche e dinamiche, (b) su singoli parametri o su aggregati di parametri, (c) su una singola vocale o su interi sistemi vocalici (a titolo puramente esemplificativo si v. Adank et al., 2004; Calamai, 2004; Fox & Jacewicz, 2017; Gahl & Baayen, 2019).

Considerata la natura esplorativa di questo contributo, non vi è spazio qui per una discussione approfondita delle possibilità di analisi e modellizzazione delle vocali (e dei sistemi vocalici), aspetti che evidentemente non sono soltanto di metodo, ma sono implicati dalle prospettive teoriche che si adottano. A fronte di queste limitazioni di intenti, riteniamo sia bene avere presente il contesto delle possibilità e dei problemi metodologici specifici che l'analisi del parlato spontaneo di un corpus pone, prima di accingersi all'analisi.

È utile anche ricordare il fine ultimo a cui si tenderà in questa e in future indagini, ovvero quello di individuare quale analisi statistica ci permetta di valutare le differenze tra sistemi vocalici in varietà di lingua diverse, o meglio ancora in una logica induttiva, quale metodo ci consenta di individuare delle varietà di lingua a partire dalla distribuzione dei valori acustici dei fonemi in un sistema vocalico in relazione a fattori esterni relativi ai parlanti e alle loro caratteristiche.

5.2 Analisi spettrale

Come è stato già anticipato al par. 4.1, l'analisi che presentiamo qui si basa su un primo sottoinsieme del corpus DIA. Il dataset analizzato contiene 2027 occorrenze vocaliche di 8 parlanti donne, 4 parlanti native dell'italiano e 4 non native con tedesco come prima lingua (in realtà si tratta del tirolese, un dialetto tedesco). Le occorrenze estratte dal corpus riguardano le 7 vocali toniche dell'italiano, ovvero quelle vocali che compaiono in sillaba accentata di parola.

Per ognuno dei tre ambiti metodologici delineati sopra abbiamo sviluppato una possibile soluzione:

- a. per la misurazione automatica dei parametri mostriamo un confronto qualitativo tra analisi per formanti e analisi spettrale basata su un insieme di coefficienti coseno discreti (*DCT coefficients*), poi ulteriormente ridotti attraverso analisi di componenti principali;
- b. per la variabilità del dato abbiamo tenuto in considerazione la variazione contestuale e individuale, trattando questi fattori come effetti casuali in un modello lineare a effetti misti;

- c. per la modellizzazione statistica abbiamo invece provato due modelli statistici relativamente elementari basati su un'osservazione statica nel tempo delle vocali (estrazione dei parametri spettrali in un punto medio) con i singoli parametri acustici come variabili dipendenti e le vocali e i fattori extralinguistici come variabili esplicative.

Il dataset analizzato ($n = 2027$) è il risultato di una prima selezione operata eliminando gli *outlier* di durata ($n = 86$), ovvero escludendo i tokens vocalici con valori inferiori a 20 ms e superiori a 1.5 volte la distanza interquartile (in questo caso > 234 ms).

Il nostro punto di partenza, in modo da avere un riferimento nell'interpretazione dei dati, è dato dall'osservazione dei valori delle prime due formanti, estratti secondo l'algoritmo di *formant tracking* di Split-Levinson per come implementato dalla funzione *forest* in emuR ($n = 1726$). Nelle Figure 1 e 2 sono rappresentate le due possibili varietà di italiano.

Figura 1 - F1 e F2 di parlanti native (ellissi di confidenza 95%)

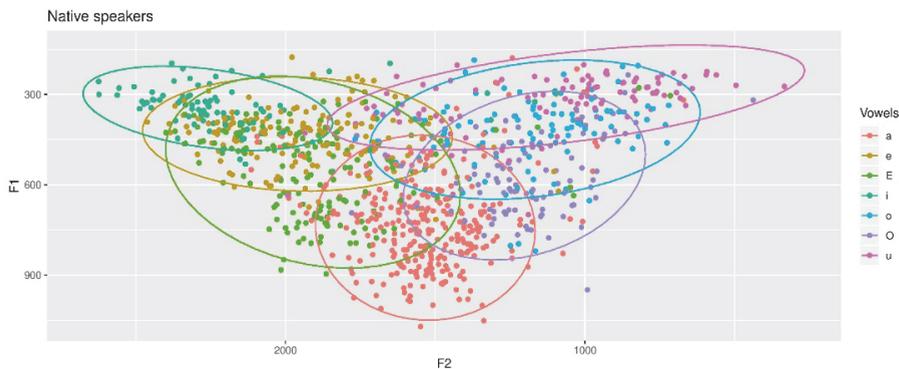
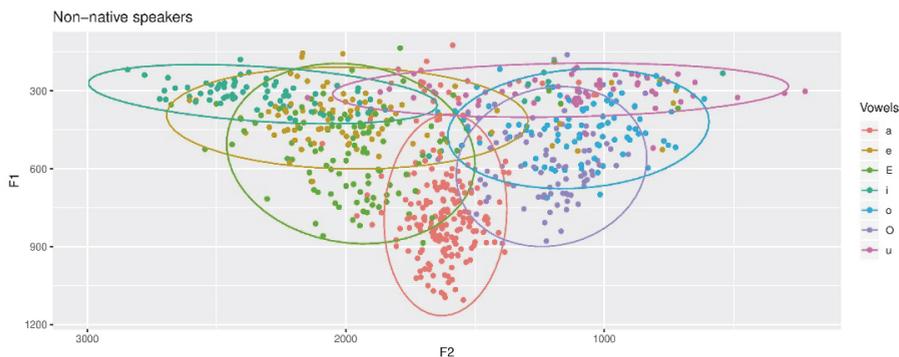


Figura 2 - F1 e F2 di parlanti tedescofone non native (ellissi di confidenza 95%)



Dai grafici si può innanzitutto osservare come, in entrambe le varietà, vi sia una notevole sovrapposizione tra le categorie, caratteristica tipica del parlato spontaneo.

Sebbene raramente univocamente individuabili, degli addensamenti attorno ai valori di tendenza centrale (moda e media) sono ben riconoscibili. Passando a una superficiale osservazione delle differenze tra le due varietà, possiamo notare in primo luogo la differenza nel *range* tra i valori massimi e minimi di F1 e F2 che porta a uno spazio vocalico complessivo più ampio nel caso della varietà di parlanti non nativi. In secondo luogo, si può constatare come la natura della sovrapposizione tra categorie muti in termini di forma e varianza delle ellissi tra le due varietà (si osservi p.e. il caso della /a/ nelle due varietà).

Nell'estrazione dei valori con *formant tracking* emerge un alto numero di occorrenze che presentano un errore nella stima della formante. L'errore più comune è l'assegnazione di un valore zero a F2 (n = 301 tokens), in particolare nel caso di /i/.

L'analisi alternativa che abbiamo deciso di sperimentare è invece un'analisi spettrale che non preveda individuazione e misurazione delle formanti, ma che consideri piuttosto l'informazione spettrale nel suo complesso. L'analisi spettrale ha come obiettivo l'estrazione dei coefficienti della trasformata discreta del coseno (*Discrete Cosine Transform*) nel punto medio della durata del segmento (Zahorian & Jagharghi, 1993). Tutte le operazioni sono state condotte in R utilizzando per lo più le funzioni di analisi del segnale disponibili nel pacchetto EmuR (Winkelmann, n.d.).

La procedura per calcolare i coefficienti è stata la seguente (Harrington 2010: 206-215):

- a. calcolo degli spettri di potenza con trasformata di Fourier veloce (FFT);
- b. trasformazione degli spettri nella scala Bark;
- c. calcolo della seconda trasformata di Fourier eseguita sugli spettri di potenza (*Discrete Cosine Transform*);
- d. estrazione da (c) dei coefficienti DCT che descrivono in modo sintetico le frequenze di risonanza.

L'esito di questo processo è analogo a quello che si applica per ottenere i coefficienti cepstrali trasformati in Mel (*Mel-Frequency Cepstral Coefficients*, Wakita, 1976); la principale differenza risiede nel fatto che la seconda analisi di Fourier viene eseguita usando unicamente funzioni coseno come basi invece che anche funzioni seno. In sintesi, il vantaggio di tale scelta risiede nel fatto che servono meno basi per ricostruire il segnale e, soprattutto, l'analisi può essere condotta utilizzando soltanto numeri reali (e non anche numeri complessi come per la trasformata di Fourier con funzioni seno).

La scomposizione cepstrale del segnale dovrebbe produrre coefficienti indipendenti, tuttavia in Iskarous (2019) si è registrata una certa ridondanza nella codifica delle informazioni relative alle risonanze per i tratti articolatori di altezza e avanzamento delle vocali. Per questo motivo, si è deciso di esaminare l'eventuale presenza di dipendenze tra i coefficienti DCT e la loro riduzione a un numero inferiore di dimensioni attraverso l'analisi delle componenti principali (PCA). L'analisi è stata condotta usando i primi 6 coefficienti DCT, in quanto da un lato sono apparsi sufficienti per ricostruire le caratteristiche salienti dello spettro originale e dall'altro in questo modo la proporzione di varianza cumulativa delle prime tre componenti ri-

sultava ottimale. Le prime due dimensioni (Dim1 = 33.2%, Dim2 = 29.4%) rappresentano infatti il 62.6% della percentuale cumulativa di varianza, mentre sommando le prime tre componenti si raggiunge l'80% (Dim3 = 17.5%). Prima di condurre l'analisi i dati sono stati ricampionati a una frequenza di 16.000 Herz ritenuta più adeguata all'analisi delle vocali.

L'esito di questa analisi è riportato graficamente nelle due figure 3 e 4.

Il risultato che si ottiene dalla PCA è una ridefinizione dei valori dei coefficienti in uno spazio geometrico di dimensioni ridotte, basato sulla struttura delle correlazioni tra le variabili di partenza (in questo caso i sei coefficienti). Si tratta perciò di uno spazio astratto, le cui unità di misura non sono più unità fisiche, ma correlazioni tra variabili. Ciò nondimeno le relazioni tra gli oggetti dovrebbero far emergere un'organizzazione basata sui parametri articolatori.

Da un esame visivo dei grafici risulta infatti che, sebbene ruotata di 100 gradi in senso antiorario rispetto alla consueta rappresentazione, la disposizione dei punti e delle categorie vocaliche appare sensata, ovvero regolata dai parametri articolatori di innalzamento/abbassamento e avanzamento/arretramento. In particolare, la Dimensione 1 (quella che cattura la maggior parte della varianza del nostro dataset) approssima l'altezza del corpo della lingua, mentre la Dimensione 2 coglie in parte la distinzione tra vocali anteriori e posteriori.

Figura 3 - PCA dei DCT coefficients parlanti nativi

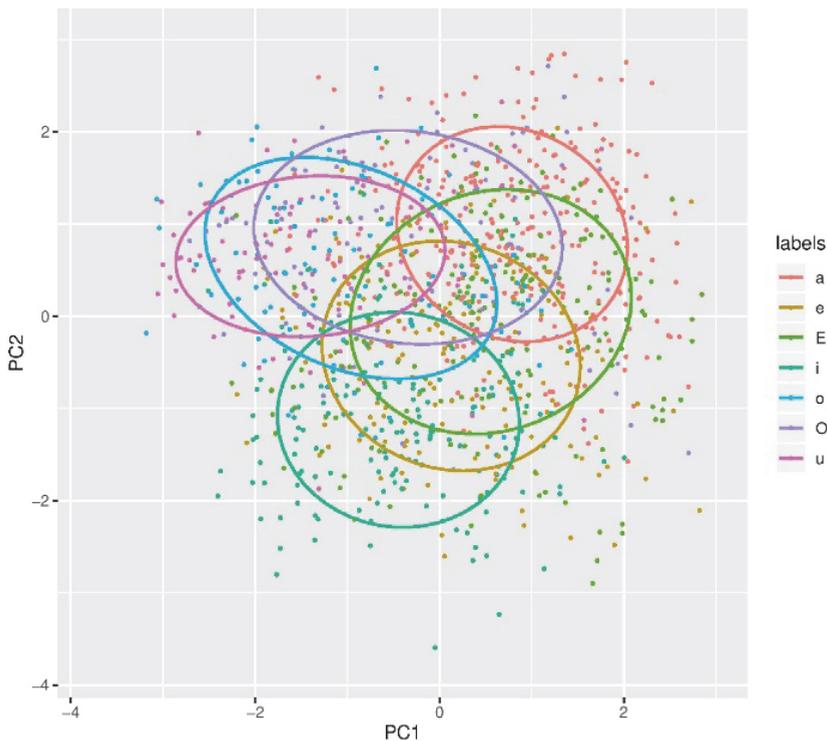
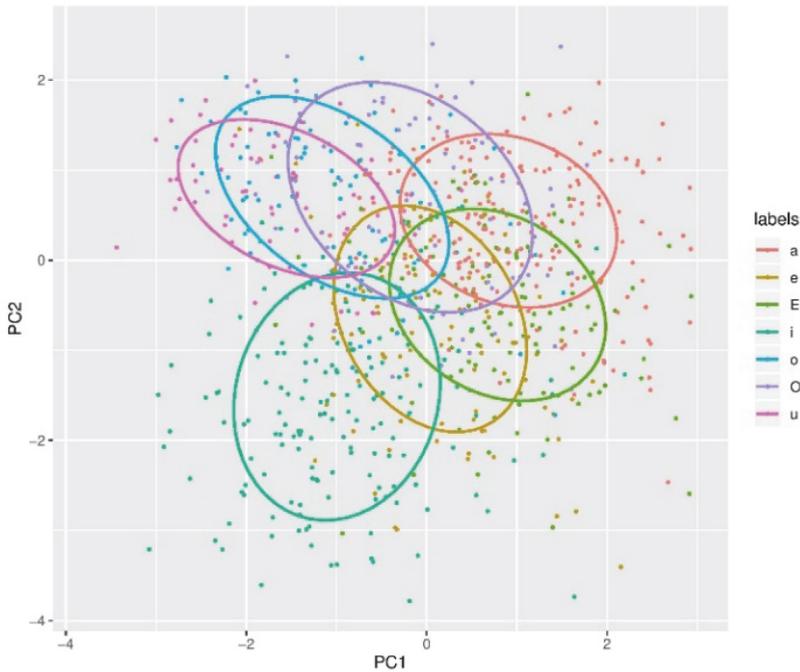


Figura 4 - PCA dei DCT coefficients parlanti tedesco non nativi



In secondo luogo, si può osservare come anche in questo caso l'*overlap* fra le categorie appaia notevole e comunque maggiore rispetto alla rappresentazione per formanti. Tuttavia, sul piano qualitativo possiamo registrare come la tendenza a una migliore discriminazione tra le vocali nella varietà di non nativi sia confermata anche dall'analisi PCA. Come abbiamo già rilevato più in generale, l'alta sovrapposibilità delle categorie non è imputabile né alla precisione della misurazione, né al trattamento quantitativo dei dati; piuttosto è da ricercare nelle condizioni di produzione del parlato che, all'aumentare della spontaneità nella comunicazione, vede diminuire, nel complesso, la definizione delle caratteristiche acustiche. Al piano della distribuzione statistica si può invece senz'altro attribuire la rappresentazione geometrica delle ellissi di dispersione, in quanto come sappiamo, forme e dimensioni (a) sono fortemente influenzate dalla presenza anche di pochi *outlier* e (b) poggiano su un assunto di gaussianità nella distribuzione bivariata tendenzialmente violato (Savy, Clemente, Lo Prejato, 2005).

5.3 Analisi statistica

L'ultima fase dell'analisi prevede l'applicazione di modelli lineari a effetti misti con lo scopo di testare in via preliminare la nostra ipotesi sulla presenza di una differenza tra i due sistemi vocalici di nativi e non nativi, tenendo sotto controllo alcune delle maggiori fonti di variabilità. Con questo obiettivo, abbiamo condotto due analisi parallele utilizzando le due prime dimensioni dell'analisi delle componenti princi-

pali come variabili dipendenti, mentre le categorie vocaliche, il retroterra linguistico dei parlanti e l'interazione tra queste due variabili sono state considerate come variabili esplicative ed effetti fissi nel modello. Sono state poi inserite tre variabili come effetti casuali: il contesto segmentale sinistro, il contesto segmentale destro⁹ e il parlante. Con le prime due variabili si cerca naturalmente di catturare l'aspetto più macroscopico della variazione contestuale connessa con la coarticolazione, e non certo controllare in modo accurato la variazione posizionale e strutturale.

Per il modello relativo alla Dimensione 1 (innalzamento/abbassamento) la differenza tra le diverse vocali è significativa ($F[6, 223.5]$, $p < 0.001$), mentre il retroterra linguistico (italofono vs. tedescofono) non è significativo in quanto effetto principale ($F[1, 0.01]$, $p = 0.95$). Tuttavia, come era prevedibile, c'è una forte interdipendenza tra le varietà di lingua e le vocali ($F[6, 7.44]$, $p < 0.001$). Questo implica che il background linguistico dei parlanti nativi e non nativi non determina due sistemi vocalici significativamente diversi per grado di apertura vocalica, ma sembra piuttosto che vi siano delle interazioni tra vocali all'interno delle due varietà di lingua.

Per la Dimensione 2 i risultati sono analoghi: le vocali sono significativamente distinte per avanzamento/arretramento ($F[6, 238.6]$, $p < 0.001$), il background linguistico dei parlanti non è significativo ($F[1, 1.2]$, $p = 0.92$), mentre vi sono significative interazioni tra le due variabili ($F[6, 7.6]$, $p < 0.001$).

Osservando la componente *random* dei due modelli, è interessante notare come i contesti segmentali sinistro e destro siano caratterizzati da una bassa varianza (attorno a 0.03), mentre la variazione intersoggettiva sia in genere più alta e maggiore nel caso dell'altezza vocalica (Dim. 1, $s^2 = 0.19$) rispetto all'avanzamento (Dim. 2, $s^2 = 0.08$). Questo significa che ci sono maggiori differenze tra parlanti in relazione al parametro articolatorio dell'innalzamento/abbassamento della lingua.

Per comprendere meglio la struttura delle interazioni tra vocali e varietà di lingua è necessario condurre un'analisi post-hoc su tutte le coppie possibili di valori. In sintesi, quest'analisi conferma l'assenza di un effetto principale del background linguistico e mette invece in evidenza come le due varietà siano organizzate in modo diverso al proprio interno lungo le due dimensioni articolatorie. Rispetto al primo aspetto, tutti i test che confrontano la stessa vocale nelle due varietà sono risultati non significativi (p.e. /a/ di nativi e /a/ di non nativi non sono statisticamente diverse). Considerata l'interazione tra vocali e varietà, ci possiamo aspettare che le vocali dei nativi e dei non nativi siano distinte in modo diverso per grado di altezza e avanzamento.

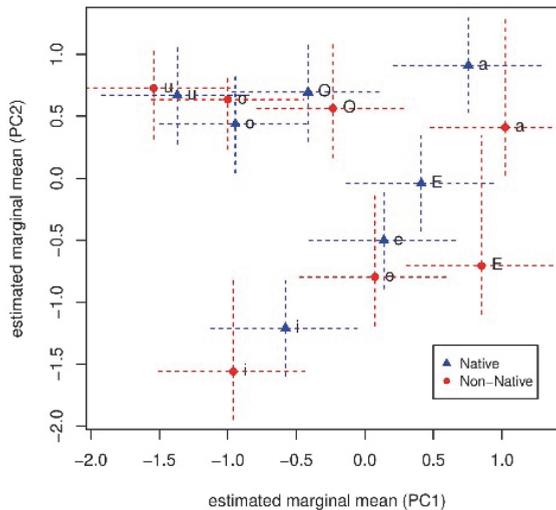
⁹ È stato preso in considerazione il fono precedente e quello seguente alla vocale.

Questo è infatti quanto emerge dall'analisi ed è riassunto nella Tabella 1.

Tabella 1 - *Vocali non distinte statisticamente per Dimensione e Varietà di lingua*

		Suoni non distinti statisticamente
Dimensione 1 (altezza)	Nativi	(u, o), (i, o), (i, ə), (e, ε)
	Non nativi	(i, o), (e, ə), (ε, a)
Dimensione 2 (avanzamento)	Nativi	(u, ə, a)
	Non nativi	(u, o, ə, a), (e, ε)

Figura 5 - *Grafico delle medie stimate dai modelli con relativi intervalli di confidenza*



Per cogliere in modo più chiaro le caratteristiche dei due sistemi vocalici ipotizzati, i valori medi dei diversi parametri così come stimati dai due modelli lineari, insieme ai relativi intervalli di confidenza, sono espressi in forma grafica nella Figura 5.

Il grafico conferma visivamente l'analisi statistica condotta: da un lato, se guardiamo alle posizioni delle diverse categorie (p.e. /o/ nativa vs. /o/ non nativa) appare evidente che non esistono due distinti sistemi vocalici; dall'altro lato, se si osservano le relazioni tra fonemi all'interno dei due sistemi, le configurazioni relative che ne risultano sono ben distinte. Appaiono qui di particolare interesse (a) i diversi rapporti tra le vocali medie nelle due varietà, anche alla luce degli studi condotti da Mioni, e (b) le (possibili) diverse specificazioni articolatorie di /i/ e /a/.

6. Osservazioni conclusive

Dall'analisi statistica emergono infine preziosi elementi di riflessione, sia di metodo che di contenuto, che verranno presi in considerazione in analisi successive. Tra questi vorremmo qui soltanto menzionarne rapidamente tre:

- a. da una rapida osservazione, la distribuzione delle due variabili continue appare decisamente non gaussiana, problema questo che pone in primo luogo seri dubbi sulla significatività dei risultati e in secondo luogo sulla reale struttura interna delle categorie;
- b. considerato il numero di occorrenze non elevato, anche la struttura elementare del modello può correre il rischio di eccedere nella specificazione causando problemi di *overfitting*, convergenza per singolarità e più in generale perdere in capacità predittiva;
- c. lo scarso ruolo della variabilità contestuale, intesa qui come effetto random, è inatteso, perciò andrebbero esaminati con maggiore attenzione i possibili effetti locali di singoli fonemi precedenti o seguenti in interazione con specifiche vocali;
- d. la variabile legata al retroterra dei parlanti potrebbe cogliere in modo molto approssimativo le differenze nella competenza dei parlanti (in futuro, considerato il contesto di bilinguismo sociale, sarebbe più opportuno utilizzare altre misure di dominanza nell'uso linguistico, come le reti sociali).

Riferimenti bibliografici

ADANK, P., VAN HOUT, R. & SMITS, R. (2004). An acoustic description of the vowels of Northern and Southern Standard Dutch. In *The Journal of the Acoustical Society of America*, 116(3), 1729–1738.

ALBANO LEONI, F. (2006). Il corpus CLIPS: presentazione del progetto. Documento scaricabile al sito <http://www.clips.unina.it> (22.05.2020).

ALBANO LEONI, F. (2007). Un frammento di storia recente della ricerca (linguistica) italiana. Il corpus CLIPS. In *Bollettino d'Italianistica*, 4, 122-130.

ALBANO LEONI, F., CUTUGNO, F. & SAVY, R. (1995). The Vowel System of Italian Connected Speech. In ELENIUS, K. & BRANDERUD, P. (Eds.), *Proceedings of XIIIth International Congress of Phonetic Sciences*, Stockholm 13-19, VIII, 1995, Stockholm, IV, 396-399.

ALBANO LEONI, F., CUTUGNO, F. & SAVY, R. (1998). Il vocalismo dell'italiano televisivo. Analisi acustica di un corpus. In RUFFINO, G. (Ed.), *Atti del XXI Congresso Internazionale di Linguistica e Filologia Romanza*, Centro di studi filologici e linguistici siciliani – Università di Palermo 18-24.IX.1995. Tübingen: Niemeyer, IV: 3-16.

ANDERSEN, N. (1974). On the calculation of filter coefficients for maximum entropy spectral analysis. In *Geophysics*, 39(1), 69–72.

BERTINETTO, P.M. (Ed.) (2001). AVIP – *Archivio di Varietà di Italiano Parlato*. 4 cd-rom. Ufficio Pubblicazioni della Classe di Lettere della Scuola Normale Superiore: Pisa.

- CALAMAI, S. (2003a). Vocali d'Italia. Una prima rassegna. In COSÌ, P., MAGNO CALDOGNETTO, E. & ZAMBONI, A. (Eds.), *Voce canto parlato. Scritti in onore di Franco Ferrero*. Padova: Unipress.
- CALAMAI, S. (2003b). The Pisan Vowel System of Read and Semi-spontaneous Speech. An Exploratory Contribution. In *Quaderni del Laboratorio di Linguistica della Scuola Normale Superiore*, 3, 72-98.
- CALAMAI, S. (2004). *Il vocalismo tonico pisano e livornese. Aspetti storici, percettivi e acustici*. Alessandria: Edizioni dell'Orso.
- CALAMAI, S., SORIANELLO, P. (2003). Aspetti stilistici del vocalismo romano. In *Quaderni del Laboratorio di Linguistica della Scuola Normale Superiore di Pisa*, 4, 27-41.
- CANEPARI, L. (1999). *Manuale di pronuncia italiana*. Bologna: Zanichelli.
- CANGEMI, F., FRÜNDT, J., HANEKAMP, H., GRICE, M. (2019). A semi-automatic workflow for orthographic transcription and syllabic segmentation. In PICCARDI, D., ARDOLINO, F. & CALAMAI, S. (Eds.), *Gli archivi sonori al crocevia tra scienze fonetiche, informatica umanistica e patrimonio digitale. Audio archives at the crossroads of speech sciences, digital humanities and digital heritage*. Milano: Officinaventuno, 419-425.
- CROCCO, C., SAVY, R., CUTUGNO, F. (Eds.) (2002). *API – Archivio del Parlato Italiano*. Dvd. Napoli: Università degli Studi di Napoli "Federico II".
- DELSARTE, P., GENIN, Y. (1986). The split Levinson algorithm. In *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 34(3), 470-478.
- ESCUDEIRO, P., BOERSMA, P., RAUBER, A.S., & BION, R.A.H. (2009). A cross-dialect acoustic description of vowels: Brazilian and European Portuguese. In *The Journal of the Acoustical Society of America*, 126(3), 1379-1393.
- FERRERO, F.E. (1968). Diagrammi di esistenza delle vocali italiane. In *Alta Frequenza*, 37, 54-58.
- FERRERO, F.E. (1972). Caratteristiche acustiche dei fonemi vocalici italiani. In *Parole e metodi*, 3, 9-31.
- FOX, R.A., JACEWICZ, E. (2017). Reconceptualizing the vowel space in analyzing regional dialect variation and sound change in American English. In *The Journal of the Acoustical Society of America*, 142(1), 444-459.
- GAHL, S., BAAYEN, R.H. (2019). Twenty-eight years of vowels: Tracking phonetic variation through young to middle age adulthood. In *Journal of Phonetics*, 74, 42-54.
- GALATÀ, V., SPREAFICO, L., VIETTI, A. & KALAND, C. (2016). An acoustic analysis of /r/ in Tyrolean. In *INTERSPEECH 2016*, 1002-1006.
- HARRINGTON, J. (2010). *Phonetic analysis of speech corpora*. Malden, MA: Wiley-Blackwell.
- KERSWILL, P., WILLIAMS, A. (2000). Creating a new town koine: children and language change in Milton Keynes. In *Language in Society*, 29, 65-115.
- KISLER, T., REICHEL, U. & SCHIEL, F. (2017). Multilingual processing of speech via web services. In *Computer Speech & Language*, 45, 326-347.
- MELUZZI, C. (2020). *Sociofonetica di una varietà di koinè. Le affricate dentali nell'italiano di Bolzano*. Milano: FrancoAngeli.

- MENDICINO A., ROMITO L., (1991). 'Isocronia' e 'base di articolazione': uno studio su alcune varietà meridionali. In *Quaderni del Dipartimento di Linguistica dell'Università della Calabria, Serie Linguistica*, 3, 49-67.
- MEREU, D., VIETTI, A. (2021). Dialogic ItAlian: the creation of a corpus of Italian spontaneous speech. In *Speech Communication*, 130, 1-14.
- MIONI, A.M. (1990a). La standardizzazione fonetico-fonologica a Padova e Bolzano (stile di lettura). In CORTELAZZO, M., MIONI, A.M. (Eds.), *L'italiano regionale. Atti del XVIII Congresso della SLI*. Roma: Bulzoni, 193-208.
- MIONI, A.M. (1990b). Bilinguismo intra- e intercomunitario in Alto Adige/Südtirol: considerazioni sociolinguistiche. In LANTHALER, F. (Ed.), *Mebr als eine Sprache. Zu einer Sprachstrategie in Südtirol – Più di una lingua. Per un progetto linguistico in Alto Adige*. Merano: Alpha&Beta, 13-36.
- MIONI, A.M. (1994). Vivere senza dialetto?. In LANTHALER, F. (Ed.) *Dialekt und Mehrsprachigkeit / Dialetto e plurilinguismo*. Merano/Meran: Alpha&Beta, 27-47.
- MIONI, A.M. (2001a). L'italiano nelle tre comunità linguistiche tirolesi (con particolare riguardo per la pronuncia). In EGGER, K., LANTHALER, F. (Eds.), *Die deutsche Sprache in Südtirol. Einheitsprache und regionale Vielfalt*. Wien: Folio, 65-76.
- MIONI, A.M. (2001b). *Elementi di fonetica*. Padova: Unipress.
- NOCCHI N., CALAMAI S. (2009). Durata e strutture formantiche nel parlato toscano: Indagini preliminari su un campione di dialoghi semispontanei. In SCHMID, S., SCHWARZENBACH, M. & STUDER, D. (Eds.), *Atti del Convegno AISV 2009 La dimensione temporale del parlato, Zurigo 4-6.II.2009*. Torriana: EDK, 195-223.
- ROMITO L., LORENZI M. (1997). Considerazioni generali sul comportamento di alcune varietà dialettali meridionali e settentrionali rispetto all'accento intensivo. In *Quaderni del Dipartimento di Linguistica dell'Università della Calabria. Serie Linguistica*, 6, 11-34.
- SAVY, R., CLEMENTE, G. & LO PREJATO, M. (2005). Per una caratterizzazione e una misura della riduzione vocalica in italiano. In COSI, P. (Ed.), *Misura dei parametri – aspetti tecnologici e implicazioni nei modelli linguistici, Atti del 1° Convegno Nazionale dell'Associazione Italiana di Scienze della Voce, Padova, 2-4 dicembre 2004*. Torriana (RN): EDK Editore, 135-160.
- SCHMID, S. (1999). *Fonetica e fonologia dell'italiano*. Torino: Paravia Scriptorium.
- TRUMPER, L., ROMITO, L. (Eds.) (1993). *Teoria e sperimentazione: parametri, tratti e segmento. Atti delle II Giornate di Studio del GFS, Calabria*, 28-29.XI.1991. Roma: Esagrafica.
- VIETTI, A. (2011). Comunità alloglotte: tedesca. In SIMONE, R. (Ed.), *Enciclopedia dell'italiano*. Roma: Istituto Enciclopedia Italiana Treccani, 482-485.
- VIETTI, A. (2017). Italian in Bozen/Bolzano: the formation of a "new dialect". In: CERRUTI, M., CROCCO, C. & MARZO, S. (Eds.), *Towards a new standard: theoretical and empirical studies on the restandardization of Italian*. Berlin: De Gruyter, 176-212.
- VIETTI, A. (2019). Phonological variation and change in Italian. In *Oxford Research Encyclopedia of Linguistics*. Oxford: Oxford University Press.
- VIETTI, A. (in stampa). Il ruolo della variabilità acustica nella costruzione del dato linguistico. In BERNINI, G. (Ed.), *Superare l'evanescenza del parlato*.

- WAKITA, H. (1976). Instrumentation for the Study of Speech Acoustics. In LASS, N.J. (Ed.), *Contemporary Issues in Experimental Phonetics*. Academic Press, 3-40.
- WINKELMANN, R. (n.d.). The EMU-SDMS Manual. Retrieved May 24, 2020, from <https://ips-lmu.github.io/The-EMU-SDMS-Manual/>
- WINKELMANN, R., HARRINGTON, J. & JÄNSCH, K. (2017). "EMU-SDMS: Advanced speech database management and analysis in R". In *Computer Speech & Language*, 45, 392-410.
- ZAHORIAN, S.A., JAGHARGHI, A.J. (1993). Spectral-shape features versus formants as acoustic correlates for vowels. In *The Journal of the Acoustical Society of America*, 94(4), 1966-1982.

ANDREA TARASI, LUCIANO ROMITO

La consonante occlusiva labiovelare sorda nei dialetti calabresi: uno studio pilota

The voiceless labiovelar occlusive consonant in Calabrian dialects: a pilot study

The objective of this research is to demonstrate with auditory and acoustic analysis the presence or absence of voiceless labiovelar stop consonant [k^w] in the dialect of Mormanno and San Pietro in Guarano through the phonetic realizations of demonstratives “quello, quelli, quella e quelle”. This pilot study is part of a larger project launched by Phonetics Laboratory which aims to define new boundaries between the Calabrian dialects and, consequently, to remodel those already existing and described in some researches. The analyzes conducted on the corpus of the research demonstrate the presence of [k^w] in Mormanno dialect. It is an important fact that to be explored in the future, because it could represent a stable feature able to determine a strong isophone to the considerable dialectological variability of Calabria.

Keywords: dialectology, Calabrian dialects, voiceless labiovelar stop consonant.

1. Introduzione

L'obiettivo di questa ricerca è dimostrare con analisi uditiva e acustica la presenza del suono occlusivo labiovelare sordo [k^w] nel dialetto di Mormanno e di San Pietro in Guarano attraverso l'analisi delle realizzazioni fonetiche dei dimostrativi “quello, quelli, quella e quelle”. Questo studio pilota si inserisce all'interno di un progetto più ampio avviato dal Laboratorio di Fonetica che mira alla definizione di nuovi confini tra i dialetti calabresi e, di conseguenza, alla rimodulazione di quelli già esistenti e descritti nelle ricerche di Trumper e Maddalon (1988, 1997 e 2016).

La letteratura attribuisce le occlusive labiovelari all'indoeuropeo perché: 1) in epoca storica, esse erano tipiche di alcune lingue indoeuropee¹; 2) in certe lingue indoeuropee, subirono un processo di semplificazione che ha prodotto per lo più il concepimento di semplici labiali oppure di semplici velari². Com'è noto, il latino è una delle lingue che ha conservato in parte le occlusive labiovelari dell'indoeuropeo infatti: a) ha preservato sostanzialmente la consonante labiovelare sorda [k^w] (ad esempio nei termini *quinque* e *aqua*); b) ha mantenuto la consonante labiovelare sonora [g^w] solo nel contesto postnasale (ad esempio nella parola *inguen*); c) ha

¹ Ad esempio, caratterizzavano il celtiberico (Lehmann 1999, Fanciullo 2007).

² È chiaro che tale percorso presuppone la formazione di elementi fonici provenienti da un'entità labiale o velare.

perso la consonante labiovelare aspirata sonora [g^wh] generando esiti differenti. Il contesto di indagine di questa ricerca, invece, non è eredità latina ma, come avviene per i dimostrativi italiani, un'innovazione romanza in quanto i segmenti nascono dall'univerbazione del latino ECCU(M) più ILLU/-A. Tale percorso dimostra, se fosse necessario, la continua trasformazione delle lingue. Questa dinamicità è frutto di mutamenti che non dipendono solo dalla sfera linguistica ma anche da variabili di tipo storico, sociale e geografico. Questi mutamenti hanno interessato e continuano a coinvolgere i dialetti italiani, attirando sempre di più l'attenzione della comunità scientifica. Molti sono gli studi che mirano a individuare fenomeni di convergenza o di divergenza tra i dialetti d'Italia con l'obiettivo di compierne una classificazione, ricorrendo a "strumenti" come isoglosse, eteroglosse e corridoi di transizione³. Quanto detto vale anche per i dialetti calabresi che sono stati suddivisi in specifiche aree dialettologiche proprio con l'ausilio dei succitati strumenti (Pellegrini, 1973; Falcone, 1976; Rohlf, 1980 e 1985; Trumper, Maddalon, 1988; Trumper, 1997⁴; Trumper, 2016). La ripartizione più articolata, e usata come riferimento in questo lavoro, è stata realizzata da Trumper e Maddalon (1988) e riproposta da Trumper (2016). Tale divisione marca le differenze tra i dialetti calabresi sulla base di quattro corridoi di transizione che si fondano su parametri di natura fonetica, fonologica, morfologica e sintattica, producendo come risultato la divisione della Calabria in cinque macroaree dialettali⁵. Poiché, gli studi condotti fino a oggi sui dialetti calabresi non hanno indagato la presenza e la distribuzione del suono labiovelare [k^w] tra le varietà dialettali della Calabria, questa ricerca punta a colmare questo vuoto.

2. *Le forme dei dimostrativi nei dialetti della Calabria: il punto*

Che alcuni dialetti italiani mostrino le labiovelari nella forma assunta dai dimostrativi, è un fatto ben noto nella produzione scientifica⁶. A tal proposito, poiché non

³ Corridoio di transizione: termine adottato nelle ricerche di Trumper, Maddalon (1988), Trumper (1997), Trumper, Chiodo (1999) e Trumper (2016) per indicare un fascio di isoglosse o di eteroglosse che delimita un'area transizionale più allargata in cui si incontrano, secondo diverse modalità, un numero differente di fenomeni linguistici.

⁴ La partizione dialettale proposta nella ricerca del 1997 si fondava su tre corridoi di transizione che comportavano, di conseguenza, la suddivisione della Calabria in quattro macroaree dialettologiche.

⁵ Trumper e Maddalon affermano che non è possibile individuare confini netti tra i dialetti calabresi in quanto essi mostrano un elevato grado di frammentarietà e divergenza (Trumper, Maddalon 1988: 245). Proprio per questi motivi, i due autori introducono la nozione corridoio di transizione (vedi nota 1) e puntano all'individuazione di macroaree dialettali.

⁶ Per una descrizione dei dimostrativi presenti nei dialetti settentrionali si consulti Rohlf (1966) pp. 205-207, Parry (1991) pp. 625-631 e Lombardi Vallauri (1995) pp. 209-225. In merito al toscano si rimanda al lavoro di Nocchi (2013). La storia dei dimostrativi romanzi è stata affrontata da Sornicola (2011) pp. 220-314, ma l'autrice non ha posto alcuna attenzione verso i dialetti della Calabria. Anche nella ricerca di Vanelli (1997) pp. 106-115 non compaiono riferimenti alle varietà dialettali calabresi. Un'analisi dell'evoluzione storica dei dimostrativi si trova nelle grammatiche storiche della lingua italiana e dei dialetti toscani di D'Ovidio, Meyer-Lübke (1978) pp. 169-170,

si dispone per i dialetti della Calabria di una bibliografia specifica su questo tema⁷, i profili dei dimostrativi attestati fino a oggi per le varietà dialettali calabresi possono essere estratti da studi che mirano: a) alla descrizione del numero di membri che costituiscono il sistema dimostrativo dei dialetti centromeridionali (Ledgeway, 2004); b) all'analisi dello sviluppo in diacronia dei dimostrativi nei dialetti italiani (Rohlf, 1966; Lausberg, 1976). Nella ricerca di Ledgeway (2004) si affronta la questione riguardante i tipi di sistemi dimostrativi che caratterizzano i dialetti centromeridionali odierni⁸. Per raggiungere tale scopo, l'autore esamina le strutture dei dimostrativi contenute in alcuni testi composti in dialetto da vari autori dell'Italia centromeridionale. In merito alle varietà dialettali della Calabria, dallo studio di Ledgeway possono essere ricavate le forme dei dimostrativi di alcuni dialetti della provincia di Cosenza e di Vibo Valentia (Ledgeway, 2004: 67 e 105):

- Amantea (CS) 'ssa (questa/costei/codesta) e chilli (quegli);
- Aprigliano (CS) chistu o chissu (questo/codesto) e chillu (quello); chisti o chissi (questi/codesti) e chilli (quegli); chista o chissa (questa/costei/codesta) e chillu (quella/colei);
- Belmonte (CS) 'sa (questa/costei/codesta) e chilli (quegli);
- Mendicino (CS) 'sa (questa/costei/codesta) e chiru (quello);
- Pedace (CS) ssu (questo/codesto) e chilli (quegli);
- San Pietro in Guarano (CS) chissu (questo/codesto) e chiddu (quello);
- Serra San Bruno (VV) chissi (questi/codesti) e chidi (quegli).

Rohlf, nel volume della sua grammatica storica dedicato alla morfologia, traccia l'evoluzione dei dimostrativi anche per i dialetti della Calabria (Rohlf, 1966: 207). L'autore afferma che i dimostrativi toscani "questo", "quello" e "codesto" sono recepiti dal "calabrese", rispettivamente, come "chistu", "chillu/chiddu" e "chissu".

Infine, nel suo manuale di linguistica romanza, Lausberg (Lausberg, 1976: 158) individua per il "calabrese" i seguenti dimostrativi: "kistu-kissu/questo; kista-kissa/questa; killu-kiđqu/quello; killa-kiđqa/quella". Inoltre, l'autore afferma che il solo IPSU, -A continua a vivere nel "calabrese" moderno attraverso "ssu/questo" e "ssa/questa" e che parallelamente si assiste anche alla sopravvivenza di ILLU, -A con "ddu/quello" e "dda/quella".

Dalla letteratura emerge un quadro netto che esclude categoricamente la presenza della consonante occlusiva labiovelare sorda nei dimostrativi dei dialetti calabresi.

Meyer-Lübke (1979) pp. 166-167, nella grammatica storica dell'italiano di Tekavčić (1980) p. 138 e nel testo di Maiden (1998) pp. 125-127.

⁷ Vedi nn. prec.

⁸ Ledgeway afferma che il sistema dimostrativo ternario, che nella citata ricerca l'autore identifica come sistema A, ha perso la propria centralità in un buon numero di dialetti centromeridionali odierni, eliminando l'antica distinzione tra CHISTO e CHISSO a favore di una generalizzazione che a volte premia il primo membro, a volte il secondo. Dunque, Ledgeway postula per i dialetti centromeridionali la presenza di due sistemi binari di dimostrativi: uno del tipo CHISTO/CHILLO, definito dall'autore sistema B, e uno del tipo CHISSO/CHILLO, indicato dall'autore come sistema C.

3. *Materiale e metodi*

Allo scopo di raggiungere l'obiettivo espresso nell'introduzione del presente contributo, è stata eseguita un'analisi uditiva e acustica delle produzioni fonetiche dei dimostrativi “quello, quelli, quella e quelle” realizzate in dialetto da parlanti di Mormanno e di San Pietro in Guarano. In questa ricerca, il modello elaborato da Trumper e Maddalon è assunto come punto di riferimento per la suddivisione dialettale della Calabria. Tale modello annovera il dialetto di Mormanno tra quelli della prima macroarea dialettale, che corrisponde alla porzione di Area Lausberg presente nel territorio calabrese, e il dialetto di San Pietro in Guarano tra quelli della seconda macroarea dialettale:

Figura 1 - (*Trumper, 2016: 227*)⁹



Il campione su cui si fonda questo studio è costituito da sei locutori per ciascun dialetto, tre di sesso femminile e tre di sesso maschile, equamente distribuiti in tre fasce d'età (18-35; 36-59; >60). I file sonori del corpus sono stati nominati con una codifica specifica contenente informazioni che riguardano sia l'intervistato che il luogo in cui è avvenuta la registrazione. Due esempi sono MM_M_21_M e CR_F_20_SPG:

- MM e CR rappresentano le iniziali del cognome e del nome del parlante;
- M e F indicano il sesso dell'intervistato;
- 21 e 20 denotano l'età dei soggetti intervistati;
- M e SPG specificano il paese in cui sono avvenute le registrazioni. La lettera “M” indica Mormanno, mentre l'acronimo “SPG” designa San Pietro in Guarano.

⁹ I triangoli di colore rosso indicano i punti interessati dall'inchiesta dialettologica e sono stati inseriti all'interno della mappa dai due autori dell'articolo.

Gli intervistati hanno letto una lista di venti frasi elaborata ad hoc per lo scopo di questa ricerca. Tale lista è stata presentata ai partecipanti in italiano e agli stessi è stato chiesto di tradurla direttamente in dialetto. Va puntualizzato che durante la stesura della lista di frasi usata per l'elicitazione del parlato non si è tenuto conto della funzione di pronomi o di aggettivo svolta dai dimostrativi in esame in quanto la loro appartenenza all'una piuttosto che all'altra categoria grammaticale non è determinata dal ricorso ad alterazioni di tipo fonetico. Pertanto, "l'etichetta dimostrativi" racchiude sia la funzione di aggettivo sia quella di pronomi.

In questo studio sono state analizzate quattrocento ottanta realizzazioni dialettali dei dimostrativi oggetto di studio. Le seguenti tabelle ne illustrano la quantità numerica estratta per ciascun segnale vocale registrato nei due comuni:

Tabella 1 - *Numero di dimostrativi analizzati per il dialetto di Mormanno*

MORMANNO				
ID_parlante	N. produzioni dimostrativo "quello"	N. produzioni dimostrativo "quelli"	N. produzioni dimostrativo "quella"	N. produzioni dimostrativo "quelle"
GMF_F_27_M	10	10	10	10
CV_M_33_M	10	10	10	10
FP_M_55_M	10	10	10	10
PR_F_57_M	10	10	10	10
GV_M_62_M	10	10	10	10
CG_F_76_M	10	10	10	10
TOT	60	60	60	60

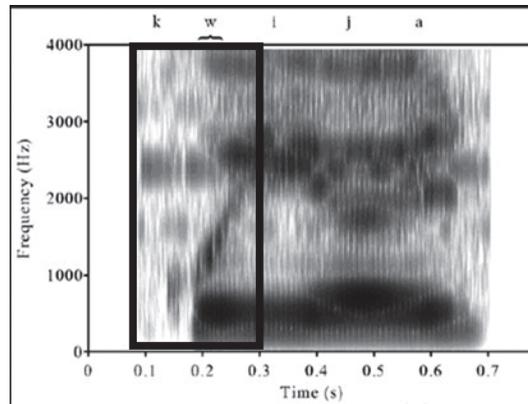
Tabella 2 - *Numero di dimostrativi analizzati per il dialetto di San Pietro in Guarano*

SAN PIETRO IN GUARANO				
ID_parlante	N. produzioni dimostrativo "quello"	N. produzioni dimostrativo "quelli"	N. produzioni dimostrativo "quella"	N. produzioni dimostrativo "quelle"
MD_M_20_SPG	10	10	10	10
RC_F_20_SPG	10	10	10	10
TA_M_35_SPG	10	10	10	10
TF_F_39_SPG	10	10	10	10
TT_M_75_SPG	10	10	10	10
MS_F_69_SPG	10	10	10	10
TOT	60	60	60	60

Su tutte le occorrenze riportate nelle Tabelle 1 e 2 è stata eseguita un'analisi uditiva e acustica. L'analisi uditiva è avvenuta attraverso l'ascolto ripetuto della porzione di segnale vocale riguardante i dimostrativi, con l'obiettivo di verificare a primo impatto la presenza/assenza della consonante oggetto di studio. Tramite l'analisi acustica, invece, sono state indagate le proprietà caratterizzanti la consonante occlusiva labiovelare sorda, in quanto l'assenza dei suoi fatti caratterizzanti consentirebbe di escluderne la presenza nei dialetti considerati da questa ricerca:

1. in genere, l'occlusiva labiovelare sorda è unita alla vocale successiva da una transizione formantica che mostra valori bassi di F2 (Ladefoged, Maddieson, 1996; Ladefoged, 2003; Kim, 2010). L'analisi del contesto appena esplicitato è avvenuto con l'ausilio del sonogramma. Sulla base dei dati acustici di letteratura, il tracciato formantico di una consonante occlusiva labiovelare sorda mostra il seguente andamento formantico:

Figura 2 - (Kim, 2010: 50)¹⁰



Inoltre, per avere contezza dell'eventuale escursione del valore di F2, sono stati estratti al cursore i valori minimi e massimi in Hertz della seconda formante rispettivamente all'attacco e al punto più stabile della vocale immediatamente successiva a [k^w];

2. di solito, la vocale che segue l'occlusiva labiovelare sorda mostra un abbassamento in Hertz del valore di F2 nella sua parte più stabile (Ladefoged, Maddieson, 1996; Ladefoged, 2003). Pertanto, in questo lavoro, per i dimostrativi oggetto di indagine sono stati misurati, nella porzione di segnale sonoro più stabile, i valori di F2 della vocale che segue [k^w], che corrisponde in tutti i casi analizzati alla vocale [i]. Allo scopo di stabilire se [k^w] produca o meno un abbassamento della

¹⁰ Il sonogramma ritratto nella Figura 2 mostra il tracciato formantico dell'occlusiva labiovelare sorda del nuuchahnulth, lingua wakashana parlata lungo la costa occidentale dell'isola di Vancouver in Canada. Questa lingua contempla la consonante occlusiva labiovelare sorda nel proprio sistema fonologico.

seconda formante di [i], per ciascun locutore sono state eseguite dieci misure dei valori formantici della succitata vocale in contesti fonetici differenti¹¹. L'analisi acustico-uditiva è stata eseguita con Multi-Speech versione 2.2.

4. *Analisi dei dati*

L'analisi uditiva prova la presenza della consonante occlusiva labiovelare sorda nelle realizzazioni dei dimostrativi eseguite dai locutori di Mormanno, mentre il risultato opposto caratterizza i parlanti rappresentativi del dialetto di San Pietro in Guarano:

Tabella 3 - *Esiti dimostrativi – analisi uditiva*

DIMOSTRATIVI	ESITI DIALETTO MORMANNO
“Quello”	[^l k ^w iɖ:u]
“Quelli”	[^l k ^w iɖ:i]
“Quella”	[^l k ^w iɖ:a]
“Quelle”	[^l k ^w iɖ:e]

Tabella 4 - *Esiti dimostrativi – analisi uditiva*

DIMOSTRATIVI	ESITI DIALETTO SAN PIETRO IN GUARANO
“Quello”	[^l ki:ru]
“Quelli”	[^l ki:ri]
“Quella”	[^l ki:ra]
“Quelle”	[^l ki:re]

Lo studio dei sonogrammi conferma in modo chiaro il risultato dell'analisi uditiva. Infatti, tutti i sonogrammi riguardanti i dimostrativi prodotti nel dialetto di Mormanno sono caratterizzati dalla dinamica transizionale di F2 tipica dell'occlusiva labiovelare sorda, mentre tutti quelli relativi ai dimostrativi esaminati per il dialetto di San Pietro in Guarano ritraggono le caratteristiche tipiche di una semplice occlusiva velare sorda:

¹¹ Chiaramente, le misure hanno riguardato la vocale “i” in contesto tonico.

Figura 3 - Dimostrativo “quelle” realizzato nel dialetto di Mormanno dal parlante GV_M_62_M

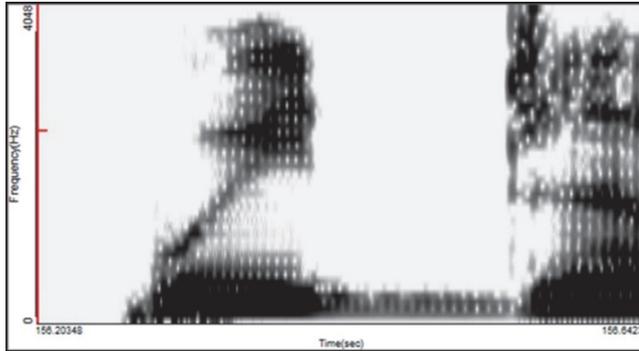
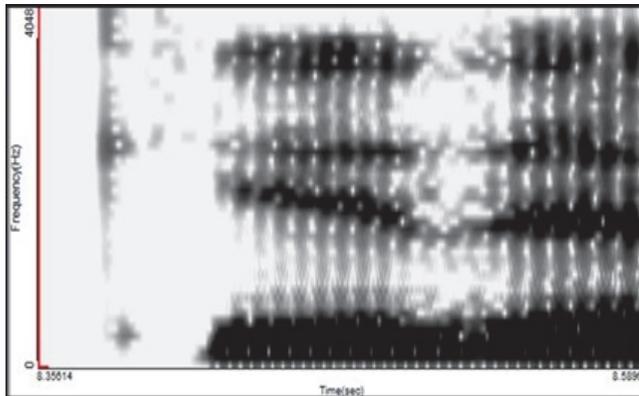


Figura 4 - Dimostrativo “quelle” realizzato nel dialetto di San Pietro in Guarano dal locutore TA_M_35_SPG



La tendenza delle transizioni formantiche riguardante la produzione dei dimostrativi dei parlanti di Mormanno mostra una differenza pari a 916,19 Hertz tra il valore minimo, che coincide con l’attacco della transizione della vocale successiva a [k^w], e il valore massimo, che corrisponde al punto stabile della vocale [i] che segue la consonante occlusiva labiovelare sorda:

Tabella 5 - *Tendenza della transizione formantica caratterizzante la realizzazione dei dimostrativi dei parlanti di Mormanno*

	MIN (Hz)	MAX (Hz)	DIFF. TRA MIN E MAX (Hz)
Media	1197,31	2113,50	916,19
Dev.St	102,35	95,40	75,35

Come dimostra la Figura 4, questa dinamica è assente dalle produzioni effettuate nel dialetto di San Pietro in Guarano, in quanto i locutori realizzano l'attacco dei dimostrativi tramite una semplice occlusiva velare sorda.

Un risultato che diverge rispetto alla letteratura (§2) è il mancato abbassamento in Hertz del valore di F2 nella parte più stabile della vocale [i]¹², l'unico vocoide che compare nel contesto di produzione dei dimostrativi sia dopo l'occlusiva labiovelare sorda che in seguito all'occlusiva velare sorda¹³. Lo stesso risultato, chiaramente, riguarda il dialetto di San Pietro in Guarano:

Tabella 6 - *Valori formantici della vocale "i" preceduta da [k^w] nella realizzazione dei dimostrativi – dialetto di Mormanno*

Vocale "i" preceduta da [k^w] nei dimostrativi			
	F0	F1	F2
Media	126,47 (Hz)	253,75 (Hz)	2023,97 (Hz)
Dev.St	1,75	3,38	37,24

Tabella 7 - *Valori formantici della vocale "i" in altri contesti – dialetto di Mormanno*

Vocale "i" in altri contesti			
	F0	F1	F2
Media	125,43 (Hz)	251,11 (Hz)	2028,77 (Hz)
Dev.St	1,96	4,44	40,12

Tabella 8 - *Valori formantici della vocale "i" preceduta da [k] nella realizzazione dei dimostrativi – dialetto di San Pietro in Guarano*

Vocale "i" preceduta da [k] nei dimostrativi			
	F0	F1	F2
Media	125,15 (Hz)	250,86 (Hz)	2020,46 (Hz)
Dev.St	1,52	3,54	38,53

5. Conclusioni

L'analisi acustico-uditiva dimostra la presenza del fono occlusivo labiovelare sordo [k^w] nei dimostrativi del dialetto di Mormanno, mentre si osserva la sua assen-

¹² Ladefoged e Maddieson (1996) evidenziano anche un abbassamento del valore di F1 come fattore dipendente dalla consonante labiovelare sorda. Tale dinamica non è stata riscontrata per i parlanti di Mormanno.

¹³ Non si ricorre ad analisi statistica in quanto i valori formantici sono simili.

za in quello di San Pietro in Guarano. Questo risultato ha una notevole valenza linguistico-storica, perché risulta evidente la corrispondenza con i tipi romanzi in cui i dimostrativi preservano la consonante oclusiva labiovelare sorda (Sornicola, 2011). Inoltre, i risultati di questa ricerca contrastano con le conclusioni degli studi presentati nel primo paragrafo di questo studio (§1), confermando ancora una volta come la Calabria sia caratterizzata da una notevole varietà ma anche frammentarietà dialettologica. L'analisi acustica, invece, conferma in parte l'esistenza di proprietà acustiche che definiscono universalmente il suono oclusivo labiovelare sordo (Ladefoged, Maddieson 1996; Kim, 2010). Nello specifico, il parametro che resta stabile è la dinamica transizionale di F2, mentre le misure formantiche non evidenziano alcun abbassamento dei valori di F2 della vocale successiva alla consonante labiovelare sorda. Infine, va sottolineato che la labiovelare si manifesta senza alcuna variabilità legata al sesso o all'età e pertanto essa è un elemento che persiste indipendentemente dai menzionati fattori sociolinguistici. In futuro, l'analisi uditiva e acustica condotta in questo studio sui dimostrativi sarà estesa a un numero maggiore di dialetti calabresi. Questo *modus operandi* consentirà di definire la diffusione di [k^w] nei dialetti calabresi, con l'obiettivo di verificare se la consonante oclusiva labiovelare sorda rappresenti un tratto stabile in grado di determinare un'isofona resistente alla notevole variabilità dialettologica della Calabria.

Riferimenti bibliografici

- D'OIDIO, F., MEYER-LÜBKE, W. (1906). *Grammatica storica della lingua e dei dialetti italiani*, trad. it. Eugenio Polcari, Modena 1978, Modena: La Goliardica.
- FALCONE, G. (1976). Calabria. In CORTELAZZO, M. (Eds.), *Profilo dei dialetti italiani*, 18, 5-108, Pisa: Pacini Editore.
- FANCIULLO, F. (2007). *Introduzione alla linguistica storica*, Bologna: il Mulino.
- KIM, E.S. (2010). The phonological and phonetic aspects of labialization: a crosslinguistic study. In *Studies in Phonetics, Phonology, and Morphology*, 16, 41-64.
- LADEFOGED, P. (2003). *Phonetic Data Analysis*, Oxford: Blackwell.
- LADEFOGED, P., Maddieson, I. (1996). *The Sounds of the World's Languages*, Cambridge: Blackwell.
- LAUSBERG, H. (1976). *Linguistica Romanza*, trad. it. Nicolò Pasero, Milano: Feltrinelli.
- LEDGEWAY, A. (2004). Lo sviluppo dei dimostrativi nei dialetti centromeridionali. In *Lingua e Stile*, XXXIX, 63-110.
- LEHMANN, W.P. (1999). *La linguistica indoeuropea: storia, problemi, metodi*, Bologna: il Mulino.
- LOMBARDI VALLAURI, E. (1995). Il sistema dei pronomi dimostrativi dal latino al piemontese (varietà torinese), una catena di trazione morfologica. *Proceedings del XXVII congresso della Società di linguistica italiana*, Lecce, IT, 28-30 October 1993.
- MAIDEN, M. (1998). *Storia linguistica dell'italiano*, Bologna: il Mulino.

- MEYER-LÜBKE, W. (1941). *Grammatica storica della lingua italiana e dei dialetti toscani*, trad. it. Matteo Bartoli e Giacono Braun, Torino: Loescher.
- PARRY, M. (1991). Le système démonstratif du cairese. *Proceedings du XVIII Congrès International de Linguistique et de Philologie Romanes*.
- NOCCHI, N. (2013). Varianti fonetiche per variabili sociali: per un'analisi sociofonetica delle labiovelari in Toscana. *Actes du Vie Dies Romanicus Turicensis*, Zurigo, ETS, 24-25 Giugno 2011.
- PELLEGRINI, G.B. (1973). I cinque sistemi dell'italo-romanzo. In *Revue roumaine de Linguistique*, 18, 105-129.
- ROHLFS, G. (1966). *Grammatica storica della lingua italiana e dei suoi dialetti*, Morfologia, trad. it. Temistocle Franceschi, Torino: Einaudi.
- ROHLFS, G. (1980). *Calabria e Salento*, Ravenna: Longo Editore.
- ROHLFS, G. (1985). *Latinità ed ellenismo nel mezzogiorno d'Italia*, Chiaravalle Centrale: Frama Sud.
- SORNICOLA, R. (2011). Per la storia dei dimostrativi romanzi: i tipi neutri [tso], [so], [ço], e la diacronia dei dimostrativi latini. In *Zeitschrift für romanische Philologie*, 127, 2: 220-314.
- TEKAVČIĆ, P. (1980). *Grammatica storica dell'italiano. Morfosintassi*, Bologna, il Mulino.
- TRUMPER, J. (1997). Calabria and Southern Basilicata. In *The Dialects of Italy*, M. Maiden & M.M. Parry (Eds.), Londra: Routledge, 355-364.
- TRUMPER, J. (2016). *Geostoria linguistica della Calabria*, Roma: Aracne.
- TRUMPER, J., Maddalon, M. (1988). Converging Divergence and Diverging Convergence: The Dialect-Language Conflict and Contrasting Evolutionary Trends in Modern Italy. In *Variation and Convergence: Studies in Social Dialectology*, Auer, P., di Luzio, A. (Eds.), 217-259.
- TRUMPER, J., CHIODO, G. (1999). La pertinenza degli eventi catastrofici naturali per la dialettologia e la linguistica romanze. *Rivista italiana di Dialettologia*, XXIII, 9-38.
- VANELLI, L. (1997). Personal pronouns and demonstratives, in *The dialects of Italy*. In Maiden M., Parry M. (Eds.), 106-115.

DALILA DIPINO

Lunghezza vocalica e consonantica in due varietà acquisite di ligure: uno studio pilota

Vowel and consonant length in two non-native Ligurian varieties: a pilot study

This pilot study investigates the realization of stressed vowel and post-stress consonant duration by two Southern Italian speakers of a Ligurian variety, Genoese and Intemelian respectively. A good approximation to native Ligurian duration patterns is achieved by both L2 speakers. However, our results suggest that the vowel length contrast is difficult to master by the non-native speaker of Genoese, especially in a less controlled context. Moreover, the younger native Ligurian speakers in the control group reveal a more complex pattern than expected, as far as consonant duration is concerned. The findings are discussed with reference to cross-linguistic patterns of vowel and consonant length acquisition in production and perception.

Keywords: Vowel Length, Consonant Length, Ligurian, L2 acquisition.

1. Introduzione

Il presente lavoro, che intende rappresentare un primo studio di carattere esplorativo, prende in esame un caso peculiare di acquisizione in lingua seconda: oggetto d'indagine non saranno infatti lingue standard, come d'uso nel settore, ma varietà dialettali, specificamente italo-romanze settentrionali.

In particolare, verranno esaminati due locutori italiani meridionali emigrati in Liguria, ognuno dei quali ha appreso una differente varietà di ligure come lingua seconda¹ in età precoce: rispettivamente il genovese e l'intemelio. Le loro produzioni saranno confrontate con quelle di parlanti nativi di ligure (oltre che di italiano) dal punto di vista dei rapporti di durata vocalica e consonantica. Tale scelta è legata al diverso trattamento della lunghezza in queste varietà, rispetto all'italiano standard e tra loro: grazie a questa ricchezza di schemi ritmici, avremo la possibilità di osservare come informatori aventi uno stesso sistema fonologico di partenza (l'italiano nel nostro caso) reagiscano a varietà L2 con una differente organizzazione temporale (genovese e intemelio).

¹ Con il termine "lingua seconda" non intendiamo riferirci all'ordine cronologico di apprendimento, quanto piuttosto al fatto che il processo di acquisizione sia avvenuto in un momento posteriore rispetto a quello della propria lingua materna. Nei casi qui esaminati, il ligure rappresenterebbe la terza varietà acquisita, dopo il dialetto locale e l'italiano, appresi all'incirca allo stesso tempo.

Nel prossimo paragrafo (§2) si tenterà di fornire una panoramica sulle conoscenze in nostro possesso riguardanti il processo di acquisizione degli schemi temporali di una lingua seconda, sia in percezione che in produzione. Seguiranno una breve descrizione dei sistemi fonologici delle varietà interessate (§3.1) e la definizione delle ipotesi di lavoro (§3.2). Al paragrafo §4 presenteremo i materiali e i metodi, infine l'analisi (§5) e la discussione dei risultati (§6) occuperanno gli ultimi due paragrafi.

2. *Produzione e percezione della lunghezza nell'acquisizione di una lingua seconda*

Negli ultimi anni, la ricerca sull'acquisizione delle caratteristiche temporali di una lingua seconda si è arricchita di numerosi contributi che hanno esplorato una serie di possibili scenari, sia dal punto di vista della produzione che della percezione.

2.1 La lunghezza vocalica e consonantica da un punto di vista tipologico

Com'è noto, le lingue del mondo possono trattare la durata diversamente: solo in alcune la sua variazione assume la funzione linguistica di segnalare lessemi di significato diverso². L'uso fonologico della lunghezza può interessare la sola classe delle vocali, come avviene ad esempio in tedesco, quella delle consonanti, come in italiano, oppure entrambi i tipi di segmenti contemporaneamente, come in giapponese.

Dal punto di vista tipologico, l'adozione della durata come tratto distintivo non è eccessivamente rara. Nel solo UPSID (UCLA Phonological Segment Inventory Database) di Maddieson (1984), su un campione di 317 lingue, ben 62, ovvero il 19,6%, esibiscono il contrasto di quantità tra vocali brevi e lunghe (Maddieson, 1984: 129); tuttavia, come nota Loporcaro (2015: 1), per il criterio di selezione adottato tale numero è "surely underestimated"³. L'opposizione di quantità consonantica risulta meno diffusa rispetto alla precedente: su 451 lingue (della versione online dell'UPSID), solo per 12 è riportata, pari al 2,66% del totale. Il grado più elevato di marcatezza⁴ è raggiunto dalle lingue in cui il contrasto non è binario, bensì ternario, ossia si sviluppa su tre gradi di lunghezza, generalmente indicati come breve, lungo e superlungo. Secondo Blevins (2004: 201), "only three to five [*scil.* languages] are reported to have a three-way contrast in vowel length (Dinka,

² Permangono, in queste come in tutte le lingue, le variazioni di durata dovute a ragioni di natura puramente fonetica, come la durata intrinseca dei segmenti, il contesto segmentale adiacente, le oscillazioni della velocità d'eloquio, la struttura di parola (v. il noto *polysyllabic shortening effect*), la posizione all'interno di frase, oppure a ragioni legate alla struttura prosodica, in particolare la segnalazione dell'accento di parola e di frase, o alla struttura dell'informazione, come l'enfasi. Su questo argomento cfr. ad esempio gli studi ormai classici di Lehiste (1970) e Marotta (1985).

³ Maddieson (1984), infatti, sceglie di escludere quelle lingue le cui opposizioni di lunghezza concernono tutti i timbri vocalici. Secondo tale criterio, anche le varietà analizzate in questo contributo sarebbero state scartate.

⁴ Sul legame tra marcatezza e acquisizione L2 si vedano ad esempio Soriano (2012: 168 e segg.) e Eckman (2008).

Coatlan Mixe, San José Paraíso Mixe, Estonian, Sami, and Yavapai); and only two are claimed to have a three-way contrast in consonant length (Estonian and Sami)”.

2.2 Il punto di vista della produzione

Relativamente alle modalità di acquisizione dell’opposizione di lunghezza in una lingua seconda, un dato che pare ormai assodato è la minore accuratezza degli apprendenti in produzione rispetto alla percezione.

Gli studi incentrati sulla produzione del contrasto di lunghezza⁵ mostrano in generale un avvicinamento da parte degli apprendenti al *target* acustico dei parlanti nativi, senza tuttavia una piena aderenza alle loro categorie fonologiche. I risultati sembrano risentire di diversi fattori: tra questi, il grado di marcatezza dei segmenti interessati (Soriano, 2009; 2014), il ruolo del tratto di lunghezza nella L1 dell’apprendente (Celata, Costamagna, 2011), il tempo di esposizione alla L2 (Kabak, Reckziegel & Braun, 2011; Hirata, 2017), il contesto in cui si realizza la comunicazione (D’Apolito, Gili Fivela, 2019), la posizione dell’accento e la frequenza lessicale (Giannini, Costamagna, 1997; Soriano, 2014).

Per quanto concerne la lunghezza vocalica molto si è scritto, in particolare sull’inglese; per le altre lingue la rosa di studi appare invece più ristretta. Mah, Archibald (2003) analizzano la durata vocalica e consonantica in una giovane parlante inglese a 4 e 6 mesi dall’inizio del processo di apprendimento del giapponese (lingua in cui sia la quantità vocalica sia quella consonantica sono distintive). Il soggetto distingue sistematicamente entrambi i tipi di segmenti, lasciando supporre che abbia acquisito le nuove categorie fonologiche della L2. Allo stesso tempo, tuttavia, l’ampia variabilità da segmento a segmento nei rapporti V/V: (così come in quelli C/C:) sembrerebbe indicare la mancanza di un controllo ancora pieno degli schemi prosodici del giapponese. Realizzazioni più accurate sono riportate da Nimz (2011) per apprendenti turchi⁶ di tedesco L2. Dopo 3 anni di istruzione, i parlanti raggiungono risultati comparabili con i nativi, nonostante la forma ortografica di alcune parole sembri influenzarne la resa⁷. Il recente contributo di Lee, Mok (2018) esamina invece le realizzazioni di vocali e consonanti giapponesi da parte di studenti cantonesi principianti e avanzati, nella cui lingua madre la quantità vocalica,

⁵ Tra la produzione, piuttosto limitata, sul tema dell’acquisizione delle geminate nella propria lingua madre, invece, si vedano ad esempio i recenti contributi di Khattab, Al-Tamimi (2015) per l’arabo libanese e di Payne, Post, Garmann & Simonsen (2017) per il norvegese e la bibliografia in essi contenuta.

⁶ Su ammissione della stessa autrice (Nimz, 2011: 1494), lo statuto linguistico della durata vocalica in turco non è ancora del tutto pacifico: sulla base dei numerosi prestiti con vocali lunghe dall’arabo e dal persiano, secondo alcuni linguisti la lunghezza vocalica dovrebbe essere considerata distintiva.

⁷ In tedesco generalmente la lunghezza della vocale è segnalata dalla grafia, come per <Spiel> [ʃpi:l] ‘gioco’ o <sehen> [ˈze:ən] ‘vedere’ (<ie> o <h>). Esistono, nondimeno, casi meno immediati, come ad esempio <Fuß> [fu:s] ‘piede’. Nell’esperimento di Nimz (2011), solo per questi ultimi sono riportate vocali significativamente più brevi. L’influenza dell’ortografia è documentata, tra l’altro, anche per altre lingue: un esempio è costituito dall’inglese L2 di italiani, i quali tendono a realizzare più lunghe le consonanti che per iscritto sono contrassegnate da un doppio grafema <CC>, p. es. <propose> vs <oppose> (Mairano, Bassetti, Sokolović-Perović & Cerni, 2018).

ma non quella consonantica, è fonologica. Gli apprendenti (di ogni livello e nella stessa misura) riescono a produrre opposizioni di lunghezza in L2 per entrambe le classi di segmenti, sebbene i rapporti di durata risultino inferiori a quelli dei nativi e manchi l'amplificazione dei contrasti tipica del parlato lento dei madrelingua. Nel caso contrario, in cui soggetti che nella loro L1 posseggano già distinzioni fonologiche di durata debbano confrontarsi con lingue in cui tale distinzione non è fonemica, i risultati vanno nella direzione opposta. Tsukada (2009) e, sulla stessa linea, De Weers, Munro (2018) studiano l'influenza dell'opposizione brevi/lunghe del giapponese (e del thailandese nel primo studio) sulla produzione delle vocali inglesi, distinte principalmente per il tratto di tensione (es. /i:/ vs /ɪ/)⁸. Come già in Munro (1993) per gli arabi, nelle produzioni degli apprendenti giapponesi si riscontra una tendenza all'amplificazione delle differenze di durata tra vocali tese e rilassate rispetto ai nativi. Tale risultato è ricondotto dagli autori ad un'acquisizione incompleta dell'opposizione di tensione dell'inglese, che si manifesta in un impiego eccessivo di un tratto che ha valore solo allofonico nella L2, ma è presente a livello fonemico nella L1 dell'apprendente (ovvero il ricorso alla durata, oltre e al di sopra delle differenze spettrali).

Anche sul versante della durata consonantica disponiamo di una ricca serie di lavori, che riguardano principalmente le geminate del giapponese (Han, 1992; Harada, 2006; Hayes-Harb, Masuda, 2008, Hirata, 2017, oltre a quelli già citati) e dell'italiano⁹.

Per l'italiano L2, gli studi incentrati sulla produzione del contrasto di lunghezza consonantica da parte di apprendenti stranieri, nella cui lingua madre manca tale opposizione (Giannini, Costamagna, 1997; Marotta, Barth, 2008; Sorianello, 2009; 2014; Kabak et al., 2011; De Clercq, Simon & Crocco, 2014) o è strutturalmente diversa (Celata, Costamagna, 2011), mostrano come gli apprendenti si avvicinino ai rapporti di durata dei madrelingua, nonostante la differenza di partenza¹⁰.

Kabak et al. (2011) forniscono indizi sul ruolo dell'esposizione alla L2 nella realizzazione dei contrasti, esaminando due gruppi di parlanti di madrelingua tedesca, il primo con nessuna esperienza in italiano, l'altro con conoscenze avanzate, a confronto con italiani nativi: come mostrato dai rapporti di durata, che si situano lungo una scala inesperti > avanzati > nativi, la produzione risente positivamente del livello di esposizione alla lingua bersaglio¹¹. Tuttavia, ad un esame più approfondito,

⁸ Nonostante qualità e quantità siano strettamente interconnesse, è stato dimostrato come nelle distinzioni vocaliche inglesi le differenze spettrali abbiano un peso maggiore rispetto alle differenze di durata, sebbene queste ultime rimangano un importante correlato secondario (Hillenbrand, Clark & Houde, 2000).

⁹ Cfr. anche i lavori sulla percezione delle geminate nei prestiti e il loro adattamento in Kubozono (2017).

¹⁰ A questi potrebbe essere aggiunto anche il recente contributo di Einfeldt, Weijer, Kupisch (2019) sulla produzione delle geminate in parlanti bilingui simultanei di italiano e tedesco, vissuti in Italia (bilingui dominanti in italiano) e in Germania (*heritage speakers*).

¹¹ Effetti dell'esperienza ugualmente positivi si rintracciano anche in Hayes-Harb, Masuda (2008) per il giapponese L2; al contrario, risultano inconsistenti i progressi documentati in Hirata (2017) per apprendenti inglesi intermedi di giapponese dopo un programma di *full immersion* di 4 mesi in Giappone.

la maggiore accuratezza degli apprendenti avanzati non è raggiunta mediante l'allungamento delle geminate, ovvero secondo la strategia temporale dei nativi, bensì attraverso la riduzione della consonante scempia, in altre parole tramite il riadattamento di una categoria già presente in L1.

I contributi di Sorianello (2009, 2014, 2019) non solo confermano la difficoltà per gli apprendenti di acquisire il tratto di lunghezza consonantica¹², difficoltà che si esprime attraverso fenomeni di degeminazione completa, la formazione di categorie temporali mediane e in qualche caso l'ipercorrezione, ma evidenziano dei *pattern* di acquisizione che possono essere ricollegati a dinamiche universali. Negli apprendenti tedeschi esaminati in Sorianello (2009), le consonanti per cui si riscontra il minor grado di accuratezza sono le ostruenti sonore e le sonoranti¹³, non a caso i segmenti posti più in basso nella scala di forza consonantica. Considerando inoltre che le occlusive sonore geminate sono alquanto rare, in italiano e nelle lingue del mondo¹⁴, l'autrice ne desume che il processo di acquisizione avvenga per "stadi progressivi, [...] da alcune consonanti prototipiche" (Sorianello, 2009: 57), per poi proseguire con i suoni meno salienti dal punto di vista percettivo, intrinsecamente meno lunghi e più marcati per diffusione e per distribuzione¹⁵.

Più recentemente, D'Apolito, Gili Fivela (2019) hanno valutato l'effetto di un altro parametro sulla geminazione in italiano L2: la quantità di informazioni disponibili nel contesto. Gli apprendenti francesi considerati non soltanto differiscono nel grado di accuratezza in base all'esperienza (con gli studenti avanzati che ottengono risultati più simili al gruppo di controllo dei nativi rispetto ai principianti), ma il loro schema temporale varia sensibilmente nel caso di un contesto di riferimento "ricco" (come una frase di senso compiuto, che specificando il significato della parola bersaglio ne facilita la disambiguazione) o "povero" (come le frasi cornice, nelle quali il riconoscimento dei membri delle coppie minime è legato alle sole caratteristiche fonetiche) (per maggiori dettagli, v. §4). Quando il contesto di disambiguazione è limitato, la maggioranza dei parlanti pronuncia le geminate in modo più accurato, ma solo i madrelingua (e, in misura minore e in modo meno sistematico, gli apprendenti avanzati) intervengono sulla vocale precedente, riducendone la durata al fine di enfatizzare i contrasti.

¹² D'altronde in linea con la maggiore marcatezza del tratto di lunghezza consonantica rispetto a quella vocalica (v. §2.1).

¹³ Sulla marcatezza delle geminate e le possibili basi fonetiche per la predilezione o meno di certi suoni (ad es. per le sonoranti, la maggiore continuità spettrale e il minore scarto di intensità rispetto alle vocali adiacenti), v. Dmitrieva (2017) e Kawahara, Pangilinan (2017).

¹⁴ Non a caso, le occlusive sonore geminate sono anche tra le più difficili da produrre dal punto di vista articolatorio, per le note ragioni aerodinamiche legate alla tenuta della sonorità (Ohala, 1983; 2011; su questo tema ci si permette di rimandare anche a Dipino, Celata, 2018 e alla bibliografia ivi contenuta).

¹⁵ Questi risultati non trovano tuttavia conferma nei dati di Kabak e colleghi (2011), i quali non evidenziano alcuna interazione significativa tra i tipi di segmenti e il livello di conoscenza degli apprendenti, ma solo tra il valore del rapporto CC/C e il tipo di consonante (differenza meno forte nel caso di consonanti sonore).

Riassumendo, sia per la durata vocalica sia per la durata consonantica i parlanti L2 sembrano in grado di costruire (o quantomeno iniziare a costruire) categorie fonologiche nuove (o ristrutturate rispetto a quelle native), con una buona approssimazione rispetto a quelle dei madrelingua. Il processo di acquisizione, tuttavia, almeno per le consonanti, sembra non essere uniforme, ma avvenire per stadi successivi, sulla base di “principi di naturalezza e di marcatezza tipologica” (Soriano, 2009: 57), conformemente alle previsioni di alcuni modelli teorici (cfr. Eckman, 1977; Major, 2001). Emergono, inoltre, categorie mediane e strategie di produzione dei contrasti temporali parzialmente diverse da quelle usate dai nativi, sia per quanto riguarda la realizzazione dei fenomeni allofonici, sia in relazione al contesto di disambiguazione. Si riscontra, infine, (seppur con qualche eccezione) un incremento dell’accuratezza in direzione delle produzioni native all’aumentare del tempo di esposizione alla lingua bersaglio.

2.3 Percezione della lunghezza

Risultati convergenti su soggetti di lingua russa (che non hanno quantità vocalica distintiva) nell’identificazione di vocali lunghe e brevi del finnico (Ylinen, Shestakova, Alku & Huotilainen, 2005) e dell’estone (Meister, Meister, 2011) sembrano suggerire l’esistenza di un *phoneme-boundary effect* (un effetto di discriminazione categoriale)¹⁶ per i parlanti nativi, contrariamente ai soggetti L2, i quali tendono a categorizzare le vocali entro un intervallo di valori più ampio e a collocare il confine categoriale in modo meno coerente e in corrispondenza di durate maggiori. Alla stessa conclusione sembra condurre Hayes-Harb (2005) per la percezione della lunghezza consonantica. Lo studio mostra inoltre come la categorizzazione possa affinarsi nel tempo, dal momento che soggetti inglesi con esperienza di un anno in giapponese percepiscono gli stimoli in modo più “categorico” di monolingui inglesi inesperti.

Sul piano dell’accuratezza, come accennato in apertura (§2.2), i risultati degli apprendenti in percezione sembrano tendenzialmente migliori rispetto alla produzione, indipendentemente dalle caratteristiche temporali della lingua di partenza. Innanzitutto, è ancora dibattuta la questione se essere sensibili al tratto di lunghezza per una categoria (le consonanti, ad esempio) garantisca il riconoscimento dello stesso tratto in un’altra categoria (le vocali). Dallo studio di Tsukada, Hirata & Roengpitya, (2014) sulla discriminazione delle vocali lunghe e brevi del giapponese da parte di apprendenti americani, italiani e thailandesi sembrerebbe di poter dare una risposta negativa, di contro ai risultati in Pajak, Levi (2014)¹⁷, Mah, Archibald (2003) e De Clercq et al. (2014). D’altra parte, lo studio comparativo di Altmann, Berger & Braun (2012) su italiano L2 e tedesco L2 ha evidenziato un’asimmetria nella facoltà di percepire contrasti di durata assenti nella propria L1. La percezione sarebbe più semplice nel caso delle vocali, mentre risulterebbe più complessa quando è coinvolta la classe delle consonanti: infatti, i parlanti nativi

¹⁶ In merito al concetto di percezione categoriale, indicata tradizionalmente come più “categorica” per le consonanti che per le vocali, si vedano tuttavia le osservazioni di Schouten, Gerrits & van Hossen (2003) sull’influenza del tipo di compito usato per la raccolta dei dati sui risultati.

¹⁷ Ma si veda a tal riguardo l’argomentazione in Lee, Mok (2018).

italiani, pur senza alcuna conoscenza del tedesco, ottengono un tasso di discriminazione delle vocali brevi e lunghe tedesche indistinguibile dal gruppo di controllo; al contrario le prestazioni dei parlanti nativi tedeschi nella discriminazione delle consonanti italiane risultano inferiori agli italiani nativi, in termini di tempi di reazione e accuratezza. Le autrici spiegano questa asimmetria con il differente carico informativo veicolato dalle vocali nella comunicazione, come l'accento, l'enfasi o l'intonazione. Questi risultati, sebbene non perfettamente in linea con quelli in Tsukada et al. (2014)¹⁸, confermano in ogni caso una tendenza già osservata in numerosi studi (cfr. tra gli altri Bohn, 1995; Escudero, Boersma, 2004; Cebrian, 2006; Nimz, 2011; Escudero, Benders & Lipski, 2009; v. Casillas, 2015: 185 per una lista più completa): pur non avendo opposizioni di lunghezza nella propria lingua madre, i parlanti tendono a basarsi sulle differenze di durata più che su quelle spettrali per discriminare le vocali di una lingua seconda (in altre parole le differenze di durata vocalica risultano più facilmente accessibili).

In sintesi, la percezione della lunghezza in L2 risulta più immediata rispetto ad altri tratti (come le differenze spettrali) e più semplice rispetto alla produzione, tanto che il suo grado di accuratezza raggiunge talvolta lo stesso livello dei nativi, relativamente alla lunghezza vocalica. Tuttavia, la discriminazione non sembra avvenire in modo categorico come nei madrelingua, ma lungo un intervallo di durate più ampio; inoltre, in linea con la maggiore sonorità e il maggior carico informativo veicolato dalle vocali da una parte e la maggiore marcatezza tipologica della quantità consonantica dall'altra, la percezione risulta più semplice quando coinvolge le vocali piuttosto che le consonanti.

Non sembra esserci ancora accordo, invece, sul ruolo e l'influenza dei tratti allofonici nel processo di acquisizione della lunghezza. Il famoso studio di McAllister, Flege & Piske (2002), in cui gli autori paragonano il grado di discriminazione della lunghezza vocalica di estoni, americani e spagnoli a quello di svedesi nativi, conclude che i fenomeni allofonici possono avere una certa incidenza sull'apprendimento della L2 e che questa incidenza è proporzionale alla centralità del fenomeno nel sistema della lingua di partenza. Difatti, i parlanti americani ottengono risultati migliori rispetto agli spagnoli,

¹⁸ In Tsukada et al. (2014) i parlanti cantonesi, malgrado abbiano nella propria L1 quantità vocalica distintiva, non ottengono un grado di riconoscimento delle opposizioni di lunghezza equiparabile ai giapponesi nativi. Tra le possibili ragioni di risultati mediamente inferiori in termini di accuratezza rispetto agli studi precedenti, gli autori citano la diversità della lingua *target* utilizzata, ovvero il giapponese, lingua moraicale e quindi ritmicamente molto diversa dalla L1, oppure l'effetto della presenza in giapponese di un fenomeno allofonico diametralmente opposto a quello dell'italiano, per cui le consonanti geminate sono precedute da vocali generalmente più lunghe (e non più brevi, come nella nostra lingua). Tale diversità ritmica potrebbe essere anche alla base dei risultati in Tsukada, Cox, Hajek & Hirata (2018), in cui italiani nativi presentano un tasso di corretta identificazione delle opposizioni di durata consonantica giapponesi molto alto, eppure simile a quello di apprendenti australiani e coreani (che non conoscono tale contrasto nelle loro lingue madri). La questione dell'accuratezza nell'identificazione dei contrasti di lunghezza in lingue familiari e non familiari è stata ulteriormente indagata in due recenti contributi dello stesso gruppo di ricerca (Tsukada, Hajek, 2019; 2020). Sottoponendo ad ascoltatori italiani e mandarini le opposizioni di durata consonantica giapponesi, sembra che la capacità di elaborazione della lunghezza consonantica in L1 si applichi positivamente in L2, sebbene sia sempre passibile di miglioramento con l'esperienza; al contrario, le scarse prestazioni dei soggetti mandarini sembrano indicare che le categorie acquisite in L2 non siano direttamente trasferibili in altre lingue.

sebbene l'inglese, come accennato in precedenza (§2.2 e nota 8), non presenti quantità vocalica contrastiva al pari dello spagnolo: ciò sarebbe dovuto al fatto che la variazione di durata in inglese è un correlato acustico della distinzione timbrica (fonologica) tra vocali tese e rilassate. Ricorrendo alle parole degli autori, “[...] the performance of these L1 groups varies as a function of the prominence of the role of duration in the L1” (McAllister et al., 2002: 256). Un analogo vantaggio percettivo è stato riscontrato ad esempio nel più recente contributo di Whang, Yazawa & Escudero (2019) per l'inglese australiano, il cui maggiore affidamento sulla durata nella distinzione delle vocali tese/rilassate rispetto all'inglese americano parrebbe agevolare la categorizzazione delle vocali giapponesi.

A dire il vero, l'influenza dei fenomeni allofonici sulla percezione dei contrasti fonologici della L2 sembrerebbe poter essere estesa all'intero spettro fonologico, al di là della sola dimensione temporale. Negli ultimi decenni ne sono stati documentati gli effetti per molteplici fenomeni¹⁹, in termini di minore accuratezza e maggior tempo di elaborazione: solo per fare alcuni esempi, si potrebbero citare gli effetti esercitati sulla percezione dalle varianti combinatorie di un fonema in L1 che nella L2 rappresentano due fonemi distinti (come i foni [d] e [ð], percepiti dagli spagnoli come meno distinti, in quanto rappresentanti di un unico fonema /d/, rispetto a soggetti inglesi in Boomershine, Hall, Hume & Johnson, 2008) oppure le conseguenze di un processo di neutralizzazione, come l'affricazione della sibilante post-sonorante in toscano (es. “ansa” [ansa] > [antsa], come “danza” [dantsa]), sulla percezione di sequenze simili in L2 (Celata, 2009). In generale, è stato messo chiaramente in luce come “la percezione non-nativa rifletta sia le proprietà fonologiche astratte, sia il dettaglio fonetico della lingua nativa” (Vayra, Avesani, Best & Bohn, 2012: 120) e oggi tutti i principali modelli teorici dell'acquisizione non nativa (SLM, Flege, 1987, 1995; PAM, Best, 1995; PAM-L2, Best, Tyler, 2007; L2LP, Escudero, 2009) concordano su questo assunto.

Resta ancora da chiarire in che direzione proceda questa influenza, in particolare nell'ambito della dimensione temporale²⁰. Come ben riassunto da Altmann et al. (2012: 391), gli studi sulla quantità vocalica da un lato suggeriscono che l'esperienza nella propria L1 di variazioni allofoniche di lunghezza faciliti la percezione del contrasto in L2 (McAllister et al., 2002), dall'altro testimoniano casi in cui il proprio sistema fonetico-fonologico potrebbe ostacolare l'acquisizione, bloccando strategie di discriminazione temporale più generali (cfr. Nenonen, Shestakova, Houtilainen & Näättänen, 2003)²¹ oppure inducendo ad assegnare ad alcuni correlati secondari più peso del necessario (Tsukada, 2009; De Weers, Munro, 2018).

¹⁹ Su questo argomento, si rimanda a Celata (2009) e Vayra, Avesani, Best & Bohn (2012).

²⁰ Per una discussione su questo tema, si vedano tra gli altri Escudero, Benders, Lipski (2009), Meister, Meister (2011), Casillas (2015).

²¹ Nenonen et al. (2003) mostrano come apprendenti russi precoci di finnico L2 ottengano percentuali di riconoscimento inferiori del contrasto di lunghezza solo per le vocali finniche assimilabili a quelle russe.

3. *Il nostro studio*

Il presente contributo si propone di esplorare alcune questioni riguardanti l'acquisizione L2 delle opposizioni di lunghezza che, come emerso dai paragrafi precedenti, rimangono ancora aperte. Tra queste troviamo la sensibilità ai contrasti di durata per una categoria diversa da quella della propria lingua madre (ad esempio consonanti *vs* vocali) oppure l'influenza positiva o negativa delle variazioni di durata allofoniche nell'apprendimento degli schemi ritmici L2 o ancora l'incidenza del contesto sull'accuratezza delle produzioni.

Come anticipato, tenteremo di rispondere ad alcuni di questi interrogativi nell'ambito della produzione, ricorrendo a varietà non standard. Come Escudero (2009) e più recentemente Williams, Escudero (2014) hanno messo in rilievo, infatti, anche la conoscenza di varietà dialettali ha importanti effetti sulla percezione delle lingue non native e va tenuta nella giusta considerazione. Oggetto d'indagine saranno le produzioni di due parlanti italiani meridionali emigrati in Liguria, ciascuno dei quali ha appreso precocemente una varietà ligure, rispettivamente il genovese e l'intemelio. I loro valori di durata saranno confrontati con quelli di parlanti nativi di ligure, oltre che di italiano, come gruppo di controllo. L'obiettivo principale è quello di esaminare il ruolo e l'influenza della L1 (italiano/varietà centro-meridionale) nell'acquisizione degli aspetti temporali della L2 (genovese o intemelio).

Come sarà chiaro dall'esposizione delle loro caratteristiche ritmico-prosodiche (§3.1), l'intrecciarsi di fenomeni fonetici e fonologici nell'organizzazione temporale di queste varietà dà vita ad una ricca casistica, che si presta ad essere studiata in prospettiva interlinguistica e potrebbe aiutarci a far luce su alcune questioni ancora in sospeso. Le specifiche domande di ricerca saranno meglio precisate al §3.2, dopo aver presentato brevemente le varietà in esame.

3.1 Sistemi fonologici a confronto

La riviera ligure si divide in molteplici gruppi dialettali (Forner, 1988), i quali differiscono, tra le altre caratteristiche, per il valore funzionale della lunghezza segmentale. Tra questi, il genovese e l'intemelio (parlato nell'estremo Ponente) si pongono agli antipodi: nel primo è distintiva la quantità vocalica, ma non quella consonantica²²; nel secondo, invece, né la durata vocalica né quella consonantica hanno valore fonologico.

La quantità vocalica in genovese coinvolge tutti i timbri vocalici e tutte le posizioni, in sillaba tonica e anche in sillaba protonica. Al quadro delle opposizioni di quantità vocalica derivato dall'allungamento in sillaba aperta tardolatino (Filipponio, 2012; Loporcaro, 2015) si aggiungono allungamenti condizionati dalle consonanti adiacenti, non ancora sufficientemente investigati: davanti a /v r ʒ/, con qualche eccezione /g/ e frequentemente ma non in modo generalizzato prima di /z/ (Toso, 1997: 16; Forner, 1975: 51-52, 250-251). Gli studi sperimentali più recenti, inoltre, hanno documentato la presenza di una parziale geminazione allofonica dopo vocale breve (v. i risultati in

²² Come del resto nelle altre varietà italo-romanze settentrionali, dove le consonanti sono generalmente scempiate (Rohlf's 1966: 321-324).

Filipponio, Garassino, 2019; Garassino, Dipino, 2019 e i tentativi di interpretazione²³ in Filipponio, Garassino & Dipino, 2019), sebbene in misura minore rispetto al dialetto ligure occidentale di Porto Maurizio (IM) (che ha un rapporto (V:)C/(V)C di 0.82).

In intemelio, la situazione è stata meno esaminata ed appare tuttora complessa. Se Forner (1988: 458) riporta l'esistenza di allungamenti condizionati dal contesto come in genovese, al contrario Azaretti (1982: 25) sulla base di valutazioni impressionistiche esclude del tutto distinzioni fonologiche di lunghezza. Quest'ultima affermazione trova conferma sperimentale nelle indagini preliminari in Garassino, Dipino (2019), che constatano l'assenza di differenze significative tra vocali etimologicamente lunghe e brevi, ma non nel più recente contributo di Garassino, Cangemi (2020). Sebbene questi ultimi risultati possano essere in parte ricondotti al diverso criterio di misurazione adottato (durata dell'intera sillaba tonica, composta da un numero variabile di segmenti, *vs* vocale tonica), si comprende quanto la realizzazione della durata in quest'area sia instabile e come la visione in nostro possesso sia ancora poco chiara²⁴, nonostante il generale accordo sull'assenza di opposizioni fonologiche di lunghezza.

Diverso e ben noto è il quadro dell'italiano standard e delle diverse varietà italo-romanze centro-meridionali, in cui la lunghezza della vocale tonica riveste un ruolo soltanto marginale, nel processo di allungamento allofonico indotto dal contesto sillabico²⁵, mentre vige la distinzione fonologica tra consonanti scempie e geminate (v. Bertinetto, Loporcaro, 2005).

Riassumendo le caratteristiche della varietà appena ricordate, si ottiene un intricato sistema di impiego della lunghezza (v. Tab. 1). Il quadro si complica ulteriormente se si tengono separati il piano della rappresentazione fonologica e il livello della sua realizzazione effettiva, il piano superficiale, quello in ultima analisi disponibile all'apprendente. Incrociando i diversi parametri, si ha un'idea degli adattamenti degli schemi prosodici della L1 che un parlante italiano centro-meridionale deve affrontare nell'acquisizione di una varietà ligure.

²³ Gli autori, sulla base del profilo di intensità e dello schema ritmico ligure, che mostra una tendenza al *controllo* (Bertinetto, Bertini, 2008), per questi allungamenti allofonici ipotizzano "maybe an incomplete degemination (as a conservative feature and not as an epiphenomen of close contact like in Bolognese), or the pressure of standard Italian" (Filipponio et al., 2019: 287). In attesa di ulteriori prove, tuttavia, la questione rimane aperta.

²⁴ Non a caso, è tuttoggi attivo presso l'Università di Zurigo un progetto di ricerca dal titolo "At the intersection between acoustic phonetics and information structure. An empirical investigation on the phonetic realization of vowel length in three Ligurian dialects (100015_178932), finanziato dal Fondo Nazionale Svizzero per la ricerca scientifica, che mira a gettare nuova luce sul ruolo e sulla distribuzione della lunghezza vocalica in quest'area.

²⁵ Il fenomeno, anche noto come *Open Syllable Lengthening*, è ben documentato in parole isolate, sotto accento frasale e in parlato di laboratorio, appare più variabile in stili d'eloquio meno controllati e a velocità d'elocuzione sostenuta (cfr. Landi, Savy, 1996). Su questo argomento, si sono succedute numerose indagini sperimentali, tra le quali ricordiamo Bertinetto (1981), Marotta (1985), Esposito, Di Benedetto (1999), cui si rimanda per ulteriori riferimenti bibliografici.

Tabella 1 - Schema della lunghezza vocalica e consonantica in genovese (GE), intemelio (VM) e italiano (IT). Le parentesi indicano variazioni di durata allofoniche

	GE		VM		IT	
	QV	QC	QV	QC	QV	QC
Livello fonologico	+	-	-	-	-	+
Livello superficiale	+	(+)	-	-	(+)	+

Come risulta chiaro dalla Tab. 1, nel caso del genovese L2, l'apprendente dovrà innanzitutto introdurre un nuovo contrasto, quello di lunghezza vocalica, non esistente come tratto fonologico nel suo sistema L1, per quanto vi si ritrovino variazioni allofoniche di durata vocalica; dovrà inoltre "disimparare" un contrasto, quello di durata consonantica, adeguandosi allo stesso tempo alle variazioni allofoniche del dialetto genovese. Nel caso dell'intemelio L2, invece, il soggetto dovrà imparare a "livellare" tutte le differenze di durata dell'italiano, in particolare le opposizioni fonologiche di lunghezza consonantica.

3.2 Domande di ricerca ed ipotesi

Come anticipato, con questo contributo intendiamo proporre uno studio esplorativo sull'interferenza a livello prosodico tra italiano/dialetto centro-meridionale e varietà liguri, nel dominio della produzione.

Le domande di ricerca si inseriscono nel solco degli studi precedenti. In primo luogo, per ognuna delle due varietà esaminate, si verificherà il grado di accuratezza dei soggetti L2 nella realizzazione del *target* acustico dei nativi. In quanto noto parametro di condizionamento sulla durata, si terrà conto anche del contesto di elicitazione, con l'intento di soppesarne il ruolo. Inoltre, grazie alla configurazione specifica dell'esperimento, come si accennava, si potrà valutare la reazione di uno stesso sistema fonologico (quello italiano meridionale) a due differenti varietà (genovese e intemelio), aventi una diversa organizzazione temporale: gli apprendenti saranno in grado di riorganizzare i propri rapporti temporali? In particolare, riusciranno a neutralizzare i contrasti di durata consonantica dell'italiano/dialetto centro-meridionale e, nel caso del genovese L2, a creare un nuovo contrasto di durata vocalica? Per quest'ultimo aspetto, la presenza di un'opposizione di durata nella lingua madre, sebbene per un'altra classe di suoni (le consonanti), facilita l'acquisizione del contrasto di durata vocalica? Più in generale, l'esistenza di fenomeni allofonici in L1 favorisce o ostacola l'acquisizione?

Alla luce dei risultati brevemente riportati ai §2.2-2.3, ci aspettiamo innanzitutto che l'apprendente di genovese riesca ad acquisire la nuova categoria di durata, realizzando il contrasto di lunghezza vocalica in genovese L2, ma in maniera non perfettamente equiparabile ai parlanti nativi in termini di accuratezza. Riguardo alla durata consonantica, i dati potrebbero andare in due direzioni: essendo sensibile alle differenze di durata consonantica nella propria lingua madre, il parlante genovese L2 potrebbe riuscire a neutralizzare il contrasto oppure, complici le differenze allofoniche presenti nel dialetto genovese, i contrasti di lunghezza potrebbero essere completamente trasfe-

riti dall'italiano al dialetto (cfr. Tsukada, 2009; De Weers, Munro, 2018). Tendenze analoghe ci aspetteremmo per il consonantismo nel parlante di intemelio L2. In merito al contesto (v. sotto, §4), invece, ipotizziamo che, in un contesto povero di informazioni, le realizzazioni L2 siano più prossime a quelle dei nativi rispetto ad un contesto lessicalmente ricco, in cui, considerata la facilità di disambiguazione dei lessemi e venendo meno la necessità di un controllo rigoroso, potremmo assistere ad una minore accuratezza.

4. *Materiali e metodi*

Al fine di verificare queste ipotesi, sono state analizzate le registrazioni di due locutori italiani meridionali (l'uno napoletano, l'altro calabrese) che hanno appreso rispettivamente il genovese e l'intemelio come lingua seconda in età precoce. Il primo informatore, GeL2, è nato a Napoli nel 1979, dove ha vissuto fino all'età di 28 anni, quando si è trasferito a Genova per insegnare nella scuola primaria. Ha acquisito come lingua materna contemporaneamente l'italiano e il dialetto napoletano, che parla tuttora regolarmente con la famiglia. I suoi contatti con il genovese risalgono ai primi anni di vita, attraverso la nonna di origine genovese, che interloquiva con lui in dialetto ligure durante le visite, fino all'età di 15 anni. Attualmente GeL2 usa con orgoglio il genovese con una cugina ligure e, quando possibile, con altre persone liguri.

Il secondo informatore, VmL2, è nato a Molochio²⁶ (RC) nel 1948. La sua famiglia emigra in Liguria, a Camproso (IM), quando ha circa un anno e mezzo. Fin dall'inizio i genitori si sforzano di insegnargli l'italiano 'standard', tant'è vero che il soggetto non è in grado di parlare nel dialetto calabrese, ma riesce a comprenderlo senza fatica. Apprende il dialetto intemelio a partire dall'età scolare, attraverso l'ambiente circostante. Usa tutt'oggi il dialetto intemelio con gli amici d'infanzia e ha contribuito ad organizzare concorsi letterari nel dialetto locale.

Le produzioni di questi parlanti sono state confrontate con due parlanti nativi (un soggetto giovane ed uno anziano) per ogni varietà considerata. I soggetti anziani hanno appreso prima il dialetto, i più giovani italiano e dialetto contemporaneamente. Il campione, dunque, ammonta nel complesso a 6 parlanti, come riassunto nella Tab. 2, ove sono indicati anche i principali dati biografici al momento delle interviste.

²⁶ Molochio appartiene oggi al gruppo dei dialetti meridionali estremi (Loporcaro 2009), mentre è probabile che fino alla seconda metà del XVI secolo vi si parlassero ancora dialetti greci (Stamuli, 2008: 48; Fanciullo, 1996: 54-7).

Tabella 2 - *Partecipanti*

		Sigla	Sesso	Età	Professione	Luogo di nascita
GE	L1	GeL1a	M	76	Ex geometra	Genova
		GeL1g	F	31	Pasticciera	Genova
	L2	GeL2	M	39	Insegnante	Napoli
VM	L1	VmL1a	F	83	Ex coltivatrice diretta	Ventimiglia (IM)
		VmL1g	M	31	Cuoco	Sanremo (IM)
	L2	VmL2	M	71	Geometra	Molochio (RC)

Il ricorso a classi d'età diverse è stato concepito per meglio valutare le conseguenze, in diacronia apparente²⁷, di eventuali pressioni dell'italiano sul dialetto e ricavare degli indizi sulla varietà che ha funto da modello nell'acquisizione della lingua seconda. Data l'esiguità del campione, tuttavia, i risultati presentati richiederanno ulteriori verifiche sperimentali su un numero più esteso di partecipanti, che riesca a cogliere anche le dinamiche sociolinguistiche del contatto.

Il materiale è stato raccolto tra la primavera e l'autunno del 2019 e proviene in larga parte da un corpus più esteso, progettato originariamente per altri scopi nell'ambito di un progetto attivo presso l'Università di Zurigo (v. nota 24). Le registrazioni sono state effettuate in ambienti silenziosi, ma non insonorizzati, con un registratore digitale Marantz PMD-661 e un microfono a cravatta Sennheiser MKE 2-P e salvate in formato .wav, con una frequenza di campionamento di 48.000 Hz.

Gli stimoli usati nell'inchiesta sono 20 parole parossitone (v. Tab. 3) condivise sia dal genovese sia dall'intemelio²⁸, che costituiscono delle coppie semi-minime distinte per quantità vocalica in genovese, mentre in intemelio dovrebbero presentare durata vocalica uniforme (v. §3.1).

²⁷ Per quanto la L1 non sia "fissata" una volta per tutte, ma si modifichi lentamente nel tempo, si considera il paradigma del tempo apparente un indice affidabile dei cambiamenti linguistici in corso. Recentemente è stata riportata l'attenzione anche su un altro tipo di evoluzione della L1: oltre a condizionare l'apprendimento della L2, la propria lingua madre ne subisce a sua volta l'influenza. Sulla cosiddetta *native language plasticity* v. de Leeuw, Celata (2019).

²⁸ Con l'eccezione dei tipi lessicali corrispondenti a "raffermo" e "fosso", che si sono rivelati poco usati al di fuori di Genova.

Tabella 3 - *Parole bersaglio*

Vocale tonica	V:	V		
a	[¹ na:zu]	‘naso’	[¹ mazu]	‘maggio’
e	[¹ ze:na]	‘Genova’	[¹ zene]	‘genere’
ɔ	[¹ pɔ:ku]	‘poco’	[¹ tɔku]	‘pezzo’
	[¹ pɔ:su]	‘raffermo’	[¹ fɔsu]	‘fosso’
œ	[¹ kœ:ʒe]	‘cuocere’	[¹ kœʃe]	‘cosce’
i	[¹ ri:zu]	‘riso’	[¹ risu]	‘riccio’
	[¹ fri:tu]	‘fritto’	[¹ situ]	‘zitto’
y	[¹ fry:tu]	‘frutto’	[¹ brytu]	‘sporco, ‘brutto’
u	[¹ spu:zu]	‘sposo’	[¹ pusu]	‘pozzo’
	[¹ du:se]	‘dolce’	[¹ duze]	‘dodici’

Tutte le vocali dell’inventario fonemico delle due varietà sono rappresentate, eccetto la vocale medio-bassa anteriore per il genovese.

Al fine di valutare l’interazione tra il contesto di occorrenza e la realizzazione dei contrasti di lunghezza, le parole sopraelencate sono state elicitate in tre diverse condizioni:

- frasi cornice, del tipo “Ho detto *riso* stavolta” (d’ora in avanti, Cornice);
- frasi di senso compiuto, con uno schema ritmico analogo, in cui la parola bersaglio appare in posizione interna di frase, come ad esempio in “La mamma prepara il *riso* con le zucchine” (Frasese);
- frasi cornice pragmaticamente marcate, con la parola *target* alternativamente in posizione focale e non focale, sul modello “Ho detto *riso*, non *riccio* stavolta” (Focus).

Per il primo ed il secondo contesto, i parlanti erano invitati a tradurre in dialetto la parola bersaglio (contesto *a*) o la frase (contesto *b*) lette loro dagli intervistatori; nel terzo compito, invece, gli stimoli erano presentati visivamente, sotto forma di carte da gioco, per consentire un’espressione del fuoco più naturale.

Due dei tre contesti (*a* e *b*) sono direttamente comparabili con quelli usati da D’Apolito, Gili Fivela (2019): il contesto Cornice rappresenta un contesto “povero” di informazioni, in cui il parlante tenderà ad utilizzare una forma fonetica chiara per favorire il riconoscimento dei lessemi; il contesto Frasese, all’opposto, collocando la parola bersaglio all’interno di una frase di senso compiuto, costituisce un contesto “ricco”, che favorisce la disambiguazione e permette al parlante di porre un’attenzione minore all’articolazione dei suoni. A questa diversità semantica si aggiunge anche quella prosodica, legata alla condizione di maggiore rilevanza informativa e dunque anche ritmico-prosodica in cui la parola bersaglio è posta nella condizione Cornice. Il contesto di Focus, infine, rappresenta un caso di focus correttivo, in cui si contrappongono due elementi lessicali che differiscono soltanto per una caratte-

ristica fonetica. Questo contesto, secondo i suggerimenti di De Jong (2004) e De Jong, Zawaydeh (2002), dovrebbe indurre il parlante ad enfatizzare i fonemi contrapposti, e nella fattispecie soltanto i tratti che sono ritenuti pertinenti.

In totale il numero di occorrenze ottenute è di 337, di cui 175 per il gruppo genovese e 162 per l'intemelio. I dati raccolti sono stati segmentati manualmente con il programma Praat, attenendosi al criterio della struttura formantica piena e ponendo i confini al più vicino *zero crossing* (Machač, Skarnitzl, 2009), e ne sono state ricavate in modo automatico le misure di durata della vocale tonica e della consonante seguente ed i rapporti reciproci. Per il calcolo della durata delle occlusive si è tenuto conto sia della fase di occlusione che di rilascio.

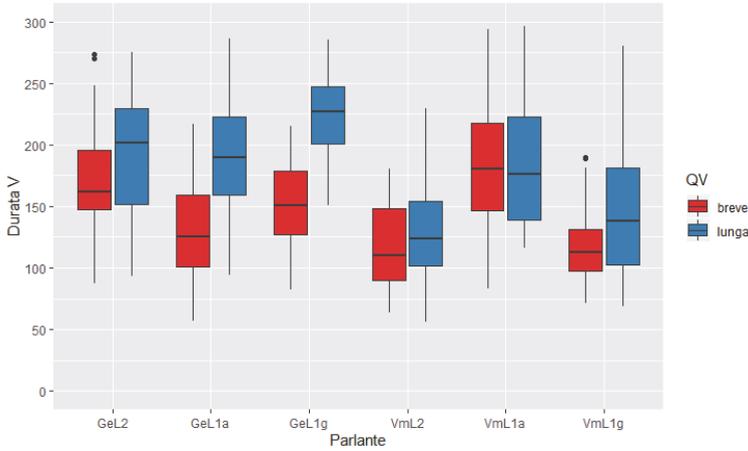
In fase di analisi si è esaminata la variazione di durata della vocale tonica e della consonante postonica in funzione della lunghezza vocalica attesa etimologicamente (breve o lunga), del contesto (Cornice, Frase, Focus) e del parlante. Considerato il numero limitato di occorrenze conseguente allo scorporo dei dati, solo ove possibile all'analisi qualitativa si è affiancata l'analisi statistica per mezzo del *Two-sample Welch t-test*, al fine di confermare la diversità delle medie tra vocali brevi e lunghe e relative consonanti postoniche. L'analisi statistica è stata svolta con l'ausilio del programma R.

5. Risultati

Poiché il nostro oggetto di studio è costituito da varietà dialettali non standardizzate, variabili in misura più o meno forte (rispetto all'esito atteso attraverso la ricostruzione diacronica e l'indicazione delle fonti) e in rapida evoluzione (v. §3.1), le produzioni dei parlanti nativi saranno mostrate nella loro variabilità e confrontate in modo analogo con quelle dei parlanti non nativi.

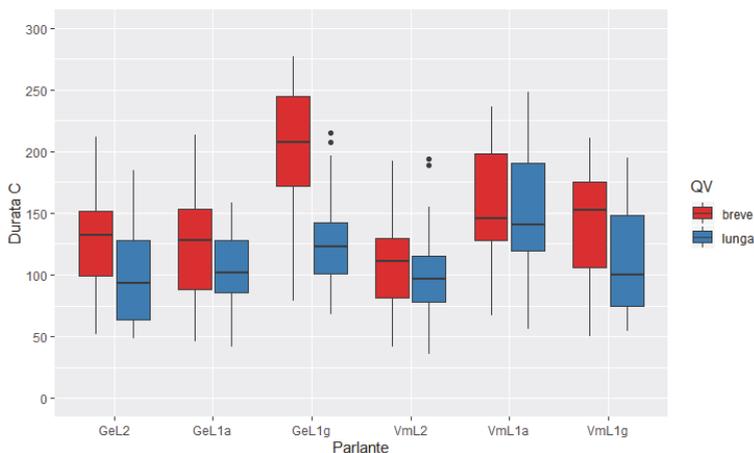
La Fig. 1 presenta i valori complessivi di durata della vocale tonica distinti sulla base della lunghezza fonologica attesa etimologicamente, mantenendo separati i dati per ogni parlante. I diagrammi mostrano come i parlanti nativi genovesi operino una distinzione netta tra vocali lunghe e brevi (soprattutto la parlante giovane GeL1g), la cui significatività è confermata dai test statistici inferenziali (GeL1a $t = -4,75$; $p < 0,001$; GeL1g $t = -6,53$; $p < 0,001$). Anche per il parlante non nativo GeL2 osserviamo differenze significative di durata della vocale tonica, sebbene di minore entità rispetto a quelle dei nativi ($t = -2,11$; $p < 0,05$). All'interno del gruppo intemelio, al contrario, il parlante meridionale sembra allinearsi completamente allo schema prosodico della parlante nativa anziana: per entrambi le vocali che nelle altre varietà sono distinte per lunghezza hanno durata simile (VmL2 $t = -0,79$; $p = 0,434$ e VmL1a $t = -1,40$; $p = 0,169$ rispettivamente, al di sotto della soglia di significatività). Il soggetto giovane VmL1g, inaspettatamente, produce variazioni di durata significative ($t = -2,92$; $p < 0,01$).

Figura 1 - Distribuzione della durata vocalica per parlante
(per le sigle, qui come in seguito, v. Tab. 1)



In riferimento alla durata consonantica, la Fig. 2 sintetizza i risultati generali, come in precedenza. Appare immediatamente evidente il contrasto tra le consonanti postoniche della parlante genovese giovane ($t = -6,80$; $p < 0,001$). Statisticamente significative sono anche le più ridotte variazioni di durata di GeL2 ($t = -3,03$; $p < 0,01$) e GeL1a ($t = -2,52$; $p < 0,05$). Al contrario, in intemelio sia il parlante anziano sia il parlante non nativo confermano anche per le consonanti l'assenza di opposizioni (VmL1a $t = -0,40$; $p = 0,693$ e VmL2 $t = -1,49$; $p = 0,143$). Ancora una volta, il parlante intemelio giovane mostra un *pattern* inatteso ($t = -2,77$; $p < 0,01$).

Figura 2 - Distribuzione della durata consonantica per parlante

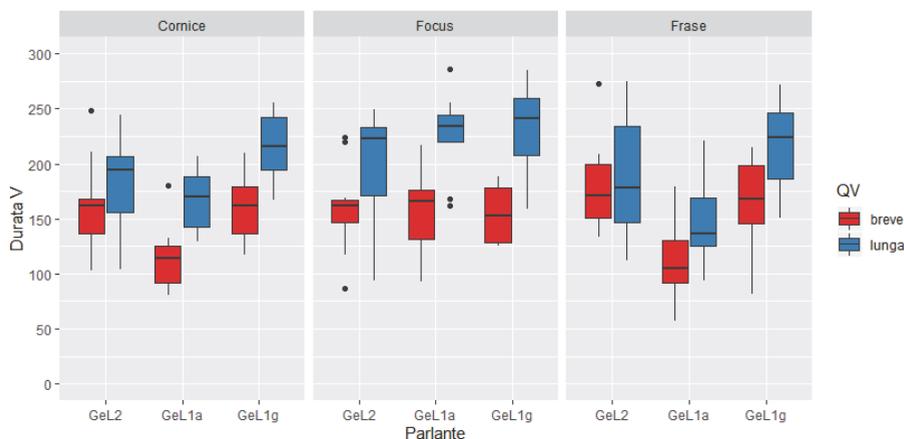


Al fine di verificare l'influenza del contesto sulla produzione dei rapporti temporali e esaminare più a fondo i dati emersi, nel seguito si presenteranno i dati scorporati

per test di elicitazione. L'analisi si limiterà alla valutazione della distribuzione dei valori di durata e al confronto delle mediane.

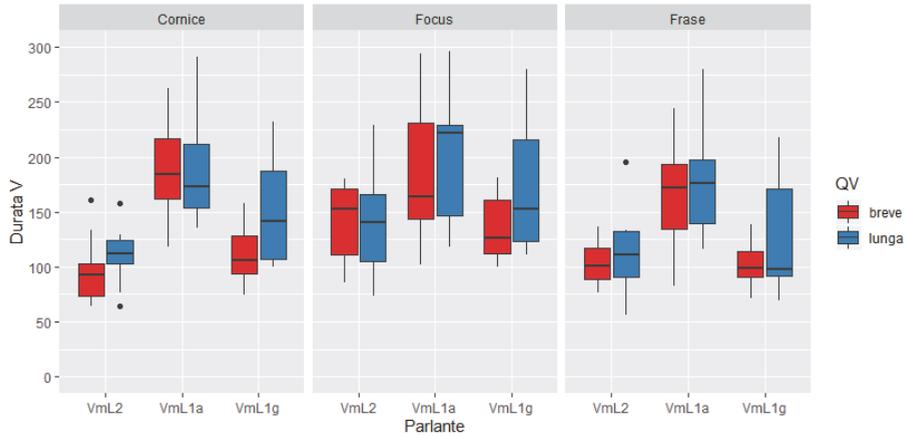
In Figura 3 sono rappresentati i dati sulla vocale tonica nel gruppo genovese, suddivisi per contesto. Tra i parlanti nativi, la distinzione di quantità è sempre realizzata, con una differenza tra vocali lunghe e brevi più netta nei test Cornice e Focus. Come previsto, il tipo di test sembra influire sulla resa dei contrasti, i quali risultano amplificati quando la parola bersaglio è in posizione focale e si riducono in presenza di un contesto ampio e maggiormente naturale. Pur mostrando segni superficiali di avvicinamento a questo schema, il soggetto non nativo GeL2 se ne distacca poi per diversi aspetti: innanzitutto, per una maggiore variabilità dei valori di durata, in particolare delle vocali lunghe; inoltre, lo scarto tra le due categorie appare più ridotto rispetto ai madrelingua nei primi due contesti, Cornice e Focus (sebbene parrebbe non imputabile al caso), e addirittura neutralizzato nel contesto di Frase.

Figura 3 - Distribuzione di durata della vocale tonica nel gruppo genovese per contesto



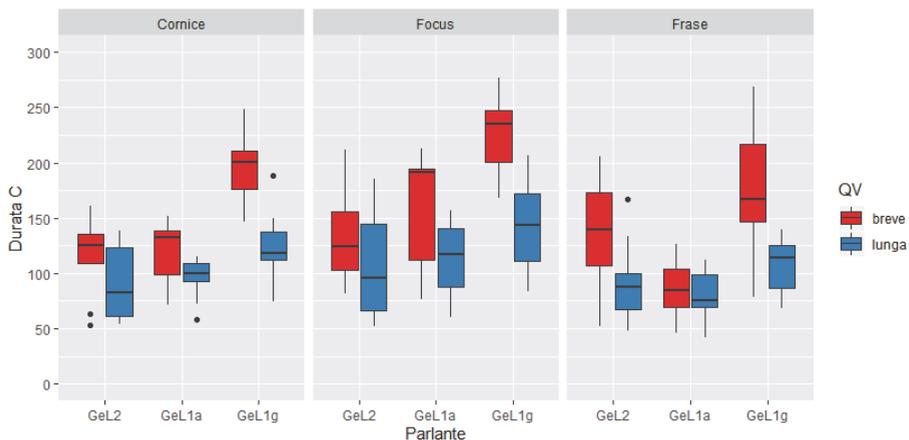
I risultati relativi alla durata vocalica in VM, riportati in Figura 4, evidenziano una situazione di partenza più instabile e complessa, segnalata dalla grande variazione delle durate. Secondo le attese, la parlante nativa anziana VmL1a sembrerebbe non differenziare tra vocali lunghe e brevi in nessun contesto, al contrario del parlante giovane VmL1g per il quale si osservano differenze notevoli tra le mediane in Cornice e Focus ed una maggiore variazione per le V lunghe. All'opposto, VmL2, il parlante non nativo, pare produrre un contrasto di lunghezza nel contesto Cornice, che si annulla negli altri contesti (crucialmente anche nel focus fonologico, in cui, se questo contrasto avesse valenza fonologica, ci aspetteremmo un'amplificazione).

Figura 4 - Distribuzione di durata della vocale tonica nel gruppo intemelio per contesto



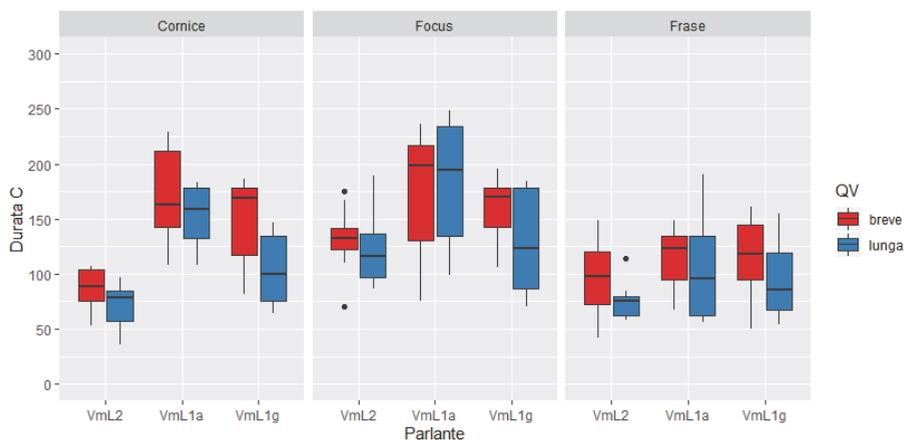
Con riferimento alla lunghezza consonantica, come si evince dai grafici in Figura 5, tendenzialmente sia nei parlanti nativi genovesi sia nel trilingue precoce le consonanti che seguono una vocale breve risultano più lunghe (tranne per GeL1a in Frase, v. infra). Dal confronto tra parlanti nativi giovane ed anziano risalta immediatamente il diverso grado di geminazione della consonante dopo vocale breve: se nel parlante anziano gli scarti di durata sono ridotti (in special modo rispetto a quelli relativi alle vocali), fino ad annullarsi in Frase, nella parlante giovane, invece, assumono dimensioni rimarchevoli, collocandosi sullo stesso piano delle variazioni di durata vocalica. Da questo punto di vista, GeL2 sembra avvicinarsi alle realizzazioni del parlante anziano, limitatamente ai contesti Cornice e Focus; in Frase, il divario aumenta ed emerge una differenza notevole tra consonanti lunghe e brevi.

Figura 5 - Distribuzione dei valori di durata della consonante postonica nel GE, in base al contesto



Il diagramma relativo alla durata consonantica nel gruppo intemelio (in Fig. 6), infine, mostra un quadro più omogeneo, in cui si oscilla dalla completa assenza di differenze di durata significative della parlante nativa anziana a fenomeni di parziale geminazione del parlante nativo giovane e del trilingue precoce.

Figura 6 - Distribuzione dei valori di durata della consonante postonica nel VM, in base al contesto



Riportando i valori assoluti appena esaminati ai rapporti paradigmatici tra segmenti lunghi e brevi, si ottiene il quadro sinottico riprodotto in Tab. 4. Limitandoci alle considerazioni più rilevanti, si osserva come il parlante genovese non nativo mostri da questo punto di vista un comportamento divergente, esibendo un rapporto V:/V di solo 1,1:1, rispetto al rapporto medio dei nativi di 1,4:1. Al contrario, il parlante non nativo intemelio si situa perfettamente al livello di VmL1a, per cui è evidente l'appiattimento dei valori verso un rapporto 1:1.

Tabella 4 - Rapporti paradigmatici, suddivisi per parlante e per contesto

		Cornice		Focus		Frasi		Tutti	
		V:/V	(V)C/ (V):C	V:/V	(V)C/ (V):C	V:/V	(V)C/ (V):C	V:/V	(V)C/ (V):C
GE	GeL1a	1,5	1,2	1,5	1,4	1,4	1,1	1,4	1,3
	GeL1g	1,4	1,6	1,5	1,6	1,3	1,7	1,4	1,6
	GeL2	1,1	1,3	1,2	1,3	1,1	1,5	1,1	1,3
VM	VmL1a	1,1	1,1	1,3	1	1,1	1,1	1,1	1
	VmL1g	1,4	1,5	1,3	1,2	1,2	1,2	1,2	1,3
	VmL2	1,1	1,2	1	1,1	1,1	1,3	1,1	1,1

La Tabella 5, relativa ai rapporti sintagmatici (in genere considerati percettivamente più salienti), appare invece più caotica, ma permette di comprendere meglio le strategie di realizzazione della lunghezza nei parlanti non nativi. GeL2, in parti-

colare, tende a mantenere la vocale tonica sempre molto lunga rispetto alla consonante successiva, anche in condizione fonologica breve (come si evince dai rapporti V/C largamente in favore della vocale più che nei nativi). Sembra inoltre che attui un'iper-correzione dell'opposizione di lunghezza, se si confronta il suo rapporto in Cornice e Frase (di ben 2,4) a quello dei parlanti nativi.

Tabella 5 - *Rapporti sintagmatici, suddivisi per parlante e per contesto*

		Cornice		Focus		Frase		Tutti	
		V:/C	V/C	V:/C	V/C	V:/C	V/C	V:/C	V/C
GE	GeL1a	1,8	1	2,3	1,1	2	1,3	1,8	1
	GeL1g	1,9	0,8	1,8	0,7	2,1	1	1,8	0,8
	GeL2	2,4	1,8	2,3	1,4	2,4	1,6	2	1,3
VM	VmL1a	1,4	1,2	1,5	1,2	2	1,5	1,4	1,2
	VmL1g	1,7	0,8	1,7	0,9	1,6	1	1,4	0,8
	VmL2	1,8	1,3	1,3	1,2	1,6	1,3	1,3	1,1

6. *Discussione e conclusioni*

Cercheremo ora di sintetizzare le linee principali fin qui emerse. Innanzitutto, un primo risultato da segnalare è che, al di là del momento di acquisizione linguistica, l'effetto del test coinvolge in modo evidente tutti i parlanti (cfr. Filipponio, Garassino, 2019; Garassino, Filipponio, 2020). Rispetto al contesto Cornice, nel Focus si nota un generale incremento delle durate e, ove presenti, anche un'enfaticizzazione dei contrasti di lunghezza. Il test Frase provoca al contrario una diminuzione delle differenze tra categorie distintive, fino alla loro scomparsa nel caso di GeL2. Si ottiene, in altri termini, una progressione dei contrasti di durata del tipo Frase < Cornice < Focus, passando cioè dal contesto più "ricco", più naturale e meno marcato, che grazie ad un contesto semantico più ampio favorisce l'interpretazione dei lessemi coinvolti, ai contesti "poveri", che necessitano di compensare le carenze del contesto con l'accentuazione dei tratti fonologici, al fine di facilitare la disambiguazione delle parole bersaglio.

La Tabella 6 rappresenta un tentativo di schematizzazione dei risultati ottenuti.

Tabella 6 - Sintesi dei risultati

		L2	L1 anziani	L1 giovani
GE	QV	+	+	+
	QC	(- in Frase)	(+ allofonica)	(+?)
VM	QV	-	-	(+allofonica?)
	QC	-	-	(+allofonica?)

Nei parlanti genovesi nativi la quantità vocalica distintiva sembra resistere ancora molto bene, nonostante le pressioni dell'italiano (cfr. Cerruti, Crocco & Marzo, 2017) sulla generazione più giovane. La conferma dello statuto fonologico della quantità vocalica appare evidente dal fatto che le opposizioni di lunghezza sono realizzate da tutti i parlanti al di sopra della soglia di significatività e in tutti i contesti, addirittura amplificate nel contesto di fuoco correttivo (Focus), il quale, se interpretato come diagnostica fonologica alla stregua di De Jong (2004; De Jong, Zawaydeh, 2002), dovrebbe fornire indizi sui tratti realmente pertinenti nelle opposizioni fonologiche. Tuttavia, se si considerano gli schemi ritmici della parlante genovese giovane nella loro globalità, la pressione dell'italiano sembra avere conseguenze più serie, che fanno presagire una condizione di transizione: nella parlante GeL1g, infatti, ritroviamo veri e propri contrasti di durata tra le consonanti, il che pare suggerire che l'informatrice abbia perso un controllo pieno sulle consonanti dal punto di vista della gestione temporale. Questo dato può essere ricondotto a due diversi meccanismi, entrambi probabilmente innescati dal pressante e prolungato contatto con l'italiano standard: da un lato, potrebbe darsi che la parlante stia trasferendo il tratto di geminazione dell'italiano nel dialetto, con la conseguenza di pronunciare le consonanti "all'italiana"; dall'altro potremmo essere in presenza di un'interferenza per *over-differentiation* (nell'accezione classica di Weinreich, 1957: 5), per cui la parlante tende ad esasperare le variazioni, qui solo allofoniche, che si riscontrano nel parlante anziano (come d'altronde suggerisce la sovraestensione della geminazione anche a parole che in italiano non contengono una consonante lunga, come [ˈzen:e] 'genere')²⁹. Sebbene allo stato attuale non si disponga evidentemente degli elementi necessari per trarre conclusioni definitive, sembra possibile supporre che la varietà di questa parlante stia attraversando una fase delicata, di instabilità, in quanto, a fronte di una buona tenuta del contrasto di lunghezza vocali-

²⁹ In un certo senso, anche ai parlanti giovani potrebbero essere applicati certi schemi di interferenza proposti dalle teorie dell'acquisizione L2. Sebbene i nostri informatori giovani abbiano dichiarato di aver appreso dialetto e italiano contemporaneamente, emergono segnali che lasciano supporre un'influenza forte dell'italiano sul dialetto e che crucialmente rispecchiano certe aspettative dei modelli teorici sul tema. In effetti, come anticipato al §2.3, non è insolito per soggetti la cui L1 ha un contrasto fonemico di durata, presente in L2 solo a livello allofonico, trasferire questo tratto alla lingua di acquisizione o perlomeno soppesare diversamente i fattori acustici in gioco (v. De Weers, Munro, 2018).

ca, nello spazio di maggiore libertà del sistema linguistico genovese, ovvero quello della variazione di durata consonantica (non essendo questa condizionata da vincoli fonologici), emergono schemi temporali estranei al dialetto (e coerenti con l'italiano, non a caso).

Per quanto concerne l'intemelio, il quadro offerto dalla parlante L1 anziana, le cui produzioni sono sostanzialmente uniformi in termini di durata, corrisponde a quello prospettato in letteratura, mentre risulta più confusa e complessa la situazione del parlante nativo più giovane VmL1g³⁰, il quale, quando si prendano in considerazione tutti i contesti nel complesso, presenta differenze significative di durata sia vocalica che consonantica. Sui suoi risultati potrebbero agire, in misura diversa, più fattori: primariamente, la pressione della lingua standard (e la maggiore esposizione ad essa), che provoca un controllo insufficiente della durata consonantica e l'applicazione degli schemi ritmici dell'italiano anche nella produzione dialettale. Un altro possibile fattore potrebbe essere l'influenza della variazione microdiatopica: uno dei due genitori di questo parlante, infatti, proviene dall'entroterra ligure, in particolare da Pigna (IM), nel cui dialetto è segnalata la persistenza di alcune vocali lunghe (cfr. Merlo, 1938); ciò potrebbe spiegare le differenze di durata della vocale tonica.

Venendo ora al confronto tra i due trilingui precoci e i rispettivi gruppi di controllo, emergono chiaramente due modelli. Il parlante GeL2 realizza i contrasti di durata vocalica del genovese (le differenze di lunghezza vocalica prese nel complesso sono infatti significative), ma il rapporto tra vocali lunghe e brevi non è prodotto in modo del tutto accurato. Questo risultato è coerente con la letteratura sull'argomento (v. §§2.2-2.3), dalla quale si evince come l'opposizione di lunghezza vocalica sia facilmente accessibile, anche da coloro che non hanno tale contrasto nella propria L1 (cfr. tra gli altri Bohn, 1995; McAllister et al, 2002; Escudero et al., 2009; Altmann et al., 2012), ma sul piano della produzione non vi sia sempre piena aderenza alle realizzazioni native (cfr. Mah, Archibald, 2003). Prendendo in esame i singoli contesti, si nota come GeL2 riesca persino a enfatizzare il contrasto nel contesto di Focus, il che testimonia un'acquisita consapevolezza dei tratti pertinenti nel contrasto e suggerisce la formazione (almeno ad uno stadio embrionale) di una categoria fonologica nuova. Tuttavia, sembra non riuscire a padroneggiare pienamente l'organizzazione temporale genovese: in primo luogo, le sue vocali sono sostanzialmente sempre molto lunghe (come suggerito dai rapporti V/C); inoltre quando il contesto favorisce il recupero delle informazioni lessicali, il controllo temporale si allenta e sembrano riemergere gli schemi prosodici della propria L1. Infatti, nel contesto di Frase, vocali etimologicamente brevi e lunghe si sovrappongono e il grado di geminazione delle consonanti postoniche dopo vocale breve aumenta. Anche in questo caso, i nostri risultati concordano con le tendenze già riscontrate in altri studi sull'influenza del contesto sull'accuratezza delle produzioni L2 (D'Apolito, Gili Fivela, 2019).

³⁰ Questo risultato andrebbe nella stessa direzione del recente contributo di Garassino, Cangemi (2020) citato al §3.1.

Il parlante intemelio non nativo, all'opposto, presenta rapporti di durata comparabili a quelli del parlante nativo anziano, mostrando un completo adeguamento agli schemi temporali della varietà locale parlata dagli anziani. Tali risultati potrebbero essere legati a svariate ragioni. È ormai ampiamente noto quanto nell'acquisizione di una lingua seconda "earlier is better", non solo per la dibattuta questione del periodo critico (a partire da Lenneberg, 1967), ma perché l'età d'inizio dell'acquisizione è inscindibilmente connessa ad altri fattori, come la dominanza linguistica, la frequenza d'uso della L1, il tipo di input ricevuto, le dinamiche sociali (Flege, 2007). Nel nostro caso, VmL2 è stato esposto molto precocemente al dialetto intemelio e, a differenza di GeL2, ciò è avvenuto nell'area in cui il dialetto è effettivamente parlato (in un'epoca tra l'altro in cui era molto più vivo in termini di domini d'uso e numero di parlanti). Inoltre, VmL2, sebbene utilizzi tutti i giorni l'italiano, non usa mai il suo dialetto d'origine, diversamente da GeL2, che ne fa un uso frequente³¹. Un'altra ragione potrebbe essere ricondotta ad una questione più astratta, che contempla le caratteristiche temporali della L1 e della L2 e il tipo di adattamento richiesto all'apprendente, come anticipato al §3.2. Nel caso dell'acquisizione dell'intemelio, al parlante meridionale è bastato apprendere la neutralizzazione di un contrasto esistente nella propria L1 (vale a dire, l'opposizione di lunghezza consonantica) o, se vogliamo, l'utilizzo di un solo parametro di quella categoria (assimilabile alle scempie italiane). All'apprendente di genovese GeL2, al contrario, era richiesto non soltanto di imparare a degeminare le consonanti geminate dell'italiano, ma anche la formazione di una nuova categoria, ovvero l'opposizione di lunghezza vocalica. Ancora una volta, il grado di *similarità* tra le categorie fonologiche della lingua nativa e non nativa si conferma la chiave di lettura fondamentale (Best, 1995; Flege, 1987; Escudero, 2009). Dalla nostra analisi, non disponiamo invece di elementi a sufficienza per affermare se l'esistenza di fenomeni allofonici in italiano abbia favorito l'apprendimento del contrasto di durata vocalica genovese. Questa domanda resta aperta per ulteriori approfondimenti.

Siamo naturalmente consapevoli del fatto che il numero limitato di soggetti e di dati, la natura essenzialmente qualitativa del presente studio e l'instabilità stessa delle varietà esaminate non ci consentano di giungere a conclusioni definitive. A tal fine saranno necessarie ulteriori conferme sperimentali, che contemplino più partecipanti e analisi quantitative più affidabili, senza per questo trascurare l'importanza delle sfumature sociolinguistiche. In ogni caso, per quanto possibile entro i limiti di questo contributo, si è cercato di mostrare alcune linee di comportamento in L2 e di suggerire direzioni di ricerca ancora aperte, disponibili ad ulteriori indagini.

³¹ Sul piano puramente uditivo, il parlante GeL2 rispetto a VmL2 conserva anche un marcato "accento meridionale". Sul tema dell'accento straniero, si veda ad esempio Marotta, Boula de Mareüil (2009). A questo riguardo, come suggeriva un revisore anonimo, sarebbe interessante e utile condurre in futuro uno studio di percezione, che proponga a liguri nativi le produzioni degli stessi parlanti usati per questo lavoro e ne valuti il giudizio.

Ringraziamenti

Questo studio è stato possibile grazie al supporto del Fondo Nazionale Svizzero per la Ricerca scientifica nell'ambito del progetto 100015_178932. Intendiamo ringraziare Lorenzo Filipponio e Davide Garassino per la proficua discussione sui temi di questo contributo. Un sentito ringraziamento va ad un revisore in particolare, i cui preziosi suggerimenti hanno permesso di migliorare enormemente il testo. Ovviamente, ogni errore o imprecisione ancora presenti sono da imputare solo all'autrice. Ci teniamo infine a ringraziare gli informatori che hanno preso parte alla ricerca per la loro disponibilità e pazienza.

Riferimenti bibliografici

- ALTMANN, H., BERGER, I. & BRAUN, B. (2012). Asymmetries in the perception of non-native consonantal and vocalic length contrasts. In *Second Language Research*, 28 (4), 387-413.
- AZARETTI, E. (1982) [1977]. *Etimologia dei dialetti liguri attraverso l'evoluzione del ventimigliese*. Sanremo: Casablanca.
- BERTINETTO, P.M. (1981). *Strutture prosodiche dell'italiano: Accento, quantità, sillaba, giuntura, fondamenti metrici*. Firenze: Accademia della Crusca.
- BERTINETTO, P.M., BERTINI, C. (2008). On modeling the rhythm of natural languages. *Proc. of the 47th International Conference on Speech Prosody*, 427-430.
- BERTINETTO, P.M., LOPORCARO, M. (2005). The sound pattern of Standard Italian, as compared with the varieties spoken in Florence, Milan and Rome. In *Journal of the International Phonetic Association*, 35(2), 131-151.
- BEST, C.T. (1995). A direct realist perspective on cross-language speech perception. In STRANGE, W. (Ed.), *Speech Perception and Linguistic Experience: Issues in Cross-language Research*, York Press, Timonium MD, pp. 167-200.
- BEST, C.T., TYLER, M.D. (2007). Nonnative and second-language speech perception: Commonalities and complementarities. In *Language experience in second language speech learning: In honor of James Emil Flege, 1334*, 1-47.
- BLEVINS, J. (2004). *Evolutionary phonology: The emergence of sound patterns*. Cambridge University Press.
- BOERSMA, P., WEENINK, D. (2009). Praat: Doing phonetics by computer (Version 5.1.05). <http://www.praat.org/>
- BOHN, O.-S. 1995. Cross-language speech perception in adults: first language transfer doesn't tell it all. In STRANGE, W. (Ed.), *Speech Perception and Linguistic Experience*. MD: York Press, 279-304.
- BOOMERSHINE, A., HALL, K.C., HUME, E. & JOHNSON, K. (2008). The impact of allophony versus contrast on speech perception. In *Contrast in phonology: Theory, perception, acquisition*, 13, 145-172.
- CASILLAS, J. (2015). Production and perception of the /i/-/I/ vowel contrast: The case of L2-dominant early learners of English. In *Phonetica*, 72(2-3), 182-205.

- CEBRIAN, J. (2006). Experience and the use of non-native duration in L2 vowel categorization. In *Journal of Phonetics*, 34(3), 372–387.
- CELATA, C. (2009). The impact of allophonic variation on L2 speech perception. In WATKINS, M.A., RAUBER, A.S. & BAPTISTA, B.O. (Eds.), *Recent research in second language phonetics/phonology: Perception and production*, 64-80.
- CELATA, C., COSTAMAGNA, L. (2011). Timing delle geminate nell'italiano L2 di apprendenti estoni delle geminate nell'italiano L2 di apprendenti estoni. In *Quaderni del laboratorio di linguistica*, 10(1), 1-27.
- CERRUTI, M.S., CROCCO, C. & MARZO, S. (2017). On the development of a new standard norm in Italian. In CERRUTI, M.S., CROCCO, C. & MARZO, S. (Eds.), *Towards a New Standard. Theoretical and Empirical Studies on the Restandardization of Italian*. Berlin-Boston: De Gruyter, 3-28.
- D'APOLITO, S., GILI FIVELA, B. (2019). L2 Pronunciation Accuracy and Context: A Pilot Study on the Realization of Geminate in Italian as L2 by French Learners. *Proc. Interspeech 2019*, 1706-1710.
- DE CLERCQ, B., SIMON, E. & CROCCO, C. (2014). Rosa versus rossa: The acquisition of Italian geminates by native speakers of Dutch. In *Phrasis: Studies in Language and Literature*, 2(2), 3-29.
- DE JONG, K. (2004). Stress, lexical focus, and segmental focus in English: patterns of variation in vowel duration. In *Journal of Phonetics*, 32(4), 493-516.
- DE JONG, K., ZAWAYDEH, B. (2002). Comparing stress, lexical focus, and segmental focus: patterns of variation in Arabic vowel duration. In *Journal of Phonetics*, 30(1), 53-75.
- DE LEEUW, E., CELATA, C. (2019). Plasticity of native phonetic and phonological domains in the context of bilingualism. In *Journal of Phonetics*, 75, 88-93.
- DE WEERS, N., MUNRO, M. (2018). The role of duration in Japanese speakers' productions of English vowels. *Proc. of the 9th Pronunciation in Second Language Learning and Teaching conference*, University of Utah, Settembre 2017. Ames, IA: Iowa State University, 41-53.
- DIPINO, D., CELATA, C. (2018). An UTI study on Italian alveolar stops. In VIETTI, A., SPREAFICO, L., MEREU, D. & GALATÀ, V. (Eds.), *Models and Methods for the Analysis of Speech Under Real Communicative Conditions*. Studi AISV, 4, Milano: Officinaventuno, 41-53.
- DMITRIEVA, O. (2017). Production of geminate consonants. Implications for typology. In KUBOZONO, H. (Ed.), *The phonetics and phonology of geminate consonants*, 34-65.
- EINFELDT, M., VAN DE WEIJER, J. & KUPISCH, T. (2019). The production of geminates in Italian-dominant bilinguals and heritage speakers of Italian. In *Language, Interaction and Acquisition*, 10(2), 177-203.
- ECKMAN, F.R. (2008). Typological markedness and second language phonology. In EDWARDS, J.G.H., ZAMPINI, M.L. (Eds.), *Phonology and second language acquisition*, 36, 95-115.
- ECKMAN, F.R. (1977). Markedness and the contrastive analysis hypothesis. In *Language learning*, 27(2), 315-330.
- ESCUDERO, P. (2009). Linguistic perception of "similar" L2 sounds. In *Phonology in perception*, 15, 152-190.

- ESCUADERO, P., BENDERS, T. & LIPSKI, S.C. (2009). Native, non-native and L2 perceptual cue weighting for Dutch vowels: The case of Dutch, German, and Spanish listeners. In *Journal of Phonetics*, 37(4), 452-465.
- ESCUADERO, P., BOERSMA, P. (2004). Bridging the gap between L2 speech perception research and phonological theory. In *Studies in Second Language Acquisition*, 551-585.
- ESPOSITO, A., DI BENEDETTO, M.G. (1999). Acoustical and perceptual study of gemination in Italian stops. In *The Journal of the Acoustical Society of America*, 106(4), 2051-2062.
- FANCIULLO, F. (1996). *Fra Oriente e Occidente: Per una storia linguistica dell'Italia meridionale*. Pisa: ETS.
- FILIPPONIO, L. (2012). *La struttura di parola dei dialetti della Valle del Reno. Profilo storico e analisi sperimentale*. Bologna: Arnaldo Forni Editore.
- FILIPPONIO, L., GARASSINO, D. & DIPINO, D. (2019). Between Phonology and Typology. Consonant Duration in Two Gallo-Italian Dialects. In PICCARDI, D., ARDOLINO, F. & CALAMAI, S. (Eds.), *Gli archivi sonori al crocevia tra scienze fonetiche, informatica umanistica e patrimonio digitale*. Milano: Officinaventuno, 269-291.
- FILIPPONIO, L., GARASSINO, D. (2019). Center and Periphery in Phonology: a “stress-test” for two Ligurian Dialects. In *Italian Journal of Linguistics* 31(2), 141-168.
- FLEGE, J.E. (1987). The production of “new” and “similar” phones in a foreign language: evidence for the effect of equivalence classification. In *Journal of Phonetics*, 15, 47-65.
- FLEGE, J.E. (1995). Second language Speech Learning: Findings, and problems. In STORAGE, W. (Ed.), *Speech Perception and Linguistic Experience: Issues in Cross-Language Research*. Timonium, MD: York Press, 233-277.
- FLEGE, J.E. (2007). Language contact in bilingualism: Phonetic system interactions. In *Laboratory phonology*, 9, 353-382.
- FORNER, W. (1975). *Generative Phonologie des Dialekts von Genua*. Buske.
- FORNER, W. (1988). Areallinguistik I. Ligurien / Aree linguistiche I. Liguria. In *Lexikon der romanistischen Linguistik* 4, 453-469.
- GARASSINO, D., CANGEMI, F. (2020). “No duration without intonation”. The interplay of lexical and post-lexical durational differences. *Proc. 10th International Conference on Speech Prosody*, Tokyo, Japan.
- GARASSINO, D., DIPINO, D. (2019). Vowel length in Intemelian Ligurian. An experimental and cross-dialectal investigation. *Proc. 19th International Congress of Phonetic Sciences*, Melbourne, Australia, 122-126.
- GARASSINO, D., FILIPPONIO, L. (2020). The impact of information and prosodic structure on the phonetic implementation of vowel length in Ligurian. In TEIXEIRA KALKHOFF, A., SELIG, M. & MOOSHAMMER, C. (Eds.), *Prosody and conceptional variation. Situational conditions of communication, degree of communicational planning, and activity types as parameters for prosodic design*. Bern: Peter Lang.
- GIANNINI S., COSTAMAGNA L. (1997). Language learning strategies in interlanguage phonology of Italian L2. A case history: the acquisition of consonant length. In *New sounds 97. Proc. of the 3rd International Symposium on the acquisition of second language speech*, University of Klagenfurt, Klagenfurt, 96-103.

- HAN, M.S. 1992. The timing control of geminate and single stop consonants in Japanese: A challenge for nonnative speakers. In *Phonetica*, 49, 102-127.
- HARADA, T. 2006. The acquisition of single and geminate stops by English speaking children in a Japanese immersion program. In *Studies in Second Language Acquisition*, 28, 601-632.
- HAYES-HARB, R., MASUDA, K. (2008). Development of the ability to lexically encode novel second language phonemic contrasts. In *Second Language Research*, 24 (1), 5-33.
- HILLENBRAND, J.M., CLARK, M.J. & HOUDE, R.A. (2000). Some effects of duration on vowel recognition. In *Journal of the Acoustical Society of America*, 108(6), 3013-3022.
- HIRATA, Y. (2017). Second language learners' production of geminate consonants in Japanese. In KUBOZONO, H. (Ed.), *The Phonetics and Phonology of Geminate Consonants*. Oxford University Press, 163-183.
- KABAK, B., RECKZIEGEL, T. & BRAUN, B. (2011). Timing of second language geminates and singletons. *Proc. of the 17th International Congress of the Phonetic Sciences*, 994-97.
- KAWAHARA, S., PANGILINAN, M. (2017). Spectral continuity, amplitude changes, and perception of length contrasts. In KUBOZONO, H. (Ed.), *The phonetics and phonology of geminate consonants*, 13-33.
- KHATTAB, G., AL-TAMIMI, J. (2015). The acquisition of gemination in Lebanese-Arabic children. *Proc. of the 18th International Congress of Phonetic Sciences*. Newcastle University.
- KUBOZONO, H. (Ed.). (2017). *The phonetics and phonology of geminate consonants*. Oxford University Press.
- LANDI, R., SAVY, R. (1996). Durata vocalica, struttura sillabica e velocità d'eloquio nel parlato connesso. In PERETTI A., SIMONETTI P. (Eds.), *Atti del XXIV Convegno Nazionale dell'Associazione Italiana di Acustica (AIA)*, Trento, 65-70.
- LEE, A., MOK, P. (2018). Acquisition of Japanese quantity contrasts by L1 Cantonese speakers. In *Second Language Research*, 34(4), 419-448.
- LEHISTE, I. (1970). *Suprasegmentals*. Cambridge: MIT Press.
- LENNEBERG, E.H. (1967). The biological foundations of language. In *Hospital Practice*, 2(12), 59-67.
- LOPORCARO, M. (2013). *Profilo linguistico dei dialetti italiani* (Vol. 275). Laterza.
- LOPORCARO, M. (2015). *Vowel length from Latin to Romance*. OUP Oxford.
- MACHAČ, P., SKARNITZL, R. (2009). *Principles of phonetic segmentation*. Albatros Media as.
- MADDIESON, J. (1984). *Patterns of sound*. Cambridge: Cambridge University Press.
- MAH, J., ARCHIBALD, J. (2003). Acquisition of L2 length contrasts. *Proc. 6th Generative Approaches to Second Language Acquisition Conference*, 208-212.
- MAIRANO, P., BASSETTI, B., SOKOLOVIĆ-PEROVIĆ, M. & CERNI, T. (2018). Effects of L1 orthography and L1 phonology on L2 English pronunciation. *Revue française de linguistique appliquée*, 23(1), 45-57.
- MAJOR, R.C. (2001). *Foreign Accent: The Ontogeny and Phylogeny of Second Language Phonology*. Routledge: New York-London.
- MAROTTA, G. (1985). *Modelli e misure ritmiche: la durata vocalica in italiano*. Bologna: Zanichelli.

- MAROTTA, G., BARTH, M. (2008). Sonoranti allo specchio: italiano L2-tedesco L1. In COSTAMAGNA, L., MAROTTA, G. (Eds.), *Processi fonetici e categorie fonologiche nell'acquisizione dell'italiano*, 93-120.
- MAROTTA, G., BOULA DE MAREÜIL, P. (2009). Persistenza dell'accento straniero. Uno studio percettivo sull'italiano L2. *Atti del V Convegno Internazionale AISV, La dimensione temporale del parlato*. Universität Zürich, 475-494.
- MCALLISTER, R., FLEGE, J.E. & PISKE, T. (2002). The influence of L1 on the acquisition of Swedish quantity by native speakers of Spanish, English and Estonian. In *Journal of Phonetics*, 30(2), 229-258.
- MEISTER, L., MEISTER, E. (2011). Perception of the short vs. long phonological category in Estonian by native and non-native listeners. In *Journal of Phonetics*, 39(2), 212-224.
- MERLO, C. (1938). Contributo alla conoscenza dei dialetti della Liguria odierna. I. Degli esiti di *r* (primario o secondario da *L*) e di *n* intervocalico nel dialetto di Pigna (Imperia). In *L'Italia Dialettale*, 14, 23-58.
- MUNRO, M.J. (1993). Productions of English vowels by native speakers of Arabic: Acoustic measurements and accentedness ratings. In *Language and Speech*, 36(1), 39-66.
- NENONEN, S., SHESTAKOVA, A., HUOTILAINEN, M. & NÄÄTÄNEN, R. (2005). Speech-sound duration processing in a second language is specific to phonetic categories. In *Brain and language*, 92(1), 26-32.
- NIMZ, K. (2011). Vowel Perception and Production of Late Turkish Learners of L2 German. *Proc. of the 17th International Congress of Phonetic Sciences*, 1494-1497.
- OHALA, J.J. (1983). The origin of sound patterns in vocal tract constraints. In MACNEILAGE, P.F. (Ed.), *The Production of Speech*. Heidelberg and Berlin: Springer Verlag, 189-216.
- OHALA, J.J. (2011). Accommodation to the aerodynamic voicing constraint and its phonological relevance. *Proc. of the 17th International Congress of Phonetic Sciences*, 64-67.
- PAJAK, B., LEVY, R. (2014). The role of abstraction in non-native speech perception. In *Journal of Phonetics*, 46, 147-160.
- PAYNE, E., POST, B., GARMANN, N.G. & SIMONSEN, H.G. (2017). The acquisition of long consonants in Norwegian. In KUBOZONO, H. (Ed.), *The Phonetics and Phonology of Geminate Consonants*, 130-162.
- ROHLFS, G. (1966). *Grammatica storica della lingua italiana e dei suoi dialetti. Fonetica*. Torino: Einaudi.
- SCHOUTEN, B., GERRITS, E. & VAN HESSEN, A. (2003). The end of categorical perception as we know it. In *Speech communication*, 41(1), 71-80.
- SORIANELLO, P. (2009). L'acquisizione del tratto di lunghezza consonantica in italiano L2. In ROMITO, L., GALATÀ, V. & LIO, R. (Eds.), *La fonetica sperimentale. Metodo e applicazioni. Atti del IV Convegno Nazionale AISV, Torriana, EDK*, 40-61.
- SORIANELLO, P. (2012). *Linguaggio e sindrome di Down*. FrancoAngeli.
- SORIANELLO P. (2014). Italian geminate consonants in L2 acquisition. In COSTAMAGNA L., CELATA C. (Eds.), *Consonant gemination in first and second language acquisition*, Pacini, Pisa, 25-46.
- SORIANELLO, P. (2019). Le consonanti affricate: strategie di acquisizione in italiano L2. In *Italiano LinguaDue*, 1, 68-88.

- STAMULI, M.F. (2008). Morte di lingua e variazione lessicale nel greco di Calabria: Tre profili dalla Bovesia. Tesi di Dottorato di ricerca in Filologia moderna. Università degli Studi "Federico II, Napoli.
- TOSO, F. (1997). *Grammatica del genovese: varietà urbana e di koinè*. Le Mani.
- TSUKADA, K. (2009). Durational characteristics of English vowels produced by Japanese and Thai second language (L2) learners. In *Australian Journal of Linguistics*, 29(2), 287-299.
- TSUKADA, K., COX, F., HAJEK, J. & HIRATA, Y. (2018). Non-native Japanese learners' perception of consonant length in Japanese and Italian. In *Second Language Research*, 34(2), 179-200.
- TSUKADA, K. HAJEK, J. (2019). Cross-language perception of Italian and Japanese consonant length contrasts: a comparison of native Italian listeners with and without Japanese language learning experience. *Proc. of the 19th International Congress of Phonetic Sciences*, Melbourne, Australia, 82-86.
- TSUKADA, K., HAJEK, J. (2020). Perception of consonant length in familiar and unfamiliar languages by native speakers of Mandarin, Italian and Japanese. *Proc. 10th International Conference on Speech Prosody*, 66-70.
- TSUKADA, K., HIRATA, Y. & ROENGPITYA, R. (2014). Cross-language perception of Japanese vowel length contrasts: Comparison of listeners from different first language backgrounds. In *Journal of Speech, Language, and Hearing Research*, 57(3), 805-814.
- VAYRA, M., AVESANI, C., BEST, C.T. & BOHN, O.S. (2012). Non solo dettagli fonetici, non solo categorie fonologiche: l'interazione tra fonetica e fonologia nella percezione di suoni non-nativi. In *Studi e Saggi Linguistici*, 119-146.
- YLINEN, S., SHESTAKOVA, A., ALKU, P. & HUOTILAINEN, M. (2005). The perception of phonological quantity based on durational cues by native speakers, second-language users and nonspeakers of Finnish. In *Language and Speech*, 48(3), 313-338.
- WEINREICH, U. (1957). On the description of phonic interference. In *Word*, 13(1), 1-11.
- WHANG, J., YAZAWA, K. & ESCUDERO, P. (2019). Perception of Japanese vowel length by Australian English listeners. *Proc. of the 19th International Congress of Phonetic Sciences*, 2228-2232.
- WILLIAMS, D., ESCUDERO, P. (2014). A cross-dialectal acoustic comparison of vowels in Northern and Southern British English. In *The Journal of the acoustical society of America*, 136(5), 2751-2761.

PARTE III

LA VARIAZIONE
INTERLINGUISTICA

STEPHAN SCHMID, KRISTIJAN RAJIC

La pronuncia delle vocali nasali nel francese L2 di apprendenti italofoeni: acquisizione e insegnamento esplicito

The pronunciation of nasal vowels in the L2 French of Italian-speaking learners: acquisition and explicit teaching

In this contribution we investigate the acquisition of French nasal vowels by Italian-speaking learners on the one hand and the impact of explicit pronunciation training on the other hand. In a classroom experiment, 10 students received detailed instructions about French nasal vowels and were recorded twice in a reading task (before and after the pronunciation lesson). A control group of 10 students without explicit phonetic training was recorded twice as well. The inspection of a few less than 1680 tokens revealed seven different solutions to cope with the remarkable difficulty of French nasal vowels for Italoophone learners. However, the group who had received explicit pronunciation training performed much better in the second recording, whereas the control group did not improve substantially (also showing greater interindividual variability).

Keywords: nasal vowels, French as a second language, pronunciation teaching.

1. *Introduzione*

Negli ultimi decenni sono stati condotti numerosi studi sperimentali su fenomeni fonetici in diverse situazioni di contatto, dove l'italiano appare sia come L1 sia come L2; v. ad esempio vari contributi apparsi negli Atti dei Convegni AISV, in particolare nei volumi *Studi AISV 1* (Vayra, Avesani & Tamburini, 2015) e *Studi AISV 2* (Savy, Alfano, 2016). Tra gli studi fonetici sull'acquisizione delle lingue straniere possiamo citare, a titolo esemplificativo, alcune ricerche sulle vocali dell'inglese (Busà, 1995; Sisinni, Grimaldi, 2010) e sull'intonazione del tedesco (Stella, 2013). Non mancano lavori sul francese L2 di italofoeni, ad esempio sulla realizzazione di gruppi consonantici (D'Apolito, Gili Fivela, 2009) o sulla pronuncia delle rotiche (De Paolis, 2019). Invece, a quanto ci risulta, non esistono ricerche sulla realizzazione delle vocali nasali nel francese L2 di italofoeni.

Il presente contributo costituisce innanzitutto un primo tentativo di colmare tale lacuna. In secondo luogo intende anche verificare l'efficacia dell'insegnamento esplicito della pronuncia in classe, seguendo la metodologia adottata in uno studio precedente (Schmid, Pedrazzini, 2016). A questo scopo è stato condotto un esperimento didattico con una classe di apprendenti italofoeni del francese.

Il contributo è articolato come segue. Nel §2 si abbozza un'analisi contrastiva dei due sistemi vocalici dell'italiano e del francese al fine di formulare delle ipotesi sulla difficoltà di pronuncia delle vocali nasali del francese, tenendo conto anche delle predizioni di alcuni modelli teorici sulla pronuncia in L2; inoltre, si discute brevemente dell'efficacia o meno di un insegnamento esplicito della pronuncia per l'acquisizione di una lingua seconda. Nel §3 si descrivono il campione, il corpus nonché il metodo adottato per la raccolta e l'elaborazione dei dati, mentre nel §4 si presenta l'esperimento didattico condotto in classe. Nel §5 si espone un'analisi qualitativa delle realizzazioni delle vocali che sfocia in una tipologia degli errori prodotti dagli allievi, dopodiché si presentano nel §6 i risultati di questa classificazione degli errori da un punto di vista quantitativo. Seguono, nel §7, alcune osservazioni conclusive.

2. L'acquisizione delle vocali nasali francesi

2.1 Analisi contrastiva tra l'italiano (L1) e il francese (L2)

Rispetto alle sette vocali dell'italiano, la lingua francese possiede un sistema molto più complesso composto da ben sedici fonemi vocalici. Come si evince dalla Tab. 1, il francese condivide con l'italiano standard le sette vocali /i e ε a ɔ o u/, a cui si aggiungono però anche lo schwa, tre vocali anteriori arrotondate orali /y ø œ/ e quattro vocali nasali /ã ɔ̃ ẽ œ̃/ (Maturi, 2014: 130-131; Mioni, 1973: 98-100):

Tabella 1 - *Il sistema vocalico del francese*

i	y		u
e	ø	ə	o
ε ẽ	œ œ̃		ɔ ɔ̃
a			ɑ ã

Le quattro vocali nasali del francese, evidenziate in grassetto nella Tab. 1, possono essere illustrate tramite le seguenti parole:

- (1) a. *lent* [lã] 'lento', *manger* [mãʒe] 'mangiare'
 b. *pont* [pɔ̃] 'ponte', *vont* [vɔ̃] 'essi vanno'
 c. *peindre* [pɛ̃dʁ] 'dipingere', *vin* [vɛ̃] 'vino'
 d. *brun* [brœ̃] 'marrone', *parfum* [parfœ̃] 'profumo'

Di queste quattro vocali, nel francese 'tendenziale' il fonema /œ̃/ viene in genere sostituito da /ɛ̃/ (Fougeron, Smith, 1999: 78-79). La rappresentazione grafematica delle vocali nasali consiste perlopiù di sequenze <VN>, come nella parola *bon* [bɔ̃] 'buono'; tuttavia, non per tutte le vocali sussiste una relazione biunivoca tra fonemi e grafemi, come viene illustrato dalla coppia di omofoni *temps* [tã] 'tempo' e *tant* [tã] 'tanto'.

Il sistema di partenza degli apprendenti analizzati in questa sede consiste nell'italiano regionale ticinese, il cui vocalismo coincide in larga misura con quello

dell'italiano regionale lombardo di tipo milanese. Questo sistema contiene le sette vocali [i e ε a o u] che possiedono però uno statuto fonologico leggermente diverso rispetto al vocalismo dell'italiano standard (cfr., ad esempio, Canepari, 1980: 9697; Schmid, 1999: 146; Bertinetto, Loporcaro, 2005: 137; Maturi, 2014: 104-105). Nello specifico, la distribuzione delle vocali medie anteriori è determinata dal contesto fonotattico: la variante medio-alta ricorre in sillaba tonica aperta non finale e davanti a consonante nasale tautosillabica (per cui si pronuncia [ˈbɛ:ne] e [ˈtempo]), mentre in sillaba chiusa e in sillaba aperta finale si ha generalmente la variante medio-bassa (quindi [ˈdɛ:tɔ] e [pɛrˈkɛ]). Invece, la distribuzione delle vocali medie posteriori non è prevedibile in base al contesto fonotattico, per cui si ha sia [ˈso:lɛ] e [ˈkɔ:zɛ] che [ˈpol:ɔ] e [ˈpɔstɛ]. Per quanto riguarda la nasalizzazione delle vocali davanti a consonante nasale tautosillabica, essa viene spesso attribuita non soltanto all'italiano regionale veneto, ma anche ad altre varietà settentrionali tra cui quella lombarda (Canepari, 1979: 207; Telmon, 1993: 105); in base alla nostra esperienza, tale fenomeno è poco saliente nell'italiano regionale ticinese, dove una lieve nasalizzazione appare caso mai come tratto individuale oppure come processo di coarticolazione nel parlato ipoarticolato.

2.2 Predizioni in base ad alcuni modelli teorici

Varie considerazioni portano a ipotizzare che le vocali nasali del francese creino delle difficoltà agli apprendenti italofofoni. Le vocali nasali possono essere considerate dei suoni difficili innanzitutto a causa della loro assenza nell'inventario fonemico dell'italiano: questa predizione scaturisce dalla classica 'ipotesi dell'analisi contrastiva' (*Contrastive Analysis Hypothesis* CAH, Lado, 1957). A ciò si aggiunge il fatto che le vocali nasali risultano essere marcate dal punto di vista tipologico, dato che esse non sono molto diffuse negli inventari fonemati delle lingue del mondo; già nella prima versione della banca dati UPSID solo 71 su 317 lingue (= 22,4%) contenevano vocali nasali (Maddieson, 1984: 130). La cosiddetta 'ipotesi del differenziale di marcatezza' (*Markedness Differential Hypothesis* MDH, Eckman, 2008) fornisce quindi una seconda predizione circa la difficoltà delle vocali nasali francesi per parlanti che non le possiedono già nella loro L1. In effetti, le quattro vocali nasali del francese sono piuttosto rare anche nella più recente banca dati PHOIBLE che comprende più di 3000 inventari fonemati (Moran, McCloy, 2019) (v. Tab. 2); nello specifico, i fonemi /ɔ̃/ e /ɛ̃/ appaiono con una frequenza lievemente superiore rispetto alla vocale /ã/, mentre la quarta vocale nasale /œ̃/ risulta essere rarissima.

Tabella 2 - *Frequenza delle vocali nasali nella banca dati PHOIBLE*

ɔ̃	8%
ɛ̃	8%
ã	2%
œ̃	<1%

Guardando invece la *token frequency* delle vocali nasali all'interno di un corpus fonologico della lingua francese (Delattre, 1965: 62; cfr. anche Mioni, 1973), si nota che la vocale nasale più frequente risulta essere [ã], seguita in ordine decrescente dalle vocali [õ], [ẽ] e [œ] (v. Tab. 3).

Tabella 3 - *Token frequency delle vocali nasali in un corpus francese (Delattre, 1965)*

ã	8%
õ	4%
ẽ	2%
œ	<1%

Viene comunque confermato il carattere doppiamente marcato della vocale /œ/ (già evidente dal suo statuto marginale nel francese odierno; cfr. *supra*); dal punto di vista della tipologia fonologica, la marcatezza di questa vocale deriva non soltanto dalla nasalità, ma anche dalla combinazione dei due tratti [+arrotondato] e [-posteriore]. In sostanza, a causa della scarsa frequenza sia interlinguistica che intralinguistica delle vocali nasali, gli apprendenti incontrano nell'input disponibile poche occasioni che permettano loro di rendersi conto del carattere specifico di questa categoria di suoni.

I due modelli teorici più influenti nella ricerca sulla pronuncia in L2 mettono in evidenza il ruolo della percezione. Secondo una delle principali predizioni dello *Speech Learning Model* (SLM, Flege, 1995) sarebbero soprattutto le differenze tra suoni simili tra la L1 e la L2 che rischiano di non essere percepite dagli apprendenti (v. ora anche la versione riveduta del modello SLM-r: Flege, Bohn, in stampa). Anche secondo il *Perceptual Assimilation Model* (nella sua versione PAM-L2), gli apprendenti tenderebbero ad 'assimilare' percettivamente i suoni nuovi della L2 ai suoni articolatoriamente più simili della loro L1 (v. Best, Tyler, 2007; cfr. anche Frontera, Romito, 2015). Non avendo noi stessi condotto un esperimento percettivo, non disponiamo di elementi empirici per poter formulare un'ipotesi del tipo 'le vocali nasali del francese L2 vengono assimilate alle vocali orali articolatoriamente più vicine dell'italiano L1'.¹

In base alla situazione di apprendimento indagata in questa sede ci sembra però opportuno insistere sulla possibile interferenza della modalità scritta sulla pronuncia di una lingua straniera studiata in ambito scolastico (cfr. Mairano et al., 2018). Se si considera la summenzionata rappresentazione grafematica delle vocali nasali come sequenze <VN> nonché l'evidente somiglianza grafica di moltissime coppie di lessemi dell'italiano e del francese (v. ad esempio it. *ponte* vs. fr. *pont* [põ], it. *vino* vs. fr. *vin* [ẽ], ecc.), possiamo ipotizzare che il peso dell'input scritto nell'insegna-

¹ Com'è noto, la percezione delle vocali nasali costituisce un fenomeno particolarmente complesso. Per la problematica in generale rimandiamo al contributo fondamentale di Beddor (1993); la percezione delle vocali nasali francesi è stata esaminata con parlanti nativi da Delvaux (2009) e da Detey, Racine (2015) con apprendenti giapponesi.

mento scolastico nonché la parentela linguistica tra L1 e L2 portino molti apprendenti italofoeni ad ‘assimilare’ percettivamente le vocali nasali del francese alle sequenze di vocali orali e consonanti nasali dell’italiano. Non a caso questo fenomeno viene spesso indicato come uno dei tratti più salienti nella pronuncia del francese L2 da parte di italofoeni (v. ad esempio, Maturi, 2014: 151; Murano, Paternostro, 2016: 152).

2.3 L’insegnamento esplicito della pronuncia

Nella didattica delle lingue straniere, la pronuncia sembra oggi ricevere meno attenzione rispetto alle altre componenti strutturali del sistema linguistico, il che può essere in parte l’effetto della cosiddetta ‘svolta comunicativa’ che ha messo l’accento su competenze di tipo pragmatico, com’è evidente negli obiettivi didattici formulati nel ‘Quadro comune europeo di riferimento per la conoscenza delle lingue’ (QCER, v. Consiglio d’Europa 2010). D’altra parte non si può nemmeno escludere che la scarsa attenzione rivolta agli aspetti fonetici delle lingue seconde derivi da una sorta di scetticismo circa le reali possibilità di migliorare la pronuncia degli apprendenti. La tesi che la fonologia sia più difficile da padroneggiare in una L2 rispetto agli altri livelli di analisi è stata a lungo riprodotta anche nella letteratura scientifica: basti citare a questo proposito il cosiddetto *Joseph Conrad Phenomenon* che prende il nome da un famoso scrittore polacco il quale, pur avendo scritto capolavori della letteratura inglese, conservava per tutta la sua vita un fortissimo accento straniero in questa lingua (Tarone, 1987: 80). Da un punto di vista più generale vanno ricordate anche le fondamentali obiezioni teoriche all’esistenza di un’interfaccia tra la competenza linguistica e la competenza metalinguistica (v. in questo senso soprattutto Krashen, 1987).

Tuttavia, in anni più recenti la questione dell’efficacia dell’insegnamento esplicito della pronuncia in una lingua straniera è stata affrontata anche dal punto di vista della ricerca empirica, come viene testimoniato da una rassegna di 75 studi, di cui ben l’82% riporta un miglioramento in seguito a un’istruzione fonetica (Thomson, Derwey, 2015: 338-339). Nel 61% degli studi presi in esame veniva adottato un tipo di insegnamento in classe che si svolgeva prevalentemente secondo il cosiddetto ‘approccio PPP’ (Presentazione – Pratica – Produzione; cfr. Thomson, Derwey, 2015: 330); notiamo, tra parentesi, che tale approccio assomiglia per certi versi al tradizionale ‘metodo fonetico’ propagato da Canepari (1979: 7, 9).

Il metodo PPP è stato anche adottato in una ricerca sulla pronuncia delle occlusive nel tedesco L2 da parte di apprendenti italofoeni nella Svizzera italiana (Schmid, Pedrazzini, 2016). In quel lavoro, l’insegnamento esplicito delle occlusive sorde aspirate sembrava produrre un effetto piuttosto scarso sulla durata del VOT, almeno quando si considerava la media dei 10 soggetti analizzati. Tuttavia, va notato anche che ben 4 allievi su 10 sono stati in grado di migliorare notevolmente la propria pronuncia: questo risultato suggerisce che il metodo PPP sembra attecchire presso apprendenti con un determinato ‘stile cognitivo’, suscettibili cioè di costruire una

specie di ‘interfaccia’ tra il sapere metalinguistico e la competenza fonologica in L2 (cfr. Sharwood Smith 1981, 1991).

Il presente contributo adotta la stessa metodologia usata nello studio precedente (Schmid, Pedrazzini, 2016), riprendendone la modalità di presentazione nella ‘lezione di fonetica’ nonché il metodo della raccolta ed elaborazione dei dati. Anche in questo caso gli apprendenti sono studenti italo-foni di una scuola media superiore della Svizzera italiana, solo che questa volta la lingua obiettivo non è la prima lingua nazionale (il tedesco), bensì la seconda (il francese); inoltre, il fenomeno segmentale analizzato non riguarda il consonantismo (il VOT delle occlusive), bensì la realizzazione delle vocali nasali del francese.

2.4 Domande di ricerca

Riassumendo quanto esposto sinora ribadiamo che il nostro contributo pone due domande di ricerca generali:

- a. Quali difficoltà pongono le vocali nasali del francese a un apprendente italo-fono?
- b. È possibile insegnare la pronuncia delle vocali nasali francesi a studenti di una scuola media superiore di lingua madre italiana?

Più specificatamente ci si chiederà come vengano realizzate le quattro vocali nasali francesi da parte degli apprendenti italo-foni del campione qui analizzato (DR1), quale sia l’impatto dell’insegnamento esplicito sulla pronuncia (DR2) e infine quale sia il grado di variabilità interindividuale tra gli apprendenti (DR3).

3. *L’indagine empirica: raccolta dei dati e procedura di analisi*

3.1 Il campione

I soggetti analizzati provengono da due classi di una scuola commerciale della Svizzera italiana. Si tratta più precisamente di 20 allieve di due terze classi che al momento della registrazione avevano un’età compresa tra i 17 e i 19 anni. Le allieve hanno iniziato a studiare il francese a partire della terza elementare e hanno avuto un insegnamento (non continuo) per sette anni.

3.2 Il corpus

Per allestire il corpus di parlato letto in francese L2 è stato preparato un questionario composto da 35 frasi di senso compiuto che contenevano almeno una parola con una vocale nasale (v. Tab. 4); in due frasi appaiono 2 vocali nasali e altre due frasi contengono 3 vocali nasali, per cui il totale delle parole da analizzare ammonta a 42. Entrambe le classi sono state registrate due volte a distanza di una settimana, con la differenza che nella prima classe l’insegnamento fonetico ha avuto luogo tra la prima e la seconda registrazione (in modo da poter verificare un eventuale miglioramento della pronuncia in base alle indicazioni fornite), mentre nella seconda classe l’istruzione esplicita è stata impartita soltanto dopo la seconda registrazione (la seconda classe funge quindi da gruppo di controllo). Di ogni classe si sono analizzate

le produzioni di 10 soggetti di sesso femminile; in totale il corpus ammonterebbe a 1680 repliche (2 registrazioni x 2 classi x 10 alunni x 42 parole), ma durante le registrazioni ci sono state numerose defezioni per cui alla fine abbiamo potuto analizzare soltanto 1361 occorrenze.

Tabella 4 - *Esempi di alcune parole (fonemi e grafemi)*

[ã]	[ɔ]	[ɛ]	[œ]
<an>	<on>	<in>	<un/um>
<i>langue</i>	<i>non</i>	<i>cinq</i>	<i>un</i>
<i>tranche</i>	<i>son</i>	<i>vingt</i>	<i>lundi</i>
<i>francs</i>	<i>pont</i>	<i>vin</i>	<i>brun</i>
<i>dimanche</i>	<i>long</i>	<i>fin</i>	<i>quelqu'un</i>
<i>manger</i>	<i>leçon</i>	<i>singe</i>	<i>parfum</i>
<en>		<ain/aim>	
<i>penser</i>		<i>pain</i>	
<i>vendredi</i>		<i>main</i>	
<i>descendre</i>		<i>train</i>	
<i>vent</i>		<i>saint(e)</i>	
<i>cent</i>		<i>faim</i>	

3.3 Raccolta ed elaborazione dei dati

Le registrazioni sono state fatte mediante il programma *Speech Recorder* (Draxler, Jansch, 2004) direttamente su un computer portatile al quale era collegato tramite un'interfaccia audio un microfono, fissato a sua volta mediante un collarino a una distanza di ca. 5 cm dalle labbra delle locutrici. La segmentazione e l'annotazione delle registrazioni sono state effettuate in *Praat* (versione 6.1.06, Boersma, Weenink, 2019).

L'analisi uditiva delle realizzazioni delle allieve è stata compiuta con l'ausilio di un *TextGrid* in cui sono stati annotati cinque *Tiers*: parola, grafia, fonema, realizzazione fonetica, categoria della realizzazione (cfr. §5).

4. La lezione di fonetica

In una delle due classi, tra la prima e la seconda registrazione si è impartita una lezione di fonetica della durata di circa 80 minuti. Le allieve hanno avuto l'opportunità di acquisire dapprima alcune conoscenze di base sulle vocali orali delle lingue romanze, con un focus particolare sul sistema vocalico francese. In seguito le allieve hanno imparato sia dal punto di vista anatomico che fisiologico la definizione di nasalità, rendendosi conto della posizione del velo nel tratto vocale e in particolare della condizione in cui il velo è abbassato. Per le tutte le vocali (orali e nasali) si è indicata la posizione nel trapezio vocalico, specificando da un lato il mantenimento della posizione articolatoria rispetto alle vocali orali e dall'altro lato la differen-

za nella trascrizione fonetica che risiede nel simbolo della *tilde* (~) sovrapposto ai quattro simboli fonetici in questione.

Si è poi proceduto con la presentazione delle quattro vocali nasali francesi, ognuna illustrata con lessemi noti alle allieve (cfr. Tab. 5).

Tabella 5 - *Le quattro vocali nasali, esempi di lessemi e i loro grafemi corrispondenti*

Simbolo	Esempi di lessemi	Grafemi corrispondenti
[ã]	<i>m<u>a</u>nger, pen<u>s</u>er, se<u>m</u>bler</i>	<an>, <en>,
[õ]	<i>po<u>n</u>t, fo<u>n</u>t, to<u>m</u>ber</i>	<on>, <om>
[ẽ]	<i>che<u>m</u>in, pa<u>i</u>n, fa<u>i</u>m, rien</i>	<in>, <ain>, <aim>, <ien>
[œ]	<i>lu<u>n</u>di, u<u>n</u>², bru<u>n</u>, par<u>f</u>um</i>	<un>, <um>

Si è accennato anche al fatto summenzionato che oggi giorno il fonema /œ/ viene sostituito da molti parlanti nativi del francese con /ẽ/. Le allieve hanno esercitato diverse volte la pronuncia di una serie di parole che contengono le quattro vocali nasali.

Successivamente si è attirata l'attenzione su tre errori tipicamente commessi da parlanti italofofoni (L1) nella pronuncia delle vocali nasali. In primo luogo si è cercato di sensibilizzare le allieve a distinguere [ã] da [õ], ricordando che [ã] corrisponde alle combinazioni grafematiche <an>, <en> e , [õ] invece ai digrafi <on> e <om>. A titolo esemplificativo si è portato l'esempio della coppia minima di *dans* [dã] vs. *dont* [dõ] (sull'utilità del ricorso a coppie minime nell'insegnamento della pronuncia in una lingua straniera v. Mairano, Calabrò, 2016). Un secondo errore tipico commesso da parlanti italofofoni di L1 consiste nell'effettiva pronuncia della lettera che corrisponde alla consonante nasale (e a volte anche quelle orali seguenti) che farebbe invece parte del digrafo <VN> della vocale nasale (ad esempio: *dimanche* 'domenica' che viene reso come *[dimãŋ] oppure come *[dimanʃ] anziché come [dimãʃ]). Anche in questo caso le allieve hanno avuto modo di pronunciare sia la versione scorretta che quella corretta delle vocali, per potersi così rendere conto della differenza delle due realizzazioni nonché di dove risiedesse l'errore che molte di loro avevano commesso durante il primo ciclo di registrazioni. La parte espositiva della lezione si è conclusa con un procedimento analogo per la messa in rilievo di un terzo errore frequente presso apprendenti italofofoni del francese, ovvero la confusione fra la nasale medio-bassa anteriore [ẽ] e la corrispondente posteriore [õ]. I soggetti hanno cercato di capire la natura di questo errore pronunciando la parola *singe* 'scimmia'. Anche qui si è sottolineata la realizzazione errata *[sõʒ] rispetto alla versione corretta [sẽʒ].

² Nelle 35 frasi lette dai partecipanti, l'articolo indeterminativo *un* /œ/ ricorre tre volte nei sintagmi *un saint* 'un santo', *un long dimanche* 'una lunga domenica' e *un ami* 'un amico'. Nell'ultimo caso dev'essere pronunciata anche la consonante nasale /n/ a causa del noto fenomeno della *liaison*: [œn ami]. Questo aspetto non è stato trattato nella lezione di fonetica, ma non ha posto nessun problema alle allieve.

Come prima applicazione pratica sono state selezionate per essere lette insieme alcune delle 35 frasi del questionario, correggendo puntualmente la pronuncia nel caso di realizzazioni scorrette delle vocali nasali e mostrando il motivo dell'errore. Come ci si aspettava, lessemi quali il sostantivo *leçon* 'lezione' o l'aggettivo *long* 'lungo' sono stati spesso realizzati con una sequenza composta da una vocale orale e da una consonante nasale [ɔn]. La seconda applicazione pratica in classe consisteva in un lavoro a coppie. Le due allieve avevano il compito di leggere ad alta voce una serie di lessemi proposti e di scegliere tramite una crocetta il simbolo fonetico corrispondente alla vocale nasale pronunciata. Un esempio viene dato nella Tab. 6 con la parola *français* 'francese' /frãse/:

Tabella 6 - Esempio tratto da una scheda di esercizi trattata durante la lezione di fonetica

<i>français</i>	<input type="checkbox"/> [ã]	<input type="checkbox"/> [õ]	<input type="checkbox"/> [ɛ]	<input type="checkbox"/> [œ]
-----------------	------------------------------	------------------------------	------------------------------	------------------------------

Anche questo esercizio ha causato diverse incertezze, in particolare per quanto riguarda la pronuncia della vocale nasale posteriore bassa (è l'esempio riportato nella Tab. 6). Da un lato la difficoltà consisteva nel riconoscimento del simbolo fonetico, dall'altro lato si sentivano spesso realizzazioni del tipo *[frõse], con la vocale medio-bassa posteriore invece di [ã]. Alla fine della lezione è stato chiesto alle apprendenti di rileggere a casa ad alta voce l'intero questionario in modo da potersi preparare per l'incontro del secondo ciclo di registrazioni. Inoltre, hanno dovuto rifare a casa l'esercizio relativo alla Tab. 6 che è poi stato corretto *in plenum* in occasione della seconda lezione.

Lo scopo della lezione di fonetica e della ripetizione a casa di due esercizi (in cui i vocaboli contenenti le vocali nasali erano noti da due occasioni precedenti, ovvero dal primo ciclo di registrazioni e dalla lezione di fonetica) era di poter fornire una prima risposta alla seconda domanda di ricerca (DR2) posta alla fine della parte introduttiva (§2.4): fino a che punto è possibile migliorare la pronuncia degli apprendenti mediante un insegnamento fonetico esplicito?

5. *Analisi qualitativa*

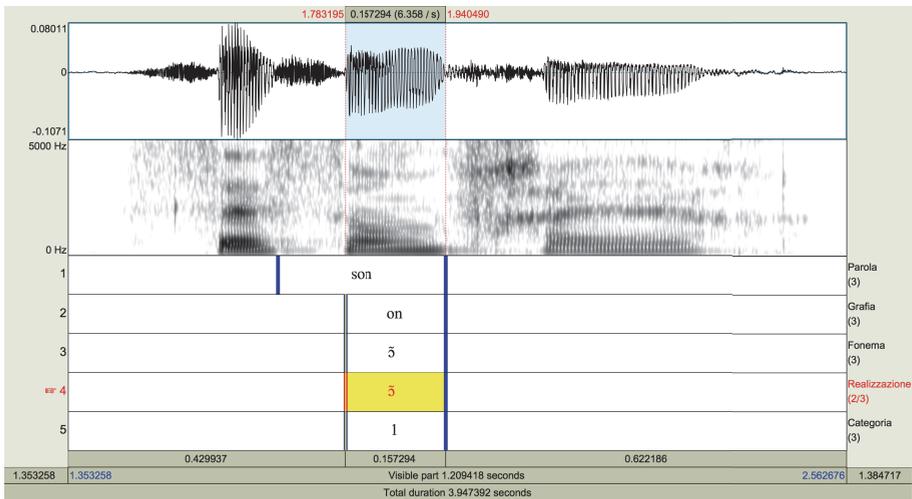
L'effettiva pronuncia delle vocali nasali è stata classificata – in base all'ascolto, alla segmentazione e all'annotazione delle registrazioni in *Praat* – con lo scopo di descrivere i processi di sostituzione fonologica nelle interlingue delle apprendenti attraverso il procedimento collaudato della 'analisi degli errori' (Corder, 1981). In questo modo sono state individuate sette categorie di realizzazione delle vocali nasali francesi da parte delle allieve:

1. vocale nasale conforme al modello del francese standard;
2. vocale nasale diversa dal modello della L2 (ad esempio la parola <vin> 'vino' pronunciata [vã] anziché [vɛ]);

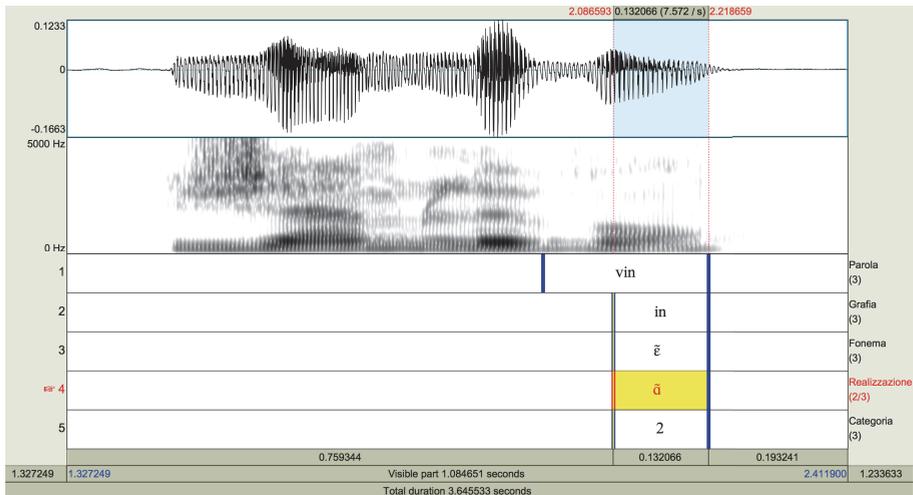
3. sequenza di una vocale nasale conforme al modello della L2 seguita da una consonante nasale (ad esempio la parola <temps> ‘tempo’ pronunciata [tā̃ŋ] anziché [tā̃]);
4. sequenza di una vocale nasale diversa dal modello della L2 seguita da una consonante nasale (ad esempio la parola <temps> ‘tempo’ pronunciata [tō̃ŋ] anziché [tā̃]);
5. sequenza di una vocale orale seguita da una consonante nasale (ad esempio la parola <temps> ‘tempo’ pronunciata [tɑ̃ŋ] anziché [tā̃]);
6. sequenza di una vocale orale seguita da una consonante non nasale;
7. semplice vocale orale (anziché nasale).

La Fig. 1 mostra un esempio di pronuncia corretta della vocale nasale [ɔ̃] nella parola *son* ‘suono’, come si vede anche osservando lo spettrogramma che presenta un solo segmento di una vocale nasale:

Figura 1 - *Vocale nasale secondo la norma del francese standard*

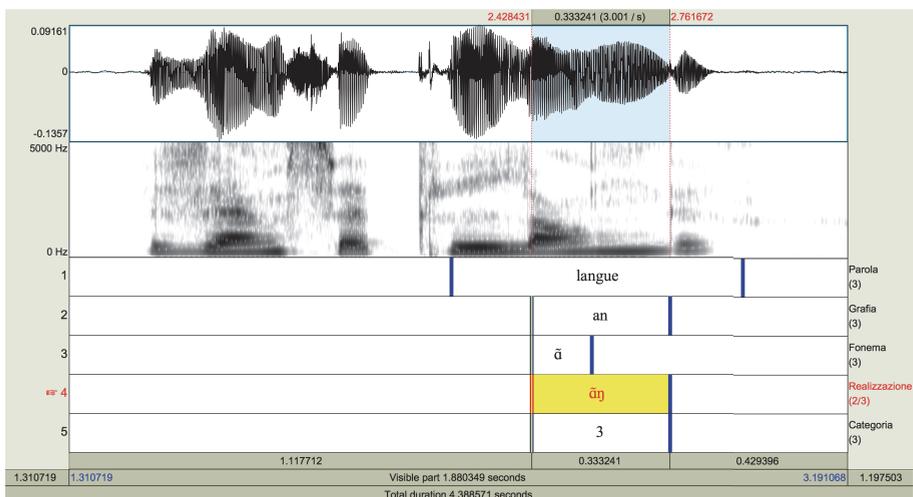


La Fig. 2 illustra invece un esempio dalla seconda categoria: il soggetto pronuncia una vocale nasale diversa da quella prevista dal francese standard (qui [ã] anziché [ẽ̃] in *vin* ‘vino’):

Figura 2 - *Vocale nasale non conforme alla norma del francese standard*

Tale fenomeno è probabilmente legato al fatto che la vocale nasale [ã] è la più frequente in francese (cfr. §2.2, Tab. 2) e probabilmente anche quella incontrata più spesso dagli apprendenti nell'input a loro accessibile.

L'esempio di *langue* 'lingua' della Fig. 3 mostra nello spettrogramma la presenza di due foni corrispondenti alla grafia di <an>; il primo segmento corrisponde alle vocale nasale attesa [ã], il secondo è la nasale velare [ŋ].

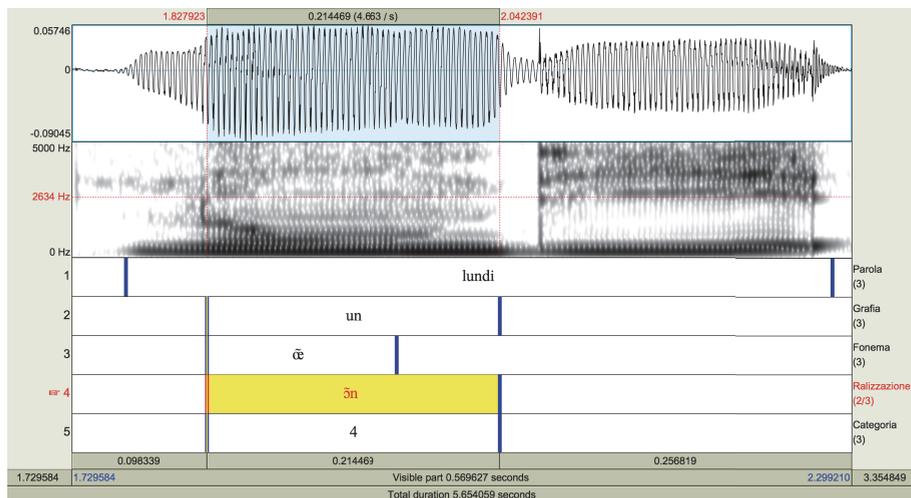
Figura 3 - *Vocale nasale seguita da consonante nasale*

La lettura del digramma <an> fa in questo caso pronunciare alla parlante italoфона la sequenza [ãŋ], dove l'elemento vocalico nella parola *langue* è reso come vocale nasale posteriore bassa, secondo la norma della lingua francese standard; tuttavia viene aggiunta una consonante nasale velare sul modello della parola italiana *lingua*.

Ciò significa che la parlante mantiene la sequenza binaria <vocale nasale + consonante nasale>.

Un esempio per la quarta categoria di realizzazione è riportato nella Fig. 4:

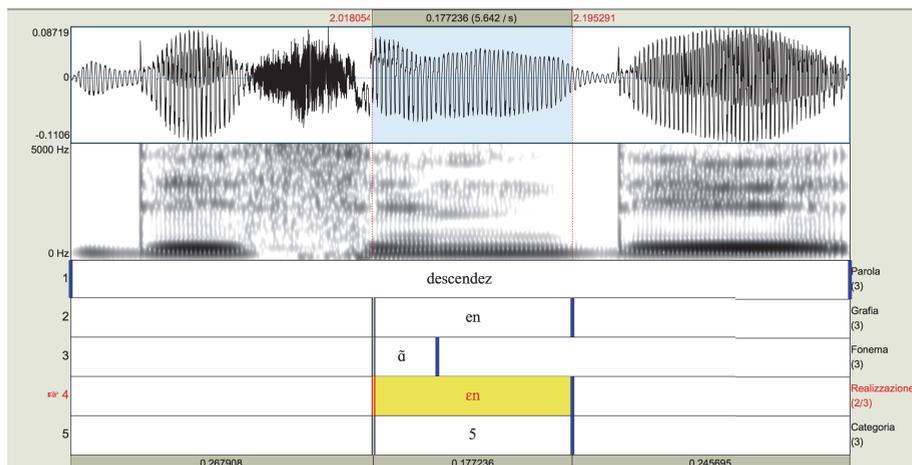
Figura 4 - *Vocale nasale non conforme alla norma del francese standard seguita da consonante nasale*



In questo caso la parola *lundi* 'lunedì' viene pronunciata *[lɔ̃ndi] anziché [lœ̃di] o [lɛ̃di]. Si nota quindi la sequenza di una vocale nasale diversa dallo standard ([ɔ̃]) seguita da una consonante nasale ([n]). Tale sequenza di due segmenti è anche visibile nello spettrogramma.

La quinta categoria presenta il caso di una vocale orale seguita da una consonante nasale:

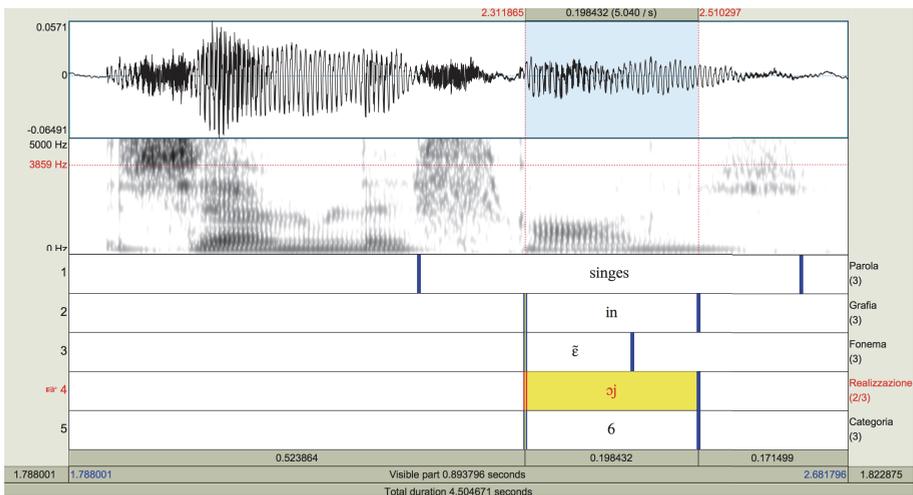
Figura 5 - *Vocale orale seguita da consonante nasale*



Nella Fig. 5 tale categoria è rappresentata dall'esempio di *descendez* 'scendete', pronunciato come [desende] invece di [desãde]; come vedremo di seguito, si tratta della categoria di errori più frequente (v. Fig. 8).

La penultima categoria, quella della vocale orale seguita da consonante non nasale, è raffigurata nella Fig. 6 tramite l'esempio *singes* 'scimmie':

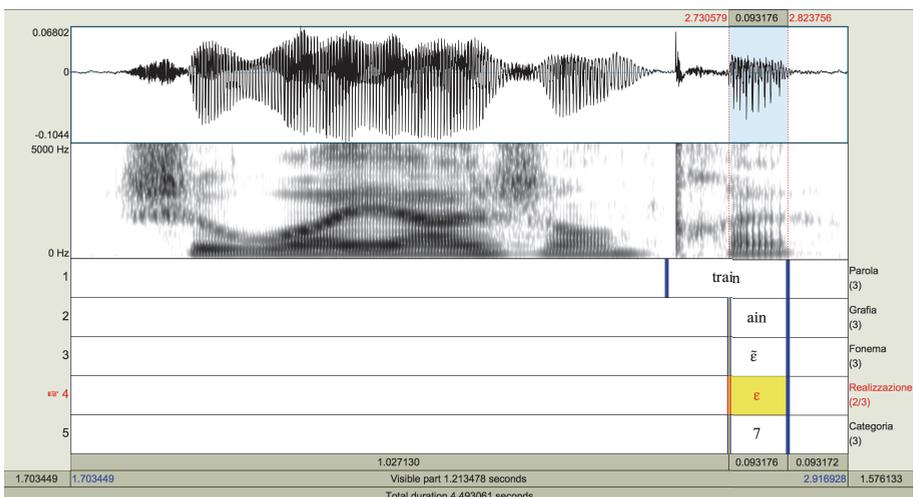
Figura 6 - *Vocale orale seguita da consonante non nasale*



La parlante pronuncia due foni distinti visibili nello spettrogramma che corrispondono alla vocale orale [ɔ] e all'approssimante palatale [j].

La categoria 7, invece, è quella della pronuncia di una sola vocale orale:

Figura 7 - *Vocale orale*



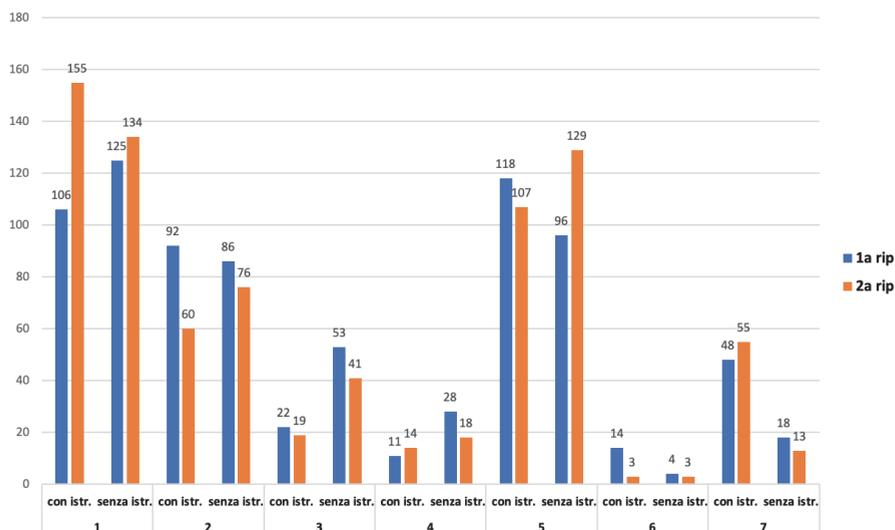
La Fig. 7 mostra l'esempio di *train* 'treno', pronunciato [trɛ] anziché [trɛ̃]. Tale fenomeno potrebbe essere dovuto al fatto che la vocale nasale da pronunciare si trova in posizione finale di parola al confine di frase, per cui la pronuncia risulta 'più debole' e il tratto della nasalità scompare.

6. *Analisi quantitativa*

6.1 Effetto dell'insegnamento esplicito

Si passa ora a una verifica quantitativa dell'effetto dell'istruzione esplicita della pronuncia. Nell'istogramma della Fig. 8 si opera per ciascuna delle sette categorie illustrate nel paragrafo precedente un confronto tra le realizzazioni delle vocali nasali prodotte dalle allieve delle due classi nella prima e nella seconda registrazione. Per ognuna delle sette categorie (disposte da sinistra a destra sull'asse delle ascisse) si presentano quattro colonne che indicano il numero delle relative realizzazioni (a sinistra i risultati della classe che ha ricevuto l'istruzione esplicita prima della seconda registrazione e a destra i risultati della classe che funge da gruppo di controllo). Le due registrazioni vengono differenziate tramite i colori delle colonne: la prima registrazione è segnalata con il colore blu, mentre i valori ricavati dalla seconda registrazione appaiono nelle colonne dal color arancione.

Figura 8 - *Le sette categorie: classi e ripetizioni*



Un primo risultato che emerge dall'istogramma della Fig. 8 è che nella classe con istruzione il numero delle pronunce corrette (categoria 1) è aumentato notevolmente dalla prima (106) alla seconda registrazione (155). Questa differenza (49 casi) corrisponde a un aumento del 46,2% delle realizzazioni corrette e porta a ritenere

positivo l'effetto della lezione esplicita impartita a tale classe fra le due ripetizioni. La classe senza istruzione vede pure un lieve aumento delle pronunce corrette dalla prima (125) alla seconda ripetizione (134), pur non avendo potuto usufruire di un insegnamento esplicito; si potrebbe comunque supporre che alcune allieve abbiano riflettuto su quanto pronunciato e abbiano cercato di rettificare durante la seconda ripetizione. Tuttavia in questa classe l'aumento è di soli 9 casi, equivalente a una crescita del 7,2 % (di fronte a 46,2% nella prima classe); questa differenza notevole tra le due classi ci porta quindi ad attribuire un effetto positivo all'insegnamento esplicito della pronuncia.

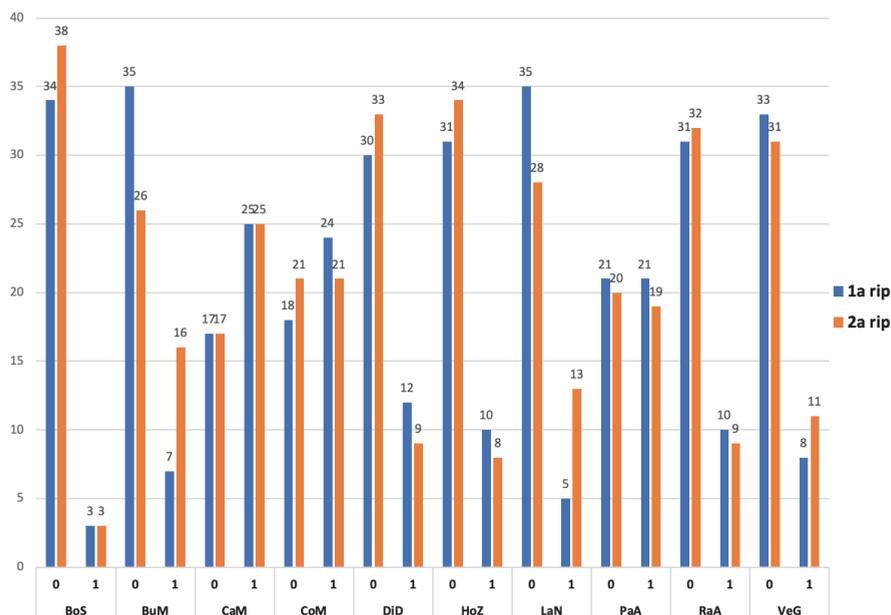
Il secondo risultato che si evince dall'istogramma della Fig. 8 è che nella classe con istruzione il numero di pronunce di vocali nasali sbagliate (categoria 2) diminuisce fra la prima e la seconda ripetizione (con un calo del tasso d'errore del 34,8%), mentre nella classe senza istruzione il tasso d'errore decresce solo dell'11,6%. Questa differenza speculare tra le due classi rispetto alla prima categoria fornisce quindi un ulteriore indizio che permette di affermare che l'insegnamento esplicito ha contribuito a migliorare la pronuncia delle vocali nasali.

Il terzo risultato interessante emerge dai dati relativi alla categoria 5 (vocale orale seguita da una consonante nasale), che costituisce il tipo di errore più frequente. Si constata nella classe con istruzione un calo del tasso d'errore del 9,3%, il che costituisce senz'altro un dato positivo se si tiene conto del fatto che nel gruppo di controllo succede il contrario: in questa classe, infatti, l'errore aumenta dalla prima alla seconda ripetizione del 34,4%. Anche in questo caso la notevole differenza tra le due classi ci fornisce dunque un terzo elemento che porta ad attribuire all'insegnamento esplicito della pronuncia senz'altro un effetto positivo.

Resta comunque da valutare, al di là dei valori assoluti ottenuti per le due classi considerate come gruppi omogenei, il livello di variazione interindividuale in ognuna delle due classi.

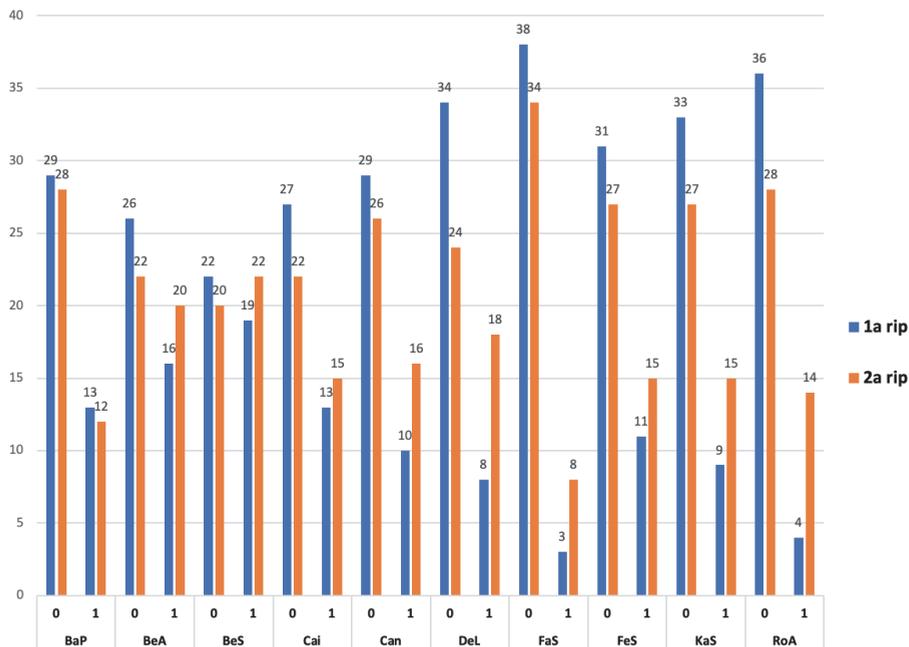
6.2 Variazione interindividuale nelle due classi

Iniziamo con il gruppo di controllo. L'istogramma della Fig. 9 mostra per ognuna delle 10 allieve (elencate da sinistra a destra) quattro colonne. Le colonne etichettate dal codice 0 rappresentano il numero complessivo delle realizzazioni 'non standard' (cioè degli errori), mentre le colonne segnalate dal simbolo 1 corrispondono alle realizzazioni che coincidono con la norma del francese standard. I due colori blu e arancione si riferiscono invece alla prima e alla seconda registrazione.

Figura 9 - *Variazione interindividuale nel gruppo di controllo (DR3)*

Il quadro generale che emerge dalla Fig. 8 mostra che in questa classe non esiste un pattern omogeneo nei risultati prodotti dalle allieve registrate. Infatti in metà della classe si è verificato addirittura un aumento del tasso d'errore dalla prima alla seconda ripetizione (BoS, CoM, DiD, HoZ, RaA); un'allieva ha mantenuto il tasso d'errore costante (CaM) e per le quattro allieve rimanenti (BuM, LaN, PaA, VeG) si ha un sorprendente miglioramento, in quanto soprattutto in un caso l'errore diminuisce del 25,7% (BuM) e nell'altro del 20% (LaN). Il fatto che il numero di pronunce sbagliate aumenti o rimanga costante non sorprende, vista la mancata istruzione nella classe di controllo; il calo delle pronunce errate nei quattro soggetti summenzionati potrebbe essere dovuto al fatto che costoro hanno percepito la difficoltà dell'esercizio e hanno successivamente cercato di migliorare. Ad ogni modo il grado di variazione interindividuale riscontrato nella classe di controllo corrisponde alle aspettative, dato che il percorso scolastico delle allieve non ha mai permesso loro di studiare da un punto di vista teorico e pratico quanto appreso dalla classe con istruzione durante la lezione di fonetica (cfr. §4).

La variazione interindividuale nella classe che ha beneficiato di un'istruzione esplicita della pronuncia viene invece illustrata nella Fig. 10. Anche in questo grafico, l'istogramma mostra per ognuna delle 10 allieve (elencate da sinistra a destra) quattro colonne, dove le colonne indicate dai codici 0 e 1 rappresentano rispettivamente il numero complessivo degli errori e delle realizzazioni corrette. Anche in questo grafico i due colori blu e arancione si riferiscono invece alla prima e alla seconda registrazione.

Figura 10 - *Variazione interindividuale nella classe con istruzione (DR3)*

A differenza della Fig. 9 relativa al gruppo di controllo, in questo grafico si osserva una tendenza piuttosto generale al miglioramento della pronuncia che si manifesta nella diminuzione degli errori sotto 0 (dove le colonne di colore arancione sono più basse rispetto a quelle blu) e nell'aumento delle realizzazioni corrette sotto 1 (dove le colonne color arancione diventano più alte rispetto a quelle blu). Questo è il caso di ben 9 allieve su 10 (BeA, BeS, Cai, CaN, DeL, FaS, FeS, KaS, RoA); l'unica eccezione è l'allieva BaP. Le nove allieve (il 90%) hanno mostrato nella seconda registrazione un notevole progresso (valori fra il 15,4% e il 25%) in seguito alla lezione di fonetica. Cinque di queste nove allieve (il 55,6% della classe) mostrano addirittura un miglioramento superiore al 50% rispetto ai propri risultati iniziali.

Si consideri a questo proposito la Tab. 7 che evidenzia i progressi del numero di pronunce corrette attraverso i valori positivi nella colonna 'differenza' (con l'unica eccezione, già menzionata, della allieva BaP per la quale il valore in questa colonna ha un valore negativo):

Tabella 7 - *Pronunce corrette per ogni allieva nella prima e nella seconda ripetizione e conseguente progresso (differenza positiva) o peggioramento (differenza negativa)*

<i>Allieva</i>	<i>1a ripetizione</i>	<i>2a ripetizione</i>	<i>Differenza</i>
BaP	13	12	-1
BeA	16	20	+4
BeS	19	22	+3
CaI	13	15	+2
CaN	10	16	+6
DeL	8	18	+8
FaS	3	8	+5
FeS	11	15	+4
KaS	9	15	+6
RoA	4	14	+10

La differenza di -1 nel caso di BaP sembrerebbe negare l'efficacia dell'insegnamento almeno per questo individuo. Inoltre, altre cinque allieve mostrano una differenza sì positiva, ma comunque assai contenuta (da +2 a +5). Per le quattro allieve rimanenti l'insegnamento si è invece mostrato effettivamente efficace; tant'è vero che per un'allieva il numero di pronunce corrette è aumentato di +10, mentre le altre differenze sono comprese tra i valori +6 e +8 (anch'essi abbastanza elevati). Tutto sommato i risultati di questo esperimento didattico portano a ritenere che un lavoro esplicito con gli allievi e le allieve sulla teoria (articolazione) e sulla messa in pratica (esercizi di pronuncia) possa avere degli effetti positivi che si traducono in una maggiore sensibilità da parte della maggior parte degli apprendenti riguardo alla percezione e alla pronuncia delle vocali nasali francesi.

7. Osservazioni conclusive

Tornando alle tre domande di ricerca formulate nel §2.4, abbiamo rilevato presso le apprendenti italofone del nostro campione notevoli difficoltà nella pronuncia delle vocali nasali francesi (DR1). Gli ostacoli si sono mostrati chiaramente nella prima registrazione e sono perdurate anche nella seconda registrazione nel gruppo di controllo. Gli errori di pronuncia si sono manifestati in ben sei categorie di realizzazione (cfr. §5). Gli errori più frequenti sono quelli corrispondenti alle categorie 2, 5 e 7 (ovvero: vocale nasale diversa da quella prevista dal modello del francese standard, vocale orale seguita da consonante nasale, vocale orale). Il tipo di errore più frequente in assoluto si manifesta nella categoria 5 (vocale orale seguita da consonante nasale) e sembra essere influenzato dall'input ortografico (come predetto in § 2.2), ovvero dalla presenza dei grafemi <n> e <m>. Inoltre, sono state soprattutto le vo-

cali /ã/ e /ẽ/ a creare delle difficoltà per via delle diverse possibilità grafiche esistenti per la rispettiva pronuncia delle due vocali nasali.

Per quanto riguarda invece la seconda domanda di ricerca (DR2), ribadiamo che in questo caso l'insegnamento esplicito della pronuncia si è mostrato efficace nel rendere le allieve della classe con istruzione più sensibili agli errori di pronuncia più frequenti. Infatti, come mostrato sopra, l'esito globale della lezione di fonetica nella classe con istruzione ha avuto ripercussioni positive in nove casi su dieci. Rispetto allo studio precedente sul VOT nelle occlusive del tedesco (Schmid, Pedrazzini, 2016), l'impatto della lezione di fonetica ha raggiunto un numero maggiore di allievi (forse perché la nasalità delle vocali francesi ha comunque una maggiore funzione comunicativa in quanto riesce a distinguere fonemi, laddove l'aspirazione delle occlusive sorde ha uno statuto meramente allofonico). Per quanto riguarda infine la terza domanda di ricerca (DR3), è interessante osservare come la variazione interindividuale sia molto più elevata nel gruppo di controllo che non nella classe che ha beneficiato della lezione di fonetica: interpretiamo questo dato come ulteriore indizio a favore dell'efficacia dell'insegnamento esplicito della pronuncia.

Ricerche future potranno aggiungere nuovi elementi alla discussione. Sarebbe senz'altro utile poter registrare un gruppo più ampio di locutori, fra cui anche parlanti di sesso maschile. Inoltre una ricerca con un maggior numero di registrazioni su un arco di tempo più esteso (ad esempio una decina di registrazioni distribuite su due semestri con più di una sola sessione d'insegnamento fonetico esplicito) potrebbe portare a progressi ancora più sostenibili per ogni categoria, permettendo di trarre anche delle conclusioni più robuste dal punto di vista empirico. Oltre al parlato letto si potrebbero registrare anche conversazioni semi-spontanee (ad esempio attraverso il compito *MapTask* o il cosiddetto 'testo delle differenze') per verificare come vengono pronunciate le vocali nasali in un contesto che non è focalizzato sulla forma linguistica. Alla mera classificazione uditiva delle realizzazioni degli apprendenti andrebbe aggiunta, ovviamente, un'analisi acustica sul piano delle formanti. Per motivi di tempo e di spazio purtroppo ciò non è stato possibile in questa sede, ma speriamo di poter condurre in un'altra occasione una ricerca più approfondita sull'acquisizione delle vocali nasali francesi da parte di apprendenti italofofoni.

Ringraziamenti

Teniamo a ringraziare l'insegnante di francese e gli allievi di due classi della Scuola Cantonale di Commercio a Bellinzona per la loro disponibilità e per l'impegno con cui hanno contribuito a rendere possibile questa ricerca. Siamo grati anche a Sandra Schwab per uno script di *Praat* che ha permesso di estrarre automaticamente il numero delle realizzazioni per ogni categoria di errore. Infine vogliamo ringraziare due revisori anonimi per i loro commenti critici che hanno permesso di migliorare la stesura del testo; ovviamente i due autori rimangono gli unici responsabili di eventuali errori e imprecisioni nel presente contributo.

Bibliografia

- BEDDOR, P. (1993). The perception of nasal vowels. In Huffman, M., Krakow, R. (Eds.), *Nasals, Nasalization, and the Velum*. San Diego: Academic Press, 171-196.
- BERTINETTO, P.M., LOPORCARO, M. (2005). The sound pattern of Standard Italian, as compared with the varieties spoken in Florence, Milan and Roma. In *Journal of the International Phonetic Association*, 35(2), 131-151.
- BEST, C., TYLER, M. (2007). Nonnative and second-language speech perception: Commonalities and complementarities. In Bohn, O., Munro, M. (Eds.), *Language experience in second language speech learning: In honor of James Emil Flege*. Amsterdam: Benjamins, 13-33.
- BOERSMA, P. WEENINK, D. (2019). *Praat: doing phonetics by computer*. Versione 6.1.06, <http://www.praat.org>.
- BUSÀ, M.G. (1995). *L'inglese degli italiani: l'acquisizione delle vocali*. Padova, Unipress.
- CANEPARI, L. (1979). *Introduzione alla fonetica*. Torino: Einaudi.
- CANEPARI, L. (1980). *Italiano standard e pronunce regionali*. Padova: CLEUP.
- CONSIGLIO D'EUROPA (2010). *Quadro comune europeo di riferimento per le lingue: apprendimento insegnamento valutazione*. Terza ristampa. Milano: La Nuova Italia-Oxford.
- CORDER, S. (1981). *Error analysis and interlanguage*. Oxford: Oxford University Press.
- D'APOLITO, S., GILI FIVELA, B. (2009). Il francese L2 come lingua straniera: l'assimilazione del tratto di sonorità. *Atti del V Convegno AISV*, 216-241.
- DELATTRE, P. (1965): *Comparing the phonetic features of English, French, German and Spanish: An Interim Report*. Heidelberg: Julius Groos Verlag.
- DELVAUX, V. (2009). Perception du contraste de nasalité vocalique en français. In *French Language Studies*, 19, 25-59.
- DE PAOLIS, B. (2019). Acquisition de la consonne rhotique en français L2 par des adultes italophones résidant en France. In *Bollettino LFSAG*, 3, 61-82.
- DETEY, S. / RACINE, I. (2015). Does perception precede production in the initial stage of French nasal vowel quality acquisition by Japanese learners? A corpus-based discrimination experiment. *Proceedings of the 18th International Congress of Phonetic Sciences*. Glasgow UK: paper n. 894.
- DRAXLER, C., JÄNSCH, K. (2004). SpeechRecorder – a universal platform independent multi-channel audio recording software. *Proceedings LREC 2004*. Lisbon, Portugal.
- ECKMAN, F. (2008). Typological markedness and second language phonology. In Hansen Edwards, J., Zampini, M. (Eds.), *Phonology and Second Language Acquisition*. Amsterdam: Benjamins, 95-115.
- FLEGE, J. (1995). Second language speech learning. Theory, findings, and problems. In Strange, W. (Ed.), *Speech perception and linguistic experience: Issues in crosslanguage research*. Timonium: York Press, 233-277.
- FLEGE, J., BOHN, O.-S. (in stampa). The revised Speech Learning Model. In Wayland, R. (Ed.), *Second language speech learning*. Cambridge: Cambridge University Press.
- FRONTERA, M., ROMITO, L. (2015). *La percepción lingüística: las oposiciones fonéticas/ fonológicas en L2*. Saarbrücken, Editorial Académica.

- FOUGERON, C., SMITH, C. (1999). French. In *Handbook of the International Phonetic Association*, Cambridge, Cambridge University Press, 78-81.
- KRASHEN, S.D. (1987). *Principles and Practice in Second Language Acquisition*. London: Prentice-Hall.
- LADO, R. (1957). *Linguistics Across Cultures*. Ann Arbor: University of Michigan Press.
- MADDIESON, I. (1984). *Patterns of sounds*. Cambridge: Cambridge University Press.
- MAIRANO, P., CALABRÒ, L. (2016). Are minimal pairs too few to be used in pronunciation classes. In SAVY, R. ALFANO, I. (a cura di), 255-268.
- MAIRANO, P., BASSETTI, B., SOKOLOVIĆ-PEROVIĆ, M., CERNI, T. (2018). Effect of L1 orthography and phonology on L2 English pronunciation. In *Revue française de linguistique appliquée*, 23(1), 45-57.
- MATURI, P. (2014). *I suoni delle lingue, i suoni dell'italiano*. Nuova introduzione alla fonetica. Bologna: il Mulino.
- MIONI, A. (1973). *Fonematica contrastiva. Note ed esercizi*. Bologna: Patron Editore.
- MORAN, S., MCCLOY, D. (Eds.) (2019). *PHOIBLE 2.0*. Jena: Max Planck Institute for the Science of Human History. (Available online at <http://phoible.org>, Accessed on 2020-08-12)
- MURANO, M. / PATERNOSTRO, R. (2016). Les italoalphons. In Detey, S., Racine, I. / Kawaguchi, Y., Eychenne, J. (Eds.), *La prononciation du français dans le monde: du natif à l'apprenant*. Paris : CLE International, 149-154.
- SAVY, R., ALFANO, I. (a cura di) (2016). *La fonetica nell'apprendimento delle lingue/Phonetics and language learning* (Studi AISV 2). Milano: Officinaventuno.
- SCHMID, S. (1999). *Fonetica e fonologia dell'italiano*. Torino: Paravia.
- SCHMID, S., PEDRAZZINI, G. (2016). La pronuncia delle occlusive nel tedesco L2 di apprendenti italofofoni: un esperimento didattico. In SAVY, R. ALFANO, I. (a cura di), 45-60.
- SHARWOOD SMITH, M. (1981). Consciousness raising and the second language learner. In *Applied linguistics*, 2(2), 159-168.
- SHARWOOD SMITH, M. (1991). Speaking to many minds: On the relevance of different types of language information for the L2 learner. In *Second Language Research*, 7(2), 118-132.
- SISINNI, B., GRIMALDI, M. (2010). Metodologie negli studi cross-linguistici. Riflessioni sulla relazione fra dati acustici e percettivi. *Atti del V Convegno AISV*, 519-548.
- STELLA, A. (2013). Coordinazione interarticolatoria nella produzione dell'intonazione del tedesco come lingua straniera. *Atti del IX Convegno AISV*, 411-426.
- TARONE, E. (1987). The phonology of interlanguage. In Ioup, G., Weinberger, S. (Eds.), *Interlanguage phonology. The acquisition of a second language sound system*. New York: Newbury House, 70-85.
- TELMON, T. (1993). Varietà regionali. In Sobrero, A. (Ed.), *Introduzione all'italiano contemporaneo. La variazione e gli usi*. Roma-Bari: Laterza, 93-149.
- THOMSON, R.I., DERWING T.M. (2015). The Effectiveness of L2 Pronunciation Instruction: A Narrative Review. In *Applied Linguistics*, 36(3), 326-344.
- VAYRA, M., AVESANI, C. & TAMBURINI, F. (a cura di) (2015). *Il farsi e il disfarsi del linguaggio. Acquisizione, mutamento e destrutturazione della struttura sonora del linguaggio* (Studi AISV 2). Milano: Officinaventuno.

FABIO ARDOLINO

Stereotypical dimensions and Second Language development in a migrant community: A pilot study

Complex forms of reciprocal influences link social, psychological, and linguistic dynamics in social integration processes. Despite the growing number of studies directly assessing the relationships between socio-psychology and language, such a matter keeps resisting a comprehensive systematization, especially for what concerns the phonetic investigation. Founding its premises on a widely adopted socio-psychological model (the *Stereotype Content Model*), the study here proposed intends to develop a preliminary analysis of the role exerted by stereotypical constructs in specific phonetic aspects of the L2 development. The research – involving a sample of Italo-French participants in a spontaneous setting – highlights the role of a particular stereotypical dimension (the *warmth*) in orienting the L2 acquisitional outputs: the result is discussed from a broader affective and behavioural perspective.

Keywords: stereotypes, SLA, migration, French L2.

1. Introduction

Anyone who has experienced forms of transfer from his/her habitual social context to another – albeit a short-term one – understands how a functional integration does not merely depend on communicative factors (active and passive competence of the dominant contextual language), but that it is embedded in a complex matrix of cognitive, psychological, social, and behavioral variables. Nevertheless, it is impossible to fully distinguish the former element from the latter: language use largely depends on so-called *identity factors* (personal representations, beliefs, values, and idiosyncrasies). As observed by Heller (1987), linguistic competence heavily influences self-identity in distinct contexts, constituting a precondition to create bonds outside the native environment. Moreover, a functional competence in the dominant vehicular language is a necessary stepping stone in achieving full integration in a community (see Cummins, 1979).

In their studies on Australian immigrants, Chiswick and Miller (1999) experimentally proved the predictor role exerted by linguistic competence on the asserted social identity: while multilingual immigrants easily declare their adhesion to more than one social group, unskilled speakers reveal a more strict identification towards their native social cluster. That said, psycho-sociological and linguistic investigations reveal a much more complex frame as compared to the naïve implication “good competence – good integration”, since migrants often face considerably complex communicative and social scenarios. The inquiries led by Goldstein (1996) on Portuguese-speaking migrants in Canada, for instance, show that the same language

can receive radically opposite evaluations in different communicative contexts: in the Torontonion low-class factory workers, the English – dominant language of the country – generates mistrust, since it is perceived as an attempt to exclude the less-competent coworkers.

Evidently, the host community has a role in determining similar dynamics. Adult learners in nonnative contexts are regularly subject to explicit or implicit negative attitudes (Perdue, 1993: 39-51). Natives often judge nonnative speakers as non-trustworthy or ignorant people, attitudes that can sometimes lead to distrust or over discrimination (Esser, 2006). Such a stigma can sometimes broaden to embrace foreign-accented varieties, even if fully intelligible – a phenomenon known as *linguicism* (see Skutnabb-Kangas, 2015). More extensive analyses, like those carried on by Belot and Ederveen (2002) and by Adsera and Pytlikova (2015), reveal that the persistence of linguistic prejudices toward nonnatives can exert wider influences on the migration fluxes.

The acquisitional *côté* lies underneath all the examined examples. Whilst successful acquisition grants a better integration, it is also true that such competence arises (at least in spontaneous acquisition) only when the learner is regularly involved in a net of interactions with target-language speakers. This state of facts also recurs in the acquisitional model postulated by Krashen (1987, 1988) in which two factors, among others, are considered: 1) the dimensional parameters of the received *inputs*, in terms of number, frequency, and quality; in other words, the number of successful interactions that a learner establishes during his acquisitional process, and 2) the so-called *affective filter*: a macro-category in which psychosocial factors (identity, attitudes, and emotions) co-occur in shaping the learner's motivation, self-confidence, communication willingness, and interactional behaviors.

Krashen's model constitutes a first attempt to connect the psychosocial sphere with the acquisition and the use of a second language. This link – although often raised in theory – remains, however, partly underinvestigated. As observed by Norton (2000: 39), in focusing such matters, scholars' attention appears to be mainly focused on didactical contexts and, therefore, generally oriented on the quality of the learning processes outputs. Twenty years after Norton's remarks, it is still possible to highlight a lack of quantitative studies exploring the role of psycho-sociality in second language spontaneous acquisition. As a matter of fact, this scarcity is also imputable to the problematic integration between socio-psychological and linguistic protocols.

The explorative study here presented precisely intends to put forth a quantitative method to compare specific socio-psychological variables and L2 outputs. This inquiry is developed within a segmental level – the phonetic one – which is intensely involved in accent perception and “foreignness” judgments (see Marotta, Boula De Mareüil, 2009).

In order to place the study in an exhaustive theoretical frame, the next paragraph (§ 2) is devoted to outlining an essential literature review for what concerns acquisitional research regarding migrant speakers. Furthermore, a brief digression (§ 3)

discusses the reference models in which the experimental protocol is developed. The study is then described in § 4, while conclusions are provided in § 5.

2. *L2 dynamics in adult migrants: a preliminary literature review*

As mentioned above, a specific research field centered on the L2 dynamics characterizing adult migrants appears still unsystematized. In fact, several distinct analyses – following divergent theoretical and methodologic approaches – have been carried on, in distinct disciplinary areas (Sociolinguistics, Communication studies, Language variation and, above all, Language Acquisition and Learning): this paragraph attempts to set up a representative review of the main disciplinary trends directly or tangentially focusing such matter.

Concerning the acquisitional field, the influential research developed by Chiswick and colleagues constitutes a crucial starting point. In Chiswick and Miller (1995), the census surveys of four distinct countries (United States, Canada, Israel, and Australia) constitute the database to determine which elements are more strongly involved in adult migrants' language fluency; in the study, four main factors are circumscribed, namely *exposition* (to L2), *efficiency* (individual predisposition to language learning), and *cost-benefit evaluation*. A fourth variable, the *linguistic distance* – the amount of structural proximity between languages – was later added (see Chiswick, Lee & Miller, 2005)¹.

Among those factors, the *cost-benefit evaluation* variable takes on a particular sociolinguistic meaning. The postulation of a direct role of the advantages/disadvantages balancement in the L2 acquisitional process addresses the idea that migrants proceed towards a “full” competence in those contexts where second language mastery guarantees real economic, social, and relational opportunities (see also Chiswick, 1998). Along the same lines moves the broad qualitative study carried on by Norton (2000) on five immigrant women in Canada. Discussing the collected interviews, the author concludes that the possibility to successfully fit in the host community fabric represents one of the most influential variables in the L2 acquisition outcomes: subjects having access to contexts in which women are encouraged to achieve financial and personal self-realization tend to manifest more native-oriented L2 competences.

In those terms, the *cost-benefit evaluation* can be easily considered as a strict correlate (if not a full synonym) of the widely known *motivation* variable: learners who associate competence depth and personal possibilities tendentially appear more motivated to invest attention and energy in the L2 acquisition.

The motivational aspects of L2 acquisition seem to be sufficiently addressed mainly in educational settings (see Gardner, Lalonde, & Moorcroft, 1985 and Dörnyei, 2009), although there is evidence of a direct role exerted by motivation in

¹ In this regard, it is to emphasize that in a previous work Dustmann and Van Soest (2001) critically addressed the issues arising from surveys and self-reports use in L2 competence assessment.

several L2-related processes, e.g., foreign accent maintenance (Moyer, 2007) and L2 prosodic accuracy (Frontera, Paone, 2018).

A comprehensive model aimed at integrating language acquisition and social integration processes is the *Acculturation model*, firstly proposed by Schumann (1978) and partly anticipated by the study carried on by Cancino, Rosansky, and Schumann (1975). The model postulates a series of principles reversing the classical relationships between the acquisition of a new language and the integration toward a host community: in detail, it states that the acquisition of a dominant language in a nonnative speaker can be seen as a function of the psychological, behavioral, and cultural assimilation among hosts' values and habits (i.e., the *acculturation* process). A significant contribution to this framework comes from the research carried on by Berry and colleagues (see, among others, Berry, Kim, Minde & Mok, 1987 and Berry, 1997), focusing on the distinct acculturation's outcomes, and from the recent work by Yilmaz and Schmid (2015), enlarging the pictures to include broader sociopolitical factors.

As mentioned in § 1, the ensemble of communicative interactions in which the learner is involved constitutes a keystone in second language dynamics' assessment. The study proposed in Wiklund (2002) and focalized on second language proficiency explicitly explores the relationships between this matter and the composition of learners' social networks. The research – involving adolescents immigrants in Sweden – highlights how better proficiencies are achieved by those learners having Swedish natives among their nearest network nodes (i.e., close friends). A more comprehensive study by Smith (2002) on American immigrants in Southeast Asia corroborates Wiklunds' findings, also highlighting how denser network – in which stronger social links already exists – play a part in inhibiting proficiency progresses; in other words, learners enjoying a satisfactory social inclusion feel less urge to achieve an advanced competence in the dominant language. Among the same lines, Danzer and Yaman (2011) focus on the detrimental role that *ethnic enclaves* (neighborhoods in which same-origin migrants concentrate) exert on the L2 abilities.

Addressing the notion of *variation* in SLA (Littlewood, 1981) allows a more in-depth look at the processes mediating between L2 acquisition and outcomes. In a comprehensive discussion regarding this topic, Tsimpli (2006: 387:8) addresses three crucial points worth mentioning: a) the production of outcomes deviating from the target language is an expected phenomenon in L2 learner, generally linked to poor expositions to L2 inputs, b) compared with pronunciation, grammar is generally more susceptible to a successful native-like acquisition, and c) individual variation in the second language may also depend from individual differences in the L2 development process, e.g., lesser ability in decoding linguistic input and in developing a trustworthy mental representation of the target language. Drummond (2012) enlarges such a frame, highlighting how variation phenomena can depend on specific L1 factors, acting as constraints in L2 acquisitional processes. Schlef (2013), in conclusion, focalizes his research on the *acquisition of variation* issue, suggesting that the acquisition of specific variation pattern does not merely depend

on the local group with which the learner interacts, but on a more complex frame of distinct personal, psychological and environmental variables.

2.1 Social grouping and stereotyping in L2 dynamics

As pointed out in the previous paragraph, the investigation of second language processes in migrant learners broadly considers social and interactional factors: in this light, it appears opportune to enlarge the discussion to the psychological constructs involved in such dynamics. In view of the experimental approaches used in this work, it appears particularly appropriate to discuss two of the most widely investigated psychosocial constructs: *attitudes* and *stereotypes*.

In social psychology, an *attitude* is (following the early definition by Allport, 1935) a form of readiness towards an object, able to psychologically or behaviourally affects the reaction toward them (or toward stimuli related to it). Even if this term is frequently used – in common language – as a synonym for *stereotype*, critical differences separate the two concepts in the disciplinary practice. A first, crucial distinction regards the constructs' spreading: while attitudes mainly depend on subjective experiences, stereotypical beliefs are typically shared – in an overt or covert form – by all the members in the same community (see McGarty, Yzerbyt & Spears, 2002: 2-5). In light of this, only stereotypes can act as forms of prior knowledge in social classification processes, by orienting beliefs and expectations about social categories perceived as homogeneous (see McGarty, 2002: 155-166). Evidently, personal attitudes can be developed on pre-existing, supra-individual stereotypes: doing research about stereotypical constructs, therefore, often means investigating attitudes toward social groups (the term of *stereotypical attitude* exactly defines this kind of scenario, see Wood, 2008).

The *Social Identity Theory* (SIT) developed by Tajfel (1974) precisely intends to systematize the role of stereotypical beliefs in social grouping processes. In the SIT model, individuals classify other social actors within two principal categories: an affiliation group (*ingroup*) and a non-belonging group (*outgroup*).

In this frame, stereotyping emerges as a double-aimed device: a) it makes categorization easier, over-schematizing complex knowledge about groups' characteristics, and b) it enhances ingroup cohesion and self-esteem (the ingroup generally receive positive attributions at the expenses of the outgroup). In these terms, the SIT offers a theoretical background in which to place experimental observations about sociological and psychological processes underlying immigration: even if a more in-depth contextualization of SIT-based analysis in migration contexts goes beyond the objectives of the present study, it seems appropriate to recall that group stereotypes have revealed their predictivity in personal motivation and cognitive capacity (Forbes, Schmander, 2010), in attentional resources mobilization (Allen, Sherman, Conrey & Stroessner, 2012), as well as in social networking characteristics (Clark, Kashima, 2008).

Literature addressing the relations between psychological constructs and second language dynamics in migrants appears to be very poor, while a more consistent cor-

pus of research explores the effects of ethnic and national stereotypes in language learning. The existence of stereotypical constructs towards specific nationhoods is addressed in El-Dash and Busnardo (2001) and Nikitina, Don, and Loh (2014): in both these studies, a positive correlation between stereotype orientation and motivation to learn a language is founded. Hammond (1990) shifts the research focus on ingroup stereotypes in spontaneous contexts, by observing that negative stereotypes about foreign accents have no incidence on pronunciation accuracy in Spanish-speaking immigrants in Miami: this finding suggests a lesser involvement of ingroup stereotypes in SLA processes.

3. Theoretical frameworks of the study

3.1 The Stereotype Content Model

The *Stereotype Content Model* (SCM; Fiske, Cuddy, Glick & Xu, 2002) constitutes a widely-known framework to systematize stereotypical evaluations. A crucial aspect of the SCM is the possibility to describe the orientation and the strength of stereotypes toward social groups using two evaluations, labeled as *warmth* and *competence* dimensions.

The *warmth* dimension encapsulates the evaluations regarding groups' solidarity attributes (tolerance, sympathy, kindness, and similar). In contrast, the *competence* considers the evaluations of groups' intellectual and practical skills (intelligence, ability, ambition). A social group receives an independent stereotypic evaluation on each of the cited dimensions: while the ingroup is generally well-evaluated on both the attributions, outgroups typically receive a negative evaluation in at least one of the judgment.

In the original SCM formulation, the authors provide a graphical representation of the *warmth/competence* evaluations in a 2×2 matrix (see Fig. 1). This schematization highlights the possibility of synthesizing social stereotyping among four possible combinations, each associated with a specific emotive response, driving the intergroup interaction (e.g., *intergroup emotion*).

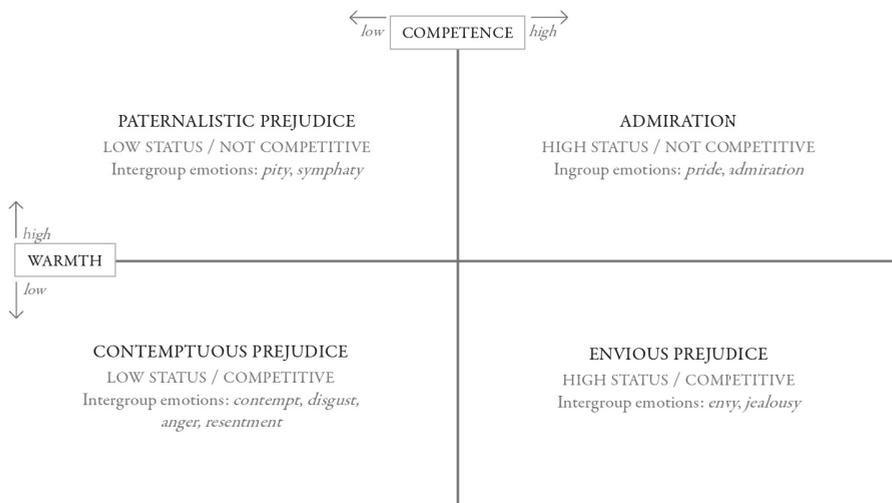
In addition, Fiske and colleagues postulate that two perceived groups' characteristics (*status* and *competition*) can predict *warmth/competence* evaluation. High-status groups (groups situated within the higher social position) receive a generally high competence evaluation, while groups pictured as competitive are generally associated with low *warmth* attributions.

Evidently, the dimensional judgments postulated by the SCM are diriment in the analysis of the socio-psychological collocations of immigrant communities towards the host group. As highlighted by Lee and Fiske (2006), when observed in terms of ingroup/outgroup dynamics, not all immigrant groups receive similar evaluations on the *warmth/competence* spectrum. The more an immigrant group will be perceived as harmoniously inserted into the host community, the closer the stereotypical beliefs will be to the ingroup evaluations. Such an evaluation appears to be strongly relative: an outgroup can be considered "accepted" in the hosting ingroup compared

to a third group, worse evaluated. Similar dynamics can also be explored from the immigrant group's perspective: immigrants involved in a more advanced integration process form better stereotypical constructs about the hosting outgroups.

In a framework in which L2 competence constitutes a precursor for (or a consequence of) a successful integration, it is consequently possible to speculate about a more direct relationship between language skills and stereotypical constructs. This eventuality is directly assessed in the hypotheses of this study.

Figure 1 - Warmth \times competence matrix with possible outgroups type combinations, typical status, and competition, corresponding to prejudice and associated intergroup emotions. The high warmth/high competence combination is associated with the ingroup (source: Fiske et al., 2002)



3.2 Phonetics perception and acquisition in SLA

The impossibility to enclose first and second language acquisition processes within the same theoretical models is widely-accepted evidence in the acquisitional field, and it depends on the existence of a series of internal (neuro-cognitive) and external (environmental) discontinuity that makes the postulation of different models necessary, even with some contact points. It is easy to guess that the presence of the L1 – a well-established linguistic system spontaneously acquired by the learner since birth – constitutes a major issue in SLA modeling.

The *Perceptual Assimilation Model* (PAM; Best, 1995) extensively addresses the role of L1 in nonnative sound perception. PAM framework – and its most recent declination PAM/L2 (Best, Tyler, 2007) – postulates that L1 phonological categories constitute the mnemonic holds to process nonnative perceived sounds: the association between “new” sounds and “old” categories are selectively triggered by the recognition of phonetic-articulatory similarities. Such processes can produce three distinct outcomes: a) the new sound will be judged identical to a native one,

and therefore *categorized* within the same phonological class; b) the new sound will be judged similar to more than a native allophone, this leads to an *uncategorized* classification; c) no similarities are found between the new sound and the stored categories, so the first one remains *unassimilated*.

The case in which two nonnative sounds are assimilated to the same native category generates instead two possible outputs: a) the two sounds are judged equally similar to the native category (*single category assimilation*, SC), and b) one of the two new sounds is judged more similar to the native category than the other (*category goodness assimilation*, CG). In general, the farther a new sound is judged from a native category, the better it will be discriminated by the listener as a completely new phonemical object.

The *Speech Learning Model* (SLM) developed by Flege (1995, 2007) focuses more directly on nonnative sound production in SLA. Similarly to the PAM, the SLM postulates that a nonnative sound is processed based on the pre-existing L1 phonetic categories and that these assimilative processes produce an articulatory outcome in speech production. In Flege's model, the relationship between a new linguistic sound and a native category is expressed along a continuum from an *identical* to a *new* (through a *similar*) attribution. If two sounds are judged as sufficiently similar, they were classified within the same category of the native one (*equivalent classification*), and the creation of a proper category for the nonnative sound will be inhibited. Segments that receive a proper classification within an *ad-hoc* category tend to be more accurately realized in the production phase. As observed by Vayra, Avesani, Best, and Bohn (2012), the main difference between SLM and PAM is the possibility that the judgment of similarity between native and nonnative sounds would produce a new category: while in the SLM this possibility is excluded, the PAM admits that a proper categorization takes place in the CG scenario for the sound judged as less similar to the native category.

4. *The pilot study*

4.1 Hypothesis and phonetic targets

The general premise of the study here presented is that stereotypical dimensions (in terms of strength and negative/positive orientation) towards the host community, given their strict involvement in several acquisitional-related factors, can predict L2 outcomes in spontaneous acquisitional contexts. In particular, we hypothesize that more positive stereotypical beliefs can be directly involved in more accurate identification of the nonnative sounds, motivating their greater distance (in terms of acoustic-articulatory production characteristics) from the L1 sounds on which their acquisition is hinged.

Due to its exploratory nature, the analysis involves a small sample of Italian-French bilingual immigrants (see § 4.1.1) and focuses on the production of a spe-

cific French oral vocalic class², the *frontal rounded* vowels /y, ø, œ/ (see Tranel, 1987: 35), absent both in Standard Italian (Bertinetto, Loporcaro, 2005) and in the participants' regional varieties (Neapolitan Italian, see Sornicola, 1997: 332). Following the principles of the SLM model, such a class can be potentially assimilated either to the Italian *frontal unrounded* vowels class (/i, e, ε/), with which it shares the [+frontal] trait, or to the Italian *back rounded* vowel class (/u, o, ɔ/), with which the [+rounded] trait is shared. Table 1 resumes the L2 target vowels and the relating potentially similarly-judged Italian sounds.

Table 1 - *Vocalic sounds involved in the study. The French targets (marked with an asterisk) and the Italian more similar segments (unmarked) are represented on the same row*

	FRONTAL -ROUNDED	FRONTAL +ROUNDED	BACK +ROUNDED
<i>closed</i>	i	y*	u
<i>closed-mid</i>	e	ø*	o
<i>open-mid</i>	ε	œ*	ɔ

On an articulatory level, the frontality trait refers to an advanced position of the tongue within the oral cavity. The consequent restriction of the oral cavity causes a rising in the F2 frequency (Lammert, Proctor & Narayanan, 2013). In contrast, the production of rounded vowels is characterized by a circular positioning of the lips, accompanied by an elongation of the vocal trait (*labialization*) that produces a lowering in all the formant values, particularly prominent in the F3 (see Maeda, 1990: 142-143). In light of this, two unidimensional parameters (F2 and F3) and a bidimensional one (F3 × F2) will be considered to characterize each vowel.

4.2 Materials and methods

4.2.1 Participants

Eleven speakers, ranging from 28 to 32 years old, took part in the data collection phase. Participants were selected from a larger sample of contributors for an Italian-French oral corpus in order to grant the greater homogeneity for what concerns sociolinguistic factors. All participants were native of the same linguistic area (Campania region) and shared the same Italian regional variety (Neapolitan Italian). Acquisition onset is comparable since all informants resided in France for a similar number of years (from 3 to 5) and since they reported a mainly spontaneous acquisition of the French language, starting from the immigration in the country. In addition, all participants inhabited within the Aix-Marseille conurbation, a vast urbanized area that, although organized around two independent cities

² A fourth sound, the back unrounded open vowel /a/, equally suitable for the analysis, has been excluded since its phonemic status in modern spoken French is disputed (see Calliope, 1989: 7). Such observation appears to be consistent with the sound distribution in the analyzed corpus.

(Marseille and Aix-en-Provence), is characterized by inextricable cultural, social, and business connections; sample constructions also consider a balance for what concerns sex (F = 5; M = 6). The participants' salient characteristics are summarized in Table 2.

Table 2 - *Participants of the study. Yrs: number of years lived in France; Acq: French acquisition type; 1 = mainly spontaneous, 2 = completely spontaneous*

N°	SEX	AGE	CITY	NATIVE OF	YRS	ACQ
1	F	33	Aix-en-Pce	Naples	4	2
2	F	28	Marseille	Salerno	3	1
3	M	29	Aix-en-Pce	Avellino	3	2
4	M	32	Aix-en-Pce	Naples	5	1
5	F	28	Aix-en-Pce	Naples	3	2
6	M	31	Marseille	Naples	3	2
7	M	32	Aix-en-Pce	Salerno	5	1
8	M	30	Marseille	Naples	3	2
9	M	32	Marseille	Naples	5	2
10	F	28	Marseille	Naples	3	2
11	F	29	Aix-en-Pce	Naples	4	1

4.2.2 The *warmth/competence* questionnaire

The first part of the data collection protocol aims at collecting psychometric data regarding participants' stereotypical beliefs regarding the French-speaking host community. With this goal, a Likert-like questionnaire was developed, containing 18 items associated with a six-point scale (from *completely disagree* to *completely agree*).

Twelve of the items contain statements finalized at obtaining the subject stereotypic evaluation: the statements are equally divided to address both the STM dimensions (*warmth* and *competence*) and having similar formulation ("I think French people are _"³). In each item, the sentence is filled with a classical attribution of *warmth* (*sympathy, kindness, friendliness, altruism, pleasantness, willingness to help*) or of *competence* dimension (*competence, intelligence, skillfulness, erudition, ability, bossiness*). Six more items were intended to act as fillers and regarded general issues about French cultural fields.

The scoring of answers' values (from 0 to 5 points) is assumed as the overall questionnaire score: therefore, each questionnaire can score up from 0 to 30 points, with an average value of 15, for each of the investigated dimensions. The questionnaire results were tested for reliability, returning a more than acceptable internal consistency ($\alpha = 0.8$)⁴. Table 3 shows scores, means, and standard deviations reported by each participant in the questionnaire.

³ In Italian "*Penso che i francesi siano _*".

⁴ It is to be reported that the Cronbach alpha test is reputed not completely trustworthy with small samples (see Yardagül, 2008).

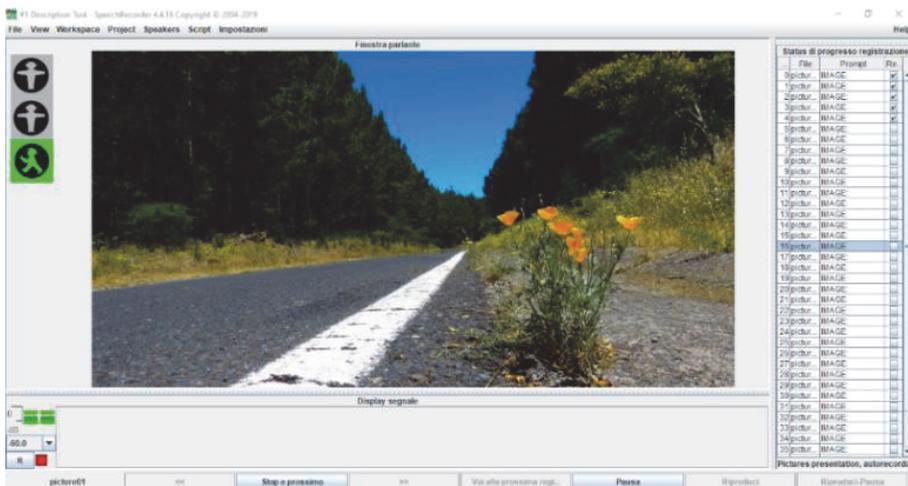
Table 3 - Scores, means, and standard deviation of the warmth/competence questionnaire for each participant and overall measures

Subject	WARMTH		COMPETENCE	
	Score	M (SD)	Score	M (SD)
1F	17	2.83 (.41)	23	3.83 (.75)
2F	13	2.17 (.41)	22	3.67 (.52)
3M	15	2.50 (.84)	24	4.00 (.89)
4M	16	2.67 (.52)	25	4.17 (.41)
5F	13	2.17 (.41)	26	4.33 (.52)
6M	15	2.50 (.55)	22	3.67 (.52)
7M	9	1.50 (.55)	22	3.67 (.51)
8M	6	1.00 (.63)	22	3.67 (.52)
9M	11	1.83 (.41)	21	3.50 (.55)
10F	9	1.50 (.55)	21	3.50 (.55)
11F	12	2.00 (.63)	25	4.17 (.75)
M (SD)	12.36 (3.38)	2.06 (.56)	23 (1.73)	3.83 (.29)

The questionnaires' results highlight two distinct tendencies for what concerns *warmth* and *competence* stereotypes. While the *warmth* scores' mean falls under the possible average value, the *competence* scores' mean remains above them: to reframe such a scenario in the SCM, participants seem to generally evaluate French hosts as a "low *warmth*, high *competence*" outgroup, perceiving it as a competitive/high-status group. That said, it is possible to observe a considerable individual variation in evaluations, especially for what concerns the *warmth* dimension.

4.2.3 Oral corpus and processing of the phonetic data

After completing the questionnaire, each participant was involved in a set of recordings aiming at collecting spontaneous speech in French and Italian. In order to buffer forms of attunement in French elicitation (an Italian-speaking experimenter attend to the protocol), the first phase of the collection consisted of a multiple picture-description task (see Mackey, Gass, 2005: 76-80): informants visualized a set of 45 pictures (randomized in each session) and were invited to describe it in French. Pictures were submitted through the SpeechRecorder platform (Draxler, Jänsch, 2004) and displayed on a 13" laptop placed in front of the informant in a quiet room (see Fig. 2). A total of 1 32' 11" of French speech was collected (~12' for each speaker).

Figure 2 - Screenshot of a single picture-description task with the SpeechRecorder platform⁵

The phase devoted to the Italian speech collection was organized as a free conversation, partially imprinted on the *network of modules* model reported in Labov (1981: 35). The main topics involved in the conversation regards: a) informants' life in France, b) work conditions in Italy and France, c) widespread beliefs towards Italians and French, d) stereotypes towards French; e) French and Italian acquaintances in Aix/Marseille, f) personal point of views on specific topics linked with the informant's work/hobbies. The total Italian speech collected amounts to 1 39' 25" (~13' for each speaker). A Logitech USB H650e served as a recording device for both settings.

The segmentation and annotation phase of the corpus was conducted in four steps: 1) automatic speech recognition and orthographic transcription with BAS's ASR utility; 2) human control and manual correction of the output; 3) phonetic segmentation of speech with BAS's WebMAUS Basic tool (Kisler, Reichel & Schiel, 2017), using the ASR transcription and the audio file as inputs; 4) human control and manual correction of the WebMAUS output, using the *TextGrid+sound* visualization in Praat (Boersma, 2001). Vowel spectral measures were then extracted using a Praat script written for the purpose. The phonetic data filtering for connected speech raises a series of issues linked with the high variability that this mode entails, both inter- (anatomical constraints, environmental conditions) and intra-speaker (speech rate, prosodic context, and voice quality) in nature. In order to grant data reliability, and to not arbitrarily intervene on the dataset, a full-automatic approach for outliers filtering was borrowed from Sandoval, Berisha, Utianski, Liss, and Spanias (2013). Sandoval and colleagues' method involves Gaussian Mixture

⁵ As an exemple, we report the speech elicited through this picture: *C'est [hesitation] c'est l'image d'une route, près du côté droit de la rue, d'un bois [pause], on voit [hesitation] on voit la ligne de démarcation en premier plan, on voit aussi des fleurs jaunes et il sont entourées par [hesitation] par des arbres.* (This is the picture of a road, taken from the right side of the road, from a wood, we see the dividing line in the foreground, we also see some yellow flowers, and they are surrounded by trees).

Model's (GMM) parameters to determine the *maximum likelihood fit* of the observed data; then, the fit probability of each datapoint is estimated: datapoints with low fit values are discarded as outliers. GMM approach was independently applied to female and male speech, with the following rejection scores⁶: Italian speech (females); /i/ = 195/1442 (13.5%), /e/ = 74/1280 (5.8%), /ɛ/ = 49/600 (8.2%), /ɔ/ = 47/445 (10.6%), /o/ = 136/1336 (10.2%), /u/ = 61/336 (18.2%). Italian speech (males); /i/ = 296/2118 (13.8%), /e/ = 238/2058 (11.6%), /ɛ/ = 91/680 (13.4%), /ɔ/ = 82/527 (15.5%), /o/ = 276/1893 (14.6%), /u/ = 141/559 (25.2%). French speech (females); /y/ = 32/331 (9.6%); /ø/ = 12/204 (5.8%); /œ/ = 1/144 (1.4%). French speech (males); /y/ = 94/552 (17%); /ø/ = 11/282 (3.9%); /œ/ = 4/172 (2.3%)⁷. Table 4 summarizes the mean values of the first three formants for each involved vowel (according to the speaker's sex), as measured from the filtered dataset.

Table 4 - Measured mean values of F1, F2, and F3 (in Hz) for each vowel involved in the study, according to speaker sex. The grayed-out text reports mean values as registered in reference literature for Italian (Cosi, Ferrero & Vagges, 1995) and French vowels (Gendrot, Adda-Decker, 2005)

	i	e	ɛ	y	ø	œ	o	ɔ	u
F1	405.39	505.59	592.23	413.55	457.04	609.57	539.41	602.04	457.08
<i>female</i>	339	436	630	371	420	436	688	506	360
F1	339.39	424.36	475.41	358.63	382.43	474.85	461.47	502.67	414.73
<i>male</i>	291	394	513	336	384	400	552	447	325
F2	2064.50	1825.65	1760.63	1775.55	1619.41	1630.92	1351.17	1277.24	1402.36
<i>female</i>	2672	2508	2302	2063	1693	1643	1115	990	838
F2	1881.13	1666.81	1670.34	1571.86	1416.78	1329.87	1167.01	1107.25	1258.23
<i>male</i>	2251	2082	1989	1803	1474	1445	949	856	789
F3	2820.89	2614.60	2569.75	2675.48	2577.83	2530.64	2620.79	2568.53	2567.58
<i>female</i>	3595	3158	2999	2745	2687	2715	2712	2606	2466
F3	2505.52	2357.58	2374.71	2350.66	2342.51	2365.77	2303.85	2278.42	2361.92
<i>male</i>	3079	2752	2669	2545	2687	2440	2569	2528	2529

4.3 Data analysis

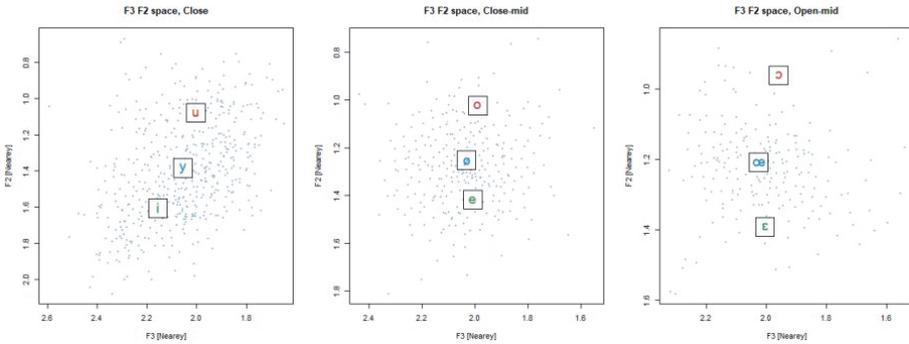
Before proceeding with the data analysis, the collected formant values were normalized with the *Nearey2* algorithm (Nearey, 1977) to make females' and males' datasets comparable in the same metric space. To better assess the categorization of Italian vowels in the F3 × F2 space, the centroid of each vowel was defined for each speaker individually. For what concerns the French target vowels, instead, it seemed appropriate to consider the data distribution in its entirety as to minimize the impact on its representativity. Similarly, in the one-dimensional space (inde-

⁶ The model was built using the *mclust* package for R.

⁷ The generally lower rejection scores of the French vowels are probably due to the semi-spontaneous picture-description setting, that elicited a more careful speech.

pendent F2 and F3 values), the Italian vowels are characterized using the F2 and F3 means, while French targets' formant values are considered in their full distributions. Figure 2 graphically summarizes the relative positions of vowels' centroid and targets' full distributions.

Figure 3 - Centroids of all the analyzed vocalic segments and French targets distribution in the $F3 \times F2$ space



To quantify the distance between the speakers' vowel categories in L1 and L2 speech, the Mahalanobis distance between Italian vowel centroids and French target distributions was computed for each speaker separately vowel class (close, close-mid, and open-mid). Mahalanobis distance was preferred to the traditional Euclidean distance since it allows a more trustworthy measure considering the distribution's shape (see Di Benedetto, 1989); the computed Mahalanobis distances are summarized in Table 5⁸.

As to explore the role of *warmth* and *competence* in language production accuracy, a linear regression model was used, taking as dependent variable the computed Mahalanobis distances (independently for each parameter) and, as independent variables, both stereotypical dimensions (*warmth* and *competence*) and main acquisitional variables (age, years in France, and acquisition type). None of the 18 models fully satisfies the significance threshold; that said, two of them present p -values that are very close to the value of significance (models far from the significance threshold will not be furtherly discussed).

⁸ The Mahalanobis distances were computed using the *StatMatch* package for R.

Table 5 - Mahalanobis distances for the involved parameters, divided by speakers and rounded up to the second decimal place

	$/y - i/_{F_2}$	$/y - i/_{F_3}$	$/y - i/_{F_3 \times F_2}$	$/y - u/_{F_2}$	$/y - u/_{F_3}$	$/y - u/_{F_3 \times F_2}$
1F	3.17	1.00	1.33	1.58	1.14	1.79
2F	0.97	1.31	1.34	3.93	1.84	3.13
3M	3.96	2.30	1.60	1.32	1.43	2.05
4M	1.91	1.36	1.47	2.09	0.99	2.47
5F	1.38	0.97	1.36	2.67	2.45	2.49
6M	2.32	1.36	1.64	2.75	1.00	2.56
7M	1.72	1.54	1.71	3.27	1.00	3.11
8M	2.96	2.20	1.81	1.33	1.23	1.98
9M	1.74	1.51	1.65	3.13	1.57	2.96
10F	1.57	1.13	1.55	2.21	1.79	2.25
11F	1.36	1.70	1.79	2.46	0.97	2.49
	$/o - e/_{F_2}$	$/o - e/_{F_3}$	$/o - e/_{F_3 \times F_2}$	$/o - o/_{F_2}$	$/o - o/_{F_3}$	$/o - o/_{F_3 \times F_2}$
1F	1.93	0.75	1.36	0.95	1.14	1.72
2F	3.66	1.20	1.33	1.58	0.78	1.61
3M	2.75	1.21	1.67	1.72	1.41	1.13
4M	2.51	2.72	1.72	1.97	0.97	2.21
5F	2.08	1.14	1.45	3.00	0.96	1.96
6M	3.12	1.04	1.55	1.01	1.14	1.81
7M	2.06	2.21	1.83	2.96	0.96	2.46
8M	0.83	1.25	1.50	1.63	2.13	2.64
9M	3.25	4.01	2.05	3.82	1.01	2.88
10F	1.72	1.11	1.55	1.79	1.08	2.07
11F	4.21	3.23	1.83	2.38	1.31	2.17
	$/\alpha - \varepsilon/_{F_2}$	$/\alpha - \varepsilon/_{F_3}$	$/\alpha - \varepsilon/_{F_3 \times F_2}$	$/\alpha - \varepsilon/_{F_2}$	$/\alpha - \varepsilon/_{F_3}$	$/\alpha - \varepsilon/_{F_3 \times F_2}$
1F	5.21	0.89	1.12	1.95	3.37	1.19
2F	4.12	1.32	1.8	2.12	2.55	1.54
3M	6.08	0.71	1.06	4.71	0.66	1.69
4M	6.13	1.21	1.42	4.1	1.87	2.66
5F	3.55	1.44	1.65	2.86	2.91	3.15
6M	4.13	0.97	1.52	2.73	0.95	2.16
7M	2.68	1.34	1.31	1.41	0.99	1.88
8M	6.34	0.74	1.42	4.11	0.77	2.78
9M	4.14	1.43	1.54	5.16	1.67	2.52
10F	3.15	2.77	1.31	4.22	2.26	6.35
11F	3.46	1.09	2.02	6.41	0.92	3.03

One of these (having as dependent variable the F2 distance between close-mid rounded front /ø/ and back /o/ vowels, $F(5) = 4.62$, $p = 0.059$, $R^2 = 0.82$, Adj. $R^2 = 0.05$) does not show a significant role of stereotypical dimensions (*warmth*, $p = 0.37$; *competence*, $p = 0.64$) in values fluctuations, and therefore it will not be further discussed.

The model involving the distance in bidimensional $F3 \times F2$ space for the same vocalic couple ($F(5) = 4.55$, $p = 0.06$, $R^2 = 0.82$, Adj. $R^2 = 0.64$) shows the most interesting results, highlighting the significant role of the independent variable “warmth” ($p = 0.02$, coef. = -0.06) in shaping the data distribution (no other variable presents significant p -values). The slight negative correlation between *warmth* dimension and Mahalanobis distances for the /ø/ and /o/ vocalic sounds is confirmed in the model having F3 distances as a dependent variable, that suggests, albeit not significantly ($F(5) = 4.55$, $p = 0.2$, $R^2 = 0.67$, Adj. $R^2 = 0.35$), a significant role of the *warmth* dimension and a negative correlation between the two variables ($p = 0.04$, coef. = -0.09).

In order to confirm the results produced through the linear regression, a more specific mixed-model analysis was performed on the relation between *warmth* dimension and /ø – o/ distance in the $F3 \times F2$ space⁹, with the acquisitional variables assumed as random intercepts. An ANOVA carried out between mixed models with and without the *warmth* variable as fixed effects produces a robustly significant p -value ($\chi^2(1) = 8.42$, $p = 0.003$), confirming the effect of *warmth* on distances between the vowels’ formant parameters.

5. Conclusions

With all due caution that an explorative work imposes, the data analysis seems to offer encouraging suggestions about the effects of migrants’ stereotypical beliefs in second language outputs. Statistical tests, in particular, confirm the existence of an effect – albeit slight – of the stereotypical evaluation of *warmth* dimension on the similarity between one of the involved French vocalic targets – the front, close-mid round vowel – and one of the two more similar Italian sounds in terms of phonetic traits (specifically, the *back close-mid* rounded vowel).

The hypothesis stated in § 4.1 – in which it is assumed that better stereotypical beliefs correspond to a more native-oriented L2 production – receive therefore only a partial satisfaction, limited to the stereotypical dimension of *warmth*. A similar result is particularly interesting in light of recent experimental findings (Futamura, 2017; Oldmeadow, 2018) emphasizing the role of *warmth* evaluations in determining more prosocial behaviors. In this frame, the significant role of *warmth* in L2 outputs can be acquisitionally interpreted as a consequence of more frequent interactions within the host community.

⁹ Mixed model analysis was performed using the *lme4* package for R.

On the other hand – and even if the statistical non-significance *per se* does not allow to exclude them from the influent variables – it is to some extent surprising that a positive evaluation toward hosts' *competence* does not produce significant effects in orienting a more target-like L2 acquisitional output. A similar incongruence can be opportunely explained by involving an additional psychosocial model: the BIAS map (*Behaviors from Intergroup Affect and Stereotypes Map*; Cuddy, Fiske & Glick's, 2008), a comprehensive framework – borrows its premises from the SCM – aimed at modeling the affective and behavioral consequences of intergroup classification. In the BIAS map systematization, it is postulated that, in the presence of contrasting *warmth/competence* evaluations, subjects' *active behaviors* (and, specifically, social interactivity) follows *warmth* judgments orientation (in contrast, *competence* evaluation are more influential on passive conducts). This considering – and emphasizing the need for broader experimental evidence – the BIAS map model seems to constitute a useful tool to interpret the acquisition-related individual variation in migrants' L2 outputs.

Acknowledgments

The author would like to acknowledge the Institute of Italian Culture of Marseille and its director, Dr. Francesco Neri, that was of great help in the participant recruitment phase and that made its spaces available to perform some of the recordings.

Bibliography

- ADSERA, A., PYTLIKOVA, M. (2015). *The role of language in shaping international migration*. London: Centre for Research and Analysis of Migration, University College London.
- ALLEN, T.J., SHERMAN, J.W., CONREY, F.R., & STROESSNER, S.J. (2009). Stereotype Strength and Attentional Bias: Preference for Confirming versus Disconfirming Information Depends on Processing Capacity. In *Journal of Experimental Social Psychology*, 45(5), 1081-1087.
- ALLPORT, G.W. (1935). Attitudes. In MURCHINSON, C. (Ed.), *A Handbook of Social Psychology*, Worcester MA: Clark University Press, 798-844.
- BELOT, M., EDERVEEN, S. (2012). Cultural Barriers in Migration between OECD Countries. In *Journal of Population Economics*, 25, 1-29.
- BERRY, J.W. (1997). Immigration, Acculturation, and Adaptation. In *Applied Psychology*, 46, 5-34.
- BERRY, J.W., KIM, U., MINDE, T., & MOK, D. (1987). Comparative Studies of Acculturative Stress. In *The International Migration Review*, 21(3), 491-511.
- BERTINETTO, P., LOPORCARO, M. (2005). The sound pattern of Standard Italian, as compared with the varieties spoken in Florence, Milan and Rome. In *Journal of the International Phonetic Association*, 35(2), 131-151.
- BEST, C., TYLER, M.D. (2007). Nonnative and second-language speech perception: Commonalities and complementarities. In BOHN, O.S., MUNRO, M.J. (Eds.), *Language*

Experience in Second language Speech Learning. In Honor of James Emil Flege, Amsterdam: John Benjamins, 13-34.

BEST, C.T. (1995). A direct realist perspective on cross-language speech perception. In Strange, W. (Ed.), *Speech Perception and Linguistic Experience: Issues in Cross-language Research*, Timonium MD: York Press, 167-200.

BOERSMA, P. (2001). *Praat, a system for doing phonetics by computer*. In *Glott International*, 5(9), 341-345.

CALLIOPE, 1989. *La parole et son traitement automatique*. Paris: Masson.

CANCINO, H., ROSANSKY, E.J. & SCHUMANN, J.H. (1975). The Acquisition of the English Auxiliary by Native Spanish Speakers. In *TESOL quarterly*, 9(4), 421-430.

CHISWICK, B.R. (1998). Hebrew Language Usage: Determinants and Effects on Earnings Among Immigrants in Israel. In *Journal of Population Economics*, 11(2), 253-271.

CHISWICK, B.R., MILLER, P.W. (1995). The Endogeneity between Language and Earnings: International Analyses. In *Journal of Labor Economics*, 13(2), 246-288.

CHISWICK, B.R., MILLER, P.W. (1999). Immigration, Language and Multiculturalism in Australia. In *Australian Economic Review*, 32, 369-385.

CHISWICK, B.R., LEE, Y.L., MILLER, P.W. (2005). Immigrant earnings: a longitudinal analysis. In *The review of Income and Wealth*, 51, 485-503.

CLARK, A.E., KASHIMA, Y. (2008). Stereotypes Help People Connect with Others in the Community: A Situated Functional Analysis of the Stereotype Consistency Bias in Communication. In *Journal of Personality and Social Psychology* 93(6), 1028-1039

COSI, P., FERRERO, F.E. & VAGGES, K. (1995). Rappresentazioni acustiche e uditive delle vocali italiane. In *Proceedings of the 23rd National Conference AIA*, Bologna, Italy, 29-31 May 1995. 151-156.

CUDDY, A.J.C., FISKE, S.T., GLICK, P. (2007). The BIAS Map: Behaviors from Intergroup Affect and Stereotypes. In *Journal of Personality and Social Psychology*, 92(4), 631-648.

CUMMINS, J. (1979). Cognitive/Academic Language Proficiency, Linguistic Interdependence, the Optimum Age Question and Some Other Matters. In *Working Papers on Bilingualism*, 19, 199-205.

DANZER, A.M., YAMAN, F. (2011). *Ethnic concentration and language fluency of immigrants in Germany*. London: Department of Economics, City University London.

DI BENEDETTO, M. (1989). Vowel representation: Some observations on temporal and spectral properties of the first formant frequency. In *The Journal of the Acoustical Society of America* 86(1), 55-66.

DÖRNEY, Z. (2009). The L2 motivational self system. In DÖRNEY, Z., USHIODA, E. (Eds.), *Motivation, language identity, and the L2 self*. Bristol: Multilingual Matters, 9-42.

DRAXLER, J., JÄNSCH, K. (2004). SpeechRecorder – a universal platform independent multichannel audio recording software. In *Proceedings of 4th International Conference on Language Resource and Evaluation*, Lisbon, Portugal, 26-28 May 2004, 559-562.

DUSTMANN, C., VAN SOEST, A. (2001). Language Fluency and Earnings: Estimation with misclassified language indicators. In *The Review of Economics and Statistics*, 83(4), 663-674.

EL-DASH, L.G., BUSNARDO, J. (2002). Brazilian attitudes toward English: Dimensions of status and solidarity. In *International Journal of Applied Linguistics*, 11(1), 57-74.

- ESSER, H. (2006). *Migration, Sprache und Integration*. Berlin: AKI.
- FISKE, S.T., CUDDY, A.J., GLICK, P., & XU, J. (2002). A model of (often mixed) stereotype content: Competence and warmth respectively follow from perceived status and competition. *Journal of Personality and Social Psychology*, 82(6), 878-902.
- FLEGE J.E. (2007), "Language contact in bilingualism: Phonetic system interactions". In COLE J., HUALDE J. (Eds), *Laboratory Phonology 9*. Berlino/New York: Mouton de Gruyter, 353-380.
- FLEGE, J.E. (1995). Second language speech learning: Theory, findings, and problems. In STRANGE, W. (Ed.), *Speech Perception and Linguistic Experience: Issues in Cross-Language Research*, Timonium MD: York Press, 167-200.
- FORBES, C.E., SCHMANDER, T. (2010). Retraining Attitudes and Stereotypes to Affect Motivation and Cognitive Capacity Under Stereotype Threat. In *Journal of Personality and Social Psychology*, 99(5), 740-754
- FRONTERA, M., PAONE, E. (2018). Prosodic accuracy and foreign accent in cultural migrants. In *Proceedings of 9th Tutorial and Research Workshop on Experimental Linguistics*, Paris, 28-30 August 2018.
- FUTAMURA, I. (2017). The influence of prosocial behaviors on evaluations of morality and warmth. *Shinrigaku Kenkyu: The Japanese Journal of Psychology*, 87(6), 590-599.
- GARDNER, R., LALONDE, R., & MOORCROFT, R. (1985). The role of Attitudes and Motivation in Second Language Learning: Correlation and experimental considerations. In *Language Learning*, 35, 207-227.
- GENDROT, C., ADDA-DECKER, M. (2005). Impact of duration on F1/F2 formant values of oral vowels: an automatic analysis of large broadcast news corpora in French and German. In *Proceedings of the 9th Interspeech*, Lisbona, Portugal, 4-8 September 2005.
- GOLDSTEIN, T. (1996). *Two Languages at Work: Bilingual Life on the Production Floor*. Berlin/New York: Mouton de Gruyter.
- HAMMOND, R.M. (1990). The Affective Filter and Pronunciation Proficiency – Attitudes and Variables in Second Language Acquisition. In ARENA, L.A. (Ed.), *Language Proficiency*. Boston, MA: Springer, 65-71.
- HELLER, M. (1987). The role of language in the formation of ethnic identity. In PHINNEY, J., ROTHERAM, M. (Eds.), *Children's Ethnic Socialization*. Newbury Park, CA: Sage, 180- 200.
- KRASHEN, S.D. (1987). *Principles and Practice in Second Language Acquisition*. Upper Saddle River, NJ: Prentice Hall.
- KRASHEN, S.D. (1988). *Second Language Acquisition and Second Language Learning*. Upper Saddle River, NJ: Prentice Hall.
- LABOV, W. (1981). Field methods used by the project on linguistic change and variation. In *Sociolinguistic Working Paper*, 81, 1-43.
- LAMMERT, A., PROCTOR, M.I., NARAYANAN, S.S. (2013). Interspeaker Variability in Hard Palate Morphology and Vowel Production. In *Journal of Speech Language and Hearing Research*, (56)6, 1924-1933.
- LEE, T.L., FISKE, S.T. (2006). Not an outgroup, not yet an ingroup: Immigrants in the Stereotype Content Model. In *International Journal of Intercultural Relations*, 30(6), 751-768.

- MACKEY, A., GASS, S.M. (2005). *Second language research: Methodology and design*. Mahwah, NJ: Lawrence Erlbaum.
- MAEDA, S. (1990). Compensatory articulation during speech: Evidence from the analysis and synthesis of vocal tract using an articulatory model. In HARDCASTLE, W.J., MARCHAL, A. (Eds.) *Speech production and speech modeling*. Amsterdam: Springer, 131-149.
- MAROTTA, G., BOULA DE MAREÜIL, P. (2009). Persistenza dell'accento straniero. Uno studio percettivo sull'italiano L2. In SCHMID, S., SCHWARZENBAC, M., & STUDER, D. (Eds.). *Proceedings of the 5th International Congress AISV*, Zürich, Switzerland, 4-6 February 2009, 475-494.
- MCGARTY, C. (1999). *Categorization in Social Psychology*. London: Sage.
- MCGARTY, C., YZERBYT, V.Y., & SPEARS, R. (2002). *Stereotypes as explanations: The formation of meaningful beliefs about social groups*. Cambridge: Cambridge University Press.
- MOYER, A. (2007). Do language attitudes determine accent? A study of bilinguals in the USA. In *Journal of multilingual and multicultural development*, 28(6), 502-518.
- NEAREY, T.M. (1977). *Phonetic Feature Systems for Vowels*. PhD Dissertation, University of Alberta.
- NIKITINA, L., DON, Z.M., LOH, S.C. (2014). Focus on Brazil: Country images held by Malaysian learners of Brazilian Portuguese. In *Calidoscopio*, 12(1), 78-82.
- NORTON, B. (2000). *Identity and Language Learning: Gender, Ethnicity and Educational Change*. Harlow: Longman.
- OLDMEADOW, J.A. (2018). Stereotype content and morality: How competence and warmth arise from morally significant interactions. In *Social Psychology*, 57(4), 834-854.
- PERDUE, C. (1993). *Adult language acquisition: Cross-linguistic perspectives. Volume 1, Field methods*. Cambridge: Cambridge University Press.
- SANDOVAL, S., BERISHA, V., UTIANSKI, R., LISS, J., & SPANIAS, A. (2013). Automatic assessment of vowel space area. In *The Journal of the Acoustical Society of America*, 134(5), 477-483.
- SCHUMANN, JH (1978). An Acculturation Model for Second Language Acquisition. *Journal of Multilingual and Multicultural Development*, 7(5), 379-392.
- SKUTNABB-KANGAS, T. (2015). Linguicism. In CHAPELLE, C.A. (Ed.), *Encyclopedia of Applied Linguistics*. Hoboken, NJ: Wiley, 67-75.
- Smith, L.R. (2002). The social architecture of communicative competence: A methodology for social-network research in sociolinguistics. *International Journal of the Sociology of Language*, 153, 133-160.
- SORNICOLA, R. (1997). Campania. In MAIDEN, M., PARRY, M. (Eds.), *The Dialects of Italy*. Londra/New York: Routledge, 330-337.
- TAJFEL, H. (1974). Social identity and intergroup behavior. In *Social Science Information*, 13(2), 65-93.
- TRANSEL, B. (1987). *The sounds of French: an introduction*. Cambridge: Cambridge University Press.
- VAYRA, M., AVESANI, C., BEST, C. & BOHN, O. (2012). Non solo dettagli fonetici, non solo categorie fonologiche: l'interazione tra fonetica e fonologia nella percezione di suoni non-nativi. In *Studi e saggi linguistici*, 5, 119-148.

- WIKLUND, I. (2002). Social networks from a sociolinguistic perspective: The relationship between characteristics of the social networks of bilingual adolescents and their language proficiency. In *International Journal of the Sociology of Language*, 2002(153), 53-92.
- WOOD, G. (2008). Gender stereotypical attitudes: Past, present and future influences on women's career advancement. In *Equal Opportunities International*, 27(7), 613-628.
- YARDAGÜL, H. (2008). Minimum sample size for Cronbach's coefficient alpha: A Monte-Carlo study. In *Hacettepe Üniversitesi Journal of Education*, 35, 397-405.
- YILMAZ, G., SCHMID, M.S. (2015). Second language development in a migrant context: Turkish community in the Netherlands. In *International Journal of the Sociology of Language*, 236, 101-132.
- LITTLEWOOD, W.T. (1981). Language Variation and Second Language Acquisition Theory. In *Applied Linguistics*, 2, 150-158.
- TSIMPLI, I. (2006). Variation in Second Language Acquisition. In BROWN, K. (Ed.), *Encyclopedia of language and linguistics*. Oxford: Elsevier, 387-394.
- DRUMMOND, R. (2012). Aspects of identity in a second language: ING variation in the speech of Polish migrants living in Manchester, UK. In *Language Variation and Change*, 24, 107-133.
- SCHLEEF, E. (2013). Migrant teenagers' acquisition of sociolinguistic variation: the variables (ing) and (t). In AUER, P., CARO, J. & KAUFMANN, G. (Eds.), *Language variation – European perspectives IV*. Amsterdam: Benjamins, 201-213.

ANCUTA BUDEANU, ANNA DE MEO, MASSIMO PETTORINO

Caratteristiche fonetiche dell'italiano di romeni in Calabria. Riflessioni su ritmo e lunghezza vocalica e consonantica

Phonetic characteristics of Italian spoken by Romanians in Calabria. Reflections on rhythm and vocalic and consonant length

The present work aimed to explore some suprasegmental features of L2 Italian of Romanian immigrants residing in Calabria region (southern Italy), compared with L1 Romanian. For the experimental analysis, a model based on two parameters was used, the percentage of vowel portion (%V) and the average value of the interval between two consecutive vowel onsets (VtoV). The analysis focused on the rhythmic-prosodic aspects, with particular attention to the vowel and consonantal length. As for L2 Italian, the results highlight how it has rhythmic characteristics different from both the target language and the mother tongue, thus excluding the possibility of a transfer. Furthermore, the data show how the lack of constant use of the mother tongue also creates changes on the rhythmic-prosodic level in one's L1.

Keywords: L2 Italian, L1 Romanian, prosody, rhythmic competence, vowel and consonantal length.

1. Introduzione

Secondo i dati Istat aggiornati al 1° gennaio 2020, la comunità romena rappresenta quasi un quarto di tutti gli immigrati residenti in Italia (23%), seguita a molta distanza dalle comunità albanese (8%), marocchina (8%) e cinese (6%).

Nonostante questa presenza abbia caratterizzato in modo significativo il fenomeno migratorio in Italia negli ultimi dieci anni, soprattutto a seguito dell'ingresso della Romania nell'Unione Europea nel 2007, gli studi specificamente dedicati a questo gruppo sono relativamente pochi, anche se la comunità romena compare in vari lavori come gruppo di riferimento. Accanto a qualche lavoro di carattere sociologico e demografico (Caritas, 2010; Cingolani, 2009; Valtolina, 2012), si possono citare anche alcune pubblicazioni di carattere linguistico, dedicate alle problematiche dell'italiano L2 di romeni o al romeno L1 in condizioni di contatto (Cohal, 2014; Chini, Andorno, 2018; Galatà et al. 2012, 2017; Romito, Tarasi, 2012; Romito et al., 2012; Romito et al., 2014).

Lo studio qui presentato si è concentrato sulla comunità romena della Locride, regione della Calabria ionica, nella quale i romeni rappresentano il 27,46% della po-

polazione straniera residente. Gli immigrati romeni di quest'area sono caratterizzati da una omogeneità di provenienza, venendo tutti da alcuni centri urbani del nord-est della Romania (Moldavia), regione comparabile a quella calabrese di riferimento per la particolare fragilità nello sviluppo demografico ed economico.

I romeni della presente indagine non hanno creato sul territorio calabrese una rete sociale ampia e coesa e mancano centri di aggregazione e associazioni, ad eccezione del piccolo e isolato monastero di San Giovanni Theristis, nel comune di Bivongi, che dal 2008 funge da centro spirituale dei romeni ortodossi.

2. *Lo studio*

2.1 Obiettivi

Il presente lavoro ha inteso esplorare alcune caratteristiche soprasedgmentali del parlato italiano e romeno prodotto da immigrati romeni adulti residenti nella Locride. I dati raccolti sono stati messi a confronto con i risultati emersi da analisi comparabili condotte sull'italiano L1 di nativi residenti nella Locride e sul romeno L1 di soggetti residenti in Romania. Oggetto specifico dell'indagine è stata la competenza ritmica e la gestione della lunghezza vocalica e consonantica.

2.2 Partecipanti

Ai fini della ricerca sono stati coinvolti 60 soggetti, maschi e femmine, con un'età compresa tra 30-35 anni, suddivisi in sei gruppi di 10 persone, sulla base di tre caratteristiche:

1. lingua materna (italiano o romeno),
2. titolo di studio (diploma o laurea),
3. immigrato/non immigrato (solo per i partecipanti romeni).

Gli italiani coinvolti nello studio, tutti residenti nella Locride (Calabria) hanno dichiarato di non avere alcuna competenza di romeno; i romeni non immigrati, tutti residenti in Moldavia (Romania), hanno dichiarato di non conoscere l'italiano. Al momento della raccolta dei dati tutti gli immigrati, anche essi originari della Moldavia (Romania), risiedevano sul territorio calabro della Locride da 10 anni. Gli immigrati già laureati nel proprio paese di origine hanno dichiarato di essere iscritti a un nuovo corso di laurea in Italia.

Tra gli immigrati diplomati e laureati emerge una differenza nella modalità di acquisizione dell'italiano, che si presenta come spontanea per il gruppo di immigrati diplomati e come mista, spontanea e guidata, per il gruppo degli immigrati laureati. I primi hanno dichiarato di non aver frequentato alcun corso di lingua e di aver imparato la lingua italiana esclusivamente attraverso l'uso, per immersione quotidiana in situazioni comunicative autentiche legate ai contesti lavorativi; i secondi, per necessità legate allo studio universitario, hanno avuto bisogno di frequentare corsi specifici di lingua italiana.

2.3 Metodologia e strumenti di ricerca

La ricerca ha previsto la somministrazione ai romeni immigrati in Italia di un questionario di rilevazione dei dati linguistici (scolarizzazione in Romania e in Italia, lingua materna, lingue conosciute, contesti d'uso delle diverse lingue).

Successivamente, tutti i soggetti coinvolti nello studio hanno letto e registrato in italiano e/o in romeno il testo *Il vento di tramontana e il sole* (40 in italiano, di cui 20 in italiano L1 e 20 in italiano L2; 40 in romeno L1, di cui 20 eseguite dai romeni residenti in Romania e 20 da quelli immigrati in Italia).

Il corpus così costruito è stato utilizzato sia per realizzare un test percettivo, finalizzato alla valutazione qualitativa dell'accento italiano e romeno da parte di ascoltatori nativi, sia per procedere a un'analisi acustica mediante software Praat (Boersma – Weenink 2020).

2.3.1 Questionario per la rilevazione dei dati linguistici

Al fine di ottenere una fotografia linguistica di ciascun soggetto immigrato coinvolto nello studio, è stato somministrato un questionario che ha rilevato le esperienze di formazione in Romania e in Italia, le lingue conosciute e quelle attivamente utilizzate, il livello di competenza posseduto (su una scala a 4 livelli: benissimo/bene/abbastanza bene/male), i contesti d'uso delle varietà del repertorio.

Tutti i 20 immigrati hanno dichiarato una eccellente competenza del romeno lingua materna (100% benissimo), un buon livello di competenza dell'italiano (bene 50%; benissimo 50%) e una competenza limitata del dialetto calabrese (50% abbastanza bene; 50% male). Tra i laureati, 7 hanno dichiarato di conoscere anche altre lingue europee, apprese a scuola e/o all'università (4 francese, 2 inglese e 1 tedesco e inglese), per le quali hanno dichiarato una competenza linguistica per tali lingue prevalentemente buona e solo in un paio di casi negativa.

Per quanto riguarda l'uso attivo delle varietà del repertorio individuale, nessuno dei soggetti intervistati ha riportato le lingue straniere della formazione scolastica, ma hanno fatto riferimento esclusivamente a quelle presenti nell'uso attivo quotidiano, ossia il romeno, l'italiano e il dialetto calabrese.

L'italiano è la varietà più utilizzata da parte dei laureati in tutti i contesti d'uso, mentre i diplomati tendono a utilizzare più codici comunicativi (romeno e italiano in famiglia; dialetto e romeno nel tempo libero; dialetto e italiano sul lavoro) con ricorso frequente al code-switching. I dati complessivi attestano una evidente regressione della lingua romena, che lascia spazio all'italiano nell'interazione comunicativa pubblica e privata (Fig. 1 e 2).

Figura 1 - *Uso delle varietà del repertorio linguistico di romeni immigrati diplomati (valori percentuali)*

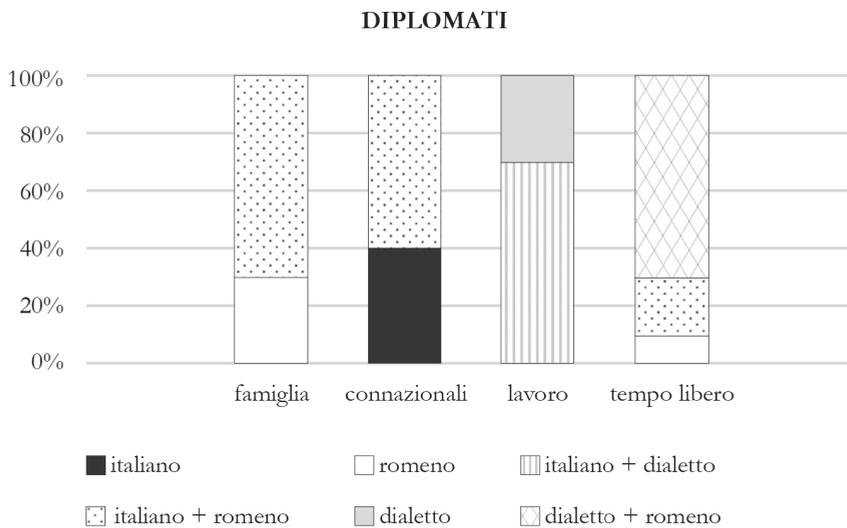
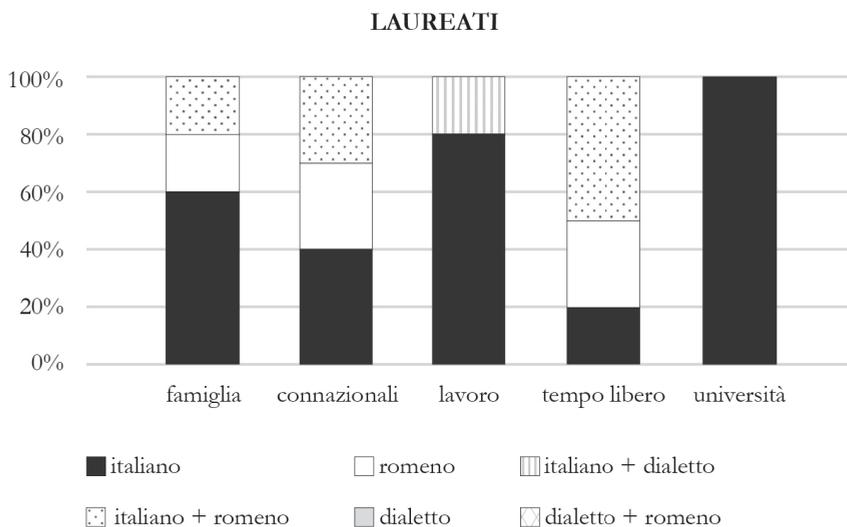


Figura 2 - *Uso delle varietà del repertorio linguistico di romeni immigrati laureati (valori percentuali)*



2.3.2 Test percettivo

Il test percettivo è stato somministrato a 10 parlanti nativi italiani laureati residenti nella Locride e 10 parlanti nativi romeni laureati residenti in Moldavia. Ciascun partecipante ha avuto il compito di ascoltare nella propria lingua materna 40 ripetizioni della frase iniziale del testo *Il Vento di tramontana e il sole* (“si bisticciavano un giorno il vento di tramontana e il sole, l’uno pretendendo d’esser più forte dell’altro”; “între vântul de Nord și soare a izbucnit o dispută, fiecare susținând că el este mai puternic decât celălalt”). Per l’italiano le 40 ripetizioni sono state prodotte da 20 italiani nativi e da 20 romeni parlanti italiano L2; per il romeno le 40 ripetizioni sono state prodotte da 20 romeni residenti in Italia e 20 romeni residenti in Romania. I test di ascolto hanno proposto ordini casuali delle voci selezionate, al fine di evitare effetti del posizionamento sul giudizio dell’ascoltatore.

Gli ascoltatori, ai quali non è stata descritta la natura delle voci che avrebbero ascoltato, hanno ricevuto il compito di attribuire un giudizio di “nativo”-“non nativo” a ciascun enunciato ascoltato e di assegnare un valore all’accento percepito su una scala a 5 punti. La scala rappresenta il grado di vicinanza percepito tra la voce ascoltata e il modello di accento nativo che ciascun ascoltatore possiede (5 valore massimo e 1 valore minimo).

I risultati del test percettivo separano nettamente i parlanti nativi italiani dai locutori stranieri, ma mostrano anche una interessante differenza tra i parlanti nativi romeni residenti in Romania e quelli emigrati da un decennio in Italia.

Per quanto riguarda l’italiano L1, i parlanti sono stati giudicati come “nativi” nella quasi totalità (100% laureati e 98% diplomati), ottenendo dei giudizi molto simili sulla qualità dell’accento, sebbene leggermente migliori nel caso dei laureati (4,1 vs 3,9 attribuito ai diplomati). Questi risultati sono comparabili con quelli ottenuti dai soggetti residenti in Romania, riconosciuti quasi tutti come “nativi” (100% laureati e 97% diplomati) e con buoni giudizi sulla qualità dell’accento, sebbene leggermente migliore per i laureati (4,5 vs 4 attribuito ai diplomati).

Dal test percettivo è emerso che nelle produzioni in italiano L2 tutti i romeni immigrati sono stati percepiti come “non nativi”, ma la vicinanza al modello nativo è stata giudicata in maniera molto diversa, con un punteggio di 3,6 per i laureati e di 1,9 per i diplomati. Questo dato evidenzia un vantaggio dell’apprendimento misto della seconda lingua, in cui si combinano processi spontanei e percorsi formali, guidati (Doughty 2003; Nuzzo-Grassi 2016).

Per quanto riguarda il romeno L1 dei residenti in Italia, la percentuale di soggetti identificati come nativi appare molto più bassa rispetto a quella dei residenti in Romania (72% nel caso dei laureati e il 61% nel caso dei diplomati), come pure la qualità dell’accento stesso (3,9 laureati; 3,1 diplomati).

Questi dati hanno sollecitato una riflessione sulle caratteristiche del parlato dei soggetti emigrati, con l’obiettivo di rilevare gli elementi devianti rispetto al modello nativo che possano aver motivato il giudizio degli ascoltatori. Poiché il parlato letto di questi soggetti risulta privo di errori lessicali, morfologici e fonetico-segmentali, la ricerca si è concentrata sul livello ritmico-prosodico.

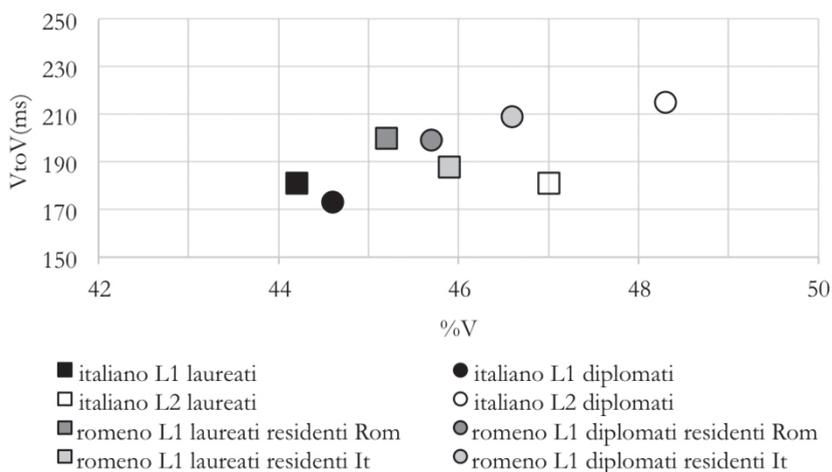
3. *Aspetti ritmico-prosodici*

3.1 La competenza ritmica

Al fine di individuare quali tratti fonetici abbiano potuto determinare il giudizio di “non nativo” per l’italiano L2 e il romeno L1 dei romeni immigrati in Italia, si è deciso di procedere a un’analisi ritmico-prosodica delle loro produzioni utilizzando come modello di riferimento quello proposto da Pettorino et al. (2013). Tale modello, che si focalizza sul modo in cui il parlato viene percepito e non su come esso viene prodotto, si basa su due parametri: la percentuale di porzione vocalica all’interno dell’enunciato (%V) e il valore medio dell’intervallo tra l’inizio di una vocale e l’inizio della vocale successiva (VtoV). Il primo parametro, quello della percentuale vocalica, risulta essere strettamente connesso alla classificazione ritmica delle lingue e tale correlazione gode di ampio consenso nella letteratura scientifica, a partire dal lavoro di Ramus et al. (1999). Il secondo parametro, il VtoV, permette di individuare i punti di discontinuità del segnale acustico, ossia quelli corrispondenti agli attacchi vocalici, caratterizzati da un brusco incremento dell’intensità e per questo motivo alla base della percezione del ritmo. Tale parametro rappresenta dunque la controparte percettiva della velocità di articolazione, generalmente espressa in sillabe al secondo, la quale, contrariamente al VtoV, risulta spesso difficilmente misurabile a causa delle difficoltà di individuazione dei confini sillabici.

I risultati delle analisi effettuate sono sintetizzati nel grafico mostrato in Fig. 3.

Figura 3 - Valori medi di VtoV in ms e %V nell’italiano L1 e L2 e nel romeno L1 in Romania (Rom) e in Italia (It) di laureati e diplomati



I valori relativi all’italiano L1 e al romeno L1 dei parlanti residenti in Romania rientrano all’interno dell’area isosillabica, con una percentuale vocalica che per l’italiano si colloca tra 44,2-44,6% e per il romeno intorno al 45,2-45,7%. Per quanto

riguarda la velocità dell'eloquio, notiamo come l'italiano L1 risulti essere più rapido ($V_{toV}=173 - 181$ ms) rispetto al romeno ($V_{toV}=199 - 200$ ms).

Per quanto riguarda il romeno L1 prodotto dai romeni residenti in Italia, è interessante notare come, sia nel caso dei laureati sia nel caso dei diplomati, siano presenti differenze in termini di %V e V_{toV} rispetto ai parlanti nativi residenti in Romania: la percentuale di vocale aumenta in entrambi i casi (+0,7 per i laureati e +0,9 per i diplomati), mentre il V_{toV} segue un andamento diverso in quanto decresce per i laureati (-12 ms) e aumenta per i diplomati (+10 ms). Questi dati potrebbero giustificare gli esiti del test percettivo che attribuiscono un'identità di "non nativo" romeno al 18% dei laureati e al 39% dei diplomati residenti in Italia.

Per quanto riguarda l'italiano L2 di romeni, è possibile osservare come la percentuale vocalica aumenti sia nel caso dei diplomati, con un valore superiore al 48%, sia nel caso dei laureati con un valore pari al 47%, collocandosi in tal modo in un'area al limite tra la zona isosillabica e quella isomoraica. Per quanto riguarda la velocità dell'eloquio, si nota come il parlato dei laureati abbia un V_{toV} uguale a quello del corrispettivo gruppo dei parlanti nativi italiani ($V_{toV}=181$ ms), mentre il parlato dei diplomati risulti essere l'eloquio più lento di tutto il corpus raccolto ($V_{toV}=215$ ms). I dati dell'italiano L2 sembrano evidenziare come l'esposizione all'input dell'italiano L1 produca caratteristiche ritmico-prosodiche dell'italiano L2 che vanno in direzione opposta non solo rispetto al ritmo della lingua target ma anche rispetto a quello della lingua materna, escludendo dunque l'ipotesi di un transfer. Questo punto verrà approfondito più avanti.

3.2 La competenza prosodico-segmentale

La seconda fase dell'analisi acustica è stata condotta a livello prosodico-segmentale, con particolare attenzione alla lunghezza vocalica e consonantica. Sono state prese in considerazione parole caratterizzate da un nesso VC (dove V è la vocale e C è la consonante breve) e un nesso VCC (dove CC è la consonante lunga). È importante sottolineare che il romeno non presenta l'opposizione funzionale breve/lunga per le consonanti, significativa invece per l'italiano.

Relativamente alla sequenza VC, è stata analizzata la parola "sole", che occorre 4 volte nel testo; le misurazioni effettuate per l'italiano L1 e l'italiano L2 hanno interessato la vocale tonica /o/ e la consonante breve /l/ allo scopo di verificare il rapporto tra la durata della consonante e la durata della vocale. Complessivamente sono state misurate 160 occorrenze della parola "sole", 40 per ciascun gruppo di soggetti coinvolti nello studio. Per operare un confronto tra le due lingue, è stata analizzata per il romeno L1 la parola "omul" (it. "l'uomo"), caratterizzata da vocale tonica /o/ in sillaba aperta seguita da consonante nasale bilabiale; per questa parola, che occorre 2 volte nel testo in romeno, sono state misurate 80 occorrenze.

Dai risultati ottenuti non è emersa una differenza significativa tra l'italiano L1 e l'italiano L2 per quanto riguarda il rapporto di durata tra consonante breve e vocale nel nesso VC (Tabella 1). Infatti, sia in italiano L1 che in italiano L2 la durata della consonante risulta essere circa la metà di quella della vocale ($C/V=0,44$). I dati

relativi al romeno L1 risultano simili, con un rapporto C/V di 0,55 per i romeni residenti in Romania, e 0,50 per quelli residenti in Italia (Tabella 2). La lieve riduzione di durata consonantica da parte dei romeni residenti in Italia è in linea con i dati emersi dall'analisi ritmica che evidenzia un incremento della percentuale vocale, soprattutto nei parlanti diplomati. Con riferimento all'italiano L2, questi dati giustificano la mancata difficoltà da parte dei soggetti romeni nel produrre questo tipo di sequenza.

Tabella 1 - Sequenza fonica VC per l'italiano L1 e per l'italiano L2 di romeni

Soggetti	Sequenza VC /ol/ "sole"		
	Durata (s)		
	V	C	C/V
italiano L1			
laureati	0,161	0,074	0,46
diplomati	0,152	0,063	0,41
valore medio	0,157	0,069	0,44
italiano L2			
laureati	0,150	0,067	0,44
diplomati	0,160	0,069	0,43
valore medio	0,155	0,068	0,44

Tabella 2 - Sequenza fonica VC per il romeno L1 in Romania e in Italia

Soggetti	Sequenza VC /om/ "omul"		
	Durata (s)		
	V	C	C/V
romeno L1 in Romania			
laureati	0,155	0,085	0,54
diplomati	0,159	0,089	0,56
valore medio	0,157	0,087	0,55
romeno L1 in Italia			
laureati	0,150	0,077	0,51
diplomati	0,166	0,080	0,48
valore medio	0,158	0,079	0,50

Per quanto riguarda la produzione della consonante lunga in italiano, sono state scelte le parole "mantello" e "allora", diversificate dalla vocale che precede la consonante lunga: tonica (/el:/) nel primo caso e atona (/al:/) nel secondo.

Le durate della consonante lunga e della vocale tonica/atona sono state misurate sia in italiano L1 sia in italiano L2. La parola "mantello" presenta 4 occorrenze nel testo, per cui le misurazioni complessive sono state 160, 40 per ciascun gruppo di

parlanti. Per quanto riguarda la parola “allora”, le misurazioni sono state complessivamente 80, poiché la parola presenta 2 occorrenze nel testo (Tabella 3).

Tabella 3 - *Sequenza fonica VCC per l'italiano L1 e per l'italiano L2 di romeni*

Soggetti	Sequenza VCC Durata (s)					
	mantello			allora		
italiano L1	V	CC	CC/V	V	CC	CC/V
laureati	0,105	0,150	1,428	0,064	0,130	2,031
diplomati	0,091	0,139	1,527	0,060	0,118	1,966
valore medio	0,098	0,145	1,479	0,062	0,120	1,935
italiano L2						
laureati	0,110	0,112	1,018	0,095	0,114	1,2
diplomati	0,135	0,091	0,674	0,105	0,100	0,952
valore medio	0,123	0,102	0,830	0,100	0,107	1,07

I dati presenti in Tabella 3 evidenziano come in italiano L1 la consonante risulti molto più lunga rispetto alla vocale, con una durata pari a una volta e mezzo quella della vocale nel caso di vocale tonica e di circa il doppio nel caso di vocale atona. Per quanto riguarda l'italiano L2 dei romeni la vocale tonica risulta più lunga della consonante e quella atona di pari lunghezza. Anche in questo caso emerge una differenza tra i laureati e i diplomati romeni in quanto la produzione della consonante lunga da parte dei laureati è migliore, con un valore che risulta essere quasi pari a quello della vocale, mentre nel caso dei diplomati la vocale presenta un valore più alto (0,135s) rispetto a quello della consonante lunga (0,091s).

I risultati ottenuti dall'analisi prosodico-segmentale sono perfettamente in linea con i risultati dell'analisi ritmica. L'aumento della percentuale vocalica nell'italiano L2 di romeni, infatti, può essere messo in relazione anche al diverso rapporto tra durata consonantica e vocalica nei nessi VCC nei quali la consonante risulta più breve della vocale tonica, o uguale alla vocale atona.

4. Conclusioni

L'indagine sperimentale condotta su parlato italiano e romeno prodotto da romeni residenti in Italia, romeni residenti in Romania e italiani residenti in Italia, ha riguardato aspetti ritmico-prosodici e prosodico-segmentali.

Per quanto riguarda il primo punto, i risultati hanno confermato la natura isosillabica delle due lingue, con il romeno situato in una zona più prossima alle lingue isomoraiche, ad es. il giapponese, caratterizzate da un valore percentuale di componente vocalica superiore al 50% (Pettorino, Pellegrino 2016). È interessante notare come il romeno L1 dei residenti in Italia sia, sul piano ritmico, diverso non solo rispetto all'italiano L1, ma anche rispetto al romeno parlato in Romania.

L'incremento della percentuale vocalica, in quest'ultimo caso, segnala un'interferenza da parte dell'italiano sulla lingua materna dei soggetti immigrati, in un contesto d'uso in cui la lingua seconda prevale quantitativamente come input ricevuto e output prodotto in contesti lavorativi e privati.

Ancora maggiore appare la distanza, sul piano ritmico, tra italiano L2 dei romeni e italiano L1, evidenziata anche essa da un più marcato aumento della percentuale vocalica nel parlato dei soggetti immigrati. Questo fenomeno andrebbe esaminato all'interno di un discorso più generale sulle seconde lingue, poiché, le difficoltà che il parlante incontra nel produrre una L2 determinano la necessità di rallentare l'eloquio per risolvere i problemi di programmazione on line (Vitale, De Meo, 2017). Tali momenti non possono che occorrere in coincidenza di elementi vocalici i quali, per loro natura statici, consentono un prolungamento di durata che nei foni consonantici, essendo dinamici, non è possibile.

Sul piano prosodico-segmentale, i risultati delle analisi hanno evidenziato la difficoltà, da parte dei parlanti romeni, ad assimilare alcune regole prosodiche proprie dell'italiano: se è piuttosto semplice conservare la isosillabicità nel romeno, lingua che non presenta l'alternanza consonantica breve/lunga, non è affatto semplice conservarla in una lingua come l'italiano, caratterizzata da un attento dosaggio temporale tra la componente vocalica e quella consonantica all'interno dei nessi VC e VCC.

Infine, i dati della ricerca hanno messo in evidenza come il contatto linguistico, associato a una riduzione marcata dell'uso della propria lingua materna, provochi cambiamenti nella L1 dal punto di vista ritmico-prosodico, tali da indurre ascoltatori romeni nativi ad attribuire un giudizio di non nativo agli enunciati ascoltati, prodotti da romeni immigrati in Italia. Il fenomeno, che potrebbe rientrare nel cosiddetto "atrito linguistico" (Schmid, Köpke, 2019), andrà esplorato con un ampliamento del corpus che presenti anche soggetti di altre fasce di età, sia più anziani sia più giovani, i quali mostrano certamente atteggiamenti diversi nei confronti del romeno, per i più piccoli non più L1 ma solo "heritage language" (De Bot, Gorter, 2005).

Riferimenti bibliografici

- BOERSMA, P., WEENINK, D. (2020). *Praat: doing phonetics by computer*. <http://www.praat.org/>
- CHINI, M., ANDORNO, C. (2018). *Repertori e usi linguistici nell'immigrazione. Una indagine su minori allogliotti dieci anni dopo*. Pavia: Franco Angeli.
- CINGOLANI, P. (2009). *Romeni d'Italia. Migrazione, vita quotidiana e legami transnazionali*. Milano: Il Mulino.
- COHAL, A. (2014). *Mutamenti nel romeno di immigrati in Italia*. Pavia: Franco Angeli.
- DE BOT, K., GORTER, D. (2005). A European Perspective on Heritage Languages. *Modern Language Journal*, 89, 612-616.
- DOUGHTY, C.J. (2003). Instructed SLA: Constraints, compensation, and enhancement, in DOUGHTY, C.J, LONG, M. (Eds.). *The handbook of second language acquisition*. Malden: Blackwell, 256-310.

- GALATÀ, V., MENEGUZZI, G., CONTER, L., ZMARICH, C. (2012). Primi dati sull'acquisizione fonetico-fonologica dell'italiano L2 in prescolari rumeni. In PAOLONI, A., FALCONE, M. (Eds.). *La voce nelle applicazioni*. Volume 8. Roma: Bulzoni Editore, 35-50.
- GALATÀ, V., ANGOSENE, G., ZMARICH, C. (2017). Italian as L2 in Romanian pre-schoolers. Evidence from a perception and production task. In BERTINI, C. et al. (Eds.), *Fattori Sociali e Biologici nella Variazione Fonetica – Social and Biological Factors in Speech Variation*. Volume Studi AISV 3. Milano: Officinaventuno, 257–280.
- NUZZO, E., GRASSI, R. (2016). *Input, output e interazione nell'insegnamento delle lingue*. Torino: Bonacci.
- PETTORINO, M., MAFFIA, M., PELLEGRINO, E., VITALE, M., DE MEO, A. (2013). VtoV: a perceptual cue for rhythm identification. In MERTENS, P., SIMON, A.C. (a cura di), *Proceedings of the Prosody-Discourse Interface Conference 2013 (IDP-2013)*, Leuven, September 11-13, 101-106.
- PETTORINO, M., PELLEGRINO, E. (2016). %V and VtoV: an acoustic-perceptual approach to the rhythmic classification of languages. In BARDEL, C., DE MEO, A., (Eds), *Parler les langues romanes/Parlare le lingue romanze/Hablar las lenguas romances/Falando línguas románicas*. Napoli: Università di Napoli L'Orientale, 13-28.
- PITTAU, F., RICCI, A., ILDIKO TIMȘA, L. (2010). *I romeni in Italia tra rifiuto e accoglienza*. Roma: Idos.
- RAMUS, F., NESPOR, M., MEHLER, J. (1999). *Correlates of linguistic rhythm in the speech signal*. *Cognition*, 73 (3), 265-292.
- ROMITO, L., TARASI, A. (2012). A rhythmical-prosodic analysis of Italian L1 and L2. In DE MEO, A., PETTORINO, M. *Prosodic and Rhythmic Aspects of L2 Acquisition: The Case of Italian*. Cambridge: Cambridge Scholar Publishing, 137-152.
- ROMITO, L., SAVY, R., TARASI, A., LIO, R. (2012). Analisi percettiva, musicale e “automatica” dell'italiano L1 e L2. In FALCONE, M., PAOLINI, A. *La voce nelle applicazioni*. Roma: Bulzoni Editore, 93-106.
- ROMITO, L., TARASI, A., LIO, R. (2014). Italian Index: rhythmical-prosodic analysis of Italian L2 produced by Albanian, Chinese, Polish and Romanian speakers. In CONGOSTO, Y., MONTERO, M. & SALVADOR, A. (Eds.), *Fonética Experimental, Educación Superior e Investigación*. Madrid: Arco Libros, 127-149.
- SCHMID, M.S., KÖPKE, B. (Eds.) (2019). *The Oxford Handbook of Language Attrition*. Oxford, Oxford University Press.
- VALTOLINA, G. (2012). *Figli migranti. I minori romeni e le loro famiglie in Italia*. Milano: Franco Angeli.
- VITALE, M., DE MEO, A. (2017). Rhythmic Differences and Second Language Acquisition. An Italian-based study. In GUDMUNDSON, A., ÁLVAREZ LÓPEZ, L., BARDEL, C. (Eds.). *Romance Languages. Multilingualism and language acquisition*. Frankfurt am Main: Peter Lang, 109-122.

MARINA CASTAGNETO, IRINA STAN

Code-switching in contesto italiano L2: una famiglia croata e due famiglie romene a confronto

Code-switching in L2 Italian context: A comparison between a Croatian and two Romanian families

This work examines the *code-switching* phenomena in one Croatian-Italian and two Romanian-Italian bilingual families. The analysis is based on two sub-corpora, one for each language combination, and it focuses on the frequent and numerically comparable *code-switching* phenomena which are dichotomously divided according to their predominant function, namely participant related vs. discourse related. The various functions covered in the two sub-corpora are linked to interactional strategies connected both to participants, such as contrast, comment, specification, repetition, etc., and to discourse, e.g. to negotiate the language of the conversation when the constellation of the participants changes. We therefore highlight the enormous functional potential of bilingual speech as an interactional and conversational strategy, also in the light of former considerations regarding *code-switching* phenomena between Italian and Italian dialects.

Keywords: bilingualism, code switching, contact linguistics, transfer

1. *Il corpus e gli informanti*¹

Questo lavoro intende studiare i fenomeni di *code-switching* a livello comunicativo-interazionale in tre famiglie bilingui: una famiglia bilingue croato-italiano e due famiglie bilingui romeno-italiano, tutte residenti in Piemonte.

La famiglia croata è composta da padre (P), madre (M), due figli: la prima figlia frequenta l'università, e ha raccolto i dati del sub-corpus croato per la sua tesi triennale (S), mentre il figlio ha 16 anni, ed è in età scolare (F). Il padre, per ragioni legate al suo lavoro, che riguarda l'imprenditoria e il commercio, è bilingue da oltre venti anni, mentre gli altri componenti della famiglia sono bilingui da 10 anni, dal momento del trasferimento del nucleo familiare in Italia.

Le due famiglie romene hanno un nucleo composto nel primo caso da una coppia di giovani nonni (bilingui soprattutto passivi in italiano) e la nipotina che vive con loro, A, che al momento della raccolta dati ha quasi 11 anni, vive in Italia da 10 anni e frequenta la terza elementare (ha iniziato gli studi più tardi per vicende fami-

¹ A Irina Stan si devono i paragrafi 1; 2; 4; 4.1.2; 4.2.2; 4.3.2; 4.4.2; 5.2.2; 6; 8; 9; 9.2, mentre Marina Castagneto è responsabile dei paragrafi 3; 4.1.1; 4.2; 4.2.1; 4.3.1; 4.4; 4.4.1; 5; 5.1; 5.2; 5.2.1; 5.3.1; 5.3.2; 7; 9.1; 10; 11.

liari). Nel secondo caso si tratta di una coppia di genitori bilingui con due figlie, L, di 5 anni, e la piccola V, di circa 2 anni (che però in ragione della sua giovane età non ha preso parte attiva alle conversazioni).

Il corpus da cui abbiamo estratto le analisi che qui presentiamo è costituito da due sub-corpora. Per il croato, abbiamo estratto 125 turni che presentavano fenomeni di *code-switching* tra italiano e croato (da un totale di oltre mille turni da 11 conversazioni libere). Per quanto riguarda il rumeno, abbiamo estrapolato 276 turni soggetti a *code-switching* da tre lunghe conversazioni all'interno delle famiglie romene (per un totale di 6 ore e 48 minuti di registrazione). A differenza del sub-corpus croato che contiene conversazioni libere tra i membri della famiglia a cena e in altri contesti simili, il sub-corpus rumeno consiste in conversazioni in cui una giovane intervistatrice rumena (I), che ha familiarità e confidenza con le due famiglie, parla di vari temi con le due bambine più grandi (A e L) come interlocutrici privilegiate; alle interazioni contribuiscono anche la nonna di A (N, nelle prime due conversazioni) e la madre di L (M). I due sub-corpora sono stati audio-registrati e trascritti pragmaticamente secondo le convenzioni AVIP (Savy, 2006)².

Le condizioni sociolinguistiche comuni alle tre famiglie di informanti sono su un piano di assoluta comparabilità: L1 non dominante in assenza di sostegno della propria comunità (Romaine, 1995: 183-187); L1 non di prestigio rispetto alla stessa L2, l'italiano; attitudine positiva verso il bilinguismo da parte dei genitori, che pure producono dei fenomeni di *code-switching* nel loro parlato e non sembrano seguire specifici criteri distribuzionali nella selezione delle due lingue.

2. Il tipo di bilinguismo

Per i nostri informanti si tratta di un bilinguismo:

- precoce (tranne nel caso delle figlie e forse del figlio della famiglia croata, quest'ultimo esposto all'italiano a sei anni, quindi sul confine fra un bilinguismo precoce e un bilinguismo tardivo);
- coordinato, cioè con due sistemi linguistici indipendenti, in cui anche le unità di contenuto restano separate (cfr. Ervin, Osgood, 1954);
- consecutivo (cioè in cui la L2 è stata appresa dopo la L1);
- additivo, considerato che le famiglie mantengono il rapporto con la L1;
- in ambiente L2, quindi l'apprendimento è accompagnato da acculturazione (cfr. Bettoni, 2006: 45);
- popolare, piuttosto che elitario;
- integrativo e assimilativo;
- (tendenzialmente) isolato (cfr. Francescato, 1981), perché le famiglie intervistate non sono inserite in comunità di parlanti con la stessa L1³;

² Un'analisi dei *code-mixing* all'interno degli stessi sub-corpora è stata proposta in Castagneto, Stan (2020). Siamo grate a M. Pejakovic per i dati relativi al croato, tratti da Pejakovic (2014).

³ La famiglia di L intrattiene qualche contatto sporadico con alcuni vicini di casa rumeni; i nonni di A intrattengono in romeno rapporti solo telefonici.

- (quasi) bilanciato, almeno per quanto riguarda i componenti più giovani delle famiglie e ciò che il nostro corpus consente di verificare, ovvero nelle interazioni in contesto familiare, diamesicamente orali e diafasicamente legate a un registro colloquiale e informale, in domini legati alla famiglia, al cibo, all'istruzione⁴.

Nel nostro corpus, i partecipanti alternano liberamente le due lingue, selezionate nella stessa quantità di parlato, e il transfer è bidirezionale per tutti i tipi individuali. Solo i prestiti in tutti e due i sub-corpora sono sempre dall'italiano nel romeno o nel croato, ma mai viceversa, probabilmente in ragione del maggior prestigio dell'italiano.

È presente una forte attivazione bilingue in tutti i parlanti, il che incide sulla quantità di transfer in produzione e percezione, sul tipo di transfer, sulla facilità di elaborazione nelle due lingue e sulla frequenza nella commutazione di codice.

Una differenza è legata alla circostanza che nella famiglia croata è presente anche un'adeguata alfabetizzazione nella lingua debole, perché in casa sono presenti libri e materiale audiovisivo in croato (dvd, giochi elettronici) e si fa uso anche di internet in questa lingua, mentre nelle due famiglie romene l'istruzione in romeno è solo orale, attraverso la abitudine di raccontare fiabe o di giocare in L1. Il diverso metodo di raccolta dati (cfr. par. 1), non ha creato significativi ostacoli alla comparabilità dei dati, molto simili per tipo di transfer e quantità dei fenomeni.

3. Definizione e classificazione dei *code-switching*

L'analisi si soffermerà in particolare sui casi di *code-switching*, che sono molto frequenti e numericamente comparabili tra i due sub-corpora (29 occorrenze nel sub-corpus croato; 36 nel sub-corpus romeno). In questo lavoro adotteremo una definizione 'ampia' di *code-switching*, nel solco dei lavori di Gumperz, per cui il *code-switching* è definibile come "juxtaposition within the same speech exchange of passages of speech belonging to two different grammatical systems or sub-systems" (Gumperz, 1982: 59). Adotteremo inoltre il punto di vista di Auer, che, all'interno di una prospettiva fortemente interazionale, sottolinea come talvolta il *code-switching* avvenga anche tra turni (Auer, 1984: 20)⁵. Saranno pertanto classificati come *code-switching* anche i casi di *code-switching* tra turni.

In questo lavoro, in prospettiva pragmatico-funzionale, verrà seguito in modo flessibile e per ragioni di tipo espositivo l'approccio di Auer (1984, 1995), adottato anche in molti lavori sul *code-switching* tra italiano e dialetto in Italia (es. Alfonzetti, 1992: 32), che si basa su una prima distinzione dicotomica delle commutazioni di

⁴ Siamo consapevoli che il bilinguismo assolutamente bilanciato è considerato una condizione rarissima, perché un bilingue perfettamente bilanciato dovrebbe avere uguale competenza in tutte le abilità linguistiche (cfr. Romaine, 1995: 12-19), cioè produzione e comprensione scritta e parlata nei diversi domini. Ma il nostro corpus non ci permette di sondare in modo così ampio la competenza nelle due lingue dei parlanti intervistati.

⁵ La possibilità che i *code-switching* avvengano anche tra turni diversi è condivisa da Romaine (1995: 122-123) e accettata come un tipo improprio di commutazione di codice in Cerruti (2004: 96, nota 4).

codice sulla base di una funzione preponderante connessa ai partecipanti o connessa al discorso⁶. Verrà dunque messa in evidenza la enorme potenzialità funzionale del discorso bilingue come strategia interazionale e conversazionale. A questi tipi di commutazione aggiungerei le commutazioni dovute a *priming*, cioè i casi in cui la commutazione è innescata da una particolare struttura conversazionale (come una parola o un sintagma) selezionata in una lingua diversa dalla lingua correntemente in uso durante quello specifico segmento di conversazione. Abbiamo deciso di non considerare le commutazioni di codice legate al dominio tra i *code-switching* con funzione discorsiva perché, come specifica Alfonzetti (1992: 16-17) l'uso alternato di due codici in relazione a differenti domini non si colloca necessariamente all'interno di uno specifico evento linguistico, ma pertiene a un livello macro-socio-linguistico. Per quanto riguarda le commutazioni dovute a *priming*, è stata operata la scelta di non suddividerle ulteriormente tra commutazioni con funzioni discorsive e commutazioni legate ai partecipanti⁷, perché si tratta di forme di commutazione interfrasale di codice che non rispondono immediatamente a istanze funzionali o pragmatiche, ma che sono descritte su base formale.

Siamo consapevoli che una classificazione di questo tipo si rende necessaria solo per ragioni di comodità espositiva, ed è operata unicamente sulla base della funzione prevalente di uno specifico *code-switching* all'interno del particolare contesto conversazionale in cui ha luogo. Ma le funzioni di uno stesso *code-switching* possono essere plurivoche, o sommarsi, e non ha senso fissare una lunga lista di potenziali funzioni della commutazione. Essendo interpretabili solo nel contesto, infatti, le funzioni sono potenzialmente infinite, come del resto lo sono i contesti.

Bisogna tuttavia tenere presente che i parlanti bilingui non estrarrebbero nell'uso le potenziali funzioni della commutazione di codice da una lista mentale, ma usano le commutazioni come strategia interazionale e conversazionale, e lo fanno, potremmo dire, in un modo 'iconico' rispetto al contesto conversazionale, di cui le commutazioni di codice costituiscono una chiave di codifica e di inferenziazione: se cambiano una o più caratteristiche del contesto o dell'interazione (come ad esempio la costellazione di partecipanti), se cambia qualche tratto della macroattività in cui i parlanti sono coinvolti (come l'evento comunicativo, nel senso di Hymes, 1974), se cambia qualcosa a livello di un'attività minore per estensione conversazionale (ad esempio un atto linguistico), allora i mutamenti improvvisi di lingua "make relevant /mantain /revise /cancel" (cfr. Auer, 1995: 123) alcuni tratti del contesto. La commutazione di lingua non ha dunque un significato referenziale o denotativo, come se fosse una parola, ma ha il significato operativo di segnalare un cambio improvviso, e indica iconicamente "otherness" (Auer, 1995: 124); il

⁶ Anche Giacalone Ramat (1995: 50-51) accetta questa dicotomia per l'analisi del corpus del suo lavoro, raccolto nelle province di Pavia e Piacenza, ma restringe le funzioni del *code-switching* legate ai partecipanti ai soli casi di convergenza e divergenza, nel rispetto della teoria dell'adattamento (Giles, St. Clair, 1979).

⁷ In questo caso, dunque, è stata compiuta la scelta di non seguire il modello di Auer (1984) che distingue tra casi di "participant related transfer" e "discourse related transfer".

suo significato può dunque essere riconosciuto solo tramite un processo di inferenziazione che nasce e si conclude nel suo contesto di produzione, e si motiva al suo interno. Nei termini di Peirce potremmo forse dire che il *code-switching* è dunque un'icona complessa di tipo diagrammatico, ovvero un'icona in accordo alla categoria di secondità, quindi al rapporto con il suo oggetto⁸. In quanto icona di tipo diagrammatico, la commutazione di codice in sé non assomiglia ai contenuti che veicola, ma esprime analogicamente una relazione segnica di cambiamento di rapporto, o di contrasto, all'interno del contesto o dell'interazione, e un conseguente cambiamento dei suoi tratti o funzioni. In quanto segno iconico, basato quindi sulla motivazione, la interpretazione del *code-switching* non può che fondarsi su un processo inferenziale di ipotesi (o di 'abduzione', nei termini della semiotica peirceiana); per inferirne la funzione non si può prescindere da un'analisi sequenziale e bisogna comprendere bene il contesto e il cotesto, soprattutto il turno che precede la commutazione, che ne costituisce il "sequential environment" (Auer, 1995: 116), la cornice (o *frame*) conversazionale.

Fatta dunque questa premessa per spiegare il modo in cui intendiamo considerare le commutazioni di codice, e cioè come strumento semiotico e come risorsa conversazionale cui possono attingere i parlanti bilingui, procediamo a presentare le funzioni prevalenti delle commutazioni riscontrate nei due sub-corpora.

4. *Code-switching con funzioni legate ai partecipanti*

Per quanto concerne i *code-switching* con funzioni legate ai partecipanti, il numero di occorrenze è fortemente comparabile in entrambi i sub-corpora: a 14 *code-switching* con funzioni legate ai partecipanti nel sub-corpus croato corrispondono 18 *code-switching* con funzioni equivalenti nel sub-corpus romeno, e in entrambi i sub-corpora le funzioni più rappresentate sono quelle di contrasto o divergenza con l'interlocutore (6 occorrenze nel sub-corpus romeno, 5 in quello croato) e di commento (4 occorrenze in entrambi i corpora). Alcuni *code-switching* sono legati alla preferenza del parlante (4 occorrenze in romeno e 1 in croato). Sono inoltre presenti le funzioni di specificazione (2 casi in romeno, 1 in croato), di ripetizione, con funzione di chiarimento (2 occorrenze in croato)⁹ e, in modo occasionale, an-

⁸ Per Peirce, infatti, "an iconic diagram is a systematic arrangement of signs, none of which necessarily resembles its referent, but whose relationship to each other mirrors the relationships of their referents" (Peirce, 1980: 156-157).

⁹ Le autrici sono consapevoli che i casi di ripetizione sono difficilmente ascrivibili in modo univoco a casi di *code-switching* con funzioni legate ai partecipanti o al discorso. In realtà la ripetizione è solo un "conversational locus" in cui è facile che avvenga una commutazione, ma la stessa etichetta metalinguistica "ripetizione" mostra che si tratta di una struttura conversazionale identificabile più sul piano formale che su quello funzionale. Come scrive Alfonzetti (1992: 109) "[la ripetizione] rappresenta pertanto una configurazione che il *code-switching* può assumere" per svolgere diverse funzioni (nel caso del corpus di Alfonzetti, soprattutto di enfasi e di elaborazione del messaggio).

Nel caso delle due occorrenze presenti nel sub-corpus croato di questo lavoro (es. 22; es. 9, turno S71, dove il *code-switching* per ripetizione segue un primo *code-switching* della stessa parlante, con funzione

che le funzioni di enfasi (1 occorrenza nel sub-corpus romeno) e di correzione (1 occorrenza in entrambi i sub-corpora).

Vediamone qualche esempio.

4.1 Contrasto

4.1.1 La funzione di contrasto nel sub-corpus croato

(1)¹⁰ DIALOGO VIII

P97: Sigurno je ona bila, isli ste u piscinu skupa.

“Sicuramente era lei, andavate in piscina insieme.”

S98: Isli smo u piscinu skupa? **Ma ne, boh, non mi ricordo.** Andavo con Sara, aspetta ah sì ma Marta è venuta solo all’inizio poche volte, poi non veniva più [...]

“Andavamo in piscina insieme? Ma no”

In questa porzione di dialogo del sub-corpus croato la figlia riprende in modo quasi integrale l’ultima frase del padre, operando solo uno spostamento deittico dell’enunciazione attraverso un diverso accordo verbale: questo segmento conversazionale costituisce un buon esempio di coesione tra turni attraverso un fenomeno di risonanza frasale ben inquadrabile all’interno della disciplina della sintassi dialogica (Du Bois, 2014; Du Bois, Giora, 2014)¹¹. La figlia prende poi le distanze dal contenuto della frase del padre sottolineando il suo disaccordo attraverso una “classica” commutazione di codice intraturno.

Nel dialogo I tra madre e figlia, uno dei più lunghi e più ricchi di commutazioni, sono presenti diversi *code-switching* in cui è evidente la funzione di contrasto, anche se molti di essi presentano ulteriori funzioni secondarie di tipo discorsivo, visto che all’interno della rapida conversazione familiare i cambi di topic sono frequenti e i contrasti spesso coincidono con tentativi di compartimentazione del dialogo. Si osservi il seguente estratto:

di correzione) le ripetizioni si sviluppano su turni diversi prodotti da parlanti diversi e si inscrivono all’interno di un “contratto pedagogico” (cfr. par. 4.4.1), per cui la loro funzione si inquadra meglio tra le funzioni legate ai partecipanti.

¹⁰ Negli es. riportati i *code-switching* analizzati sono in grassetto, e per ogni turno viene riportata in corsivo una traduzione dei segmenti in croato o in romeno.

¹¹ La recente disciplina della sintassi dialogica mostra l’*engagement* dei parlanti nel costruire l’enunciazione sulla base del turno dialogico immediatamente precedente del proprio interlocutore, riproducendone selettivamente parole e strutture. Con le parole di Du Bois, Giora: “dialogic syntax investigates the linguistic, cognitive and interactional processes involved when language users reproduce selected aspects of a prior utterance, and when recipients respond to the parallelisms and resonances that results, drawing inferences for situated meanings” (Du Bois, Giora, 2014: 352). Nel caso specifico del segmento analizzato ci troviamo di fronte a una risonanza, cioè a una “activation of affinities across utterances” (Du Bois, 2014: 11) che potremmo includere tra i *frame grab*, cioè “a linguistic structure that calls on a speaker to combine a capacity to orient to the prior linguistic production of a co-participant and to make grammatically appropriate adaptations, while still allowing for creative elaboration in the articulation of the speaker’s own perspective” (Hobson *et al.*, 2012: 2720).

(2) DIALOGO I

M11: Oko jedanaest sati imam pullman i onda sam tamo navečer u pola deset u Zagreb. <pb> Kupila sam svasta za jest ima da kuhate, ti lunedì nemoj da ides.

“Verso le 11 prendo il pullman e verso le nove e mezza di sera sono a Zagabria. Ho comperato tante cose da mangiare così potete cucinarvi da mangiare, tu [si rivolge alla figlia] lunedì non andare [al lavoro]”

S12: **Ma se mi hanno detto di venire.**

M13: <eh> se ti fa male il dente cosa devi fare? Chiamalo e gli dici che non riesci andare

S14: Ma ho tanto lavoro da fare, **kako ne idem?**

“Come faccio a andarci?”

Questo segmento conversazionale, in cui il *code-switching* al turno S14 ha anche la evidente funzione di segnalare un tentativo di chiusura di una porzione conversazionale, è anche un caso patente della altissima reversibilità della direzione di mutamento tra i due codici. La figlia, infatti, ai turni S12 e S14 compie due commutazioni, entrambe con funzioni di contrasto, ma commutando una prima volta verso l'italiano, la seconda volta verso il croato.

La condizione di reversibilità della direzione, che è particolarmente leggibile nel sub-corpus croato, mostra come la funzione pragmatica di contrasto sia sussunta dal *code-switching* in sé, a prescindere da quale sia il codice verso cui viene operata la commutazione. Questa condizione è particolarmente vera per le situazioni in cui i parlanti sono fortemente bilingui (per cui non è possibile parlare di una vera e propria preferenza o di una decisa dominanza di uno dei due codici da parte degli interagenti), quando entrambi i codici vengono accettati, così come il loro uso alternato, e se l'evento linguistico è ascrivibile a situazioni informali¹². Negli esempi di commutazione nel sub-corpus croato qui analizzato, è evidente come la direzione del *code-switching* dipenda unicamente dal codice che è momentaneamente la lingua base del discorso e non rivesta un particolare significato sociale o identitario, del resto poco pertinente all'interno di un contesto familiare. La reversibilità della direzione della commutazione nel sub-corpus croato è ancora più evidente quando i *code-switching* hanno funzioni di tipo discorsivo come segnalare i cambiamenti di topic, l'inaugurazione di nuovi segmenti conversazionali o l'apertura/chiusura di *side-sequences* all'interno della conversazione. Anche Cerruti (2004: 123) aveva comunque osservato come fenomeni di commutazione di codice siano particolarmente rilevanti nella scansione in sequenze costitutive¹³.

4.1.2 La funzione di contrasto nel sub-corpus romeno

La direzione del *code-switching* sembrerebbe invece essere rilevante nel sub-corpus romeno. Nei 6 casi presenti nel sub-corpus, A usa il *code-switching* con funzione

¹² Le stesse osservazioni sono presenti in molti lavori relativi alla commutazione di codice tra italiano e dialetti italiani (cfr. par. 10).

¹³ Cerruti in questo caso fa riferimento soprattutto alla progressione interna nella narrazione, ma l'osservazione è valida anche per la progressione tematica che caratterizza l'interazione dialogica.

di contrasto per opporsi alla nonna, e commuta sempre in direzione dell'italiano, come nel seguente caso:

(3) DIALOGO II

N37: După ce vezi emisiunea asta te hotărăști dacă-ți pregătesc, ce-ți pregătesc de mâncare.

"Dopo che vedi questa trasmissione ti decidi se ti preparo, cosa ti preparo da mangiare"

A38: Nimic

"Niente"

N39: <eh>

A40: **Cosa ti aspettavi da me?**

N41: Știu de ce: ai mâncat salam când ai venit și de-aia nu ți-e foame

"So io perché: hai mangiato del salame quando sei arrivata e perciò non hai fame"

A42: Dar nu mi-e foame de loc, așa în general

"Ma non ho proprio fame, così in generale"

N43: Am înțeles

"Ho capito"

A44: **No, no, non mi viene proprio. Cioè non ce la faccio.**

Ai turni 40 e 44 la bambina si oppone alla nonna cambiando il codice e usando l'italiano. Qui è lecito pensare che la direzione del *code-switching* possa ricoprire almeno in parte dei significati sociali o identitari, considerando che la bambina ha quasi 11 anni, vive da 10 in Italia, frequenta la scuola italiana e ha probabilmente l'italiano come lingua preferenziale, come si può capire dalla circostanza che durante il gioco parla ai suoi pupazzi usando l'italiano.

4.2 Commento

Il secondo tipo di funzione legata ai partecipanti per numero di occorrenze è quella di commento (4 occorrenze per ciascuno dei due sub-corpora).

4.2.1 La funzione di commento nel sub-corpus croato

Ecco due esempi dal sub-corpus croato:

(4) DIALOGO VII

F83: Ah, era devastante, mi sono stancato parecchio, **nista ne radim od proslog ljeta**, quindi riprendere è stato pesante.

"non faccio niente dall'estate scorsa"

(5) DIALOGO II

F33: čitam ovu knjigu iz engleskog, **che è molto interessante.**

"Sto leggendo il libro di inglese"

Come si nota, anche nel caso di *code-switching* con funzioni di commento la direzione del *code-switching* operata nei due esempi dallo stesso parlante è del tutto irrilevante, e dipende dalla lingua in uso.

4.2.2 La funzione di commento nel sub-corpus rumeno

La stessa osservazione non vale per il sub-corpus romeno, in cui la bambina A usa per i commenti esclusivamente il romeno, deviando dall'italiano:

(6) DIALOGO II

A08: E uno... e due... e tre. **Uh, da' cât este?**

"Uh, ma quanto è?"

4.3 Preferenza del parlante

4.3.1 Preferenza dei parlanti nel sub-corpus croato

Nel dialogo VIII in un'occorrenza si riscontra lo stesso modello individuato da Alfonzetti come ricorrente (Alfonzetti, 1992: 44-49) costituito dalla successione in sequenza, nello stesso turno, di due *code-switching*: il primo per adeguarsi al codice usato dall'altro partecipante nel turno precedente, il secondo di ritorno al codice precedentemente in vigore nella conversazione.

(7) DIALOGO VIII

M94: Ma mislim da se zove Marta, mama radi joj u onoj farmaciji di fianco alla casa della zia. Ma imala je gips na nozi, mislim da je imala incidente. *"ma penso che si chiami Marta, sua mamma lavora in quella farmacia di fianco alla casa della zia. Ma aveva il piede ingessato, credo che abbia avuto un incidente"*

S95: O mio Dio! Ma sei sicura che sia lei?

M96: Si c'era anche papà [rivolta al padre] **i vidje-li smo ju obodvoje, jel Niko?**

"la abbiamo vista entrambi, vero Niko?"

La conversazione tra madre, padre, figlia si svolge prevalentemente in croato, con enunciazioni mistilingui. Ma quando la figlia, per *priming*, passa all'italiano (turno S95) la madre fornisce la seconda parte della coppia adiacente, la risposta alla domanda, nella stessa lingua, per poi ritornare al croato.

4.3.2 Preferenza dei parlanti nel sub-corpus rumeno

Nel sub-corpus romeno abbiamo 4 occorrenze di *code-switching* che segnalano la preferenza dei partecipanti, tra cui un caso piuttosto complesso di un lungo turno di A che presenta ben 3 *code-switching* con questa funzione:

(8) DIALOGO I

A100: [rivolta all'orsetto] Baciolinoo, ti dobbiamo cucire che sei tutto rotto. [1, rivolta a I] **Nu, că am văzut că s-a descusut aicea și trebuie să-i zic lu' Maia să îl coasă dar pentru moment stă si așa.** [2, rivolta all'orsetto] **Giusto?** [Lo bacia] Amore mio! [3, rivolta a I] **Vorbesc cu ei și fac finta că sunt veri**

[1, rivolta a I] *"No, che ho visto che si è scucito qui e devo dire a Maia di cucirlo ma per il momento sta anche così"*

[3, rivolta a I] *"Parlo con loro e faccio finta che sono veri"*

In questo turno vediamo come A parli in italiano con il suo orsacchiotto, un interlocutore un po' sui generis perché non può rispondere. Alla intervistatrice però si rivolge in romeno, in quanto amica della nonna, la cui lingua preferenziale è il romeno, e perché ha condotto tutta la conversazione con lei in romeno. I segmenti in romeno hanno anche le funzioni discorsive rispettivamente di commento (1) e di giustificazione (3) rispetto alla scelta di parlare con l'orsacchiotto, e fungono pertanto da *background* alla conversazione.

4.4 Correzione

Un caso particolare è dato dai *code-switching* con funzione di correzione, presenti in un solo caso in entrambi i sub-corpora.

4.4.1 La funzione di correzione nel sub-corpus croato

Nel sub-corpus croato si verifica un'intera sequenza con cui la figlia corregge un errore in italiano della madre:

(9) DIALOGO VII

M68: Ma dai, uzmi makar uova ili bananu.

“prendi almeno le uova oppure la banana”

S69: **Mamma si dice uva e non uova** <risata>

M70: Io non capisco mai, **uvijek confondam uva e l'uova**, mi sembrano uguali

“confondo sempre uva e le uova”

S71: **Uova su jaja, uva je voce, to jest grozdje.**

“Uova sono le uova (jaja), uva è la frutta, cioè uva (grozdje)”

M72: Ma uvijek se confondam, **non riesco a pronunciare bene**

“Ma sempre mi confondo”

Al di là dei *code-mixing* prodotti dalla madre, in cui possiamo osservare la presenza di *-am*, “outside late system morpheme” (cfr. Myers-Scotton, 2002, 4-M Model) su un prestito dall'italiano, circostanza che mostra inequivocabilmente come il croato sia la lingua matrice di questa parlante, notiamo che tra le due parlanti si è instaurato ciò che Franca Orletti (2004: 31) ha definito “stipulazione di un contratto pedagogico”. Anche se questo “contratto” si verifica in genere nelle interazioni tra parlante nativo e non nativo, qui la figlia assume il ruolo del parlante nativo italiano e si propone come docente; la madre ne accetta il ruolo, il che consente alla figlia di potere correggere l'errore della madre-allieva senza atti di mitigazione o riparazione, se si eccettua la risata alla fine del turno S69. Anzi è la madre a cercare di giustificare il suo errore attraverso delle commutazioni con funzione di commento verso la sua lingua madre, che sembra costituire un'oasi di affidabilità.

4.4.2 La funzione di correzione nel sub-corpus rumeno

Più tipicamente, in romeno la sequenza di pedagogia naturale si instaura tra la piccola L, di 5 anni, e sua madre:

(10) DIALOGO III

L99: Ma me l'ha detto la maestra dell'asilo dei bambini. Che quando piange poi è sposata

M100: Sposata, non sposata

L101: <mmh>?

M102: **Sposata e când te măriți, spossata e când ești obosit.** Capito?
 "Sposata è quando ti sposi, spossata è quando sei stanco"

L103: Capito.

I104: **Cred că ai înțeles, doar ești o fetiță isteată**
 "Ci credo che hai capito, sei una bimba sveglia"

L105: **Già, mi hai preso la parola dalla bocca. Cioè, mica sono scema**

In questo caso è interessante come la spiegazione della madre, in romeno, si concluda con una *tag-question* in funzione di *checking* (*capito?*) anche senza mitigazione, seguita addirittura da un complimento per incoraggiare l'apprendimento della "discente" da parte della intervistatrice (turno I04), che partecipa al dialogo e fa da sostegno all'azione pedagogica della madre.

5. Code-switching con funzioni connesse al discorso

Prescindendo ora da altre funzioni legate ai partecipanti, presenti in modo più sporadico nei due sub-corpora, presentiamo le più importanti funzioni del *code-switching* connesse al discorso all'interno dei due sub-corpora.

Tra le funzioni del *code-switching* connesse all'organizzazione del discorso (9 occorrenze nel sub-corpus croato, 4 in quello romeno) osserviamo come frequentemente i *code-switching* possano chiudere/aprire un segmento conversazionale (4 occorrenze esclusivamente nel sub-corpus croato) o inaugurare un nuovo topic (1 occorrenze nel sub-corpus romeno, 2 nel sub-corpus croato). Inoltre non sono rari i casi in cui i *code-switching* tra turni si rivelano utili nella organizzazione ed esplicitazione della struttura del discorso, ad esempio per passare dalla porzione di conversazione relativa al tema che giustifica la stessa interazione a una sequenza marginale (cfr. par. 5.2) relativa alla attività che i partecipanti stanno svolgendo, introducendo un vero e proprio "sequential contrast" (2 occorrenze nel sub-corpus romeno, 1 nel sub-corpus croato). I *code-switching* con funzione discorsiva possono inoltre servire a negoziare la lingua della conversazione quando cambia la costellazione dei partecipanti (2 nel sub-corpus croato), sia che si voglia fare entrare un nuovo partecipante (o che un partecipante si autoselezioni per la presa di turno), sia che si voglia restringere la costellazione selezionando un partecipante specifico. In un caso, nel sub-corpus romeno, la funzione discorsiva è di citazione (1 occorrenza)¹⁴. Vediamo qualche esempio.

¹⁴ L'unico caso di citazione presente nel corpus di questo lavoro è comunque un tipo di citazione peculiare, perché non corrisponde a un enunciato prodotto da un altro parlante in una situazione diversa, ma a ciò che è scritto sull'orsacchiotto di A (cfr. es. 20, turno A95), in cui la bambina dichiara in romeno di avere chiamato il suo orsetto *Baciolino* perché sull'orso è scritto "ti meriti un bacio". È

5.1 Inaugurazione di un nuovo topic, e apertura/chiusura di segmenti conversazionali

Nel caso del sub-corpus croato (4 occorrenze), che coinvolge tutti parlanti adulti e quindi pragmaticamente consapevoli dei meccanismi di gestione della conversazione, il *code-switching* viene usato anche come segnale di chiusura di segmenti della conversazione; la chiusura può essere poi seguita da un nuovo segmento informativo, e non è detto che il parlante che apre un nuovo segmento non decida di ricorrere nuovamente al *code-switching* come risorsa pragmatico-comunicativa, collaborando nel merito e nel metodo con l'interlocutore precedente, es.:

(11) DIALOGO IV

F58: No, mi servono domani, perciò me li puoi dare domani.

P59: Va bene, **dat cu ti sutra**.

“(i soldi) te li do domani”

F60: S, **lo sai chi ho visto l'altro giorno? Non ci crederai mai!**

Al *code-switching* con funzione di chiusura del padre al turno P59 fa dunque eco un nuovo *code-switching* con funzione di apertura di un nuovo segmento conversazionale da parte del figlio al turno F60.

5.2 Sequenze marginali

Su un piano di segmenti più ampi di singoli atti linguistici o di coppie di turni, notiamo come in entrambi i sub-corpora (in 2 occorrenze nel sub-corpus romeno, 1 nel sub-corpus croato) i *code-switching* servano anche per sottolineare pragmaticamente l'ingresso o l'uscita da una sequenza marginale (*side-sequence*, nei termini di Auer, 1984: 32), così come il ritorno alla lingua precedentemente usata nel dialogo può indiziare il ritorno al tema principale della conversazione. Come scrive Alfonzetti (1992: 67-68), queste sequenze marginali consentono l'esecuzione di attività secondarie rispetto allo svolgimento principale della conversazione, e i *code-switching* consentono di segnalare la sospensione del discorso così come, eventualmente, anche il ritorno alla sequenza conversazionale sovra-ordinata.

Si veda l'esempio seguente.

5.2.1 Sequenze marginali nel sub-corpus croato

(12) DIALOGO I (i partecipanti sono a tavola)

M17: Dai chiamalo. <pl> Ho fatto la spesa, per tutta settimana vi basta.

S18: **Jest dobar ovaj formaggio.**

“Che buono questo formaggio”

M19: Jest, to je onaj Grana Padano.

ovvio che la funzione di questo *code-switching* non sia paragonabile a quella dei frequentissimi *code-switching* per citazione riportati ad esempio da Alfonzetti (1992: 119-137) nel parlato mistilingue tra italiano e dialetti; già la circostanza che il *code-switching* accada non dopo un *verbum dicendi*, ma dopo un *verbum scribendi*, fa sì che la funzione non possa essere quella di messa in rilievo, frequente in questi casi, e che la prosodia non venga modificata. È però interessante osservare come, grazie al meccanismo di *triggering* (Clyne, 1969), la commutazione si prolunghi oltre il discorso riportato.

“*Si è quel Grana Padano*”

P20: Nemoj te sad navalit samo na suhovinu, I pasta e riso sve moramo est.
[rivolto ai figli] “*Non dovete mangiare solamente gli affettati, anche pasta e riso bisogna mangiare*”

M21: **Adesso che me ne vado in Croazia, dovete cucinare voi**

Nei turni precedenti la madre ha annunciato la sua imminente partenza per Zagabria. I turni 18-20 rappresentano quindi una sequenza inserto e, come spesso accade (cfr. Auer, 1984: 39-40), la sequenza inserto è spesso associata a un'attività secondaria svolta durante la conversazione, in questo caso il pranzo, con relativo commento sul cibo. Ma la ripresa del tema precedente da parte della madre, e cioè la spesa per l'alimentazione della famiglia in sua assenza, riporta anche nella conversazione la lingua italiana (turno M21).

5.2.2 Sequenze marginali nel sub-corpus rumeno

La seguente occorrenza in romeno presenta una sequenza a latere della conversazione principale ma è correlata anche a mutamenti di costellazione dei partecipanti e a mutamenti di tratti del contesto (la madre, parlando dalla cucina, si inserisce nella “conversazione” di L con la sorellina di due anni), per cui si tratta più di una alternanza di codice che di un *code-switching* in senso stretto:

(13) DIALOGO III, (le bambine stanno disegnando)

L65: V, V, fagli gli occhi. Aspetta li faccio io. [verso I]

Sai, io sono più brava [*la sorellina si mette a piangere*]

<P>

L66: [le ridà la penna] Tieni. Non piangere più. [*la sorellina piange ancora*]

M67: [dalla cucina] **V! L, de ce plânge V?**

“*L, perché sta piangendo V?*”

L68: Vrea să deseneze dar nu știe. Vrea să se măsoare cu mine.

“*Vuole disegnare ma non sa. Vuole misurarsi con me*”

[arriva la madre]

M69: [a V] **Cosa c'è? Sei gelosa?**

5.3 Cambio di costellazione dei partecipanti

5.3.1 Cambio di costellazione dei partecipanti nel sub-corpus croato

Nel sub-corpus croato sono presenti 2 occorrenze di *code-switching* di questo tipo.

Nell'ultimo turno conversazionale dell'esempio 11 notiamo, tra l'altro, in corrispondenza con un cambio di topic, come al turno F60 il figlio abbia deciso di rivolgersi direttamente alla sorella, che già partecipava al dialogo, restringendo la costellazione dei partecipanti. Anche con questa funzione la direzionalità del *code-switching* sembra reversibile:

(14) DIALOGO I

P26: Tko ce mene vozit di mattina?

“e chi mi porta di mattina?”

M27: Sorella. Devi chiamare lei che ti porta al mattino. E S te / **ti odvezi F ujutro u Casate i onda ga vozi mama od Mattea ali ako ti budes isla radit onda ne znam, morat cu zamolit nju da ga vraca ona kuci.**

Dovete prendere tutti ognuno la sua chiave.

“te porta F domani a Casate e al ritorno lo porta la mamma di Matteo, ma se tu andrai a lavorare allora non so, dovrò chiedere alla mamma di Matteo di portarlo tutti i giorni.”

Nell'es. 14 si nota come la madre operi un *code-switching* verso il croato per restringere la costellazione dei partecipanti alla sola figlia, e un secondo *code-switching* verso l'italiano per tornare alla costellazione precedente. In casi come questo non è possibile attribuire il *code-switching* alla preferenza verso un codice di uno o più partecipanti, tanto più che la madre nei turni precedenti ha interagito in italiano nonostante sia l'unico membro della famiglia per cui possiamo parlare sicuramente di un bilinguismo con croato dominante, e si sia rivolta alla figlia in croato, nonostante si tratti di una parlante bilingue quasi perfetta.

5.3.2 Alternanza di codice

Per alcune occorrenze del sub-corpus croato piuttosto che di vera e propria commutazione di codice, seguendo l'impostazione di Berruto (1990: 109-110), siamo di fronte a uno specifico tipo di commutazione, che lo studioso definisce “alternanza di codice”. Osserviamo la seguente porzione di dialogo:

(15) DIALOGO VII

S82: Allora fratellino kako je bilo u palestri danas, **che esercizi ti ha fatto fare Mauro?**

“come è andata in palestra”

F83: <ah>, era devastante, mi sono stancato parecchio, **nista ne radim od proslogljeta**, quindi riprendere è stato pesante.

“non faccio niente dall'estate scorsa”

S84: E vedi poi ridevi quando tornavo io a casa stanca morta. Guarda chi c'è.

F85: Ma ciao V! **Kako si**

“Come va?”

V86: **Bene, grazie.**

S87: **Sjedni V, hoces sladoled, vidi kako F jede, hoces?**

“Siediti V, vuoi un gelato, guarda come mangia F, lo vuoi?”

V88: **Si il gelato mi piace tanto.**

F89: Come lo vuoi, vaniglia oppure cioccolato?

V90: A me piace cioccolato.

A prescindere dai *code-switching* ai turni 82-83, che hanno funzione di commento e mostrano l'ampia reversibilità dei *code-switching* nelle conversazioni analizzate, notiamo come il *code-switching* interfrasale al turno 85 segnali simultaneamente un cambio nella costellazione dei partecipanti, dovuto all'ingresso nella conversazione

di V, la giovane cuginetta di 8 anni dei due parlanti, ma anche un cambio di dominio nella conversazione. Siamo di fronte a un'alternanza di codice, e non a una commutazione di codice in senso stretto, perché il passaggio di lingua, oltre a segnalare il cambio di interlocutore "intende sottolineare un mutamento all'interno della situazione comunicativa" (Regis, 2005: 11), una rottura rispetto all'interazione in atto legata più al contesto che all'identità del destinatario. La funzione discorsiva di questa alternanza è assolutamente evidente poiché il *pattern* conversazionale atteso prevede che con i bambini nelle interazioni sia usato soprattutto l'italiano (e anche in questo caso notiamo un parallelismo con quanto è stato osservato per i *code-switching* nelle interazioni bilingui italiano/dialetto, cfr. Alfonzetti, 1992); anche in questo caso la preferenza di codice per l'italiano della cuginetta si rivela immediatamente dalla circostanza che tutti i turni successivi di sua pertinenza siano in italiano, nonostante i due fratelli le si rivolgano in croato. Non avrebbe dunque senso attribuire ai *code-switching* di F (turno F85) e di S (turno S87) una funzione di convergenza, o tantomeno di divergenza, visto che il *code-switching* avviene all'interno di un saluto che consente alla nuova arrivata di esser riconosciuta come partecipante ratificata ed entrare nella conversazione.

Si potrebbe inoltre osservare che, una volta adempiuta la funzione pragmatica di segnalare l'alternanza di codice e una variazione di tratti nella situazione comunicativa, termina la negoziazione della lingua e la conversazione procede nella lingua preferenziale della bambina, l'italiano.

6. Dominio

Nel sub-corpus romeno, in cui le interlocutrici privilegiate sono due bambine, le funzioni più specificamente pragmatiche del *code-switching* sono meno sfruttate, ma più spesso il *code-switching* segnala un cambio di dominio (13 occorrenze), ad esempio nei casi in cui le bambine inaugurano segmenti conversazionali sui giochi o sulla scuola.

Non sorprende, del resto, che A codifichi elementi relativi al gioco (cfr. turno A100, es. 8) o alla scuola in italiano, visto che studia in un contesto scolastico italiano e gioca con bambini italiani, per cui è il contesto ad attivare il *code-switching*. Cfr. il seguente esempio:

(16) DIALOGO II

N22: Astea ce-s? Hârtii... Lasă că mâine ești libera. Deci luni vă duceți la școala, și după-aia vă mai duceți...marți?

"Cosa sono queste? Cartacce... Tanto domani sei libera. Quindi lunedì andate a scuola, e poi ci andate ancora...martedì"

A23: <ah> stai așa. **Lunedì non si può perché è Halloween. Poi martedì c'è sciopero. Quindi lunedì e martedì zero; poi domani il maestro e la maestra non si sa... il maestro dipende a che ora si alza e bla, bla, bla. Quindi sabato, domenica, lunedì, martedì, festa.**

"Ah, aspetta"

N24: Ieșim în oraș că o să fie cu măști, înțelegi? E frumos și în oraș, n-are sens să te duci în casa cuiva că nu cunoaștem. Astea, ce faci cu ele?

“Usciamo che ci saranno le maschere, sai? È bello anche in città, non ha senso andare a casa di qualcuno che non conosciamo. Di queste che te ne fai?”

A25: Dar dar n-am costum... Nu pot să mă duc goală ca un vierme. E poi dau timp urât.

“Sì, ma non ho un costume. Non posso andare nuda come un verme. E poi danno brutto tempo.”

Qui si nota come la bambina usi l'italiano per rispondere a una domanda che riguarda le attività scolastiche nel periodo di Carnevale (turno A23), ma torni a usare il romeno per rispondere sullo stesso tema relativamente a un'attività programmata al di fuori dei confini della scuola (turno A25). Ciò che è cambiato veramente è, nei termini di Hymes (1974), la “scena”, cioè i confini fisici, ma anche culturali, dell'evento linguistico.

7. Priming

Per 6 *code-switching* nel sub-corpus croato, e in un solo caso nel sub-corpus romeno, il *code-switching* è indotto da un *priming*. In numerosi casi nella frase che precede il *code-switching* è presente un *code-mixing* che coinvolge una parola, un sintagma, o un costituente di estensione ancora maggiore in un codice diverso da quello momentaneamente in uso, che poi innesca il *code-switching* verso questo codice.

Ecco un esempio di *code-switching* tratto dal sub-corpus croato dovuto a *priming* relativo a una commutazione di codice interna allo stesso turno, compiuta dallo stesso parlante¹⁵:

(17) DIALOGO I

M15: Kako ces ic kad te tako zub boli? Sta ti je rekao **dentista?**

Hai comperato Aulin?

“Come fai ad andarci se ti fa così tanto male il dente. Cosa ti ha detto il dentista? Hai comperato l'Aulin?”

e un esempio tra turni di parlanti diversi:

(18) DIALOGO V

M66: S, zasto nisi pojela taj **pezzettino di torta?**

“S, perché non hai mangiato questo pezzettino di torta?”

S67: **Eh non ce la faccio più, sono piena**

In queste occorrenze il tipo di *code-mixing* che si innesca è un caso di “congruent lexicalization” (Muysken, 2000, cap.5), che, nel caso dei *code-mixing* tra italiano e croato, coinvolge due lingue in cui non vi è un lessico simile, ma sussiste un'equivalenza strutturale di tipo generale, sia categoriale che di ordine lineare (cfr. Muysken,

¹⁵ In questo paragrafo anche i *code-mixing* che innescano un *code-switching* sono stati evidenziati in grassetto.

2000: 123); anche la somiglianza strutturale tra due lingue può infatti funzionare da “trigger” del *code-mixing*, perché in un contesto fortemente bilingue entrambe le lingue sono vicine alla soglia di attivazione. Notiamo come in tutte queste strutture ibride il verbo tenda a essere in croato, e il complemento che poi innescherà il *code-switching* in italiano. La direzionalità del *code-switching* in questo caso sembra dunque essere rilevante, ed è anche significativo che i verbi in croato con complemento in italiano abbiano tutti un significato molto generico (come “dire”, “fare”, “avere”). Si tratta, nella quasi totalità dei casi, di costruzioni a verbo supporto (cfr. Ježek, 2005: 181 sgg.), perché la base, il nome che è testa del complemento, determina il significato del collocato (il verbo); la funzione di predicato è inoltre assunta dal nome, mentre il verbo funge da supporto esprimendo valori tipicamente verbali come il tempo e il modo.

È il caso della seguente occorrenza:

(19) DIALOGO III

S46: kako nemas **fame**, **vieni!**
 “Come non hai fame”

8. *Catene di code-switching*

Concluso questo excursus sulle funzioni del *code-switching* riscontrate nei nostri sub-corpora, mostriamo come, nello stesso contesto, diversi *code-switching* possano susseguirsi in rapida sequenza per esprimere più funzioni. Mostriamo quindi un turno di A che presenta 4 occorrenze di *code-switching* (e alcuni *code-mixing* dovuti a *congruent lexicalization*, cfr. Muysken, 2000)

(20) DIALOGO I

A95: Și ăsta se chiamă **Baciolino**. **Sai perché l’ho chiamato così?** Pentru că aicea scrie **Ti meriti un bacio**. **Anche a Baciolino piace Linda**. Pentru că e bella. Și Linda è la mia preferita in tutte. **Și când se sporcă o spăl**.
 “E questo si chiama.... Perché qui c’è scritto.... Perché. E Linda è...
 E quando si sporca la lavo”

il primo *code-switching* (verso l’italiano) è un caso di *priming*, innescato dal nome dell’orsetto in italiano, Baciolino; il secondo *code-switching* (verso il romeno), “pentru că aicea scrie”, è un *code-switching* con funzione di specificazione; il terzo *code-switching*, “ti meriti un bacio” è un caso di citazione. L’ultimo *code-switching* verso il romeno, “și când se sporcă o spal”, che contiene anche un evidente *code-mixing* (“se sporca”), ha invece la funzione di commento.

9. *Code-switching e coerenza tra turni*

Un ultimo aspetto che vogliamo mettere in evidenza è come l’emersione dei *code-switching* sia legata allo sviluppo della conversazione e a problemi di gestionale locale tra turni. Si può osservare come la probabilità di occorrenza di *code-switching* tra

turni sequenziali prodotti da parlanti diversi sia condizionata dal grado di coerenza. Con le parole di Auer “coherence, in our sense, is accomplished by pointing back from one turn to another, or from one turn component to another, or forward to some utterance which is still supposed to follow. It is not restricted to grammatical structures [...] but includes sequentiality and topical cohesion as well. In addition, we need to include consistency of language choice among the factors producing coherence” (Auer 1984: 51). Considerando solo i casi di *code-switching* tra turni prodotti da parlanti diversi, i soli in cui è possibile verificare il grado di coerenza conversazionale (12 occorrenze su 29 nel sub-corpus croato; 21 su 36 nel sub-corpus romeno), notiamo come sia significativo che i *code-switching* tra turni avvengano soprattutto in casi di bassa coerenza tra due turni conversazionali (Auer, 1984: 51-52). Il *code-switching*, cioè, si verifica raramente nei casi in cui è direttamente richiesto un contributo comunicativo del co-partecipante alla conversazione (come nelle sequenze domanda-risposta o nel caso di coppie rigidamente adiacenti), mentre è più frequente nei casi in cui chi prende il turno è posto di fronte a più scelte possibili dal punto di vista della replica conversazionale.

9.1 Coerenza tra turni nel sub-corpus croato

Nel sub-corpus croato il *code-switching* avviene perlopiù in casi di bassa coerenza e con funzione discorsiva, tra cui l'inaugurazione di un nuovo segmento conversazionale, come al turno F60 dell'es. 11, che inaugura un nuovo topic, oppure nella apertura/chiusura di sequenze inserto, come nel caso della sequenza all'es. 12 che viene inaugurata (turno S18) e conclusa (turni P20/M21) con due *code-switching*.

Altrimenti il *code-switching* può essere usato in risposta a un ordine, con funzione di contrasto, come si può osservare dai primi due turni all'es. 2. Un ordine costituisce una mossa “forte” all'interno della conversazione, e in quanto tale esige una risposta dall'interlocutore, ma non genera alta coerenza con il turno successivo. Ciò è dovuto a una serie di motivi: un ordine non prevede necessariamente un'accettazione come risposta preferenziale; la risposta può essere sia linguistica che non linguistica (perché l'ordine chiede a chi lo riceve di “fare” qualcosa, non di “dire” qualcosa); la probabilità di negoziazione della forza illocutoria dell'atto è abbastanza alta.

Nel caso dell'es. 2, dunque, il *code-switching* ha la funzione di esprimere un contrasto tra le interagenti.

La funzione di contrasto sembra però innescare il *code-switching* anche nell'unico caso in cui la commutazione di codice avviene in condizioni di alta coerenza, ovvero dopo una domanda sì/no:

(21) DIALOGO I

S24: Ma jel ja moram F sad vozit svaki dan u skolu?

“Ma io devo portare F tutte le mattine a scuola?”

M25: Certo.

Ancora una volta la direzionalità del *code-switching* non sembra rilevante, considerando che in una metà delle occorrenze di *code-switching* con bassa coerenza la commutazione avviene verso l'italiano, e nell'altra metà verso il croato.

I casi di correzione dovuta a contratto pedagogico non possono essere analizzati relativamente al loro grado di coerenza, perché la lingua in questi casi è simultaneamente oggetto e strumento della comunicazione:

(22) DIALOGO II

F37: Ma sta su le zamjenice, S?

"Ma cosa sono i pronomi, S?"

S38: **Sono i pronomi, "zamjenica" si dice in croato.**

9.2 Coerenza tra turni nel sub-corpus rumeno

Anche nel sub-corpus romeno i *code-switching* si verificano soprattutto in casi di bassa coerenza conversazionale, e cioè in tutte le occorrenze di cambiamento di topic o di inaugurazione di una *side-sequence*, come nel caso del turno M67 all'es. 13.

Nel sub-corpus romeno i 6 *code-switching* con funzione di divergenza (o di convergenza), tutti tra due turni diversi, sono espressi in condizione di bassa coerenza conversazionale. Il seguente segmento conversazionale mostra un esempio della funzione di divergenza, ancora dopo un ordine:

(23) DIALOGO II

N9: După aia îți faci geanta că după aia te duci în față la televizor și uiți de ea.

"Poi ti prepari lo zaino che poi vai davanti alla televisione e te ne dimentichi"

A10: **Uffa, da, da. Va bene.**

"sì, sì"

Mentre l'es. 24 mostra l'unico *code-switching* del corpus con funzione di convergenza:

(24) DIALOGO II

N17: Nu. Nu cunoaștem, nu cunosc familia, nu...

"No, non conosciamo, non conosco la famiglia, non..."

A18: **No, ma in fondo hai ragione.**

Nell'es. 25 la bambina, al turno A44, usa addirittura il *code-switching* in funzione di contrasto per riaprire una sequenza conversazionale strutturata tipicamente "a tripletta" (Sinclair. Coulthard, 1975), e quindi sigillata dal terzo turno (N43: "ho capito"):

(25) DIALOGO II

N41: Știu de ce: ai mâncat salam când ai venit și de-aia nu ți-e foame

"So io perché: hai mangiato del salame quando sei arrivata e perciò non hai fame"

A42: Dar nu mi-e foame de loc, așa în general

"Ma non ho proprio fame, così in generale"

N43: Am înțeles

"Ho capito"

A44: **No, no, non mi viene proprio. Cioè non ce la faccio**

Quando il *code-switching* è invece dovuto al dominio, questa funzione macro-sociolinguistica sembra prevalere rispetto al peso della coerenza conversazionale. In tutte le sequenze di intervista in romeno da parte della intervistatrice, quando il dominio è la scuola o l'abbigliamento, si verifica un *code-switching* anche in risposta a domande dirette:

(26) DIALOGO I

I41: *Și așa de îmbracaminte, de culori ce-ți place? Care-i culoarea ta preferata?
"E così, da vestire, di colori, cos'è che ti piace? Qual'è il tuo colore preferito?"*

A42: **Chiari, colori chiari**

I43: *Azzurro, așa?
"Azzurro, così?"*

A44: **Si. Azzurro, fucsia, bianco...**

I45: *Aha. Și ce-ți place să-ț pui mai mult, pantaloni sau fustițe?
"E cosa ti piace indossare di più, pantaloni o gonnelline?"*

A46: **Ugual**

10. *Parallelismi con i code-switching italiano/dialetto*

In conclusione vorremmo sottolineare come possano valere per i nostri sub-corpora le stesse considerazioni che gli studiosi italiani hanno esplicitato relativamente ai loro corpora di enunciazioni mistilingui tra italiano e dialetti italiani, anche se i parlanti la cui produzione è stata analizzata in questo lavoro presentano un bilinguismo tra italiano e un'altra lingua (romeno, croato) con una distanza strutturale maggiore di quella che esiste tra l'italiano e i suoi dialetti.

Anche per i nostri parlanti:

- non esiste una forte compartimentazione sociale tra i due codici, e il diverso prestigio dei due codici o una loro diversa collocazione nel repertorio non sembra rilevante in termini di significato sociale;
- entrambi i codici sono considerati accettabili e il loro uso alternato viene effettivamente accettato come pratica continua, diffusa (cfr. Cerruti, Regis, 2005: 182) e automatizzata nella conversazione ordinaria, per cui il discorso bilingue si configura come "scelta non marcata nella maggior parte delle situazioni non formali" (ivi: 186);
- esiste un alto grado di sovrapposibilità funzionale e di intercambiabilità tra i due codici (Berruto, 1990: 118), il che implica anche un alto grado di reversibilità nella direzione del *code-switching*, che assume valore pragmatico in sé e diviene segnico delle intenzioni comunicative dei parlanti a prescindere dal codice verso cui la commutazione avviene (Cerruti, Regis, 2005: 182; Cerruti, 2004: 123);
- la dimensione informale dello scambio verbale e i rapporti familiari tra gli interlocutori favoriscono l'impiego congiunto dei due codici in uno stesso evento dialogico;
- la scelta di usare un codice diverso dall'italiano sembra dipendere più da abitudini familiari che da un grado diverso di conoscenza di uno dei due codici;

- laddove i parlanti mantengono scelte di lingua preferenziali parzialmente divergenti, le scelte divergenti non sono fonte di tensione tra gli interlocutori (cfr. Cerruti, 2004: 97);
- nelle micro-comunità si creano pattern conversazionali come quello per cui il codice più prestigioso, l'italiano, viene usato più spesso con i bambini o con i giovani: le commutazioni di codice sembrano quindi essere condizionate più da fattori sociolinguistici come l'età che da altri tipi di fattori di linguistica interna (cfr. Alfonzetti, 1992; Berruto, 2005: 10); la eventuale preferenza linguistica per un codice dipende dunque dalla biografia linguistica dei parlanti, e l'età si rivela un fattore particolarmente significativo (Giacalone Ramat, 1995: 52);
- esattamente come accade nella conversazione bilingue in italiano e dialetto, i *code-switching* operanti nei sub-corpora analizzati apparentemente non vengono usati in funzione identitaria, e non pare possibile assegnare a ciascun codice l'espressione di atteggiamenti e sentimenti tipicamente connessi all'opposizione gumperziana tra *we-code* e *they-code* (cfr. Alfonzetti 1992: 249). I due codici, infatti, non sembrano esprimere un sentimento di appartenenza a un gruppo etnico in opposizione alla comunità di maggioranza e alla sua lingua ufficiale.

11. Conclusioni

In questo lavoro sono stati analizzati i fenomeni di *code-switching* in tre famiglie bilingui (una famiglia bilingue croato-italiano, due famiglie bilingui romeno-italiano) residenti in Piemonte, adottando un approccio pragmatico-funzionale che considera i *code-switching* come strategie interazionali e conversazionali, interpretabili solo nel contesto.

I *code-switching* riscontrati nei due sub-corpora croato-italiano e romeno-italiano rivestono sia funzioni legate ai partecipanti (soprattutto con funzione di contrasto e di commento, cfr. par. 4) sia funzioni connesse al discorso (come l'inaugurazione o la chiusura di un topic, la gestione di sequenze marginali, o il cambio di costellazione dei partecipanti, cfr. par. 5). Nel solo sub-corpus romeno i *code-switching* sembrano correlare nel parlato delle due bambine anche a specifici domini (la scuola, il gioco, cioè attività che gestiscono in italiano, cfr. par. 6), e in alcuni casi i *code-switching* non sembrano corrispondere a istanze funzionali ma dipendono da un'alta attivazione funzionale di entrambi i codici, e sono spesso preceduti da casi di *code-mixing* dovuti a "congruent lexicalization" (cfr. par. 7).

I *code-switching* tra turni di parlanti diversi dei due sub-corpora rivelano fenomeni di risonanze di frame diattico e frasali ben inquadrabili nel modello teorico della sintassi dialogica (Du Bois, 2014), e si nota inoltre come i *code-switching* tra parlanti diversi avvengano tendenzialmente in casi di bassa coerenza tra turni, cioè in cui non sussiste un forte condizionamento nella selezione della mossa conversazionale dovuta alla formulazione del turno precedente (come accadrebbe nel caso delle domande sì/no, che richiedono appunto una risposta che equivalga a un *sì* o a

un *no*), ed è invece possibile avere più scelte dal punto di vista della replica conversazionale (cfr. par. 9).

Anche nei nostri sub-corpora, come per le enunciazioni mistilingui italiano-dialetto esaminate in molti studi (cfr. par. 10), si nota un alto grado di sovrapposibilità funzionale e di intercambiabilità tra i due codici, e come il loro uso alternato sia accettato e automatizzato nella conversazione senza rivestire un particolare significato sociale o identitario, del resto poco pertinente all'interno di un contesto familiare.

Riferimenti bibliografici

- ALFONZETTI, G. (1992). *Il discorso bilingue. Italiano e dialetto a Catania*. Milano: Franco Angeli.
- AUER, P. (1984). *Bilingual Conversation*. Amsterdam: Benjamins.
- AUER, P. (1995). The pragmatics of code-switching: a sequential approach. In MILROY, L. & MUYSKEN, P. (Eds.) *One speaker, two languages. Cross-disciplinary perspectives on code-switching*. Cambridge: Cambridge University Press, 115-135.
- BERRUTO, G. (1990). Italiano regionale, commutazione di codice e enunciati mistilingui. In CORTELAZZO, M.A. & MIONI, A.M. (Eds.) (1990), *L'italiano regionale, Atti del XVIII Congresso SLI*, Padova-Vicenza, Italia, 14-16 settembre 1984, Roma: Bulzoni, 105-130.
- BERRUTO, G. (2005). Che cosa ci insegna il 'parlare due lingue'? Commutazione di codice e teoria linguistica e sociolinguistica. In *Rivista di Linguistica* 17/1, 3-14.
- BETTONI C. (2006). *Usare un'altra lingua*. Roma-Bari: Laterza.
- CASTAGNETO, M. & STAN, I.S. (2020). Cross-Linguistic Influence in L2 Italian Bilingual Families: A Comparison between Conversations in a Croatian and two Romanian Families. In *Italiano Lingua Due* 12/2, 155-168.
- CERRUTI, M. (2004). Aspetti pragmatico-funzionali della commutazione di codice italiano-dialetto: un'indagine a Torino. In *Vox Romanica* 63, 94-127.
- CERRUTI, M. & REGIS, R. (2005). 'Code switching' e teoria linguistica: la situazione italo-romanza. In *Rivista di Linguistica* 17/1, 179-208.
- CLYNE, M. (1969). Switching between Language Systems. In GRAUR, A. (Ed.) *Actes du Xe Congres International des Linguistes* (August 1967). Bucharest: Editions de l'Academie la Republique Socialiste de Roumanie, 343-349.
- DU BOIS, J. (2014). Towards a Dialogic Syntax. In *Cognitive Linguistics* 25/3, 359-410.
- DU BOIS, J. & GIORA, R. (2014). From cognitive-functional linguistics to dialogic syntax. In *Cognitive Linguistics* 25/3, 351-357.
- ERVIN S.M. & OSGOOD C. (1954). Second language learning and bilingualism. In *Journal of Abnormal and Social Psychology* 58, 139-145.
- FRANCESCATO G. (1981). *Il bilingue isolato*. Bergamo: Minerva Italica.
- GIACALONE RAMAT, A. (1995). Code-switching in the context of dialect/standard language relations. In MILROY, L. & MUYSKEN, P. (Eds.) *One speaker, two languages. Cross-disciplinary perspectives on code-switching*. Cambridge: Cambridge University Press, 45-67.
- GILES, H. & ST. CLAIR, R. (1979). *Language and social psychology*. Oxford: Blackwell.

- GUMPERZ, J.J. (1982). *Discourse Strategies*. Cambridge: Cambridge University Press.
- HOBSON, R.P.; HOBSON, J.A.; GARCÍA-PÉREZ, R. & J. DU BOIS (2012). Dialogic Linkage and Resonance in Autism. In *Journal of Autism and Developmental Disorders* 42/12, 2718-2728.
- HYMES, D. (1974). *Foundations of Sociolinguistics: An Ethnographic Approach*. Philadelphia: University of Pennsylvania Press.
- JEŽEK, E. (2005). *Lessico. Classi di parole, strutture, combinazioni*. Bologna: il Mulino.
- MUYSKEN, P. (2000). *Bilingual Speech. A Typology of Code-Mixing*. Cambridge: Cambridge University Press.
- MYERS-SCOTTON, C. (2002). *Contact linguistics. Bilingual Encounters and Grammatical Outcomes*. Oxford: Oxford University Press.
- ORLETTI, F. (2004). *La conversazione diseguale*. Roma: Carocci.
- PEIRCE, CH.S. (1931-1935). *Collected Papers*. Cambridge (Mass.): Belknap Press of Harvard University Press. Trad. it.: BONFANTINI, M.A.; GRASSI, L. & GRAZIA, R. (Eds.) *Semiotica*. Torino: Einaudi, 1980.
- PEJAKOVIC M. (2014). *Fenomeni di interferenza linguistica tra italiano e croato nell'interazione familiare*. Tesi di laurea triennale, Università degli Studi del Piemonte Orientale.
- REGIS, R. (2005). *Appunti grammaticali sull'enunciazione mistilingue*. München: Lincom Europa.
- ROMAINE, S. (1995). *Bilingualism*. Oxford: Blackwell.
- SAVY, R. (2006). Specifiche per la trascrizione ortografica annotata dei testi raccolti. In ALBANO LEONI, F. & GIORDANO, R. *Italiano Parlato. Analisi di un dialogo*. Napoli: Liguori, 1-37.
- SINCLAIR, J. & COUTHARD, R.M. (1975). *Toward an Analysis of Discourse. The English used by Teachers and Pupils*. London: Oxford University Press.

FRANCESCA PEZZELLA

La geminazione in un gruppo di apprendenti albanofoni: percezione e produzione di una nuova categoria fonologica

Geminate consonants by a group of Albanian learners of L2 Italian: perception and production of a non-native contrast

The non-native consonantal length contrast is difficult to acquire due to its articulatory complexity and typological markedness. This study aims at investigating the perception and the oral and written production of Italian geminates in a small group of Albanian adult learners with a different competence in L2 Italian. A set of tasks with various levels of complexity, including words and non-words in isolation and in context were used. The consonant perception and production errors were calculated in all tests crossing the interlinguistic and tasks variables. The results confirmed the major difficulty in oral production, whereas the perception and written production reached a high level of accuracy. Particularly, the intermediate-advanced learners showed lower performances than the beginners in the completion of complex task demands including non-words.

Keywords: gemination, Italian L2, Albanian, perception, production.

1. Introduzione

Nel percorso di avvicinamento alla fonologia di una lingua seconda (L2) da parte di un apprendente adulto, rispetto a quello che accade per il lessico, la morfosintassi e la morfologia, è più forte e resistente il *transfer* dalla L1, la cui azione in ogni caso si interseca con i principi generali di marcatezza operanti anche nella costruzione del sistema fonologico della L1 (Chini, 2005). Sebbene la presenza dell'accento straniero sia indipendente dal buon esito della comunicazione (*intelligibility*) e possa avere su di esso un peso maggiore o minore in base alla natura dei fenomeni fonetico/fonologici e ai contesti comunicativi coinvolti, la competenza fonologica nella L2 necessita di essere acquisita in maniera funzionale ai bisogni dell'apprendente e all'efficacia comunicativa (D'Apolito, Gili Fivela, 2019; Derwing, Munro, 2015).

La geminazione è uno degli aspetti del consonantismo dell'italiano responsabile di numerose difficoltà nella pronuncia degli apprendenti di italiano L2. Si tratta di un fenomeno fonologico di rafforzamento, articolatoriamente complesso, e marcato a livello implicazionale e tipologico. L'UPSID (Maddieson, 1984), il più aggiornato *database* fonetico, registra la presenza delle geminate solo in 11 delle 317 lingue del campione rappresentativo. Nella letteratura internazionale le ricerche

sull'acquisizione delle geminate riguardano perlopiù il giapponese L2 di parlanti inglesi e si concentrano maggiormente sulle occlusive e sulla produzione orale (Mah, Archibald, 2003; Han, 1992).

Oggetto di questo studio è l'analisi del contrasto di lunghezza consonantica nelle abilità di percezione uditiva e di produzione orale e scritta di un piccolo gruppo di apprendenti adulti di italiano con madrelingua albanese. Il lavoro è il risultato della mia tesi di Laurea Magistrale in Linguistica¹, svolta in collaborazione con la Fondazione Ugo Bordoni e con il Centro di Studi Albanesi² della Sapienza Università di Roma, presso cui ho avuto la possibilità di interagire con alcuni studenti albanofoni. Lo studio prende principalmente le mosse dalla volontà di approfondire l'interazione tra gli aspetti percettivi e produttivi dell'acquisizione fonologica e di intervenire su una comunità come quella albanese ampiamente rappresentata nel nostro Paese³.

“Falsi principianti”, così in genere molti insegnanti definiscono gli albanofoni (Favaro, 2002: 235), facilitati nella comprensione dell'italiano grazie alla grande diffusione di cui gode la nostra lingua nel Paese delle Aquile (Serianni, 2013; Di Giovine, 2008). Sul piano fonetico-fonologico, se da un lato la presenza di tratti simili e più complessi nella lingua albanese⁴ agevola l'apprendimento dell'italiano, l'assenza della categoria delle geminate a livello segmentale e l'accento generalmente piano a livello suprasegmentale (non rappresentato graficamente) determinano numerosi errori nella pronuncia e nella scrittura dell'italiano (Dashi, 2013: 18; Favaro, 2002: 236).

L'analisi condotta in questo studio sui dati relativi alle competenze fonologiche di alcuni apprendenti albanesi non consente di individuare la significatività dei risultati dal punto di vista statistico per l'esiguità del campione sperimentale e per il mancato bilanciamento del *corpus* utilizzato. Dall'interpretazione dei dati emergono tuttavia delle tendenze sul grado e sulle modalità di acquisizione fonologica della L2 che potrebbero altresì costituire un punto di partenza per riflettere su future implicazioni in ambito didattico.

¹ La tesi è stata realizzata sotto la supervisione della relatrice Maria Roccaforte (Sapienza Università di Roma), della correlatrice Brunilda Dashi (Sapienza Università di Roma) e del relatore esterno Mauro Falcone (Fondazione Ugo Bordoni).

² Nell'ateneo romano gli studi di Albanologia sono stati inaugurati dalla cattedra di Ernesto Koliqi, per merito del quale con il decreto del Presidente della Repubblica del 1957 la cattedra è diventata Istituto (Gradilone, 1974: 270). Al maestro Koliqi è subentrato Giuseppe Gradilone e in anni più recenti Elio Miracco. Tra il 2008 e il 2010 l'Istituto è diventato Centro di Studi Albanesi, conflueno nel Dipartimento di Studi Glotto-Antropologici e Linguistici della Sapienza.

³ Secondo i dati ISTAT aggiornati al 1° gennaio 2019, gli albanesi rappresentano l'8,4% dei cittadini stranieri aventi dimora abituale in Italia, preceduti dai romeni (23%) e seguiti dai marocchini (8,0%). Inoltre, stando ai dati del MIUR aggiornati a marzo 2018 e relativi all'anno 2017/2018, dei circa 842.000 alunni con cittadinanza non italiana il 13% è rappresentato dagli albanesi, preceduti dai romeni (18,8%) e seguiti dai marocchini (12,3%).

⁴ La lingua albanese possiede 36 fonemi (7 vocali e 29 consonanti) e strutture sillabiche più complesse rispetto a quelle dell'italiano standard: in attacco di sillaba possono ricorrere più di tre consonanti, come in *shndrit* “brillare”.

2. Fattori rilevanti nell'acquisizione fonologica di una L2

Numerosi fattori di svariata natura interagiscono nel processo di acquisizione di una categoria nuova e marcata, come la quantità consonantica distintiva. La fonologia tradizionale e i primi studi sul contatto linguistico (Lado, 1957; Weinreich, 1953; Trubetzkoy, 1939), basati sull'ipotesi interferenzista, sostenevano che le differenze acustiche della L2 non rilevanti a livello fonemico nella L1 sono ignorate e adattate alla fonologia della L1. La versione moderata della *Contrastive Analysis* (Oller, Ziahosseiny, 1970) fa un passo avanti affermando che il *transfer* dalla L1 non è solo negativo, ma agisce positivamente in caso di lingue simili. Dagli anni '80 gli studi di natura cognitivo-funzionale riconsiderano la nozione di *transfer* unitamente ai fattori universali a cui sono legati gli errori evolutivi operanti nell'acquisizione della L1 e della L2. Il modello della *Categorical perception* (Wode, 1992), applicato alla percezione interlinguistica, afferma che la discriminazione fonologica è più accurata se i foni della L2 sono classificati come istanze di una categoria diversa dai fonemi della L1; se invece i suoni della L2 sono classificati come appartenenti alla stessa categoria fonologica del fonema della L1 più vicino e simile, le differenze fonetiche vengono ignorate e di conseguenza non acquisite.

Sulla scia di questi studi, James Emil Flege elabora lo *Speech Learning Model* (SLM; Flege, 1995) che, forte dell'evidenza sperimentale delle numerose ricerche condotte dall'autore e dai suoi collaboratori, dimostra che i processi adibiti all'acquisizione della L1 restano intatti per tutta la vita, permettendo agli apprendenti adulti di accedere ai tratti nuovi attraverso la reinterpretazione e il riordino categoriale del proprio inventario fonologico. L'adeguatezza dell'*input*, l'età precoce di apprendimento (*age of learning*, AOL) e di arrivo (*age of arrival*, AOA) contribuiscono all'esito positivo dell'acquisizione della L2, insieme alla distanza fonetica percepita tra L1 e L2. Maggiori sono le differenze fonetiche tra due lingue, più è probabile la formazione di nuove categorie, bloccata invece nell'acquisizione di fonemi simili per intervento del meccanismo cognitivo dell'*Equivalence Classification*. La *perceived phonetic cross-language distance* è anche al centro del *Perceptual Assimilation Model* (PAM; Best, 1991) che stabilisce come ulteriore fattore positivo nell'acquisizione l'esperienza allofonica nella L1 dei contrasti non-nativi della L2.

Lo SLM e il PAM non tengono conto della marcatezza, un altro predittore delle difficoltà nei processi di acquisizione di una L2, posto alla base della *Markedness Differential Hypothesis* (Eckmann, 1977) e del più recente *Ontogeny Phylogeny Model* (OPM; Major, 2001). Major osserva che lo sviluppo cronologico dell'interlingua, determinato dall'interrelazione tra L1, L2 e proprietà universali (U), procede più lentamente quando incontra nella L2 fenomeni simili o marcati. Nell'acquisizione di tratti marcati, U ha un'influenza più forte rispetto al *transfer* dalla L1.

Il lessico costituisce un altro fattore determinante: l'introduzione dei dettagli fonetici specifici di una lingua nell'interlingua (IL) dell'apprendente va di pari passo con la ridondanza e la frequenza di essi nel lessico d'uso (modello *usage-based*: Bybee, 2000). In uno dei primi studi sulle geminate nell'italiano L2 (Giannini, Costamagna, 1998), la teoria della *Diffusione lessicale* (Wang, 1969), applicata ini-

zialmente all'introduzione di tratti innovativi nell'evoluzione storica delle lingue naturali, viene estesa su basi psicolinguistiche alla diacronia individuale dell'apprendente. Lo studio dimostra che il tratto di lunghezza consonantica tende a stabilizzarsi prima in classi lessicali dotate di condizioni prosodiche e morfologiche favorevoli all'elaborazione del tratto (parole bisillabiche di elevata frequenza, con la consonante geminata in post-tonia e con la funzione di verbi o modificatori); poi, una volta immagazzinato il modello, l'apprendente lo applica al resto del lessico in cui riconosce tutte o qualcuna delle caratteristiche dello *schema* iniziale.

Anche le relazioni inter-ortografiche tra L1 e L2 svolgono un ruolo determinante nella percezione e nella produzione del contrasto di lunghezza consonantica. In una ricerca condotta nel 2017 da Bene Bassetti su studenti italiani di inglese, si evidenzia che durante la produzione orale di parole inglesi in contesto di frase, con e senza *input* ortografico, gli apprendenti, attivando simultaneamente la rappresentazione ortografica e fonologica, risentono maggiormente dell'influsso del sistema ortografico della L1.

Naturalmente nell'interpretazione dei dati sperimentali si è tenuto conto dei fattori psicologici e sociali, in particolare la motivazione e l'integrazione nella comunità italiana, considerati decisivi nell'acquisizione fonologica in chiave cognitivo-funzionale e ambientalista (De Bot, Lowie & Verspoor, 2005).

3. *La geminazione*

3.1 Alcune considerazioni teoriche

Sulla natura fonologica mono- o bifonematica delle geminate vi è un lungo dibattito. Nella tradizionale fonologia lineare le prime analisi risalgono a Swadesh (1937) e Trubetzkoy (1939) che considerano le geminate come segmenti lunghi; mentre secondo Hockett (1955), la geminazione rappresenta una sequenza di due consonanti uguali. Tali ipotesi contrastanti sono dovute al fatto che le consonanti geminate si comportano come un unico segmento rispetto ad alcuni fenomeni e come due segmenti uguali rispetto ad altri. Anche le analisi fonologiche di natura generativa (Chomsky, Halle, 1968) colgono l'ambiguità della geminazione ma non la superano.

Un'interpretazione più adeguata alla natura e al comportamento delle geminate deriva dalla fonologia autosegmentale (Goldsmith, 1976) basata su una rappresentazione multilineare. Le caratteristiche articolatorie di un messaggio sonoro sono rappresentate su livelli diversi, piani indipendenti l'uno dall'altro ma collegati tramite linee di associazione alla base comune delle unità temporali (*skeletal tier*). Stando a questo modello, le consonanti geminate vengono considerate come due unità sul piano temporale e associate ad un'unica specificazione sul piano dei tratti. Il comportamento diverso delle geminate come uno o due segmenti è legato ad uno dei due livelli, quello dei tratti per fenomeni che riguardano la qualità dei segmenti e quello delle unità temporali per fenomeni relativi alla struttura soprasegmentale. Un'interpretazione del genere, di natura sillabico-prosodica (Perlmutter, 1995), consente di conservare le due caratteristiche fondamentali della geminazione, con-

traddette da una spiegazione mono- o bifonematica: l'integrità, per cui tra le consonanti doppie non possono inserirsi fenomeni di epentesi, e l'inalterabilità, per cui in presenza delle geminate le regole che riguardano i segmenti singoli si bloccano.

Il principale correlato acustico e percettivo delle geminate è l'aumento della durata, un tratto fonetico linguo-specifico, che varia in dipendenza di numerosi fattori prosodici contestuali: tipo di consonante, struttura sillabica della parola, accento lessicale (post-tonia, pre-tonia, tra vocali atone), accento frasale, velocità di elocuzione, accorciamento *vs.* allungamento della vocale precedente. Varia da lingua a lingua anche l'organizzazione temporale dei gesti articolatori nella produzione delle geminate, oggetto di studi di carattere cinematico (Smith, 1995).

A livello distribuzionale, oltre che in posizione intervocalica, alcune lingue, come levantino e ungherese (Ham, 2001), presentano la geminazione in posizione iniziale e finale di parola. Il rafforzamento consonantico si può verificare anche a livello fonotattico nell'incontro tra due parole che iniziano e terminano con la stessa consonante, come in [al:a'vo:ro] (*false geminates*; Payne, 2005: 154-155), e, in italiano, tra due parole in specifiche condizioni prosodiche [e'v:ɛ:ro] o morfologiche [a'k:a:sa] innescanti il rafforzamento fonosintattico (Mioni, 1993: 131-132).

3.2 Le geminate in italiano L1 e L2

In italiano, l'opposizione fonologica di quantità si verifica in 15 fonemi consonantici /p b t d k g tʃ ɕ f v s m n l r/; la lunghezza non è distintiva nelle consonanti intrinsecamente lunghe, affricate alveodentali /t:s d:z/ e sonoranti palatali /ʃ: ʎ: ɲ:/; la sibilante sonora /z/ è sempre scempia.

Il rapporto di durata tra consonanti scempie (C) e geminate (CC) è stato indagato in numerosi studi sull'italiano L1⁵, relativi soprattutto alle occlusive. In queste ultime si osserva che il rapporto CC:C è compreso tra circa 1,8 e 1,9 (Esposito, Di Benedetto, 1999; Cerrato, Falcone, 1998). Le geminate fricative, caratterizzate dal tratto [+ continuo], tendono in genere ad essere più lunghe rispetto alle occlusive (Giovannardi, Di Benedetto, 1998; Bertinetto, Vivalda, 1978) e il rapporto tra scempie e geminate aumenta passando dalle nasali (circa 2,3; Mattei, Di Benedetto, 2000) alle laterali (2,7; Payne, 2005). Nelle affricate palatali il rapporto CC:C si ferma intorno a 1,5 (Faluschi, Di Benedetto, 2012). Nello studio acustico condotto da Payne (2005) su sette fonemi dell'italiano /p t b d m l f/, si evidenzia per /f/ e /p:/ la durata più lunga e per /l:/ la durata più breve. Le geminate sonore sono più brevi rispetto alle sorde corrispondenti, data la difficoltà aerodinamica nel mantenere per un tempo prolungato la sonorità attraverso la chiusura sopra-laringea (Kabak, Reckziegel & Braun, 2012: 997). Tale difficoltà articolatoria rende le geminate sonore meno frequenti nelle lingue rispetto alle sorde. Inoltre, come atteso, Payne

⁵ Gli studi citati si basano sulle varietà centrali dell'italiano (parlanti di area romana e toscana). In questa sede non sono stati considerati studi sulle varietà settentrionali dell'italiano in cui in modo variabile a livello diafasico, diastratico e diatopico si può assistere per influenza del dialetto a processi di degemina-zione o alla produzione di geminate con una durata più breve (Mairano, De Iacovo, 2020; Mioni, 1993).

rileva una durata più lunga nelle geminate in post-tonia rispetto alla posizione di pre-tonia e di atonia.

Tra i primi studi di carattere cinematico sulle geminate dell'italiano, oltre al già citato Smith (1995), vi è quello condotto da Zmarich & Gili Fivela (2005) che, in particolare, ha rilevato per la geminazione correlati simili ai *cluster* consonantici, considerando ciò come prova dell'eterosillabicità delle geminate, e ha evidenziato la maggiore stabilità dei parametri cinematici rispetto a quelli acustici in condizioni di velocità di eloquio accelerata.

Le ricerche sulla geminazione in italiano L2 sono aumentate soprattutto in tempi recenti e riguardano perlopiù la produzione orale di consonanti occlusive. In Soriano (2009, 2014) si osserva che apprendenti spagnoli, tedeschi e inglesi, nel contesto di una conversazione spontanea, producono correttamente le geminate in circa il 50% dei casi, mostrando una forte variabilità intersoggettiva e un rapporto CC:C più basso nelle occlusive rispetto alle sonoranti. La geminazione si stabilizza prima nelle occlusive sorde, consonanti frequenti e forti, dunque più propense a subire un fenomeno di rafforzamento, poi in consonanti più deboli, come le sonoranti. In un altro studio condotto da Soriano nel 2019 relativamente alla produzione di affricate scempie e geminate in una conversazione guidata, un gruppo di albanofoni realizza le affricate geminate come degeminate nell'84% dei casi.

I risultati di un test di ripetizione orale condotto su apprendenti tedeschi di italiano in fase avanzata e su tedeschi privi di conoscenze in italiano (Kabak, Reckziegel & Braun, 2011), mostrano che entrambi i gruppi sperimentali non raggiungono il livello *target*, ma i parlanti esperti si avvicinano ad esso. Inoltre, come sostenuto anche da Han (1992), gli apprendenti tedeschi tendono a pronunciare le consonanti scempie dell'italiano con una durata più lunga rispetto al gruppo che non conosce l'italiano, andando quindi a "riaggiustare" una categoria già presente nella L1, invece di migliorare la produzione della nuova categoria.

D'Apolito e Gili Fivela (2019) mettono a confronto due gruppi di francofoni in fase iniziale e avanzata di apprendimento dell'italiano, nella produzione di geminate inserite in frasi caratterizzate da un contesto povero e da un contesto ricco di informazioni. I parlanti inesperti presentano risultati molto oscillanti, mentre l'altro gruppo tende ad avvicinarsi alla realizzazione nativa. Come atteso, la produzione è più accurata nel contesto povero rispetto a quello ricco, contenente molte pause ed esitazioni.

Dal punto di vista percettivo, in generale la letteratura mostra un alto grado di accuratezza nella discriminazione del tratto di lunghezza consonantica. Tsukada, Cox, Hajek & Hirata (2017) hanno sottoposto gruppi di parlanti coreani, australiani, giapponesi e italiani ad un esercizio di discriminazione fonologica tra scempie e geminate in giapponese e in italiano. I coreani e gli australiani avevano il giapponese come FL (*foreign language*) ed erano privi di conoscenze in italiano; i giapponesi non conoscevano l'italiano, così come il giapponese non era noto agli italiani. Nella percezione della FL e della lingua sconosciuta, i partecipanti raggiungono più dell'80% di risposte corrette. Sia i coreani che gli australiani sono in grado di trasferire la conoscenza del nuovo tratto (le geminate) della FL ad una terza lingua; giapponesi e italiani, pur

possedendo il tratto nelle loro L1, non raggiungono il livello nativo nella discriminazione delle geminate della lingua sconosciuta. Si conferma dunque la difficoltà nel percepire la quantità consonantica. Inoltre, gli apprendenti commettono errori di diverso tipo in giapponese e in italiano, a dimostrazione del fatto che la categorizzazione della lunghezza consonantica avviene in modi differenti, non applicando il concetto generale di lunghezza in tutte le lingue aventi il medesimo tratto distintivo.

Altman, Berger & Braun (2012) mettono a confronto italiani, tedeschi privi di conoscenze in italiano e tedeschi in fase avanzata di apprendimento dell'italiano nella discriminazione della lunghezza vocalica e consonantica di non-parole. Si nota che gli apprendenti tedeschi di italiano, pur commettendo meno errori nel riconoscimento delle geminate, presentano i medesimi tempi di reazione lunghi dell'altro gruppo di germanofoni. Gli italiani mostrano gli stessi tempi di reazione nel riconoscimento di consonanti e vocali lunghe, nonostante le geminate siano parte della L1, e, in generale, in tutti i gruppi si registrano tempi di reazione più brevi nella discriminazione della lunghezza vocalica, più facile da riconoscere.

4. *Obiettivi e ipotesi*

Nel presente studio è stato analizzato il contatto tra i sistemi fonologici di una L1 e di una L2 in presenza di una categoria assente nella lingua materna (albanese). Attraverso la somministrazione di test formali, si è cercato di stabilire il grado e le strategie di acquisizione della quantità consonantica distintiva in un piccolo gruppo di adulti albanofoni. In modo specifico, sono stati messi a confronto apprendenti con un differente livello di competenza in italiano, iniziale (I) e intermedio-avanzato (IA), e per avere una visione più completa sullo sviluppo delle geminate nell'IL, si è deciso di considerare sia le abilità di produzione orale e scritta sia la capacità di percezione. Inoltre, sono state selezionate più variabili di compito, al fine di verificare se e come la percezione/produzione del contrasto non-nativo varia in situazioni differenti: compiti più o meno complessi che richiedono differenti livelli di attenzione e l'uso di una o più abilità per essere svolti e che includono parole e non-parole poste in isolamento e in contesto di frase.

Incrociando il livello interlinguistico con le abilità e con le variabili di compito considerate, sulla base dell'inquadramento teorico e della letteratura di riferimento, si presuppongono le seguenti ipotesi:

- in produzione orale, risultati oscillanti nel gruppo IA ed assenza quasi totale del nuovo tratto nel gruppo I;
- in produzione scritta, un elevato grado di errori da parte degli apprendenti I e pochi errori nel gruppo IA;
- in percezione uditiva, un elevato grado di accuratezza da parte di tutti gli apprendenti, soprattutto di quelli esperti (IA);
- in tutte le abilità, rendimenti migliori nei contesti di parole e di isolamento rispetto alle non-parole e al contesto di frase;
- in base al tipo di consonante e alla frequenza delle parole, risultati differenti.

5. Metodologie

5.1 Partecipanti

I partecipanti sono stati reclutati presso la Sapienza Università di Roma. Si tratta di 7 studenti albanofoni (5F, 2M) di età compresa tra i 19 e i 35 anni, esposti alla varietà centrale romana e aventi un diploma di scuola superiore secondaria conseguito in Albania. Fatta eccezione per una dottoranda (AB06), tutti i soggetti frequentavano un Corso di Laurea Triennale. Parallelamente al gruppo degli albanofoni, sono state selezionate 4 studentesse provenienti da Roma e dintorni, di età compresa tra i 22 e i 28 anni. Le informazioni di carattere sociolinguistico (Tabella 1) sono state raccolte attraverso un questionario conoscitivo⁶, realizzato su <https://www.surveymonkey.com/en/> e sottoposto ai partecipanti prima dello svolgimento dei test, al fine di conoscere caratteristiche individuali utili all'interpretazione dei dati sperimentali. Il questionario prevedeva 23 domande suddivise in tre parti: la prima relativa alle informazioni anagrafiche e al grado di istruzione (genere, età, città dell'Albania di provenienza, lavoro, ultimo titolo di studio conseguito); la seconda sulla competenza linguistica generale (lingue conosciute oltre l'italiano e l'albanese, certificazioni linguistiche in italiano e/o in altre lingue, autovalutazione sul livello di conoscenza per ognuna delle lingue indicate); la terza sulla competenza linguistica in italiano (periodo di permanenza in Italia, primo approccio con la lingua italiana, modalità di apprendimento dell'italiano, uso a livello quantitativo da 0 a 100 della L2 a confronto con la L1, contesti d'uso della lingua italiana, motivazione nell'apprendimento dell'italiano in una scala di valutazione da 1 a 5 stelle).

In base ad alcuni parametri omogenei rintracciabili tra i partecipanti mediante il questionario, il gruppo sperimentale è stato suddiviso in due sottogruppi differenziati per livello di competenza. Tenendo conto in particolare del conseguimento non ancora avvenuto di una certificazione nella lingua italiana e del breve periodo di permanenza in Italia (tre mesi) AB01 e AB02 sono stati individuati come apprendenti iniziali (I); mentre i restanti 5 partecipanti (AB03 – AB07), considerando in particolare il conseguimento della certificazione di livello B2 nella lingua italiana (CELI3)⁷ e il lungo periodo di permanenza in Italia (da 1 a 5 anni), sono stati individuati come apprendenti di livello intermedio-avanzato (IA). Di seguito in Tabella 1 vengono riportate alcune informazioni essenziali sugli apprendenti, che insieme ad altre specifiche differenze individuali emerse dal questionario verranno descritte nella discussione dei risultati (§6).

⁶ Il questionario conoscitivo è consultabile al seguente link: <https://www.flipbookpdf.net/web/site/8396013a9d33cc0db3925a4ca80e8a8fa457ddad202108.pdf.html>

⁷ CELI è il Certificato di conoscenza della lingua italiana differenziato per livelli rilasciato dall'Università per Stranieri di Perugia o da centri con essa convenzionati in Italia o all'estero.

Tabella 1 - *Informazioni significative sugli albanofoni tratte dal questionario sociolinguistico*

Livello	ID partecipanti	Sesso	Età	Permanenza in Italia	Lavoro	Lingue conosciute	Età primo approccio (anni)	Uso italiano 0-100	Uso albanese 0-100	Motivazione 1-5
I	AB01	F	19	3 mesi	No	alb, ita, ing, cor	<6	45	55	5
I	AB02	M	19	3 mesi	No	alb, ita, ing	18 - 26	0	100	2
IA	AB03	F	25	5 anni	Si	alb, ita, ing	12 - 18	40	60	5
IA	AB04	M	22	5 anni	Si	alb, ita, ing, spa	12 - 18	50	50	4
IA	AB05	F	20	1 anno	Si	alb, ita, ing, gr	18 - 26	50	50	4
IA	AB06	F	35	2 anni	No	alb, ita, ing, spa, ted	18 - 26	50	50	5
IA	AB07	F	21	2 anni	Si	alb, ita, ing, spa	6 - 12	80	20	5

5.2 Materiali, procedure e analisi

I test, della durata complessiva di circa 15 minuti, sono stati svolti individualmente e in forma cartacea tra gennaio e marzo 2019, in un'auletta silenziosa della Sapienza Università di Roma. Tutto il materiale documentale è stato digitalizzato e conservato in un DVD insieme al materiale audio. Per ognuna delle abilità considerate è stato selezionato un tipo di test suddiviso in parti, differenziate in base alle variabili di compito (Tabella 2). Le prove sono state eseguite nel seguente ordine: produzione orale, produzione scritta, test integrato (percezione uditiva + produzione scritta), percezione uditiva, produzione semi-spontanea (prova non considerata in questo studio).

La produzione orale controllata consiste nel parlato letto di liste, di frasi più o meno lunghe, vicine alle condizioni realistiche del parlato, e di testi (due dialoghi e un brano in prosa per le parole e una poesia metasemantica per le non-parole). Nella selezione delle parole *target* si è fatto riferimento in particolare al *Grande dizionario italiano dell'uso* (GRADIT) di Tullio De Mauro, al *Dizionario italiano multimediale d'Ortografia e di Pronuncia* (DOP) di Migliorini, Tagliavini & Fiorelli (2009) e ad alcuni siti per la generazione di parole, come <https://www.parolecon.it/>. La scelta è ricaduta su parole del lessico di base⁸, con la funzione di nomi, aggettivi o verbi, comprese nel contesto vocalico /a/-/a/ e contenenti tutte le consonanti dell'italiano soggette alla geminazione e le loro rispettive scempie, fatta eccezione per la vibrante e per le affricate palatali e alveodentali (foni più complessi da analizzare acusticamente). In totale, la prova prevedeva 48 parole (P) e 25 non-parole (NP), caratterizzate da un numero differente di sillabe (perlopiù bisillabi e trisillabi), con la consonante scempia e geminata corrispondente poste in pre-tonia o in post-tonia. Anche dal punto di vista dell'intonazione frasale, il test non è bilanciato, dal momento che le parole *target* occupano posizioni differenti all'interno delle frasi.

⁸ Nel caso di consonanti poco frequenti nel lessico dell'italiano, come /b/ e /g/, sono state selezionate le parole *abaco* (lessico tecnico) e *tagga* (lessico del web).

Il materiale audio, registrato con ZOOM H1, è stato sottoposto all'analisi acustica attraverso il *software* PRAAT 5.3.51 (Boersma, Weenink, 2001); i *files* audio utilizzati per la segmentazione fonetica sono monofonici e in formato WAV, con una dinamica di ampiezza di 16bit ed una frequenza di campionamento di 32kHz, in grado di generare un segnale vocale di alta qualità. L'analisi acustica ha previsto la misurazione della durata delle consonanti scempie e geminate, considerando nelle occlusive prima la durata totale della consonante, poi la fase di tenuta e la fase di scoppio. La misurazione prosodica è stata effettuata per tutte le parole pronunciate con l'accento richiesto in italiano: alcune parole sdruciole pronunciate dagli albanofoni come piane sono state eliminate; inoltre, [ˈma:ga] non è stata considerata se pronunciata con l'affricata palatale [ˈma:ɟa]. Tale errore potrebbe essere dovuto alla scarsa padronanza del sistema grafematico e fonetico dell'italiano e all'interferenza con la L1 che in questo caso condurrebbero gli apprendenti a confondersi tra i due fonemi, l'occlusiva velare sonora /g/ e l'affricata palatale sonora /ɟ/. Essi sono rappresentati in italiano dallo stesso grafema <g>, mentre in albanese vi è una corrispondenza biunivoca tra <g> /g/ e tra <xh> /ɟ/.

Passando alla produzione scritta, il test si compone di 10 frasi, accompagnate da immagini ad esse relative, contenenti in totale 16 *items* dotati di uno spazio vuoto da completare. Sono state selezionate parole con la funzione di nomi, aggettivi o verbi, contenenti le seguenti consonanti: /d d:ɟ l l: n n: r r: s t t: t:s/, comprese indifferentemente in vari contesti vocalici e sillabici.

Il test integrato, consistente prima in un'elaborazione percettiva dello stimolo uditivo e poi in un compito produttivo di scrittura, comprende le seguenti consonanti /k: f: p: l l: n: r r:/ in vari contesti vocalici e sillabici, distribuite in totale in 8P e 6NP. Nel test di sola percezione uditiva sono state considerate le seguenti consonanti: /b b: k k: d:ɟ l m: r r: s t t: v: t:s/ in vari contesti vocalici e sillabici, distribuite in totale in 19P e 9NP. Le ultime due sezioni descritte del test sono state svolte facendo indossare ai partecipanti delle cuffie collegate al pc, da cui l'esaminatore gestiva le prove organizzate in un documento Microsoft Power Point. Gli stimoli uditivi utilizzati, tratti da esercizi di fonetica sulle geminate presenti online⁹, sono stati normalizzati per volume tramite il software LevelOne al fine di garantire ai partecipanti una percezione coerente.

⁹ Le tracce audio usate nelle prove sono tratte da <http://italianoperstranieri.loescher.it/> per le non-parole in isolamento e da https://www.youtube.com/results?search_query=claudio+gibellato+le+consonanti+doppie per le parole in isolamento. Per le non-parole in contesto sono state utilizzate alcune strofe della poesia metasemantica *Il coratto magiglioso*, recitata dall'attore comico Claudio Gregori, detto Greg, (<https://www.youtube.com/watch?v=nXXTyW37esA>); per le parole in contesto è stato scelto un brano tratto dall'audiolibro *Le città invisibili* di Italo Calvino, recitato da Walter Zanardi (<https://www.youtube.com/watch?v=qkAz4nX4Wrc&t=234s>).

Tabella 2 - Test differenziati per abilità e variabili di compito con esempi (S = scempia, G = geminata)

Abilità	Tipologia	Compito	Materiali			
			Parole (P)		Non-parole (NP)	
			Non contesto	Contesto	Non contesto	Contesto
Produzione orale	Parlato letto	Lettura ad alta voce	Lista 12P (6S, 6G)	Frase e brano 36P (18S, 18G)	Lista 12NP (6S, 6G)	Brano 13NP (6S, 7G)
			<i>Palla</i>	<i>La giraffa si libera della sbarra allo zoo</i>	<i>Dapa</i>	<i>Non si sanna nella spavva</i>
Produzione scritta	Riempimento spazi vuoti	Scrivere S o G mancanti	Frase + immagini 15P (11G, 5S)			
			<i>Mirko gioca con il pa__one al parco <ll></i>			
Test integrato	Discriminazione fonemica + riempimento spazi vuoti	Ascolto + scrittura di S e G mancanti	Lista 8P (2S, 6G)		Lista 6NP (6G)	
			<i>A__anto</i> <cc>		<i>Fagu__a</i> <pp>	
Percezione uditiva	Discriminazione fonemica	Sottolineare forma corretta in scelta binaria	5P (5G)	Brano 14P (2S, 12G)	5NP (3S, 2G)	Brano 4NP (4G)
			A. <i>Babo</i> B. <i>Babbo</i>	<i>Per questi porti non saprei tracciare la rota/rotta</i>	A. <i>Satiba</i> B. <i>Satibba</i>	<i>Sei forse tu, corato/coratto</i>

Nell'analisi dei dati relativi alla produzione scritta, al test di tipo misto e di percezione uditiva, si è proceduto all'individuazione e al calcolo della media delle risposte corrette ed errate.

6. Risultati

6.1 Produzione orale

In accordo con la letteratura, si conferma la forte difficoltà nella resa delle geminate in produzione orale. I partecipanti fanno registrare in media il 76% di risposte errate. Sono state considerate errate le realizzazioni di geminate che nel confronto con le rispettive scempie presentino una differenza inferiore ai 40ms. Tenendo per ora da parte il rendimento di AB07, come si può vedere dalle colonne in grigio scuro della Tabella 3, i soggetti di livello iniziale (I) in tutte le variabili di compito non distinguono quasi mai le geminate dalle scempie, mentre gli apprendenti di livello intermedio-avanzato (IA) oscillano tra circa il 68% e l'82% di errori, fatta

eccezione per AB06 che mostra risultati molto simili al sottogruppo I. Stando alle informazioni desunte dal questionario, il rendimento di AB06, dottoranda da due anni in Italia, può essere giustificato dall'età (35 anni) e dall'ampio uso della L1: AB06 usa soltanto la L1 con il suo compagno albanese. Passando al rendimento di tutti i partecipanti in ognuna delle variabili (righe in grigio chiaro della Tabella 3), tra i differenti contesti si osserva soltanto un 10% circa di errori in meno nelle non-parole in isolamento e in contesto, dovuto probabilmente al maggiore controllo sulla pronuncia e sull'ortografia di parole non note.

Tabella 3 - *Produzione orale. Valori medi in percentuale degli errori negli apprendenti di livello iniziale (I) o intermedio-avanzato (IA) e nei differenti contesti: parole e non-parole non in contesto (NCP, NCNP), parole e non-parole in contesto (CP, CNP)*

Livello	Partecipanti	NCP	NCNP	CP	CNP	Totale errori
I	AB01	83,3%	100%	100%	66,6%	91,2%
I	AB02	100%	100%	93,7%	83,3%	94,1%
IA	AB03	100%	50%	93,7%	66,6%	82,4%
IA	AB04	83,3%	66,6%	62,5%	66,6%	67,6%
IA	AB05	83,3%	16,6%	93,7%	33,3%	67,6%
IA	AB06	100%	83,3%	93,7%	100%	94,1%
IA	AB07	16,6%	66,6%	31,2%	16,6%	32,4%
		80,9%	69%	79,1%	61,9%	

Osservando il rapporto CC:C, negli italofoeni (Tabella 4) esso corrisponde in media a 1,9 in tutte le consonanti e i contesti considerati. Non sono state selezionate le stesse consonanti in ogni contesto, quindi non è possibile fare un confronto bilanciato, ma, considerando le parole in contesto (CP), si può notare che le seguenti geminate fricative, nasali e liquide /s: v: m: n: l:/ presentano una durata più lunga (CC:C superiore a 2) rispetto alle occlusive /p: t:/ (CC:C inferiore a 1,2); mentre, in accordo con le caratteristiche della varietà romana (Marotta, 2005), l'occlusiva bilabiale sonora scempia /b/ viene rafforzata. Si può notare che passando dall'isolamento al contesto di frase il rapporto CC:C tende a diminuire, come è possibile osservare in /p t g f l/; mentre non si registrano differenze nella produzione delle parole e delle non-parole.

Tabella 4 - *Produzione orale negli italiani. Durata in millisecondi della consonante scempia e geminata (valori rispettivamente a sinistra e a destra) e rapporto tra le consonanti geminate e scempie (CC:C, valore riportato tra parentesi) utilizzate in ogni contesto: parole e non-parole non in contesto (NCP, NCNP), parole e non parole in contesto (CP, CNP)*

ITALIANI													
C	CC:C	NCP		NCNP		CP		CNP					
/p p:/		80	213	(2,7)	122	251	(2,1)	92	137	(1,5)			
/b b:/		123	142	(1,2)	143	201	(1,4)	127	106	(0,8)			
/t t:/		97	210	(2,2)	90	237	(2,6)	64	110	(1,7)			
/d d:/									62	144	(2,3)		
/k k:/								69	110	(1,6)	82	177	(2,1)
/g g:/		88	173	(2)	84	200	(2,4)	53	90	(1,7)			
/f f:/		111	231	(2,1)	139	261	(1,9)	90	147	(1,6)			
/v v:/								42	125	(2,9)	67	140	(2,1)
/s s:/								75	181	(2,4)	118	183	(1,5)
/l l:/		68	183	(2,7)	71	198	(2,8)	34	80	(2,3)			
/m m:/								64	143	(2,2)	95	190	(2)
/n n:/								53	117	(2,2)	78	140	(1,8)

AB07 presenta un rendimento anomalo rispetto al gruppo sperimentale, per questo è stato considerato separatamente (Tabella 5). In AB07 il rapporto CC:C equivale in media a circa 2,4 in tutte le consonanti e i contesti. Questo partecipante produce il contrasto di lunghezza consonantica, raggiungendo i risultati del gruppo di controllo. Tuttavia, nei contesti privilegiati di parole isolate e soprattutto di non-parole isolate, la durata molto lunga di alcune geminate contro la durata più breve delle rispettive scempie rappresenta una spia della scarsa padronanza nella gestione del nuovo tratto. Nei contesti maggiormente controllati, il soggetto in questione tende a “forzare” la differenza tra geminate e scempie. L’ottimo risultato raggiunto da AB07 è il riflesso di fattori individuali e socioculturali favorevoli alla piena integrazione nella comunità italiana. Si tratta di una ventunenne da due anni e mezzo nel nostro Paese, che vive con italiani, possiede certificazioni di livello B2 in italiano e in inglese, tiene lezioni private di inglese ai bambini, ha avuto un approccio precoce con la nostra lingua (6 - 12 anni) e tra tutti i partecipanti ha indicato il valore più alto nell’uso quotidiano dell’italiano (80 su 100) e più basso in quello dell’albanese (20 su 100).

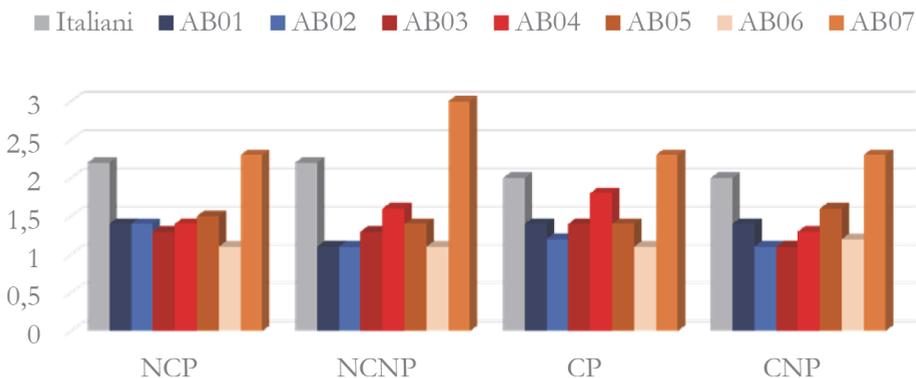
In tutti gli apprendenti albanofoni, anche in AB07 ma solo nel contesto di frase (poco controllato), si rileva una durata delle scempie maggiore di circa il 30% rispetto alla produzione degli italo-foni. Ciò si spiega con l’influsso della L1 in cui, stando all’analisi fonetico/fonologica contrastiva tra italiano e albanese condotta da Belluscio (2014), le consonanti singole tendono ad avere una durata intermedia tra le scempie e le geminate dell’italiano.

Tabella 5 - *Produzione orale negli apprendenti di livello iniziale (I) e intermedio-avanzato (IA) e in AB07. Durata in millisecondi della consonante scempia e geminata (valori rispettivamente a sinistra e a destra) e rapporto tra le consonanti geminate e scempie (CC:C, valore riportato tra parentesi) utilizzate in ogni contesto: parole e non-parole non in contesto (NCP, NCNP), parole e non-parole in contesto (CP, CNP)*

	Albanesi I				Albanesi IA				AB07			
	NCP	NCNP	CP	CNP	NCP	NCNP	CP	CNP	NCP	NCNP	CP	CNP
/p	95	185	150		114	186	141		86	79	113	
p:/	162	196	66		176	224	109		202	249	163	
	(1,7)	(1,1)	(0,4)		(1,5)	(1,2)	(0,8)		(2,3)	(3,2)	(1,4)	
/b	102	143	81		105	154	102		111	97	151	
b:/	98	208	164		134	185	118		237	276	127	
	(0,9)	(1,5)	(2)		(1,3)	(1,2)	(1,2)		(2,1)	(2,8)	(0,8)	
/t	191	174	108		180	185	103		115	54	68	
t:/	158	154	91		178	188	122		218	206	147	
	(0,8)	(0,9)	(0,8)		(0,9)	(1)	(1,2)		(1,9)	(3,8)	(2,2)	
/d				88				108				81
d:/				117				122				198
				(1,3)				(1,2)				(2,4)
/k			137	-			123	221			105	
k:/			171	126			114	202			218	
			(1,2)	-			(0,9)	(0,9)			(2,1)	
/g	-	125	91	116	141	98			82	102	64	
g:/	122	141	109	128	157	98			220	241	153	
	-	(1,1)	(1,2)	(1,1)	(1,1)	(1)			(2,7)	(2,4)	(2,4)	
/f	138	176	147		151	178	144		103	122	140	
f:/	209	167	85		206	233	186		215	224	167	
	(1,5)	(0,9)	(0,6)		(1,4)	(1,3)	(1,3)		(2,1)	(1,8)	(1,2)	
/v			68	89			80	64			65	82
v:/			100	49			67	73			167	176
			(1,5)	(0,5)			(0,8)	(1,2)			(2,6)	(2,1)
/s			129	182			100	191			78	144
s:/			167	154			164	196			226	283
			(1,3)	(0,8)			(1,6)	(1)			(2,9)	(2)
/l:	147	140	43		135	118	98		68	59	50	
l:/	143	134	43		150	156	94		190	230	240	
	(0,9)	(0,9)	(1)		(1,1)	(1,3)	(0,9)		(2,8)	(3,9)	(2,8)	
/m			87	105			82	129			64	101
m:/			77	99			93	138			123	190
			(0,9)	(0,9)			(1)	(1,1)			(1,9)	(1,9)

	Albanesi I				Albanesi IA				AB07			
	NCP	NCNP	CP	CNP	NCP	NCNP	CP	CNP	NCP	NCNP	CP	CNP
/n		76	113	134	134	74	113			56	69	
n:/		54	127	151	151	105	144			189	224	
		(0,7)	(1,1)	(1,1)	(1,1)	(1,4)	(1,3)			(3,4)	(3,2)	

Figura 1 - Produzione orale. Valori medi del rapporto CC:C in tutti i partecipanti e nei differenti contesti: parole e non-parole non in contesto (NCP, NCNP), parole e non-parole in contesto (CP, CNP). Considerazione solo dei casi in cui le geminate hanno una durata di 40ms superiore alle scempie



Ritornando alle geminate, nella maggior parte dei casi i partecipanti di livello iniziale (I) non le realizzano (Tabella 5): le producono con una durata inferiore alle scempie corrispondenti (CC:C inferiore a 1) o superiore ad esse (geminate con una differenza di durata superiore o inferiore ai 40ms rispetto alle scempie); fanno eccezione solo /b:/, /p:/ ed (f:/ in isolamento). Anche negli apprendenti IA (Tabella 5), si registrano risultati migliori nella produzione delle medesime consonanti nello stesso contesto, e naturalmente nelle restanti consonanti vi è un rendimento migliore rispetto al gruppo I, ma non viene raggiunto il livello nativo. Nel gruppo IA, la differenza media di durata tra geminate e scempie, se presente, è pari al 30% circa; negli italo-foni invece si aggira intorno al 57%.

Tra gli apprendenti di livello intermedio-avanzato (IA), si evidenzia un certo grado di variabilità (Figura 1): AB03, AB04 e AB05 raggiungono risultati migliori in più contesti. Si tratta di studenti e lavoratori in Italia i primi due da cinque anni e il terzo da un anno; mentre, come già detto, AB06 ha un rendimento simile ad AB01 e AB02, da soli tre mesi in Italia.

6.2 Produzione scritta

Nella produzione scritta, consistente in 16 spazi bianchi da riempire (11G e 5S) in 15 parole poste all'interno di frasi accompagnate da immagini, si registra il minor numero di errori (19,6%), data la forte attenzione posta sulla forma e la maggiore

attivazione del *monitor* (Chini, 2005). I due apprendenti in fase iniziale (I) presentano rendimenti differenti: AB01 ha completato tutti gli *items*, sbagliando circa il 50% delle geminate e mostrando una buona competenza lessicale; AB02 ha completato 10 spazi bianchi su 16, usando la geminata solo in *rosso* e mostrando una mancata padronanza nell'uso dell'alfabeto italiano: in *legge*, al posto di <gg>, usa <c> che in albanese rappresenta l'affricata alveodentale sorda /ts/. I risultati ottenuti dai due partecipanti si possono giustificare mediante le informazioni desunte dal questionario conoscitivo: AB01 ha avuto il primo approccio con l'italiano a meno di sei anni, AB02 tra i 18 e i 26 anni; AB01 ha indicato 45 su 100 nell'uso quotidiano dell'italiano contro lo 0 su 100 di AB02 e, in ultima analisi, ma fondamentale nell'acquisizione di una L2, vi è la motivazione: AB02 ha dichiarato di essere scarsamente interessato ad apprendere l'italiano (2 su 5), a differenza di AB01 (5 su 5).

Tabella 6 – *Produzione scritta. Valori medi in percentuale degli errori negli apprendenti di livello iniziale (I) e intermedio-avanzato (IA), nelle consonanti scempie (S) e geminate (G)*

Livello	Partecipanti	S	G	Totale errori
I	AB01	0%	54,5%	37,5%
I	AB02	20%	90,9%	68,8%
IA	AB03	0%	9,1%	6,3%
IA	AB04	20%	9,1%	12,5%
IA	AB05	20%	9,1%	12,5%
IA	AB06	40%	9,1%	18,8%
IA	AB07	0%	0%	0%
Totale errori		14,3%	26,0%	

Gli errori poco numerosi compiuti dai partecipanti di livello intermedio-avanzato (IA) riguardano *rossetti*, una parola con due geminate in successione, e la resa geminata delle consonanti scempie presenti in *cruDo*, *matiTa*, *paTatine*. Tali errori di ipergeneralizzazione, che vanno in una direzione opposta rispetto al gruppo I (I geminate scempie *vs.* IA scempie geminate), riflettono l'avanzamento del nuovo tratto nell'interlingua degli apprendenti, in grado di scardinare la presenza della scempia anche in una parola ben consolidata nel lessico della L1, *patatine*, un italianismo dell'albanese (Dashi, 2013: 322). Si nota, inoltre, un minor numero di errori in forme verbali di alta frequenza in italiano, *rotto*, *legge*, *correre*; in parole del lessico dei colori e del cibo *rosso*, *colori*, *crudo*, *panino*.

6.3 Test integrato

Il test integrato richiede un livello più alto di attenzione da distribuire nello svolgimento di due compiti simultanei, elaborazione percettiva (ascolto di uno stimolo sonoro) + produzione scritta (riempimento di spazi bianchi). La maggiore complessità giustifica il numero più alto di errori presenti in questa prova (Tabella 7). In particolare, se negli apprendenti di livello iniziale (I) non si registrano differenze significative tra i contesti di parole e di non-parole isolate (Figura 2), lo stesso non

accade nell'altro gruppo. Complessivamente gli apprendenti IA tendono a compiere più errori nella discriminazione e nella resa del materiale non lessicale (rispettivamente 33,3% *vs.* 15%).

Il differente rendimento dei due gruppi potrebbe essere spiegato dal fatto che “le capacità acustico-discriminatorie subiscono una regressione, con il progredire della competenza generale in L2” (Celata, 2004: 55). Stando all'ipotesi verificata da Chiara Celata in uno studio sulla produzione delle affricate nell'italiano L2 di apprendenti greci e portoghesi, i soggetti in fase iniziale di apprendimento fanno ricorso ad una strategia psico-acustica nell'elaborazione del materiale lessicale e non lessicale, mentre in fase avanzata ci si avvale di una strategia linguistica più completa e complessa che chiama in causa la frequenza d'uso e le caratteristiche semantiche delle parole. Dunque, in presenza delle non-parole, non potendo usufruire dell'aiuto della competenza lessicale e tralasciando le fini variazioni fonetiche, gli apprendenti esperti fanno emergere i propri punti deboli. Le difficoltà, riflesso di una padronanza ancora non piena nella discriminazione del nuovo tratto, si evidenziano in un compito complesso, come il presente test di tipo misto, e non nel test di percezione uditiva in cui l'attenzione viene posta sulla realizzazione di un solo compito.

Tabella 7 - *Test integrato. Valori medi in percentuale degli errori negli apprendenti di livello iniziale (I) e intermedio-avanzato (IA) e nei differenti contesti: parole e non-parole non in contesto (NCP, NCNP)*

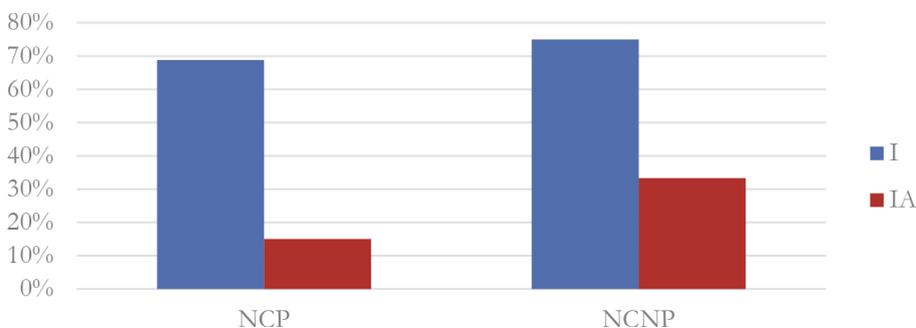
Livello	Partecipanti	NCP	NCNP	Totale errori
I	AB01	62,5%	66,7%	64,3%
I	AB02	75%	83,3%	78,6%
IA	AB03	12,5%	66,7%	35,7%
IA	AB04	25%	16,6%	21,4%
IA	AB05	0%	33,3%	7,1%
IA	AB06	37,5%	50%	35,7%
IA	AB07	0%	0%	0%
Totale errori		30,4%	45,2%	

Tra le consonanti presenti in questo test, /k: f: p: l: n: r r: /, si è registrato un minor numero di errori in /r r:/ e un maggior numero di errori in /l:/. Una spiegazione a ciò potrebbe essere ricercata nell'interazione tra le caratteristiche fonetiche e grafemiche della L1 e della L2. In albanese vi è una distinzione tra <r>, realizzata come monovibrante alveolare [r] (cfr. inglese *very*) o come retroflessa [ɾ] (cfr. norvegese *blad*), e il digramma <rr>, realizzata come polivibrante alveolare e corrispondente alla scempia /r/ dell'italiano (es. rana). Si potrebbe ipotizzare che gli apprendenti rendano /r:/ dell'italiano come <rr> perché ha caratteristiche fonetiche più vicine alla polivibrante /r/ dell'albanese e non alla monovibrante /r/ della L1 rappresentata da <r>. Quindi durante lo svolgimento delle prove, i partecipanti, attivando contemporaneamente il sistema fonologico e ortografico (Bassetti, 2017) e trovando corrispondenze fonetiche maggiori tra /r:/ della L2 e /r/ della L1 e non tra /r:/ e /r/ monovibrante alveolare dell'albanese, tenderebbero a fare meno errori nel-

la discriminazione e nella resa di /r/ come doppia rispetto ad altre consonanti. La presunta maggiore velocità nell'acquisizione della vibrante geminata /r:/ potrebbe essere dovuta anche alla sua durata più lunga rispetto ad altre consonanti doppie, che la renderebbe più facilmente percepibile.

In direzione opposta, la resistenza nella discriminazione e nella resa della laterale geminata /l:/, soprattutto negli apprendenti iniziali (AB01, AB02) e in AB06, troverebbe una giustificazione nel fatto che in albanese il digramma <ll> corrisponde alla laterale velare /L/, fonema assente in italiano, mentre il grafema <l> corrisponde a /l/, la laterale alveolare presente anche nella L2. Facendo riferimento alla loro L1, gli albanofoni noterebbero una maggiore somiglianza fonetica tra /l:/ dell'italiano e la /l/ della L1 rappresentata da <l> e non tra /l:/e /L/ della L1 rappresentata da <ll> e contraddistinta dal tratto velare che non appartiene alla geminata della L2.

Figura 2 - Test integrato. Valori medi in percentuale degli errori negli apprendenti di livello iniziale (I) e intermedio-avanzato (IA) e nei differenti contesti: parole e non-parole non in contesto (NCP, NCNP), parole e non-parole in contesto (CP, CNP)



6.4 Percezione uditiva

A livello percettivo, come atteso, il numero di errori si abbassa (26%; Tabella 8). Nel test di percezione uditiva, consistente nell'individuazione (sottolineandolo) dello stimolo uditivo che si pensa di aver ascoltato, gli apprendenti, fatta eccezione per quelli in fase iniziale (48% circa di errori), raggiungono un elevato grado di accuratezza, mostrando qualche errore soltanto nel contesto di frase.

Come è già stato discusso per il test integrato (§6.3), anche tra le consonanti presenti in questa prova, /b b: k k: d:ʒ l: m: r r: s: t t: v: ts:/, si è registrato un minor numero di errori in /r/ e un maggior numero di errori in /l:/. Se si considerano le parole selezionate, *metterò* e *babbo* di elevata frequenza nel lessico dell'italiano hanno ricevuto una risposta corretta da tutti i partecipanti.

Tabella 8 - *Percezione uditiva. Valori medi in percentuale degli errori negli apprendenti di livello iniziale (I) e intermedio-avanzato (IA) e nei differenti contesti: parole e non-parole non in contesto (NCP, NCNP), parole e non-parole in contesto (CP, CNP)*

Livello	Partecipanti	NCP	NCNP	CP	CNP	Totale errori
I	AB01	40%	60%	57,1%	50%	53,6%
I	AB02	20%	20%	64,3%	50%	42,9%
IA	AB03	0%	0%	7,1%	25%	7,1%
IA	AB04	0%	0%	42,9%	75%	28,6%
IA	AB05	0%	0%	21,4%	0%	10,7%
IA	AB06	40%	40%	28,6%	0%	28,6%
IA	AB07	0%	0%	14,3%	0%	7,1%
Totale errori		14,3%	17,1%	33,7%	28,6%	

7. Conclusioni

Nel presente studio è stata condotta un'analisi sull'acquisizione delle geminate da parte di un piccolo gruppo di apprendenti adulti albanofoni, mettendo a confronto differenti abilità, livelli interlinguistici e variabili di compito.

Come atteso, la produzione orale risulta l'abilità in cui il processo di acquisizione procede più lentamente. Gli apprendenti di livello iniziale non producono quasi mai il contrasto non nativo, fatta eccezione per alcune consonanti poste in isolamento; il medesimo contesto favorisce un rendimento migliore anche nei partecipanti di livello intermedio-avanzato, che in generale presentano risultati molto oscillanti.

A livello percettivo, gli apprendenti mostrano di riconoscere il nuovo tratto, ma in compiti più complessi, contesto di frasi e test integrato, si evidenzia la non piena padronanza nella discriminazione delle geminate. In particolare, è stata rilevata una maggiore difficoltà nel contesto delle non-parole per i partecipanti più esperti, dovuta alla loro migliore competenza lessicale in italiano.

Anche la direzione degli errori del gruppo degli esperti in produzione scritta (prevalentemente scempie geminate) è un indizio dell'avanzamento della L2. Inoltre, la relazione tra sistema fonologico e ortografico della L1 e della L2 si è dimostrata particolarmente influente sulla realizzazione di alcune consonanti (<ll>, <rr>).

Dunque, in accordo con i modelli di acquisizione fonologica di riferimento per questo lavoro (SLM, PAM, OPM), si rivela possibile accedere ad una categoria fonologica assente nella propria L1, in relazione a svariati fattori linguistici ed extra-linguistici. Se la marcatezza e la difficoltà articolatoria sottese ad un fenomeno di rafforzamento come la geminazione ne rallentano l'acquisizione, una forte motivazione nell'apprendimento della L2, una costante e ricca esposizione all'*input* e l'età precoce del primo approccio favoriscono il raggiungimento di un livello quasi nativo in uno dei partecipanti (AB07). L'assenza di questi fattori positivi penalizza invece l'apprendimento del contrasto non nativo in uno degli apprendenti di livello intermedio-avanzato (AB06), che mostra un rendimento molto simile al gruppo di livello iniziale (I).

Naturalmente le conclusioni tratte da questo studio andrebbero verificate su campioni più ampi di partecipanti con differenti lingue di partenza e attraverso l'utilizzo di materiale lessicale e non lessicale, adeguatamente bilanciato e l'applicazione di analisi più approfondite. Pertanto, si spera di continuare il lavoro in questo settore, ponendo l'attenzione su tutte le abilità coinvolte nell'acquisizione linguistica, *in primis* la percezione da cui il processo ha inizio, e cercando di intervenire concretamente a livello didattico mediante la progettazione di *training* mirati ad una maggiore focalizzazione sulle geminate nelle differenti abilità. A questo proposito, come è già stato sperimentato, si potrebbero costruire attività pratiche e di analisi sulle coppie minime (Calabrò, Mairano, 2017).

Riferimenti bibliografici

- ALTMAN, H., BERGER, I. & BRAUN, B. (2012). Asymmetries in the perception of non-native consonantal and vocalic length contrasts. *Second Language Research*, 28 (4), 387-413.
- BASSETTI, B. (2017). Orthography affects Second Language Speech: Double Letters and Geminate Production in English. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 43 (11), 1835-1842.
- BELLUSCIO, G. (2014). Fonematica contrastiva albanese-italiano. In GESUATO, S. & BUSÀ, M.G. (A cura di), *Festschrift in onore di Alberto Mioni*. Padova: CLEUP, 277-292.
- BERTINETTO, P.M. & VIVALDA, E. (1978). Recherches sur la perception des oppositions de quantité en italien. *Journal of Italian Linguistics*, 3, 97-116.
- BEST, C.T. (1991). The Emergence of Native-Language Phonological Influences in Infants: a Perceptual Assimilation Model. *Haskins Laboratories Status Report on Speech Research*, SR, 107/108, 1-30.
- BOERSMA, P. & WEENINK, D.J.M. (2001). PRAAT, a system for doing phonetics by computer. *Glott International*, 5 (9-10), 341-347.
- BYBEE, J.L. (2000). The Phonology of the Lexicon: evidence from Lexical diffusion. In BAROW, M. & KEMMER, S. (Eds.), *Usage-based Models of language*. Stanford, CA: CSLI.
- CALABRÒ, L. & MAIRANO, P. (2017). Insegnare la pronuncia a studenti sinofoni di italiano L2 con il Minimal Pair Finder. *Italiano LinguaDue*, 1, 83-97.
- CELATA, C. (2004). *Acquisizione e mutamento di categorie fonologiche. Le affricate in italiano*. Milano: FrancoAngeli.
- CERRATO, L. & FALCONE, M. (1998). Acoustic and perceptual characteristics of Italian stop consonants. *5th International Conference on Spoken Language Processing (ICSLP)*, Sydney, Australia, 30 November-4 December 1998.
- CHINI, M. (2005). *Che cos'è la linguistica acquisizionale*. Roma: Carocci Editore.
- CHOMSKY, N. & HALLE, M. (1968). *The Sound Pattern of English*. Cambridge, MA: MIT Press.
- D'APOLITO, S. & GILI FIVELA, B.G. (2019). L2 Pronunciation Accuracy and Context: a Pilot Study on the Realization of Geminate in Italian as L2 by French Learners. *Proceedings of Interspeech 2019, 1706-1710, DOI: 10.21437/Interspeech.2019-2934*, Graz, Österreich, 15-19 settembre 2019.

- DASHI, B. (2013). *Italianismi nella lingua albanese*. Roma: Edizioni Nuova Cultura.
- DE BOT, K., LOWIE, W. & VERSPOOR, M. (2005). *Second Language Acquisition: an advanced resource book*. Abingdon, Oxon: Routledge.
- DE MAURO, T. (1999-2000). *Grande dizionario italiano dell'uso*. Voll. 6. Con aggiornamenti: (2003) *Nuove parole italiane dell'uso*. Con CD-Rom. (2007). *Nuove parole italiane dell'uso*, 2. Torino: UTET.
- DI GIOVINE, P. (2008). Un millennio di storia linguistica albanese: l'influsso lessicale della lingua italiana. *L'Italia dialettale*, 59, 107-139.
- ECKMANN, F.R. (1977). Markedness and the contrastive analysis hypothesis. *Language Learning*, 27, 315-330.
- DERWING, T.M. & MUNRO, M.J. (2015). *Pronunciation Fundamentals: Evidence-based Perspectives for L2 teaching and research*. Amsterdam; Philadelphia, Pa: John Benjamins.
- ESPOSITO, A. & DI BENEDETTO, M.G. (1999). Acoustical and perceptual study of gemination in Italian stops. *Journal of the Acoustical Society of America*, 106 (4), 2051-2062.
- FALUSCHI, S. & DI BENEDETTO, M.G. (2012). Acoustic Analysis of Singleton and Geminate Affricates in Italian. *WEB-SLS, The European Student Journal of Language and Speech*. August 2012, from http://newyork.ing.uniroma1.it/Papers/J16-Faluschi_al-EJLS01.pdf
- FAVARO, G. (2002). *Insegnare l'italiano agli alunni stranieri*. Milano: RCS Libri.
- FLEGE, J.E. (1995). Second language Speech Learning: Findings, and problems. In STRANGE, W. (Ed.), *Speech Perception and Linguistic Experience: Issues in Cross-Language Research*. Timonium, MD: York Press, 233-277.
- GIANNINI, S. & COSTAMAGNA, L. (1998). Acquisizione di categorie fonologiche e diffusione lessicale del mutamento linguistico: affinità strutturali. *Archivio Glottologico Italiano*, 83, 150-187.
- GIOVANARDI, M. & DI BENEDETTO, M.G. (1998). Acoustic analysis of singleton and geminate fricatives in Italian. *WEB-SLS, The European Student Journal of Language and Speech*. August 2012, from http://acts.ing.uniroma1.it/Papers/J11-Giovanardi_al-EJLS98.pdf
- GOLDSMITH, J.A. (1976). *Autosegmental phonology*. Bloomington: Indiana University Linguistics Club.
- GRADILONE, G. (1974). *Altri studi di Letteratura Albanese*. Roma: Bulzoni.
- KABAK, B., RECKZIEGEL, T. & BRAUN, B. (2011). Timing of second language singletons and geminates. In LEE, W.S. & LEE, E. (Eds.), *Proceedings of the 17th International Congress of Phonetic Science (ICPhS XVII)*, Hong Kong, 17-21 August 2011.
- HAN, M.S. (1992). The timing control of geminate and single stop in Japanese: a challenge for non-native speakers. *Phonetica*, 49, 102-127.
- HAM, W. (2001). *Phonetic and Phonological Aspects of Geminate Timing*. London: Routledge.
- HOCKETT, C. (1955). *A Manual of Phonology*. Bloomington: Indiana University Publications in Anthropology and Linguistics, Memoir 11.
- LADO, R. (1957). *Linguistics across cultures: applied linguistics for language teachers*. Ann Arbor: University of Michigan Press.

- MAH, J. & ARCHIBALD, J. (2003). Acquisition of L2 length contrasts. In LICERAS, J.M. et al. (Eds), *Proceedings of the 6th Generative Approaches to Second Language Acquisition Conference*, 208-212.
- MAJOR, R.C. (2001). *Foreign Accent: The Ontogeny and Phylogeny of Second Language Phonology*. Routledge: New York-London.
- MADDIESON, J. (1984). *Patterns of sound*. Cambridge: Cambridge University Press.
- MAROTTA, G. (2005). Il consonantismo romano. Processi fonologici e aspetti acustici. In ALBANO LEONI, F. & GIORDANO, R. (A cura di), *Italiano parlato. Analisi di un dialogo*. Napoli: Liguori Editore, 1-24.
- MATTEI, M. & DI BENEDETTO, M.G. (2000). Acoustic analysis of singleton and geminate nasals in Italian. WEB-SLS, *The European Student Journal of Language and Speech*. August 2012, from http://151.100.8.164/Papers/J15-Mattei_al-EJLS00.pdf
- MIONI, A.M. (1993). Fonetica e fonologia. In SOBRERO, A.A. (a cura di), *Introduzione alle strutture dell'italiano contemporaneo*. Roma-Bari: Laterza.
- MIGLIORINI, B., TAGLIAVINI, C. & FIORELLI, P. (1999). *Dizionario d'Ortografia e di Pronuncia*. Roma: Rai-Eri.
- OLLER, J.W. & ZIAHOSSEINY, S.M. (1970). The contrastive analysis hypothesis and spelling errors. *Language learning*, 20, 183-189.
- PAYNE, M.E. (2005). Phonetic variation in Italian consonant gemination. *Journal of the Phonetic Association*, 35 (2), 153-181.
- PERLMUTTER, D. (1995). Phonological Quantity: from a feature to Multiple Association. In GOLDSMITH, J.A. (1995), *The Handbook of Phonological Theory*. Cambridge Mass: Blackwell. 305-317.
- SERIANNI, L. (20013). Prefazione. In DASHI, B., *Italianismi nella lingua albanese*. Roma: Edizioni Nuova Cultura.
- SMITH, C.L. (1995). Prosodic patterns in the coordination of vowel and consonant gestures. In CONNELL, B. & ARVANITI, A. (Eds.), *Phonology and phonetic evidence. Papers in Laboratory Phonology IV*. Cambridge University Press: Cambridge, 205-222.
- SORIANELLO, P. (2009). L'acquisizione del tratto di lunghezza consonantica in italiano L2. In ROMITO, L., GALATÀ, V. & LIO, R. (Eds.), *La fonetica sperimentale. Metodi e applicazioni. Atti del IV Convegno Nazionale dell'AISV*, Rende, CS, 3-5 dicembre 2007. CD-Rom. Torriana: EDK, 41-69.
- SORIANELLO, P. (2014). Italian geminate consonants in L2 acquisition. In COSTAMAGNA, L. & CELATA, C. (Eds.), *Consonant gemination in first and second language acquisition*. Pisa: Pacini, 25-46.
- SORIANELLO, P. (2019). Le consonanti affricate: strategie di acquisizione in italiano L2. *Italiano LinguaDue*, 1, 68-88.
- SWADESH, M. (1937). The phonemic interpretation of long consonants. *Language*, 13, 1-10.
- TRUBETZKOY, N.S. (1967). *Principes de phonologie*. (Ed. or. 1939). Paris: Klincksieck.
- TSUKADA, K., COX, F., HAJEK, J. & HIRATA, Y. (2018). Non-native Japanese learners' perception of consonant length in Japanese and Italian. *Second Language Research*, 34 (2), 179-200.
- WANG, W. (1969). Competing changes as a cause of residue. *Language*, 45, 9-25.

- WEINREICH, U. (1953). *Languages in contact*. New York: Linguistic Circle of New York.
- WODE, H. (1992). Categorical Perception and Segmental Coding in the Ontogeny of Sound Systems. A Universal Approach. In FERGUSON, C., MENN, L. & STOEL-GAMMON, C. (Eds.), *Phonological Development. Models, Research, Implications*. Timonium, MD: York Press, 605-631.
- ZMARICH, C. & GILI FIVELA, B. (2005). Consonanti scempie e geminate in italiano: studio cinematico e percettivo dell'articolazione bilabiale e labiodentale. In *Atti del 2° Convegno Nazionale AISV (Associazione Italiana di Scienze della Voce) "Misura dei parametri"*, Padova, 2-4 dicembre 2004.

PARTE IV

VARIAZIONE
E PROSODIA

MAYARA DA SILVA NETO, BARBARA GILI FIVELA, ELISABETTA SANTORO,
FLAVIANE ROMANI FERNANDES SVARTMAN

Do mitigation strategies affect prosodic correlates? An investigation on orders and requests in Italian

This paper aims to identify whether and to what extent mitigation procedures affect the intonational pattern of orders and requests produced by Italian speakers (from Lecce), within different social distances contexts. Data were collected using a variant of the Discourse Completion Task (DCT) and analyzed within the Autosegmental Metrical framework, on the basis of auditory perception, analysis of the phonological function and exploration of the F0 contour. The results indicate that 1) the intonational patterns of orders and requests differ especially when the latter are expressed by the interrogative form and specifically, as far as intonation is concerned, by a different boundary tone. Moreover, 2) the presence of lexical means of mitigation and the low social distance interact in a complex manner in orders and request, but often favors more preemptory, less neutral patterns.

Keywords: intonation, orders, requests, Lecce Italian, Autosegmental Metrical framework.

1. Introduction

From a pragmatic perspective, investigating orders and requests leads us immediately to the Speech Acts Theory and its subsequent developments. According to Austin (1962), author of the aforementioned theory, orders and requests would both be illocutionary acts carried out by means of *exercitive* verbs, that is, verbs that express the speaker's power, right or influence over the listener. Searle (1979), in his turn, proposes a classification where orders and requests are considered a *directive* type of illocutionary act, whose goal – or illocutionary point – is to make the listener perform an action. The two authors admit, therefore, that we are facing similar speech acts, since in their classifications both belong to the same categories. However, while Austin does not state whether there is any kind of differences between orders or commands and requests, Searle does it when he clarifies that illocutionary force and illocutionary point are two different concepts:

The illocutionary point is part of but not the same as illocutionary force. Thus, e.g., the illocutionary point of requests is the same as that of commands: both are attempts to get hearers to do something. But the illocutionary forces are clearly different. In general, one can say that the notion of illocutionary force is the resultant of several elements of which illocutionary point is only one, though, I believe, the most important one (Searle, 1979, p. 3).

Thus, orders and requests are speech acts through which speakers try to get the listeners “to do something”, which can be trivial or demand a great effort, other than have material or discursive consequences. We could assume that such speech acts may force listeners out of their comfort zone, as they need to evaluate if they can, want or must take an action and in what way. Besides, the same listeners, as stated by Brown & Levinson (1987), can be close or distant from the speaker on a horizontal axis of social distance and/or on a vertical axis in terms of hierarchy or power. In our view, this is also an important variable to be considered. Indeed, several studies in pragmatics - focused on production or perception of different speech acts, in language teaching and learning - take into account the social distance, considering it as an independent variable, a contextual element that influences linguistic choices of speakers (e.g. Takimoto, 2007; Nuzzo, 2013; Spadotto e Santoro, 2019; Santoro, Kulikowski & Silva, 2017; Silva Neto, 2018). Further, studies concerning intonation have also showed that social distance plays a role in affecting intonation (e.g. for Catalan, Astruc, Vanrell, & Prieto 2016 on requests and offers and Borràs-Comes, Sichel-Bazin & Prieto 2015 on vocatives).

Speakers are somehow aware of the importance of all contextual factors (including social distance) involved in the realization of these speech acts, as well as they are aware, at least to a certain extent, that they may have undesirable effects and/or require the use of politeness strategies, according to the social rules of a specific culture. Therefore, speakers adopt different means to mitigate the illocutionary force of the abovementioned speech acts. Innumerable studies have investigated mitigation strategies used in orders, requests and other speech acts in different languages, being such strategies subject to intra and cross variation across languages and cultures. Morphosyntactic, lexical or discursive modifiers, supportive moves (preparators, grounders, etc.) and the use of indirectness are quite recurrent means of achieving this purpose (cf. Blum-Kulka, House & Kasper, 1989; Santoro, 2017; among others). It is capital to remember, though, that besides a typical linguistic structure (interrogative form for requests and imperative form for orders, for example), also prosodic elements are conventionalized (Escandell-Vidal, 2011), and convey the specific meaning of different types of speech acts (Ladd, 2008), as well as politeness strategies (Gili Fivela, Bazzanella, 2014). Therefore, the elements used to mitigate or reinforce them will somehow interact with the prosodic makeup of the utterance.

The relation between prosodic features and other information available to listeners is an intriguing matter, which has been investigated with special regards to multimodal communication (see the integration of audio and visual information, e.g., Swerts, Krahmer, 2008). As far as unimodal investigations are concerned, much attention has been paid to the prosody-syntax-pragmatics interface, thus to the impact that syntactic and pragmatic information may have on prosodic features and correlates conveyed through speech (Nespor, Vogel, 1986; Selkirk, 1984; Féry, 2013), for instance also in relation to politeness (Hidalgo, 2003; Hübscher, Borràs-Comes & Prieto, 2017; Caballero, Vergis, Jiang, & Pell, 2018). Nevertheless, according to the authors' knowledge, less attention has been paid to the correlation between

lexical and prosodic information, even though Frota & Prieto (2015) recall that in Sardinian and, to some extent, in Friulian, the poor use of intonation strategies may be correlated to the use of lexical (and syntactic) means to mark sentence modality (Vanrell, Ballone, Schirru & Prieto, 2015; Roseano, Vanrell & Prieto, 2015). As for the impact of social distance on prosody, systematic investigations are few, besides the specific attention paid in some experimental protocols to control changes in the speaker's role, and therefore of speaker's power (Albano Leoni, 2003).

As far as the variety of Italian investigated in this paper is concerned, that is Lecce Italian, previous studies offered results on the main intonation patterns found in requests, orders, and imperative requests (Gili Fivela, Avesani, Barone, Bocci, Crocco, D'Imperio, Giordano, Marotta, Savino & Sorianello, 2015). Specifically, they showed that, as far as phonological patterns are concerned, requests expressed by an imperative form show similarities with orders, especially as for the boundary (i.e., differently from other types of requests, they show a final low boundary), and orders may show similarities with broad focus sentences (i.e., a H+L* L% nuclear pattern). However, previous investigations also showed that yes-no questions, for instance information seeking yes-no questions, show a completely different pattern (H*+L LH%), and it is well known that requests may be expressed by means of an interrogative form, being therefore intonationally different from requests expressed by imperative forms. However, previous works offered no description of the patterns used when either mitigation strategies or differences in social distance are at play.

2. Goals and hypotheses

This paper describes an investigation that is part of a wider PhD project concerning the production of orders and requests by Italian and Brazilian speakers in different social distance conditions (Silva Neto, in preparation). The goal of this paper is to investigate the impact of mitigation strategies (such as adverbs expressing politeness) on the prosodic characteristics of requests and orders in (Lecce) Italian in different social distance conditions (the aspects under investigation are necessarily intertwined with politeness modulation, though politeness *per se* is not going to be directly investigated here). Based on previous descriptions of Lecce Italian intonation (see §1), the main hypothesis is that requests and orders may be prosodically different from each other, even though requests expressed by an imperative form may show similarities with orders, especially as for the boundary tones; requests expressed by an interrogative form are rather expected to show similarities with questions (here, information seeking yes-no questions). However, as already suggested by other works in the literature (Frota, Prieto, 2015), on a regular basis, orders and requests are expected to be different from statements and questions. Moreover, in line with the discussion in §1, we expect both social distance and mitigation expressed through explicit lexical choices to affect intonation from either the phonological (tonal composition, e.g., in the case of mitigation strategies; e.g., Vanrell et al., 2015) or the phonetic point of view (e.g., F0 or intensity range), if not in both respects (Gili Fivela, Bazzanella,

2014). Specifically, we expect the presence of lexical means of mitigation and the low social distance to favor more peremptory, less neutral (and polite) patterns. By the way, throughout this text, we use the term “peremptory” to characterize a firm way of expressing oneself, which leaves little room for denial or refusal. On the other hand, we also expect the lack of lexical means of mitigation and the high social distance to favor more neutral and less peremptory patterns.

3. *Method*

3.1 Corpus and subjects

Five speakers from Lecce were audio recorded using a variant of the Discourse Completion Task tasks (DCT – Blum-Kulka et al., 1989). During the experiment, rather than producing only a spontaneous reaction to the situations, subjects were also asked to read a given target sentence with reference to the same contexts (Gili Fivela et al., 2015). In this paper, we focus on the read speech productions by three out of the five recorded speakers.

Target words and sentences were selected in order to allow for the best comparison of phonological and phonetic prosodic characteristics across utterances. In the subset analyzed here, the target word is *indovina* (“guess”). The target sentence types elicited in this experiment were orders (with the verb in the imperative mood), requests expressed by an imperative form, and requests expressed by means of interrogative form. Further, various productions were included in relation to mitigation strategies. There was a non-mitigated production, a production mitigated by means of *per favore* (“please”), and another production where a different adverb was included, in order to get a non-mitigated production corresponding to the mitigated one as for the structure and number of syllables (*questa volta* – “this time” – in the case of both types of requests and *alla svelta* – “quickly” – in the case of orders). Broad focus statements and yes-no information seeking questions were also included as control sentences, for a total of 11 sentences.

In order to elicit these data, target sentences were inserted in brief contexts whose function was suggesting specific modalities and pragmatic interpretations as well as simulating high (HD) and low (LD) social distance situations between the subject and the hypothetical listener. In all the HD subjects were induced to understand that their hypothetical hearer was someone unknown (with whom there is no frequent interaction or exchange of material or non-material goods, to return to the terms of Brown & Levinson 1987). For the sake of clarity, though, it is important to mention that, even among this kind of sentences, it is possible to find cases where the informant uses *tu* (a pronoun typically used in low social distance interactions) instead of *Lei* (used in high social distance interactions), as there can be an indication that the listener has the same age as the speaker. In the case of LD, the hypothetical hearers of the subjects were identified as being friends. All subjects were selected according to the criteria used for Italian in the investigation of Romance intonation project (Gili Fivela et al. 2015). Accordingly, those analyzed

here are university students, two women (speakers 2Gf, 3Ef) and a man (4Rm), aged between 22 and 26 years. They were born and raised in Lecce and, at the time of data collection, had not lived anywhere else for more than a year. In addition, their parents were also born, raised and living in Lecce.

3.2 Experimental procedure and analysis

The contexts used to elicit the data were randomized and presented to informants via a PC monitor. The audio signal was captured using a professional microphone and recorded with the SoundRecorder function of Praat, which was installed on a computer equipped with a Realtek onboard audio card. A brief explanation and training session preceded the recording session, and a native speaker of Italian ensured that subjects were giving the correct interpretation to the contexts and, therefore, the target utterances.

Each subject was asked to produce 5 repetitions of each context/sentence. Only read target utterances are analyzed here, that is a sample of 330 sentences (11 sentences x 5 renditions x LD and HD x 3 speakers).

Target utterances were phonologically analyzed and labelled for further phonetic investigation. The phonological analysis, which is the focus of this paper, was performed within the Autosegmental-Metrical framework (Pierrehumbert 1980; Ladd, 2008) by identifying pitch accents and edge tones that characterized the different sentence types, also with reference to previous analysis of the Lecce Italian variety. The labelling procedure regarded the main segmental and prosodic boundaries (syllables, phrases) and intonational events (tonal targets). Details concerning alignment and scaling of target tones were impressionistically observed thanks to the phonetic labelling, which is not crucial at this stage, but will rather be used for further acoustic investigation.

4. Results

The phonological patterns observed in the data are discussed in the following sections, starting with those found in control contexts, i.e. broad focus statements and information seeking yes-no questions. As for orders and requests, results regard both simple and mitigated forms, and, as for the latter, a further control is considered by taking into account adverbial phrases that do not perform a mitigating function (see 4.1).

4.1 Broad focus statements

Broad focus statements are mostly expressed by means of a H+L* L% pitch accent in both HD and LD contexts (see table 1, where “n. 15” indicates the result of 5 renditions of our 3 subjects), in line with the pattern previously found in several varieties of Italian (Gili Fivela et al., 2015; Gili Fivela & Nicora, 2018). Besides the expected variability in the realization of the H leading tone on the prenuclear, a

considerable variability concerning L* on the nuclear syllable is also found, being the latter performed both as a gradual fall within the vowel and as low target aligned quite early in it - that continues at the same level until the last syllable.

Observing the data in Table 1 it is possible to notice that there are some cases of a H*+L L% nuclear pattern. All of them correspond to the renditions of only one speaker (2Gf), who use it more often in LD than in HD contexts (where H*+L is produced only within a double pitch accent pattern involving two pitch accents on the target words, H* H*+L L%). We cannot exclude that this is due to the speaker's interpretation of the context as if a narrow focus statement was required. However, as the interpretations were checked during recordings, and there is a decrease in the number of instances in HD contexts, we think that this pattern may also be perceived as appropriate, especially in LD contexts. However, H+L* L% remains the most used pattern in both social distance contexts.

Table 1 - *Nuclear patterns found in broad focus statements – HD and LD contexts*

High social distance (HD) – n. 15					Low social distance (LD) – n. 15					
Nuclear patterns	Speakers				Total %	Nuclear patterns	Speakers			
	2Gf %	3Ef %	4Rm %	Total %			2Gf %	3Ef %	4Rm %	Total %
H+L* L%	26,6	33,3	33,3	93,2	H+L* L%	0	33,3	33,3	66,6	
H* H*+L L%	6,6	0	0	6,6	H*+L L%	33,3	0	0	33,3	

4.2 Information seeking yes-no questions

In the case of information seeking yes-no questions, the most frequent nuclear pattern is H*+ L LH%, with a peak aligned to the first half of the nuclear vowel and a pronounced fall phase (in line with previous analyses, Gili Fivela et al. 2015), both in HD (79,9% of cases) and LD (66,6% of cases) contexts – see table 2. As for the boundary tone LH%, in some renditions the rise is compressed and the LH% tone is performed with minimal modulation. At the present stage of analysis, this kind of realization was provisionally annotated as L[!H]%, and reported in italics in the table as a way to recognize the frequency of this phenomenon for the purpose of the paper. We think it is purely phonetic and does not lead to a different interpretation of the sentence (fig. 1 left *vs* right), but could possibly play a different role in the two social distances contexts (or be related to mitigation processes – see §5.4.2). Further, as shown in Table 2, this pattern is only found in productions by the speaker (2Gf), who also showed peculiarities in broad focus statements.

Even though the most used accent for information seeking yes-no questions is a rise-fall including a peak aligned to the first half of the nuclear vowel (H*+L LH%), speaker 4Rm used also a L+H* L!H% pattern to utter a few renditions both in HD and LD contexts, that is, a contour with a rise phase throughout the nuclear syllable, a peak placed in the second half of the nuclear vowel and a final rise showing a very reduced F0 decrease. This specific pattern was already found in Lecce for

counter-expectational yes/no questions (Gili Fivela et al., 2015). Indeed, in these particular utterances, a nuance of doubt seems to be at place and subject 4Rm, according to our perception, sometimes, sounds incredulous.

Further, another pattern is used by one of our speakers, that is 3Ef, who uses an L+H* H!H% pattern in both social distance conditions but especially in LD sentences. Such pattern is already attested in Lecce and several other Italian varieties (Gili Fivela et. al. 2015 and following works), though for a different type of sentence, i.e. the vocative. Here, it corresponds to an information seeking question characterized by a particularly chanting end. Further analysis will show if, besides the similar chanting quality, the pattern differs from the vocative one as for phonetic details such as the F0 range or the intensity level.

Figure 1 - *Information seeking yes-no question Indovina?, 'Does he/she guess it?', produced in contexts of HD (left, speaker 2Gf) and in LD (right, speaker 2Gf), where L[!H]% is highlighted for clarity sake, though it is not part of the phonological inventory*

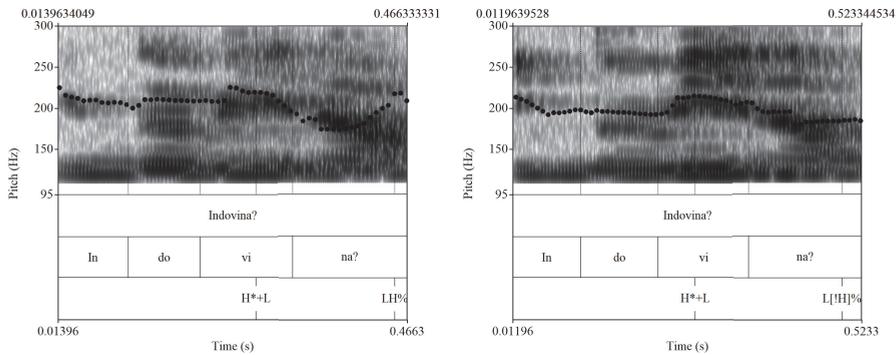


Table 2 - *Nuclear patterns found in information seeking yes-no questions – HD and LD contexts*

High social distance (HD) – n. 15						Low social distance (LD) – n. 15					
Nuclear patterns	Speakers			Partial %	Total %	Nuclear patterns	Speakers			Partial %	Total %
	2Gf %	3Ef %	4Rm %				2Gf %	3Ef %	4Rm %		
H*+L LH%	20	20	20	60	79,9	H*+L LH%	13,3	13,3	20	46,6	66,6
H*+L L[!H]%	13,3	6,6	0	19,9		H*+L[!H]%	20	0	0	20	
L+H* L!H%	0	0	13,3	13,3	13,3	L+H* L!H%	0	0	13,3	13,3	13,3
L+H* H!H%	0	6,6	0	6,6	6,6	L+H* H!H%	0	20	0	20	20

4.3 Orders

In several varieties of Italian, including that spoken in Lecce, orders are usually uttered with a falling pattern H+L* L% or, alternatively, with a rising-falling pattern H*+L L% (Gili Fivela et al., 2015). Our data are in line with these observations, since in the case of both social distance contexts the most recurrent pattern is H+L* L% (59,9%), which was produced, to some extent, by all speakers (see table 3); al-

ternatively, two out of three speakers used H*+L L% in some of their renditions (33,2 % of cases in HD and LD, specially by 2Gf, who produced it in almost all the orders). For speaker 3Ef, a secondary strategy is represented by the use of a double pitch accent H* H*+L L%.

Table 3 - Nuclear patterns found in orders – HD and LD contexts

High social distance (HD) – n. 15					Low social distance (LD) – n. 15				
Nuclear patterns	Speakers			Total %	Nuclear patterns	Speakers			Total %
	2Gf %	3Ef %	4Rm %			2Gf %	3Ef %	4Rm %	
H+L*L%	6,6	20	33,3	59,9	H+L*L%	6,6	20	33,3	59,9
H*+L L%	26,6	6,6	0	33,2	H*+L L%	26,6	6,6	0	33,2
H* H*+L L%	0	6,6	0	6,6	H* H*+L L%	0	6,6	0	6,6

As for the mitigated vs. non-mitigated contexts, unlike the other subjects, speaker 2Gf used the accent H*+L L% in most of non-mitigated orders (see table 3), and the same occurred in mitigated productions, both in HD and LD contexts (see the nuclear patterns in table 4, upper part). Speakers 3Ef and 4Rm, though, slightly changed the accent of the target word *indovina* in presence of mitigation, and in HD contexts both speakers split their productions between the two possible contours. Therefore, H*+L L% is favored (46,5% of the cases, although especially speaker 4Rm still prefers the use of an H+L* L% pattern). In LD, the preference for H*+L L% is even more clear (66,5% of the cases in table 4), as one of the two speakers (3Ef) started to use it twice as more than in HD contexts. A secondary strategy is still represented by the use of a double pitch accent for speaker 3Ef and for speaker 4Rm mainly in HD contexts. In the case of mitigation, then, a slightly different strategy is observed in comparison to that found in orders with no modulation of the illocutionary force.

If we observe the patterns associated with the mitigator *per favore* (see table 4, lower part), in HD contexts, two speakers prioritize an H*+L L-, which is, therefore, the most used one (46,6% of the cases). Only the 4Rm speaker always uses a rising accent L+H*, with a boundary tone realized either as high, H-, or low, L-. So, for this specific speaker, most of the time, we have a (L+H* H-/L-) H+L* L% combination of patterns for mitigator and target word, while for the other two speakers we have (H*+L L-) H*+L L%. In LD contexts, speaker 4Rm maintains exactly the same behavior as in HD. In the case of the two other speakers, however, one (2Gf) uses an H+L* L- pattern more frequently, which lead us to have the same percentage for the two falling accents for the adverbs (33,3% of the cases). Thus, in terms of a higher frequency in contexts of LD, we have for speakers 2Gf and 3Ef a H*+L L% pattern on the target word, while the mitigator can be either H*+L L- or H+L* L- (fig. 2).

Figure 2 - *Mitigated order* Per favore, indovina!, "Please, guess it!", produced in HD (left, speaker 3Ef) and LD (right, speaker 3Ef) contexts

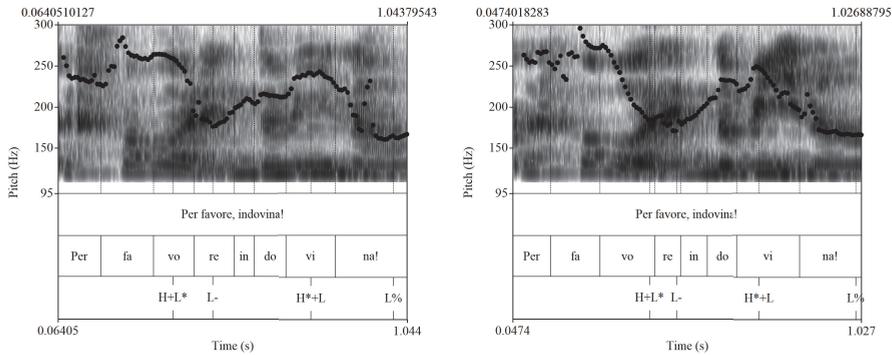


Table 4 - *Patterns in mitigated orders* Per favore, indovina!, "Please, guess it!" – HD and LD contexts

High social distance (HD) – n. 15					Low social distance (LD) – n. 15				
Nuclear patterns	Speakers			Total %	Nuclear patterns	Speakers			Total %
	2Gf %	3Ef %	4Rm %			2Gf %	3Ef %	4Rm %	
H*+L L%	26,6	13,3	6,6	46,5	H*+L L%	26,6	26,6	13,3	66,5
H+L* L%	6,6	13,3	20	39,9	H+L* L%	6,6	0	20	26,6
H* H*+L L%	0	6,6	6,6	13,2	H* H*+L L%	0	6,6	0	6,6

High social distance (HD) – n. 15					Low social distance (LD) – n. 15				
Per favore	Speakers			Total %	Per favore	Speakers			Total %
	2Gf %	3Ef %	4Rm %			2Gf %	3Ef %	4Rm %	
H*+L L-	26,6	20	0	46,6	H*+L L-	13,3	20	0	33,3
H+L* L-	6,6	13,3	0	19,9	H+L* L-	20	13,3	0	33,3
L+H* H-	0	0	20	20	L+H* H-	0	0	20	20
L+H* L-	0	0	13,3	13,3	L+H* L-	0	0	13,3	13,3

In orders with a non-mitigating adverb (see table 5), a further increase of the nuclear pattern H*+L L% compared to that of orders with *per favore* is found. There is, though, a very clear prevalence especially in LD (86,5% of cases; see table 5, upper part), mainly due to the fact that, in comparison to mitigated sentences, speaker 4Rm changes the preferred accent from H+L* to H*+L. Finally, 3Ef uses a double pitch accent as a secondary strategy, both in HD and in LD.

With regard to the patterns associated with the adverb *alla svelta*, both in HD and in LD contexts (table 5, lower part), we observed that although speaker 4Rm uses more often a falling contour H+L* L-, the other two prioritize a H*+L L- pattern, this being, therefore, the most recurrent one (59,9% in HD and 53,3% in LD). Thus, in HD contexts, for speaker 4Rm, we have a (H+L* L-) H*+L L% combination of ad-

verb and target word, for speaker 2Gf we have (H*+L L-) H+L* L% and, for speaker 3Ef, (H*+L L-) H*+L L%. In LD contexts, there is a change only as for speaker 2Gf, which uses more often the combination (H*+L L-) H*+L L%, with a different pitch accent choice in comparison to the target word of HD renditions (fig. 3).

Figure 3 - *Order with non-mitigating adverb Alla svelta, indovina!, "Quickly, guess it!" produced in HD context (speaker 3Ef)*

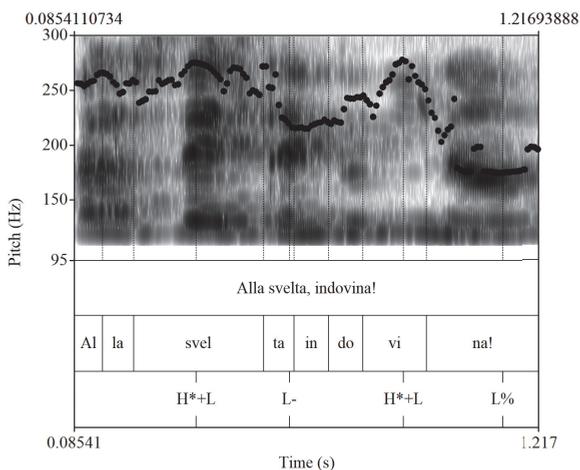


Table 5 - *Patterns found in orders with non-mitigating adverb Alla svelta, indovina!, "Quickly, guess it!" – HD and LD contexts*

High social distance (HD) – n. 15					Low social distance (LD) – n. 15				
Nuclear patterns	Speakers			Total %	Nuclear patterns	Speakers			Total %
	2Gf %	3Ef %	4Rm %			2Gf %	3Ef %	4Rm %	
H*+L L%	13,3	20	33,3	66,6	H*+L L%	26,6	26,6	33,3	86,5
H+L* L%	20	0	0	20	H+L* L%	6,6	0	0	6,6
H* H*+L L%	0	13,3	0	13,3	H* H*+L L%	0	6,6	0	6,6
High social distance (HD) – n. 15					Low social distance (LD) – n. 15				
<i>Alla svelta</i>	Speakers			Total %	<i>Alla svelta</i>	Speakers			Total %
	2Gf %	3Ef %	4Rm %			2Gf %	3Ef %	4Rm %	
H*+L L-	20	33,3	6,6	59,9	H*+L L-	20	33,3	0	53,3
H+L* L-	13,3	0	26,6	39,9	H+L* L-	13,3	0	33,3	46,6

Thus, H*+L L% is more frequent than H+L* L% as a nuclear pattern in more complex orders, both in HD and LD contexts and independently of the mitigating function of the adverb; further, it is more frequent in LD than in HD. Besides, phonetic implementation details are probably relevant in differentiating utterances

sharing the same main pattern (H+L* L%), such as unmodulated orders and broad focus statements, as well as unmodulated requests in imperative form.

4.4 Requests

4.4.1 Imperative form

Requests expressed by means of imperative form are mostly uttered using a regular falling pattern H+L* L% (figure 4, table 6) both in HD and LD contexts, with a particularly high leading tone on the prenuclear of the renditions of only one female speaker (3Ef). A previous study on Lecce Italian (Gili Fivela et al., 2015) reported an H*+L L% pattern. Indeed, our data showed that this kind of accent is the second most used in this type of sentence (see table 6) in both social distance contexts, but notice that it appears only in the renditions of one single speaker (2Gf). Other options are a double accented target (H* H*+L L%) for speaker 3Ef in both social distance contexts and, for speaker 2Gf and only in LD contexts, the use of a regular L+H* L%, which sounds less preemptory than the requests uttered with an H*+L accent (especially when in a lower range).

Figure 4 - *Request in imperative form* Indovina!, "Guess it!", produced in a HD context (speaker 4Rm)

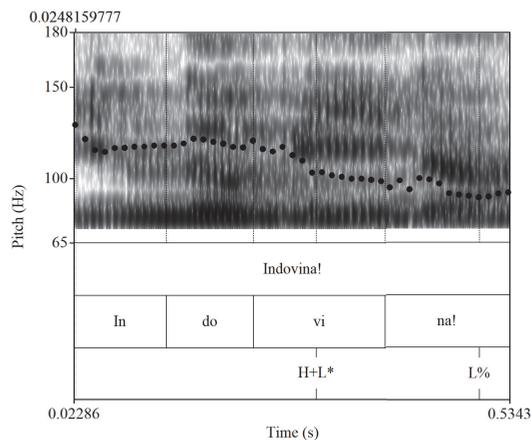
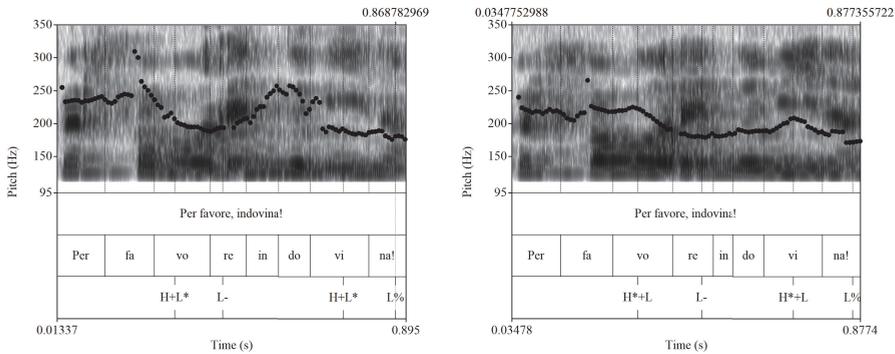


Table 6 - *Nuclear patterns found in requests expressed by means of imperative form*

High social distance (HD) – n. 15					Low social distance (LD) – n. 15				
Nuclear patterns	Speakers			Total %	Nuclear patterns	Speakers			Total %
	2Gf %	3Ef %	4Rm %			2Gf %	3Ef %	4Rm %	
H+L* L%	6,6	20	33,3	59,9	H+L* L%	0	20	33,3	53,3
H*+L L%	26,6	0	0	26,6	H*+L L%	20	0	0	20
H* H*+L L%	0	13,3	0	13,3	H* H*+L L%	0	13,3	0	13,3
-----	---	---	---	---	L+H* L%	13,3	0	0	13,3

When it comes to the mitigated requests in imperative form (table 7), results show that nuclear patterns (table 7, upper part) used in HD contexts equally split between the main options mentioned above, that is, $H+L^*L\%$ and $H^*+L\%$ (around 46,5% each; fig. 5), with a consequential increase in the number of $H^*+L\%$ nuclear patterns; comparing to non-mitigated requests. This is mainly due to productions of one of the speakers (3Ef), who used mostly a nuclear $H+L^*L\%$ pattern on non-mitigated requests and started to use more often a different one in presence of mitigation. This particular phenomenon repeats in LD contexts, where the above mentioned increase is even clearer as $H^*+L\%$ is the most used contour (59,9% of cases). It is important to highlight, though, that one of the speakers (4Rm) always used a $H+L^*L\%$ nuclear pattern on mitigated and simple requests, in both social distances.

Figure 5 - *Mitigated request in imperative form* Per favore, indovina, "Please, guess it!", produced in HD (left, speaker 2Gf) and LD (right, speaker 2Gf) contexts



On the mitigator *per favore* (table 7, lower part), two accents are found in HD, one involving a falling phase ($H^*+L\ L^-$ and $H+L^*L^-$), the other one including a rise ($L+H^*L^-$) and being used almost exclusively by speaker 4Rm, just like in the case of mitigated orders (cf. table 4). Thus, in HD contexts, for speaker 2Gf we have a combination of patterns for mitigator and target word ($H+L^*L^-/H^*+L\ L^-$) $H^*+L\%$, for speaker 3Ef the combination is ($H+L^*L^-$) $H^*+L\%$, while for 4Rm we have ($L+H^*L^-$) $H+L^*L\%$. In the case of LD sentences, $H^*+L\ L^-$ is used on *per favore* in 59,9% of the cases, but speaker 4Rm still uses an $L+H^*L^-$ pattern. Consequently, for this specific speaker the combination of patterns on mitigator and target word is ($L+H^*L^-$) $H+L^*L\%$, while for speakers 2Gf and 3Ef we have ($H^*+L\ L^-$) $H^*+L\%$.

Table 7 - Patterns found in mitigated requests expressed by means of imperative form

High social distance (HD) – n. 15					Low social distance (LD) – n. 15				
Nuclear patterns	Speakers			Total %	Nuclear patterns	Speakers			Total %
	2Gf %	3Ef %	4Rm %			2Gf %	3Ef %	4Rm %	
H*+L L%	26,6	20	0	46,6	H*+L L%	33,3	26,6	0	59,9
H+L* L%	6,6	6,6	33,3	46,5	H+L* L%	0	0	33,3	33,3
H* H*+L L%	0	6,6	0	6,6	H* H*+L L%	0	6,6	0	6,6

High social distance (HD) – n. 15					Low social distance (LD) – n. 15				
Per favore	Speakers			Total %	Per favore	Speakers			Total %
	2Gf %	3Ef %	4Rm %			2Gf %	3Ef %	4Rm %	
L+H* L-	6,6	0	33,3	39,9	L+H* L-	0	0	33,3	33,3
H+L* L-	13,3	20	0	33,3	H+L* L-	0	6,6	0	6,6
H*+L L-	13,3	13,3	0	26,6	H*+L L-	33,3	26,6	0	59,9

As for the imperative requests with a non-mitigator adverb (table 8), in HD contexts the most frequent pattern associated with *indovina* is H+L* L% (59,9% of cases; table 8, upper part), as the one used in imperative requests with no modulation of the illocutionary force, being H*+L L% a secondary option for speakers 2Gf and 3Ef and a double accented pattern H* H*+L L% only for speaker 3Ef.

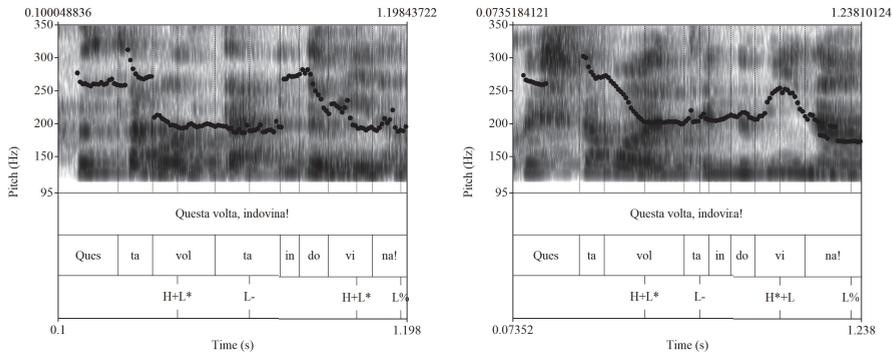
Figure 6 - Requests in imperative form with non-mitigating adverb *Questa volta, indovina!*, “This time, guess it!”, produced in HD (left, speaker 2Gf) and LD (right, speaker 3Ef) contexts

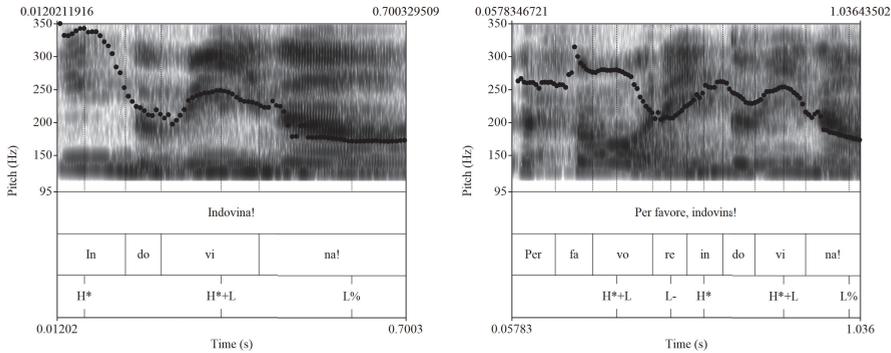
Table 8 - *Patterns found in requests with non-mitigating adverbs expressed by means of imperative form*

High social distance (HD) – n. 15					Low social distance (LD) – n. 15				
Nuclear patterns	Speakers			Total %	Nuclear patterns	Speakers			Total %
	2Gf %	3Ef %	4Rm %			2Gf %	3Ef %	4Rm %	
H+L* L%	20	6,6	33,3	59,9	H+L* L%	0	0	20	20
H*+L L%	13,3	13,3	0	26,6	H*+L L%	33,3	20	6,6	59,9
H* H*+L L%	0	13,3	0	13,3	H* H*+L L%	0	13,3	0	13,3
-----	---	---	---	---	H* H+L* L%	0	0	6,6	6,6
High social distance (HD) – n. 15					Low social distance (LD) – n. 15				
<i>Questa volta</i>	Speakers			Total %	<i>Questa volta</i>	Speakers			Total %
	2Gf %	3Ef %	4Rm %			2Gf %	3Ef %	4Rm %	
H+L* L-	26,6	20	0	46,6	H+L* L-	20	33,3	0	53,3
L+H* H-	0	0	33,3	33,3	L+H* H-	0	0	33,3	33,3
H*+L L-	6,6	13,3	0	19,9	H*+L L-	6,6	0	0	6,6
-----	---	---	---	---	L+H* L-	6,6	0	0	6,6

As for the non-mitigating adverb *questa volta* (table 8, lower part), H+L* L% is once more the most frequent contour (figure 6, left), with 46,6% of the instances of HD contexts. An L+H* H- pattern is also used in a few productions, but only by speaker 4Rm. Thus, in HD contexts, the dominant combination of patterns for adverb and target word for speaker 2Gf is (H+L* L-) H+L* L%, for speaker 3Ef it is (H+L* L-) H*+L L% or H* H*+L L% and (L+H* H-) H+L* L% for speaker 4Rm. In LD contexts, even though the pattern associated with the adverb is also most frequently H+L* L- (53,3% of cases), with speaker 4Rm using again a different contour (L+H* H-), there is a shift in the pattern most frequently associated with the target word *indovina* (59,9% of cases, table 8, upper part), which once more is H*+L L%, as in LD mitigated imperative requests (figure 5, right). This shift happened because both speakers 2Gf and 3Ef focused their productions on this pattern, even though speaker 4Rm still used an H+L* L%. We have, then, for adverb and target word in LD, a (H+L* L-) H*+L L% combination for speakers 2Gf and 3Ef, while for speaker 4Rm the patterns are (L+H* H-) H+L* L%.

It is interesting to notice that, as in the case of orders, there are a few occurrences of a nuclear double pitch accent among the requests expressed by means of imperative form without lexical modulation of illocutionary force, but also in those mitigated and in those with a non-mitigating adverb, always in the sentences produced by one particular speaker (3Ef). In those cases, the pattern is H* H*+L L%, as shown in figure 7.

Figure 7 - *Requests in imperative form* Indovina!, “Guess it”, produced in HD context (left, speaker 3Gf) and Per favore, indovina!, “Please, guess it”, produced in LD context (right, speaker 3Ef)



4.4.2 Interrogative form

Both in HD (99,7% of cases) and LD (86,5% of cases) contexts, the pattern most frequently associated with requests in interrogative form is H*+L LH% (see table 9), which corresponds to the one found in information seeking yes-no questions (see §5.2). Another similarity with the latter is that the boundary tone LH% can also be performed with minimal modulation, without changing, though, the meaning of the sentence, or the speech act type. The pattern H*+L L[!H]% is, thus, a representation of an exclusively phonetic variation of the H*+L LH% pattern itself, which just seems slightly more frequent in interrogative requests (and in LD contexts) in comparison to information seeking questions (fig. 8). Besides, in the case of LD contexts, speaker 3Ef used an L+H* H!H% pattern, which is typically found on vocatives, but in this case clearly indicates a request, possibly with a chanting end.

Figure 8 - *Requests in interrogative form* Indovini?, “Do you guess it?”, produced in contexts of HD (left, speaker 3Gf), and LD (right, speaker 4Rm) – L[!H]% is not part of the phonological inventory (see text)

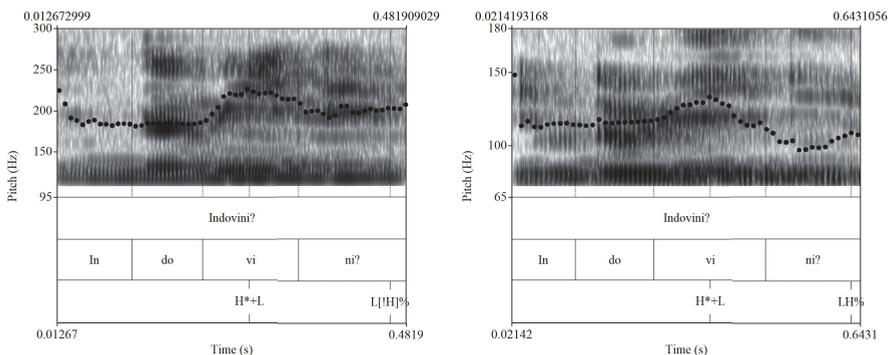


Table 9 - Nuclear patterns found in requests expressed by means of interrogative form

High social distance (HD) - n. 15						Low social distance (LD) - n. 15					
Nuclear patterns	Speakers			Partial %	Total %	Nuclear patterns	Speakers			Partial %	Total %
	2Gf %	3Ef %	4Rm %				2Gf %	3Ef %	4Rm %		
H*+L LH%	33,3	6,6	26,6	66,5	99,7	H*+L LH%	20	6,6	20	46,6	86,5
H*+LL[!H]%	0	26,6	6,6	33,2		H*+LL[!H]%	13,3	13,3	13,3	39,9	
-----	---	---	---	---		L+H* H!H%	0	13,3	0	13,3	

Also in mitigated requests in interrogative form (table 10, figure 9), in HD (72,2% of cases) and LD (88,6% of cases) contexts, the target word is mostly uttered with an H*+L LH% (or L[!H]%) contour and, even though we do not consider the difference between LH% and L[!H]% as phonological, we observe the higher frequency of the latter rendition in LD (53,3%) than in HD (26,6%) contexts (table 10, upper part). Further, once more, speaker 3Ef used a chanting pattern L+H* H!H% in some of the sentences. On the other hand, there is a HD sentence of speaker 4Rm with a L+H* L% contour (which, by the way, matches the one used on the mitigator), which does not correspond to a prototypical interrogative request pattern, and presents, according to our interpretation, a nuance of impatience or insistence that could be related to a particular interpretation of the context in that specific rendition. As for the mitigator *per favore* (“please”, table 10, lower part), specifically those found in HD contexts, very close percentages were found, with a minor prevalence of H+L* L- (39,9% of cases), since each speaker showed a preference for a different type of contour: speaker 4Rm always used an L+H* L- (as in most of his complex sentences already analyzed), speaker 3Ef preferred an H+L* L- contour, while 2Gf, most of the time, used a regular H*+L L-. In LD contexts, only speaker 3Ef presented a different behavior regarding the scenario we have just described, as she more often used the H*+L L- pattern, which presents, then, a higher frequency, with a more significant percentage (59,9% of cases).

Figure 9 - Mitigated requests in interrogative form *Per favore, indovini?*, “Please, do you guess it?”, produced in HD (left, speaker 2Gf) and LD contexts (right, speaker 2Gf) – L[!H] is not part of the phonological inventory (see text)

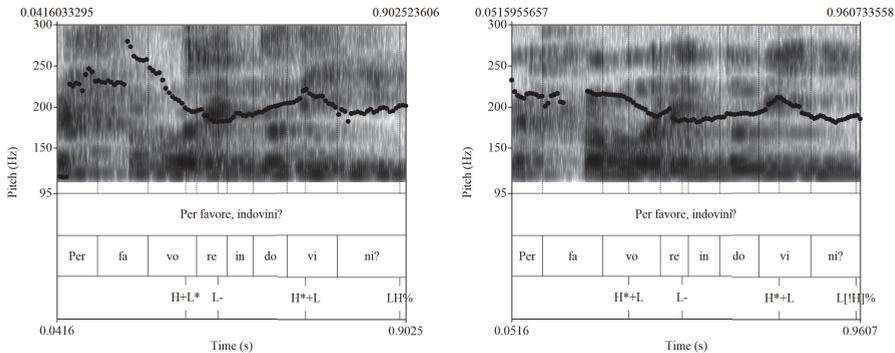


Table 10 - *Patterns found in mitigated requests expressed by means of interrogative form*

High social distance (HD) – n. 15						Low social distance (LD) – n. 15					
Nuclear patterns	Speakers			Partial %	Total %	Nuclear patterns	Speakers			Partial %	Total %
	2Gf %	3Ef %	4Rm %				2Gf %	3Ef %	4Rm %		
H*+L LH%	20	0	26,6	46,6	72,2	H*+L LH%	13,3	0	20	33,3	88,6
H*+LL[!H]%	13,3	13,3	0	26,6		H*+LL[!H]%	20	20	13,3	53,3	
L+H* H!H%	0	20	0	20	20	L+H* H!H%	0	13,3	0	13,3	13,3
L+H* L%	0	0	6,6	6,6	6,6	-----	---	---	---	---	

High social distance (HD) – n. 15					Low social distance (LD) – n. 15				
Per favore	Speakers			Total %	Per favore	Speakers			Total %
	2Gf %	3Ef %	4Rm %			2Gf %	3Ef %	4Rm %	
H+L* L-	13,3	26,6	0	39,9	H+L* L-	0	6,6	0	6,6
L+H* L-	0	0	33,3	33,3	L+H* L-	0	0	33,3	33,3
H*+L L-	20	6,6	0	26,6	H*+L L-	33,3	26,6	0	59,9

As for requests with a non-mitigator adverb, in both social distance contexts (72,2% of cases in HD and 73,2% of cases in LD, see upper part of table 11) the most recurrent pattern on the target word is, again, H*+L LH%, with a smaller number of L[!H]% renditions specially in LD contexts (6,6%) in comparison to mitigated productions, and with speaker 3Ef using the contour L+H* H!H%, but more pronouncedly in LD contexts. However, there are a few alternative productions among HD utterances which do not sound prototypical. Speaker 3Ef used a L+H* H% pattern in one of her renditions, both on the target word and the adverb; the pattern recalls that of an elliptic question on the adverb and seems to be repeated on the target. Besides, there is a single case of a H+L* LH% (by speaker 2Gf, also with a similar contour in the adverb, except for the L- intermediate boundary tone), which has a strong nuance of availability check.

The adverbs (table 11, lower part), specifically in HD sentences, presented the same percentage of instances with H+L* L- and L+H* L- contours (46,6% of cases). This is due to the fact that speaker 4Rm used L+H* L- in all of his productions and that speaker 2Ef chose this contour a few times; further, because subjects 2Gf and 3Ef used a regular H+L* L- contour more consistently. In LD sentences, even though the others maintained the same behavior that we described for HD sentences, speaker 2Gf realized all her productions with the H+L* L- contour. It became, then, the most frequent one (59,9% of cases).

Table 11 - *Patterns found in requests with non-mitigating adverb expressed by means of interrogative form*

High social distance (HD) – n. 15						Low social distance (LD) – n. 15					
Nuclear patterns	Speakers			Partial %	Total %	Nuclear patterns	Speakers			Partial %	Total %
	2Gf %	3Ef %	4Rm %				2Gf %	3Ef %	4Rm %		
H*+L LH%	26,6	0	20	46,6	72,2	H*+L LH%	33,3	0	33,3	66,6	73,2
H*+L L[!H]%	0	13,3	13,3	26,6		H*+L L[!H]%	0	6,6	0	6,6	
L+H* H!H%	0	13,3	0	13,3	13,3	L+H* H!H%	0	26,6	0	26,6	26,6
H+L* LH%	6,6	0	0	6,6	6,6	-----	---	---	---	---	---
L+H* H%	0	6,6	0	6,6	6,6	-----	---	---	---	---	---

High social distance (HD) – n. 15					Low social distance (LD) – n. 15				
Questa volta	Speakers			Total %	Questa volta	Speakers			Total %
	2Gf %	3Ef %	4Rm %			2Gf %	3Ef %	4Rm %	
H+L* L-	20	26,6	0	46,6	H+L* L-	33,3	26,6	0	59,9
L+H* L-	13,3	0	33,3	46,6	L+H* L-	0	0	33,3	33,3
L+H* H-	0	6,6	0	6,6	H*+L L-	0	6,6	0	6,6

5. Discussion

Observing data related to sentences composed only of a target word, and specifically the intonational characteristics of orders and requests in imperative form, we noticed that they share the same main pattern, H+L* L%, which was also the one identified for broad focus sentences, included in the experiment as a control. These results corroborate those previously found in Lecce and are indeed in line with our initial hypothesis, except in the case of requests in imperative form, for which H*+L L% had been previously described as the main pattern (though not in one word sentences). They share the same intonational pattern and future phonetic analysis may show that they are differentiated by phonetic properties. As for the requests in interrogative form, we identified H*+L LH% - with a strictly phonetic variant showing a L[!H]% edge tone - as the main pattern, the same found in information seeking yes-no questions. Also in this case, phonetic features might differentiate the information seeking question and the request speech acts, in particular the lower final fundamental frequency of the L[!H]% boundary tone found in interrogative requests. Thus, as for our initial hypotheses, results show that there is a clear difference between the intonational patterns of short orders and requests when the latter are expressed by the interrogative form and thus, as far as intonation is concerned, by a different boundary tone.

Further, in the case of short sentences, we only varied the social distance factor and could observe that, besides the above mentioned changes, patterns in orders are not affected by social distance, while patterns in broad focus statements and requests expressed by imperative forms are affected: they show a slightly lower number of H+L* pitch accents in LD conditions in comparison to HD ones (in favor

of the H*+L and double accents H* H*+L in LD conditions). As for requests in interrogative form, results showed that they were affected by social distance as information seeking questions were, that is, they both showed a higher number of H*+L LH% pattern in HD contexts than in LD (adding H*+L LH% and H*+L L!H% realizations). However, requests also showed a slightly higher number of realizations including a lower final high boundary tone (L[!H]%). In our view this may correspond to a lighter cue to prototypical questioning, which is known to play a mitigation role.

However, to answer the main question of this paper, results related to complex sentences (mitigator/non-mitigator adverbs + target word *indovina* “guess it”) need to be discussed. In the case of orders produced in both social distance contexts, the preferential pattern used on the target word changed in presence of mitigation (*per favore* “please”) and a non-mitigating adverb (*alla svelta* “quickly”). Specifically, adding adverbs, mitigators or not, the pattern changes from H+L* L% to H*+L%, which is considered as more peremptory (Gili Fivela, 2008). Such pattern, which is already an option in both broad focus and orders according to previous investigations, is the most frequent contour in both social distance contexts though it is used much more often in the case of LD sentences with non-mitigating adverb (which is in line with Gili Fivela, Bazzanella, 2014). Therefore, with regard to social distance, it is reasonable to state that if the speakers are socially close, the tendency of using a more peremptory contour on the target word increases. That is, on the target word in orders our expectations are confirmed, as the presence of lexical means of mitigation and the low social distance favor more peremptory, less neutral patterns. Another important aspect to be noticed is that, in general, there is a tendency to use the H*+L L% pattern both on the target word and on the adverbs *per favore* “please” and *alla svelta* “quickly” (in phonological terms, we may hypothesize a pattern copying). However, the H*+L L% pattern on the adverbs is slightly more often used in HD contexts than LD contexts. In terms of the impact of social distance, then, the situation on adverbs in orders is the opposite than expected, as a more peremptory contour H*+L is slightly more often found in HD contexts.

Requests in imperative form including adverbs differ from orders with adverbs only in HD contexts: in the case of the non-mitigating adverb, the H+L* L% pattern is attested more, while in the case of mitigating adverb, it is used as much as the more peremptory one, almost as a way of indicating more clearly, from the intonational point of view, the imperative character of the mitigated sentence. Besides, in requests in imperative form with a mitigating adverb an increase in the use of the H*+L L% pattern on the target word is observed in comparison with imperatives with no preceding adverb. As for the social distance, in HD the above mentioned pattern is used as much as H+L* L%, but in contexts of LD represents the most used contour; a phenomenon that we also identified in the orders and which is in line with previous investigations on Lecce Italian. In other words, here, too, a change is observed in relation to simple imperative requests. As for the patterns used on the mitigating adverb, in HD contexts there is no clear prevalence, though L+H*

L- is used slightly more often, but in LD we see that, as in orders, there seems to be a pattern copying, since the predominant pattern is H*+ L L- on the adverb too. Thus, the tendency to use a peremptory pattern more frequently in LD situations is also found here. On the other hand, when it comes to imperative requests with non-mitigating adverbs, in HD productions, the main pattern of target words and adverbs is the same as for the simple sentences, that is, H+L* L, considered here as less peremptory. Our idea is that, in the absence of a term that clearly performs a mitigating function, mitigation would be carried out by prosodic means, considering that the interlocutor is an unknown person and it would be necessary to use more strategies to protect the participants' faces, in Brown and Levinson's (1987) terms. Indeed, if we look at the data related to LD contexts, in which there is greater intimacy between speakers, a more peremptory pattern H*+L L% on the target word is more frequent, although on the adverb an H+L* L- contour prevails. As already mentioned, if we compare the complex forms of orders and imperative requests, we notice that the latter differs from orders mainly in contexts with a non-mitigating adverb: in mitigated requests, the pattern is often H*+L L% both in HD and LD (just like in mitigated orders); in the ones with *questa volta* "this time", though, a potentially less peremptory contour is more usual in HD contexts. The non-mitigating adverbs of orders and requests are different (*alla svelta* "quickly" and *questa volta* "this time" respectively), but both suggests a solicitation, therefore, we believe that their being different is not what motivated the intonational difference described above.

Finally, complex requests in interrogative form, mitigated or with non-mitigating adverbs and in both social distance contexts, present a basic pattern H*+L LH%, as information seeking yes-no questions. Therefore, they are quite different from requests in an imperative form, being mitigated by the very modality of the sentence. It may be important to consider, however, how the phonetic variation with minimal modulation of the boundary tone occurs (here highlighted by H*+L L[! H]%, where [!H]% is not meant to be interpreted phonologically). In LD contexts, a nuclear pattern with a less clearly rising boundary tone is more frequent, especially in the case of the presence of a mitigating adverb, and, on that adverb, a more incisive pattern is found (H*+L L-), which may represent a lower need for prosodic mitigation among close subjects. On the contrary, the mitigating adverb does not seem to have a strong impact in HD (as the patterns are the same independently of the adverb function). In the case of interrogative requests with a non-mitigating adverb, for HD and especially for LD contexts, the combination of patterns H+L* L-, on the adverb, and H*+L LH% on the target word is more often used. As such combination might have a greater mitigating potential, it may be chosen because there is no clear lexical mitigation.

6. *Conclusions*

The aim of this paper was to investigate whether and to what extent mitigation procedures affect the prosodic characteristics of orders and requests made by speakers of Lecce Italian in different social distance conditions. The corpus analyzed here was collected using a variant of the Discourse Completion Task, and the analysis was carried out within the Autosegmental Metrical framework, that is aimed at identifying the phonological patterns.

Our first hypotheses was that orders and requests could be prosodically diverse, but also tightly related, with orders and requests in imperative form being similar on the one hand, and information seeking yes-no questions and requests in interrogative form on the other. Results confirmed such hypothesis and showed that there is a change in the intonational patterns of orders and requests especially when the latter are expressed by the interrogative form and thus, as far as intonation is concerned, by a different boundary tone. Further, while only phonetic analysis will possibly confirm the existence of implementation differences when the same phonological pattern has been found (e.g. orders, broad focus statements and some imperative requests), in some cases available observations point to the existence of differentiating cues. This is the case of the lower rising in L[!H]%, which was found slightly more often in short requests in interrogative form in comparison to questions. In our view, this may correspond to a lighter cue to prototypical questioning, possibly helping in differentiating requests from information seeking questions (in line with other observations in the literature, e.g. Frota & Prieto 2015).

As for the impact of mitigation, we assumed that explicit lexical mitigators could have an impact on intonation patterns, favoring more peremptory, less neutral patterns. Similar patterns were also expected in low social distance contexts. Indeed, in presence of mitigation (and of a non-mitigating adverb), in both social distances, the preferential nuclear pattern of orders changed from a less incisive to a more peremptory one. Regarding requests in imperative form, mitigation seems to be carried out by prosodic means in the absence of a lexical mitigation, with a preference of less peremptory patterns especially in HD; the tendency to use a peremptory pattern more frequently in LD situations is also found. Considering requests in interrogative form, which may be naturally considered a more polite way of uttering a request, the presence of a mitigating adverb does not seem to have a strong impact in HD contexts (as the nuclear patterns are the same independently of the adverb function), while in LD the lexical mitigation seems to be accompanied by an increase of overall mitigating (question like) intonation patterns. However, such patterns also correspond to a less clearly and potentially less mitigating L[!H]% rising boundary tone that appears more often both in LD non-mitigated and mitigated contexts, as if in LD the prosodic mitigation could be weaker.

Thus, though mitigation is clearly intertwined with social distance, results show that lexical mitigation interferes with intonational cues in line with our hypotheses (lack of lexical cue favors intonational ones) in requests expressed by imperative forms. In requests in interrogative form, the impact is found in LD contexts only,

where it apparently goes contrary to our expectations (presence of lexical cues is accompanied by intonational one), even though it does not when considering the L[!H]% phonetic variant, which is a less modulated boundary tone and could be interpreted as being more incisive. Thus, our initial hypothesis on interference of mitigating lexical material was at least partly confirmed, and, further, once more data will be analyzed the interplay of the patterns found on adverbs will probably shed more light on the issue.

Finally, social distance most of the times seems to act in line with expectations. In the case of short sentences, results are consistent in broad focus statements, with a slightly higher number of the more incisive pattern H*+L (even though it is only the secondary one) in LD conditions in comparison to HD ones.

Turning to interrogative forms, both requests in interrogative form and information seeking questions were affected by social distance as they both showed a higher number of H*+L LH% patterns in HD contexts than in LD, where, at least in requests, they showed a slightly higher number of L[!H]%. Thus, our hypothesis on the impact of low social distance to favor more peremptory, less neutral patterns was at least partly confirmed.

All in all, it is quite clear that we cannot generalize results, as, besides the low number of subjects and observations, it is evident that differences are found across speech acts. In any case, those reported in this paper are preliminary observations only, regarding a subset of our subjects. Further, a phonetic analysis is expected to provide important clues about the phenomena we described.

Bibliography

- ALBANO LEONI, F. (2003). Tre progetti per l'italiano parlato: AVIP, API, CLIPS. In MARASCHIO, N. & POGGI-SALANI, T. (eds.). *Atti del XXXIV Congresso Internazionale di Studi della Società di linguistica italiana (SLI)*, 19-21 ottobre 2000. Firenze: Bulzoni, 675-683.
- AUSTIN, J.L. (1962). *How to do things with words*. Oxford: Clarendon Press.
- BLUM-KULKA, S., HOUSE, J. & KASPER, G. (1989). Investigating Cross-Cultural Pragmatics: an Introductory Overview. In BLUM-KULKA, S., HOUSE, J. & KASPER, G. (eds.). *Cross-Cultural pragmatics: Requests and Apologies*. Norwood, NJ: Ablex, 1-28.
- BROWN, P. & LEVINSON, S.C. (1987). *Politeness: some universals in language use*. Cambridge: CUP.
- CABALLERO, J.A., VERGIS, N., JIANG, X. & PELL, M.D. (2018). The sound of Im/politeness. In *Speech Communication*, 102, 39-53.
- FÉRY-VIDAL, M.V. (1996). *Introducción a la Pragmática*. Barcelona: Editorial Ariel.
- FÉRY, C. (2013). Focus as prosodic alignment. In *Natural Language and Linguistic Theory*, 31, 683-734.
- FROTA, S. & PRIETO, P. (2015). *Intonation in Romance*. Oxford: Oxford University Press.
- GILI FIVELA, B. (2008). *Intonation in production and perception: The case of Pisa Italian*. Torino: Edizioni dell'Orso.

- GILI FIVELA, B. & BAZZANELLA, C. (2014). The relevance of prosody and context to the interplay between intensity and politeness. An exploratory study on Italian. In *J. Politeness Research* 10, 97–126.
- GILI FIVELA B., AVESANI, C., BARONE, M., BOCCI, G., CROCCO, C., D'IMPERIO, M., GIORDANO, R., MAROTTA, G., SAVINO, M. & SORIANELLO, P. (2015). Intonational phonology of the regional varieties of Italian. In FROTA, S., PRIETO, P. (eds.). *Intonation in Romance*. Oxford: OUP, 140-197.
- GILI FIVELA, B. & NICORA, F. (2018). Intonation in Liguria and Tuscany: checking for similarities across a traditional isogloss boundary. In VIETTI, A., SPREAFICO, L., MEREU, D. & GALATÀ, V. (eds.). *Speech in the natural context*. Studi AISV 4, Milano: Officinaventuno, 131-156.
- HIDALGO, A. (2003). Cortesía y prosodia: un estudio de la frase cortés en el español de Mérida (Venezuela). MARTIN BUTRAGUEÑO, P. & HERRERA, Z.E. (eds.). *La tonía. Dimensiones fonéticas y fonológicas*. México: El Colegio de México, 319-330.
- HÜBSCHER, I., BORRÀS-COMES, J. & PRIETO, P. (2017). Prosodic mitigation characterizes Catalan formal speech: The Frequency Code reassessed. *Journal of Phonetics*, 65, 145-169.
- LADD, R. (2008). *Intonational Phonology*. 2^a edition. Cambridge: Cambridge University Press.
- NESPOR, M. & VOGEL, I. (1986). *Prosodic Phonology*. Dordrecht: Foris publications.
- NUZZO, E. (2013). La pragmatica nei manuali d'Italiano L2: una prima indagine sull'atto linguistico del ringraziare. *Revista de Italianística*, XXVI(1), 5-29.
- PIERREHUMBERT, J. (1980) The phonology and phonetics of English intonation. Massachusetts Institute of Technology, Dept. of Linguistics and Philosophy.
- ROSEANO, P., VANRELL, M.M. & PRIETO, P. (2015). "Intonational phonology of Friulan and its dialects". In: FROTA, S., PRIETO P. (eds.). *Intonation in Romance*. Oxford: OUP, 101-139.
- SANTORO, E. (2017). Richieste e attenuazione: un confronto tra italiano e portoghese brasiliano. *Normas*, vol. 7, n. 2, 179-204.
- SANTORO, E., KULIKOWSKI, M.Z., SILVA, L.A. (2017). Pragmática sociocultural: a elaboração de um corpus. In CABRAL, A.L.T., SEARA, I.R., GUARANHA, M.F. (orgs.). *Descortesia e cortesia: expressão de culturas*. São Paulo: Cortez Editora.
- SEARLE, J. (1979). *Expression and Meaning. Studies in the Theory of Speech Acts*. Cambridge: CUP.
- SELKIRK, E. (1984). *Phonology and syntax: the relation between sound and structure*. MIT Press.
- SILVA NETO, M. (2018). Requests in Italian: a study on the perception of politeness by native speakers and Brazilian learners. Master's degree dissertation, University of São Paulo.
- SILVA NETO, M. (in preparation). "Intonational patterns of orders and requests in Brazilian Portuguese and Italian: comparisons and possible effects on the perception of linguistic politeness". PhD Dissertation, University of São Paulo.
- SPADOTTO, L.N. & SANTORO, E. (2019). Ordens e pedidos em língua italiana: um estudo da percepção de falantes nativos e aprendizes brasileiros. *Revista Letrônica*, v. 12, n. 4.

SWERTS, M. & KRAHMER, E. (2008). "Facial expressions and prosodic prominence: comparing modalities and facial areas". In *Journal of Phonetics*, 36(2), 219-238.

TAKIMOTO, M. (2007). The Effects of Input-Based Tasks on the Development of Learners' Pragmatic Proficiency. *Applied Linguistics*, 30, Issue 1, 1-25.

VANRELL, M.M., BALLONE, F., SCHIRRU, C. & PRIETO, P. (2015). Sardinian intonational phonology: Logudurese and Campidanese varieties. In: FROTA, S., PRIETO, P. (eds.). *Intonation in Romance*. Oxford: Oxford University Press, 317-349.

GLEND A GURRADO

Indici fonetici dell'esclamativa wh nell'italiano parlato a Bari

The prosodic cues of wh-exclamatives in Bari Italian

The communicative function of exclamatives is based on the concept of salience. Specifically, wh-exclamatives convey a scalar implicature of highest degree: their value is always surprising. Although intonation plays a key role in disambiguating the sentence types, the literature on the prosody of exclamatives is still scarce.

The present research has two main aims: 1) verifying the phonetic cues of wh-exclamatives in Bari Italian, by means of a comparison with wh-interrogatives; 2) verifying if the emotional connotation affects the prosody of exclamative sentences. The findings revealed that some cues were more involved in the characterization of wh-exclamatives than others and that exclamatives with opposite emotional connotations did differ at a suprasegmental level.

Keywords: wh-exclamatives, prosody, acoustic analysis, emotional connotation, sense of surprise.

1. *Introduzione*

La frase esclamativa è considerata un atto linguistico espressivo di struttura variabile che fa uso di specifiche risorse sintattiche, lessicali e prosodiche. Questa tipologia frasale è molto usata nella lingua parlata e appartiene soprattutto a un registro informale; nella forma scritta è contrassegnata dal punto esclamativo.

La forza illocutiva della frase esclamativa risiede nella predisposizione del parlante a credere fortemente nel contenuto del messaggio e a voler comunicare quanto esso sia sorprendente (Rett, 2008); la reazione emotiva originata da un evento inaspettato e trasmessa dall'esclamativa può avere carattere positivo o negativo, a seconda del contesto. Il significato esclamatorio può essere comunicato mediante una serie di strutture sintattiche, risorse lessicali, espressioni tratte dal parlato emotivo e una certa intonazione.

Il tipo esclamativo è sempre stato considerato un atto minore dalla comunità scientifica, soprattutto in ambito sintattico. Nella maggior parte delle grammatiche tradizionali, infatti, si distingue fra categorie maggiori, a cui appartengono le frasi dichiarativa e interrogativa, e categorie minori, in cui si annoverano, fra le altre, le frasi esclamative. Generalmente, la frase esclamativa è definita un enunciato espressivo che ricalca in toto le caratteristiche sintattiche dei tipi frasali maggiori assumendo, così, un ruolo secondario che non le permette di aspirare a costituire una tipologia frasale indipendente (Sadock, Zwicky, 1985; König, Siemund, 2007). Con

riferimento all'italiano, alcune grammatiche dedicano poco interesse alla modalità esclamativa (tra le altre, Dardano, Trifone, 1995; Serianni, 1997); diversamente, nel terzo volume della *Grande Grammatica di Consultazione* (1995) la frase esclamativa è analizzata dal punto di vista sintattico, semantico e intonativo in un lungo paragrafo a cura di Paola Benincà.

L'interesse per la frase esclamativa è cresciuto soltanto negli ultimi decenni e ha avuto come oggetto di indagine soprattutto la dimensione sintattica e semantica, mentre le osservazioni di carattere prosodico sono ancora poche, nonostante la prosodia rappresenti, in molti contesti, l'unico mezzo a cui ricorrono i parlanti per distinguere l'enunciato esclamativo dagli altri tipi frasali. Ad oggi, anche gli studi specificamente dedicati alla frase esclamativa dell'italiano sono in crescita, tuttavia alcune questioni sono ancora dibattute: la definizione dello statuto dell'esclamativa e dei criteri che ne permettono l'identificazione, nonché il rapporto esistente con le frasi interrogative (Soriano, 2010).

Il presente studio è volto ad approfondire l'osservazione delle caratteristiche prosodiche della frase esclamativa *wh* nell'italiano di Bari. L'obiettivo è duplice: in primis, si intende individuare gli indici fonetici che concorrono a distinguere la frase esclamativa *wh* dall'interrogativa *wh*; in secondo luogo, ci si propone di osservare la misura in cui la connotazione affettiva (positiva o negativa) convogliata dalle frasi esclamative impatta sul carattere gradiente dell'enunciato e quindi sul suo contorno intonativo.

2. *Le strutture esclamative*¹

Nella maggior parte delle lingue si distingue fra esclamative *wh* ed esclamative non *wh*². Le prime sono introdotte da quegli elementi *wh* che veicolano un contenuto graduabile di qualità o quantità (3. *Che bel vestito!* 4. *Quanto sei alto!*). In queste strutture l'elemento introduttore è posizionato obbligatoriamente a inizio enunciato: si tratta di un processo definito *fronting*³. Gli elementi esclamativi italiani più diffusi sono *che*, *come*, *quanto*, *quale*. *Che* è l'introduttore che più spesso identifica le esclamative italiane, esso può modificare un aggettivo (5. *Che buon pranzetto mi hai preparato!*), un nome (6. *Che risposta ti ha dato!*) o un intero sintagma nominale (7. *Che casa grande ha!*), può avere funzione pronominale se seguito da verbo flessso (8. *Che ha combinato!*) e modificare un sintagma avverbiale in funzione predicativa (9.

¹ Gli esempi riportati in questo paragrafo sono tratti da Benincà (1995).

² Relativamente all'italiano, Benincà (1995) distingue fra esclamative totali ed esclamative parziali: a seconda che il senso di sorpresa si estenda all'intero enunciato o a parte di esso. Nella prima categoria rientrano le frasi esclamative con struttura assertiva e una serie di altri enunciati meno comuni; alla seconda categoria appartengono le esclamative con introduttore *wh* e altre espressioni in cui il sintagma focalizzato non è un elemento *wh*. Per una disamina delle strutture esclamative italiane più completa di quella proposta nel presente studio si rimanda, quindi, a Benincà (1995).

³ Termine con cui, in grammatica inglese, si intende la dislocazione a inizio frase di un gruppo di parole che generalmente occupa una posizione interna (Huddleston, Pullum, 2002)

Che strano che è tutto ciò!). *Come* è il secondo introduttore esclamativo più diffuso, ha sempre funzione avverbiale e non può essere seguito da un costituente frasale ma esclusivamente da un verbo flessivo (10. *Come sei diventato grande!*). *Quanto* può avere funzione di aggettivo (11. *Quanti libri ha scritto!*), di aggettivo sostantivato (12. *Quanti ha invitato!*), di avverbio (13. *Quanto corre!*) o di aggettivo quantificativo sostantivato (14. *Quanto ha mangiato!*). *Quale*, che appartiene a un registro stilistico alto, ha sempre funzione avverbiale (15. *Quali amici frequentava!*). Le esclamative wh possono subire l'elisione del verbo, in questo caso si parla di esclamative ellittiche o nominali (16. *Che bello!*), che spesso possono perdere l'introduttore *che* (17. *Bello!*). Nelle esclamative wh la negazione è contemplata tramite *non* espletivo (18. *Quanto intelligente non è!* 19. *Quali capricci non ha fatto!*).

Le esclamative non wh più diffuse sono quelle che ricalcano in toto la struttura assertiva (20. *È arrivato Mario!*); tale identificazione sintattica suggerisce che, in questi casi, la disambiguazione è totalmente affidata all'intonazione. Il senso di sorpresa può essere veicolato anche da un solo sintagma o da una singola parola, in occasione di un ricordo improvviso o del riconoscimento di qualcosa/qualcuno (21. *Il gas!* 22. *La mia vecchia casa!*).

3. Aspetti semantico-pragmatici

Come è stato accennato, la variabilità che caratterizza la frase esclamativa trova una soluzione nella sua funzione comunicativa, che si basa su due aspetti fondamentali: la salienza e il significato graduabile. Tutte le frasi esclamative comunicano un contenuto saliente, in quanto ritenuto sorprendente e fuori norma dal parlante; il senso di sorpresa può originarsi tanto dalle aspettative personali del locutore quanto da quelle legate al contesto⁴ (Gutiérrez-Rexach, 2008). Il concetto di significato graduabile riguarda esclusivamente l'esclamativa wh (considerata, per questo, la costruzione prototipica): nelle esclamative wh l'elemento che veicola il senso di sorpresa è graduabile su una scala di giudizio in cui si posiziona sempre sul gradino più elevato, per questo motivo il significato esclamatorio è straordinario ed eccede la norma (Zanuttini, Portner, 2003; Villalba, 2003; Gutiérrez-Rexach, 2008; Rett, 2008)⁵.

Alcuni aspetti pragmatici contribuiscono a distinguere le esclamative dalle interrogative. In primo luogo, le esclamative sono intrinsecamente fattive, nel senso che il loro contenuto è presupposto⁶; di conseguenza, possono essere subordinate soltanto

⁴ Michaelis (2001) sostiene, tuttavia, che il senso di sorpresa associato alla frase esclamativa non risieda semplicemente nella reazione a un evento o un fatto inaspettato: ad esempio, potrei non aver previsto di incontrare un collega in corridoio, ma ciò potrebbe non sorprendermi. Secondo l'autrice il concetto di sorpresa implica che il parlante giudichi una situazione come non canonica; per quest'ultima si intende una situazione diversa da quella solitamente predetta dallo speaker in base a una serie di ipotesi, che consistono in stereotipi, norme o modelli sociali.

⁵ Ad esempio, la frase esclamativa *How expensive this wine is!* implica le espressioni *This wine is extremely expensive* e *This wine is expensive to an extreme degree* (Villalba, 2008).

⁶ Concetto esposto per la prima volta da Elliot (1971, 1974) e ripreso e rielaborato da Grimshaw (1979).

a predicati fattivi (23. *Mary knows/*thinks/*wonders how very cute he is!*)⁷. Le esclamative sono, quindi, incompatibili con uno stato di ignoranza del locutore (Elliott, 1971, 1974), diversamente dalle frasi interrogative sincere che per definizione presuppongono ignoranza. In secondo luogo, si assume che le esclamative presentino un'implicatura scalare convenzionale (Zanuttini, Portner, 2003): il valore estremo espresso dall'esclamativa non è esplicito ma viene inferito e interpretato dall'ascoltatore⁸. Gli enunciati esclamativi, infine, non possono introdurre una domanda sincera né fungere da risposta a una domanda⁹ (Zanuttini, Portner, 2000, 2003).

4. *La sorpresa: un'emozione base*

Come è stato più volte ribadito, l'esclamativa esprime sorpresa. Il concetto di senso di sorpresa, positivo o negativo, è solitamente associato all'idea di cogliere qualcuno impreparato, attraverso un'azione inaspettata, che desta stupore. La sorpresa è annoverata fra le emozioni fondamentali¹⁰ insieme a gioia, tristezza, paura, disgusto, sorpresa e rabbia, selezionate in base ai dati relativi al riconoscimento universale delle emozioni per mezzo delle espressioni facciali e vocali (tra gli altri, Tomkins, 1962, 1963; Izard, 1994; Ekman, 2003). Le emozioni possono, quindi, essere espresse sia attraverso mezzi espressivi corporali (espressioni facciali, gestualità, movimenti del corpo) che vocali: alcune emozioni prediligono un mezzo anziché un altro a causa del diverso modo di reagire agli stimoli esterni del nostro organismo; tali reazioni si consolidano a livello culturale e sociale, portando a una cristallizzazione delle associazioni fra emozioni e risposte fisiologiche¹¹. Ogni emozione restituisce una realizzazione acustica unica¹², tuttavia è possibile scorgere alcuni pattern comuni. A seconda del livello di attivazione delle alterazioni fisiologiche¹³ che stimolano, si distingue fra emozioni ad alta attivazione ed emozioni a bassa attivazione¹⁴: le prime

⁷ Esempio tratto da Zanuttini, Portner (2003).

⁸ Al fine di comprendere gli aspetti solitamente associati all'esclamazione, come il senso di sorpresa, l'inaspettato e il grado estremo, Zanuttini, Portner (2003) elaborano il concetto di *widening*, che indica un processo di ampliamento semantico.

⁹ Tuttavia, in alcune lingue è possibile (cfr. Zanuttini, Portner, 2003).

¹⁰ A seconda dell'approccio adottato, il numero delle emozioni ritenute fondamentali può, tuttavia, variare.

¹¹ Vi è evidenza neuroscientifica che la comunicazione delle emozioni rappresenti un fenomeno multidimensionale e altamente complesso poiché fa uso di diversi canali di espressione. Di grande interesse i lavori incentrati sui correlati neurali delle espressioni emotive verbali e non verbali (per una rassegna di tali studi cfr. Kotz, Paulmann, 2011).

¹² "Different types of emotion are actually characterized by unique patterns or configuration of acoustic cues" (Scherer, 2003).

¹³ Le reazioni fisiologiche coinvolgono la respirazione, la fonazione e l'articolazione (per una descrizione accurata cfr. Scherer, 1986).

¹⁴ Tuttavia, studi più recenti propendono per un approccio multidimensionale che individua tre dimensioni: attivazione, valutazione e potere. La dimensione di attivazione è associata al grado di importanza dello scopo. La dimensione di valutazione riguarda il raggiungimento dello scopo o la compromissione dello stesso: emozioni con connotazione positiva si oppongono a emozioni con connotazione negativa

(paura, gioia, sorpresa e rabbia) sono generalmente caratterizzate da alti valori di F0, un'estensione tonale ampia, intensità alta, una maggiore velocità di eloquio e quindi una riduzione del numero di pause; diversamente, le seconde (tristezza e disgusto) sono veicolate da bassi valori di F0, un range tonale compresso, intensità bassa, un eloquio più lento e lunghe pause. Pur tuttavia, è stato messo in luce che disgusto e sorpresa sono difficilmente identificate tramite la voce¹⁵: probabilmente, nel corso dell'evoluzione la difficoltà di comunicare emozioni di questo tipo ha stimolato un tipo di espressione per lo più connessa alla mimica e meno al coinvolgimento delle risorse vocali (Jonhstone, Scherer, 2000). Fino ai primi anni 2000, nella maggior parte degli studi dedicati ai correlati acustici delle emozioni¹⁶ la sorpresa è, per questo, poco indagata o addirittura esclusa. Tuttavia, in alcune indagini più recenti, la sorpresa compare fra le emozioni analizzate a livello di produzione e percezione. In Maffia *et al.* (2014), l'analisi acustica ha confermato che, in quanto emozione ad alta attivazione, la sorpresa innesca un innalzamento dei valori di F0 e un'estensione del range tonale. Lo studio di De Marco, Paone (2014), che è incentrato sulla codifica e la decodifica delle sei emozioni primarie in apprendenti italiano L2, ha confermato che la sorpresa presenta le stesse caratteristiche prosodiche delle altre emozioni ad alta attivazione; inoltre, lo studio ha mostrato che la sorpresa è identificata correttamente nella maggior parte dei casi (si rilevano, però, difficoltà più o meno significative a seconda della lingua nativa degli uditori).

Alla luce degli studi effettuati sorge, tuttavia, un dubbio: la sorpresa, rispondendo a stimoli di natura diversa, può avere una connotazione positiva o negativa; a seconda che l'evento sia considerato vantaggioso o dannoso per il parlante, la sorpresa può essere associata a caratteristiche psicologiche distinte ed essere espressa tramite mezzi diversi in relazione all'evento scatenante e al modo in cui il soggetto lo valuta. Ci si chiede, dunque, se a connotazione affettiva differente corrisponda una diversa realizzazione acustica: ciò potrebbe creare ulteriore confusione nella definizione dei pattern acustici che veicolano questa emozione.

5. *L'intonazione dell'esclamativa*

Nonostante l'intonazione svolga un ruolo di fondamentale importanza ai fini tanto dell'identificazione della frase esclamativa quanto della corretta interpretazione del senso di sorpresa e dell'emozione da essa convogliati, fino alle soglie del 2000, la quasi

(la collera, ad esempio, sarà caratterizzata da un aumento dell'intensità e da un rallentamento della velocità di eloquio). Infine, la dimensione della potenza riguarda l'idea di sottomissione o dominanza derivante dal raggiungimento di uno scopo, quindi quelle emozioni associate alla dominanza presentano un'intensità maggiore rispetto alle emozioni connesse alla sottomissione (per una rassegna degli studi dedicati all'argomento cfr. Scherer, 2003).

¹⁵ In Juslin, Laukka (2003) e in Scherer (2003) queste due emozioni hanno percentuali di riconoscimento molto basse.

¹⁶ Per un'analisi puntuale dei correlati acustici delle emozioni base cfr., fra gli altri, Juslin, Laukka (2003).

totalità degli studi riservati all'esclamativa ha fornito per lo più un'indagine di stampo puramente sintattico e semantico.

Nei primi studi di ambito internazionale in cui si nomina l'intonazione esclamativa, Delattre (1966) e O'Connor, Arnold (1973), con riferimento rispettivamente al francese e all'inglese britannico, delineano un contorno intonativo esclamativo caratterizzato da un picco melodico iniziale alto e una discesa finale.

Studi più recenti, supportati da un'analisi di tipo acustico e fonologico, hanno confermato questa tendenza. L'esclamativa wh ungherese presenta dei pitch accent ascendenti, un tono di confine iniziale tendenzialmente medio %M o basso %L e un tono di confine finale medio M%. Inoltre, si rilevano una frequenza fondamentale massima più bassa e livelli di Onset inferiori rispetto all'interrogativa (Gyuris *et al.*, 2013). In uno studio teso all'individuazione dei correlati acustici della miratività¹⁷ (Rett, Sturman, in stampa), nelle esclamative inglesi si individuano tre principali aspetti: pitch accents ascendenti di tipo L+H*, extra-high targets e accenti di sintagma aggiuntivi rispetto a quanto riscontrato nei filler.

In riferimento al portoghese brasiliano, Moraes (2008) osserva che esclamative e interrogative wh sono accumulate da un tono iniziale alto seguito da una discesa continua che segue tutto l'enunciato. Tuttavia, le due modalità differiscono a livello delle sillabe finali pretonica e tonica, che sono rispettivamente più bassa e più alta nelle esclamative; infine, le frasi esclamative risultano nel complesso più lunghe delle frasi interrogative. Una discesa finale si rileva anche nelle esclamative spagnole, che presentano toni ascendenti L+H* allineati alle sillabe prenucleare e nucleare e un tono finale basso L% (Estebas Viloplana, Prieto, 2010). Il contributo di del Mar Varnell *et al.* (2015) dedicato a due varietà di sardo (logudorese e il campidanese) conferma la tendenza generale, rivelando che, generalmente, le frasi esclamative presentano una serie di toni ascendenti L+H* posti in corrispondenza delle *content words*, seguiti da un tono di confine discendente. Una discesa finale è riscontrata anche nelle esclamative francesi, in cui si osserva una salita iniziale Hi e un tono nucleare del tipo (L) H* L%. (Delais-Roussaire *et al.*, 2015). Nel caso del romeno, in Jitcă *et al.* (2015) si osserva che la posizione dell'accento nucleare, allineato all'elemento wh o al costituente che lo segue, è indice dell'entusiasmo con cui può essere prodotto uno stesso enunciato esclamativo. Nel friulano, osservato da Roseano *et al.* (2015), le esclamative wh sono caratterizzate da un contorno prenucleare ascendente L+H* seguito da un tono nucleare del tipo L+H* L%. Infine, nella maggior parte delle esclamative occitane si rilevano un accento iniziale Hi, un pitch accent prenucleare di tipo ascendente L+H* e un andamento nucleare ascendente-discendente¹⁸ (Sichel-Bazin *et al.*, 2015).

Per quanto riguarda l'italoromanzo, i primi riferimenti all'intonazione della frase esclamativa confermano la presenza di un contorno finale discendente e di un picco iniziale particolarmente alto negli enunciati molto enfatici (D'Eugenio, 1976). Altre informazioni sono deducibili dai contributi di taglio sintattico. In particolare, in Benincà

¹⁷ Con il termine 'mirativity' si intende la codifica implicita del senso di sorpresa e del superamento delle aspettative del parlante (DeLancey, 1997). Tale processo sfrutta differenti strategie di tipo morfologico, sintattico e prosodico.

¹⁸ Tuttavia, si rilevano diversi tipi di allineamento (cfr. Sichel-Bazin *et al.*, 2015).

(1995) si sostiene che nelle esclamative non wh il contorno melodico sia interamente alto e il picco più elevato corrisponda all'ultimo costituente della frase. Diversamente, nella esclamativa wh il pitch accent è generalmente allineato all'elemento introduttore. In uno studio di taglio fonetico, Canepari (2003) pone a confronto frasi assertive e frasi esclamative non wh, osservando un tono iniziale più elevato nelle seconde e un comune contorno finale discendente. Alcuni studi di orientamento fonologico si sono soffermati brevemente sull'enunciato esclamativo italiano, in particolar modo sulla varietà toscana. In Avesani (1995), si analizza l'esclamativa nominale *Massimiliano!* registrando un tono iniziale particolarmente alto (%H) che si protrae lungo tutte le sillabe atone del nome, a cui segue un contorno nucleare alto (H*) e una discesa finale (L-L%). Anche in Grice *et al.* (2005), nelle esclamative fiorentine si rileva un tono iniziale alto.

Negli ultimi dieci anni, l'intonazione della frase esclamativa italiana è stata oggetto di alcune ricerche basate su test acustici e percettivi. In Soriano (2010), è stata osservata la prosodia delle frasi esclamative wh e non wh nell'italiano di Bari, poste a confronto con la loro controparte interrogativa e assertiva. L'analisi acustica ha dimostrato che le frasi esclamative non wh sono più brevi e più intense delle assertive e presentano un attacco frequenziale più alto; i valori di Offset sono invece comparabili. Il confronto fra esclamative e interrogative wh risulta, invece, più complesso: le prime sono più lunghe delle seconde nel 64% dei casi, nel restante 36% i risultati sono approssimabili. A livello di intensità media le due modalità sono comparabili. Le esclamative presentano, inoltre, sempre un contorno finale discendente, mentre le interrogative mostrano un'ascesa finale; infine, l'Onset è comparabile nelle due modalità¹⁹. La seconda parte della ricerca, che ha previsto la manipolazione dei parametri acustici più rappresentativi dell'esclamativa (F0x e durata), ha mostrato che nel 95% dei casi gli uditori sono capaci di riconoscere la tipologia frasale. In un lavoro di poco successivo, l'autrice torna a osservare l'intonazione dell'esclamativa non wh nell'italiano di Bari (Soriano, 2012), in cui si dimostra che quest'ultima presenta valori di F0x, Onset e pitch range più alti rispetto all'assertiva. Inoltre, le due modalità sono accumulate da una discesa finale ma, mentre le assertive presentano un contorno nucleare H+L*, le esclamative mostrano un tono alto H*. Sulla base di questi dati, è stato svolto un perception gating test che ha rivelato che la parte iniziale degli enunciati costituisce un importante indice di disambiguazione dei due tipi frasali. Con riferimento al cosentino, in Soriano (2011) si analizzano esclamative non wh, wh e nominali. Lo studio ha dimostrato che tutte le esclamative presentano un contorno finale discendente. Rispetto alle altre strutture, le esclamative wh mostrano un innalzamento del contorno iniziale rispetto alla F0x, dovuto alla presenza dell'elemento introduttore; tuttavia, il pitch range è più ampio nelle esclamative non wh. Queste ultime rivelano, inoltre, valori di intensità media più alti delle altre costruzioni esclamative. Infine, esclamative wh e nominali sono più lente rispetto alle assertive; diversamente, le esclamative non wh presentano valori simili alla loro controparte assertiva. Tramite ana-

¹⁹ Questo dato non conferma l'assunto secondo cui le esclamative wh sarebbero accumulate da un contorno melodico iniziale alto: in Soriano (2010) si dice che numerosi sono gli enunciati che presentano un contorno interamente alto e una discesa frequenziale finale che vede il suo inizio in corrispondenza della vocale nucleare.

lisi fonologica, Sorianello (2011) conferma nelle esclamative *wh* la presenza di un incipit alto seguito da un contorno finale discendente. Le esclamative non *wh* sono accumulate da un andamento discendente che interessa solo la parte finale della frase che mostra per buona parte un plateau melodico posizionato su livelli frequenziali alti che produce un diffuso senso di enfasi. Sempre con riferimento al cosentino, in Kellert *et al.* (2015) si rileva che l'accento prenucleare H* allineato con l'elemento introduttore è più alto in frequenza nelle domande che nelle esclamative; diversamente il secondo tono H, che subisce il fenomeno del downstep, è riscontrato solo nelle esclamative. Inoltre, si attesta un contorno finale ascendente nelle interrogative e discendente nelle esclamative. Infine, le frasi esclamative sono più lunghe rispetto alle interrogative. Gili Fivela *et al.* (2015) incentrano il loro contributo sull'intonazione di alcuni italiani regionali. Il quadro delineato dagli autori è abbastanza complesso. In alcune varietà come quella milanese, lucchese, romana, leccese e torinese la frase esclamativa è realizzata con un tono leggermente enfatico che corrisponde a un contorno nucleare di tipo L*+> H L%; quest'opzione solitamente si alterna ad altri contorni²⁰. Un range tonale ampio caratterizza la maggior parte delle varietà²¹.

Le ricerche finora condotte restituiscono un quadro di difficile interpretazione: la maggior parte degli studi identificano nell'esclamativa un picco iniziale alto e una discesa finale ma, come si è potuto notare, le tendenze cambiano a seconda non solo delle varietà indagate ma anche delle strutture esclamative analizzate e dalle metodologie impiegate. Inoltre, mentre la letteratura incentrata sull'analisi fonologica dell'enunciato esclamativo è in crescita, ancora scarsi sono i dati relativi agli indici fonetici più rappresentativi di questa modalità. Con specifico riferimento alla varietà barese, finora solo due ricerche si sono occupate della frase esclamativa, di cui soltanto una ha indagato gli aspetti prosodici e intonativi della struttura con introduttore *wh*.

6. La ricerca

Alla luce degli studi finora incentrati sull'intonazione della frase esclamativa, la presente ricerca è volta ad approfondire l'analisi delle caratteristiche prosodiche dell'esclamativa *wh* nell'italiano di Bari. In particolare, ci si chiede:

1. A parità di struttura sintattica, la modalità frasale influenza gli indici fonetici dell'enunciato?
2. In che misura la connotazione emotiva (positiva o negativa) convogliata dalle frasi esclamative *wh* impatta sul carattere gradiente dell'enunciato e quindi sul suo contorno intonativo?

²⁰ In particolare, Torino, Firenze e Siena presentano un pattern nucleare L+H* L%. Diversamente a Pisa, Lucca, Roma, Salerno, Lecce e Pescara si attesta la presenza di un tono nucleare di tipo H*+L L%. Infine, a Cosenza si rileva il tono H* L% (Gili Fivela *et al.*, 2015).

²¹ Tuttavia, rimane ancora in dubbio che tale caratteristica sia giustificata dalla presenza di un tono iniziale (%H) così come riscontrato da Avesani (1995) e Grice *et al.* (2005) per il fiorentino, nonché da Sorianello (2010, 2011, 2012) per il barese e il cosentino.

La frase esclamativa wh è stata selezionata in quanto considerata la struttura esclamativa prototipica, nonché la più diffusa nell'italiano parlato. La scelta di confrontare la frase esclamativa con la sua controparte interrogativa deriva da un lato dalla somiglianza sintattica e morfologica fra le due modalità in numerosi contesti e, dall'altro, dalla estrema differenza fra i due tipi frasali a livello di forza illocutiva. Per quanto concerne la seconda domanda di ricerca, si intende verificare se la natura della sorpresa – positiva o negativa – influenzi in qualche modo gli aspetti prosodici dell'enunciato, incidendo sull'implicatura scalare della frase esclamativa wh.

I materiali impiegati nella ricerca constano di 20 sceneggiature preconfezionate, ognuna delle quali contenente uno specifico contesto e una esclamativa wh. Gli introduttori wh impiegati (*che, quanto, come*) sono stati selezionati in base alla frequenza d'uso che hanno nell'italiano parlato. Le frasi target si differenziano per connotazione affettiva, che può essere positiva o negativa: pur nella condivisione della trasmissione di uno stato di sorpresa, alcuni contesti descrivono un fatto triste o spiacevole, mentre altri un evento gradito; ogni esclamativa esprime, quindi, una reazione empatica diversa a seconda della situazione esposta. Al fine di individuare i tratti intonativi distintivi, è stato predisposto un campione di controllo composto da 20 enunciati interrogativi sintatticamente e morfologicamente identici alle frasi esclamative inserite nel corpus. Di seguito si riportano due esempi di sceneggiature, una contenente un'esclamativa con connotazione emotiva positiva (a) e una con connotazione emotiva negativa (b):

- (a) Rivedi dopo tanto tempo il figlio di una tua cara amica. È cresciuto tantissimo.
Esclami:
“Quanto sei alto!”
- (b) I tuoi amici ti raccontano la bella serata che hanno trascorso il giorno prima.
Tu purtroppo non sei potuto uscire perché avevi la febbre. Esclami:
“Che cosa mi sono perso!”

I soggetti che hanno preso parte al test sono nati e vissuti a Bari. Gli informatori, 5 uomini e 5 donne fra i 25 e i 35 anni, sono studenti universitari o laureati di pari provenienza ed estrazione socioculturale. Ogni informatore è stato coinvolto singolarmente in una sessione di registrazione di un'ora. Durante l'incontro il parlante è stato invitato a leggere le sceneggiature proposte, prima in silenzio e poi ad alta voce, nel modo più naturale possibile: al fine di evitare di influenzare la lettura delle frasi target, non è stata fornita alcuna informazione relativa alla finalità della ricerca. L'audioregistrazione dei materiali è avvenuta in un ambiente insonorizzato tramite un Tascam DR-40 (acquisizione formato WAV, con frequenza di campionamento pari a 44Khz e risoluzione a 24 bit). Al fine di assicurare un livello di intensità corretto e costante, a ciascun informatore è stato chiesto di posizionarsi a una distanza di circa 20 cm dal microfono. Il test ha prodotto 400 enunciati bersaglio, di cui 200 esclamative wh e 200 interrogative wh. Gli stimoli rientrano nella categoria di parlato semi-recitato. I 400 stimoli ottenuti sono stati sottoposti ad analisi acustica, considerando i parametri più usati dagli studi dedicati all'argomento: valore medio ($F0x$), minimo ($F0min$) e massimo ($F0max$) della frequenza fondamentale (Hz); escursione melodica convertita in semitoni (ST); valore frequenziale

dell'Onset e dell'Offset dell'enunciato²² (Hz); valore medio dell'intensità (dB); durata totale dell'intero enunciato (ms); durata dell'ultima vocale tonica (ms); velocità di eloquio (sill/s). La significatività statistica, impostata per $p < 0.05$, è stata calcolata tramite Paired-Sample T Test e One-Way ANOVA.

7. I risultati

Gli esiti acustici mostrano che le esclamative *wh* differiscono dalla controparte interrogativa in riferimento alla maggior parte dei parametri considerati. In secondo luogo, è stato rilevato che le esclamative con connotazione positiva mostrano valori frequenziali e di intensità maggiori rispetto a quelle con connotazione negativa; tuttavia, quest'ultima tendenza si registra solo nei soggetti femminili. I diversi parametri saranno prima discussi in termini generali (§ 7.1) al fine di cogliere le tendenze comuni, poi commentati con riferimento alla dimensione emotiva prodotta (§ 7.2) in modo da poter individuare eventuali aspetti divergenti imputabili alla trasmissione della sorpresa²³.

7.1 Esclamative *wh* vs Interrogative *wh*

7.1.1 Frequenza fondamentale

Per quanto concerne il parametro della F_0 , i risultati hanno rivelato che le frasi interrogative presentano valori più alti rispetto alle esclamative. Tuttavia, si rileva che, mentre nelle donne interrogative ed esclamative differiscono in modo significativo per tutti i parametri di F_0 , negli uomini il quadro è meno definito.

Tabella 1 - Valori medi di F_0x , F_0min , F_0max , Onset, Offset (Hz) e pitch range (ST) relativi alle frasi esclamative *wh* e alle interrogative *wh* corrispondenti separati per genere. Tra parentesi la deviazione standard

	Femmine		Maschi	
	Esclamative	Interrogative	Esclamative	Interrogative
F0x	210 (34.8)	248.2 (30)	146 (32.7)	156.9 (28.6)
F0min	156.8 (16.5)	159.3 (17.9)	96.3 (18.8)	88.2 (17.9)
F0max	295.6 (63.4)	383.1 (68.2)	210.7 (55.3)	238.1 (47.7)
PR	10.7 (3.7)	15.4 (4)	13.2 (4.8)	17 (4.1)
Onset	265.8 (61.2)	308.8 (67.6)	184.2 (49)	195.7(43)
Offset	166.9 (20.4)	222 (38.6)	103.4 (24.2)	113.9 (36.5)

²² Per estrarre i valori di Onset e Offset sono stati misurati rispettivamente i livelli frequenziali medi dei primi e degli ultimi 30 ms del contorno intonativo.

²³ Tra gli interessi primari del presente studio non rientra quello di verificare la misura in cui il genere influenza gli indici acustici dell'atto esclamativo. Questo è il motivo per cui il genere non rappresenta una variabile indipendente nel nostro *research design*. Tuttavia, ci è apparso doveroso mettere in luce la diversità fra locutori e locutrici riscontrata in relazione al comportamento di alcuni parametri acustici.

Per quanto concerne la F0x, nelle donne la differenza media fra le due modalità è di -38.27 Hz²⁴ [-3 ST ($t=-8.911$; $df=99$; $p<.001$)], mentre negli uomini è di -10.9 Hz [-1.3 ST ($t=-2.947$; $df=99$; $p<.01$)]. Diversamente, i valori minimi di F0 sono comparabili per le donne [-2.4 Hz, -0.3 ST ($t=-1.304$; $df=99$; $t>.05$)], mentre per gli uomini si rileva una sottile discrepanza [+8 Hz, +1.5 ST ($t=3.201$; $df=99$; $p<.01$)]. La tendenza relativa ai valori massimi di F0 è più definita: le esclamative presentano valori medi significativamente inferiori rispetto alle interrogative sia nelle donne che negli uomini [rispettivamente -87.5 Hz, -4.6 ST ($t=9.312$; $df=99$; $p<.001$); -27.4 Hz, -2.3 ST ($t=-4.751$; $df=99$; $t<.001$)²⁵. Di conseguenza, nelle domande il range tonale è più esteso: la differenza media tra le due modalità è di -4.3 ST nelle donne ($t=-8.470$; $df=99$; $p<.001$) e di -3.8 ST negli uomini ($t=-6.954$; $df=99$; $p<.001$).

Al fine di approfondire il ruolo del contorno iniziale e finale nelle esclamative, i valori di Onset e Offset sono stati sia verificati in riferimento al tipo frasale sia messi a confronto con i valori medi di F0. Le differenze sono significative, sia per le donne ($F=116.521$; $df=5$; $p<.001$) che per gli uomini ($F=101.453$; $df=5$; $p<.001$). Il test post-hoc di Tukey raggiunge la soglia di significatività ($p<.001$) ma per gli uomini le differenze fra esclamative e interrogative non sono significative in riferimento né ai valori di Onset né di Offset ($p>.05$). Si analizzino dapprima i dati relativi al confronto fra i due tipi frasali: nelle donne il contorno melodico iniziale dell'esclamativa è più basso rispetto a quello dell'interrogativa di 43 Hz (2.5 ST); come accennato, per gli uomini questa differenza non è significativa (-11.5 Hz, -1.2 ST). Relativamente all'Offset, i risultati mostrano che nelle donne le interrogative presentano valori significativamente più alti rispetto alle esclamative di 55 Hz (4.7 ST); anche in questo caso, per gli uomini la differenza è poco rilevante (-10.5 Hz, -1.4 ST). Nella tabella 2 sono riportati le differenze medie in ST fra F0x e rispettivamente i valori iniziali e finali del contorno intonativo.

Tabella 2 - Valori medi in ST relativi alla differenza fra F0x e rispettivamente Onset e Offset separati per genere

	Femmine		Maschi	
	Onset/F0x	Offset/F0x	Onset/F0x	Offset/F0x
Esclamative	+ 3.9	- 3.9	+ 3.8	- 5.9
Interrogative	+ 3.5	- 2.1	+ 3.7	- 5.9

Come si può notare, in entrambe le modalità il contorno iniziale di F0 è interessato da un innalzamento rispetto al valore medio di F0x. Diversamente, per quanto

²⁴ Nei nostri risultati la differenza fra le due modalità frasali è espressa sempre in base al rapporto esclamative wh/interrogative wh.

²⁵ Con F0max ci si riferisce al valore massimo di F0 dell'intero enunciato. Essendo di tipo wh, in ambedue le tipologie frasali indagate il livello frequenziale massimo è generalmente raggiunto all'inizio del contorno: la presenza dell'elemento wh favorisce un attacco prominente.

concerne il punto finale del contorno, i valori medi di Offset sono sempre inferiori rispetto ai valori di F0x.

Alla luce dei risultati raccolti, è possibile delineare alcune tendenze interessanti. I dati relativi all'Onset suggeriscono che le due modalità indagate condividono la presenza di un contorno iniziale elevato²⁶: i dati riportati nella tabella 2 confermano, quindi, l'assunto secondo cui la frase esclamativa sarebbe generalmente caratterizzata da un picco iniziale particolarmente alto (cfr. § 5), ma allo stesso tempo suggeriscono che tale comportamento prosodico è ugualmente messo in atto dalle frasi interrogative. Per quanto concerne i valori di Offset, il dato relativo alle esclamative non era disatteso, difatti nella maggior parte delle varietà italiane questi enunciati rivelano un contorno finale discendente (Soriano, 2010; Gili Fivela *et al.*, 2015). Le tendenze osservate vanno però valutate con cautela e necessitano di essere verificate tramite analisi fonologica²⁷.

A scopo esemplificativo, nelle figure seguenti sono riportate le versioni esclamativa (fig.1) e interrogativa (fig.2) della frase *Quale ha scelto* pronunciate dal parlante GR. Gli enunciati mostrano un contorno intonativo simile. In primo luogo, entrambi presentano un Onset superiore e un Offset inferiore alla F0x. Tuttavia, mentre l'interrogativa è caratterizzata da un tono periferico %H particolarmente alto, a cui segue una discesa che si conclude con un contorno nucleare H+L* L%, nell'esclamativa la parte iniziale del contorno è interessata da un sensibile innalzamento del livello frequenziale realizzato in corrispondenza dell'elemento wh; segue una morbida discesa, che si traduce nel medesimo andamento H+L* L%. Passiamo, quindi, a considerare le differenze prosodiche: l'interrogativa presenta valori frequenziali maggiori rispetto all'esclamativa; le differenze medie rilevate sono di 3.8 ST per la F0x e di circa 9 ST per l'estensione tonale così come per l'Onset.

²⁶ Nelle donne, tuttavia, la frase interrogativa presenta valori maggiori rispetto alla controparte esclamativa.

²⁷ La presente ricerca ha una finalità fonetica, l'analisi fonologica, che è in corso, sembra confermare la presenza di una diversa configurazione nucleare nelle due tipologie poste a confronto. In entrambe è stato rilevato l'innalzamento del contorno iniziale nella maggior parte delle occorrenze (circa il 75%). Nelle esclamative i pitch accents nucleari più diffusi sono H*+L e L*, entrambi seguiti da un tono di confine basso. Nell'interrogativa, in percentuale maggiore si rilevano i toni nucleari H+L* e L*; a seconda dei contesti, entrambi possono essere seguiti da un tono basso (L%) o alto (H%). Da un rapido controllo delle percentuali di occorrenza, si è dedotto che le possibilità di coincidenza dei contorni intonativi delle due modalità sono minime. Nei casi in cui esclamative e interrogative presentino profili intonativi simili, le differenze prosodiche intervengono nella disambiguazione degli enunciati.

Figura 1 - Forma d'onda, spettrogramma a banda larga, curva di F0, trascrizione ortografica e ToBI della frase esclamativa wh "Quale ha scelto!" pronunciata dal locutore GR

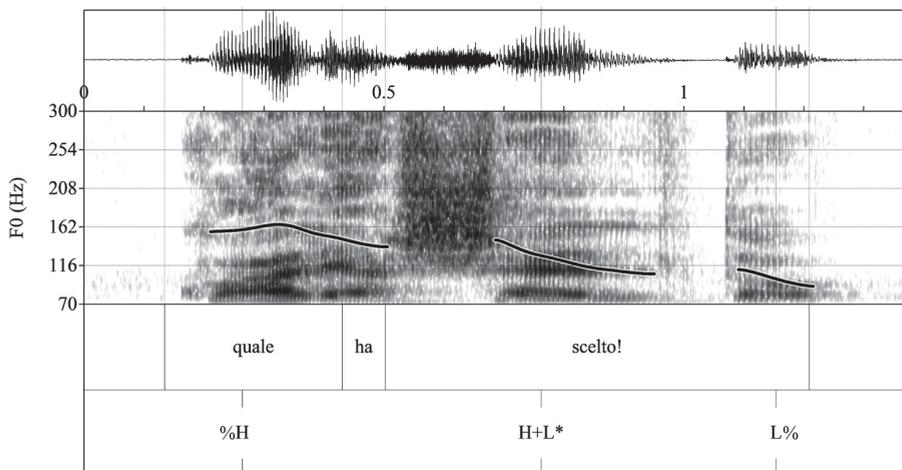
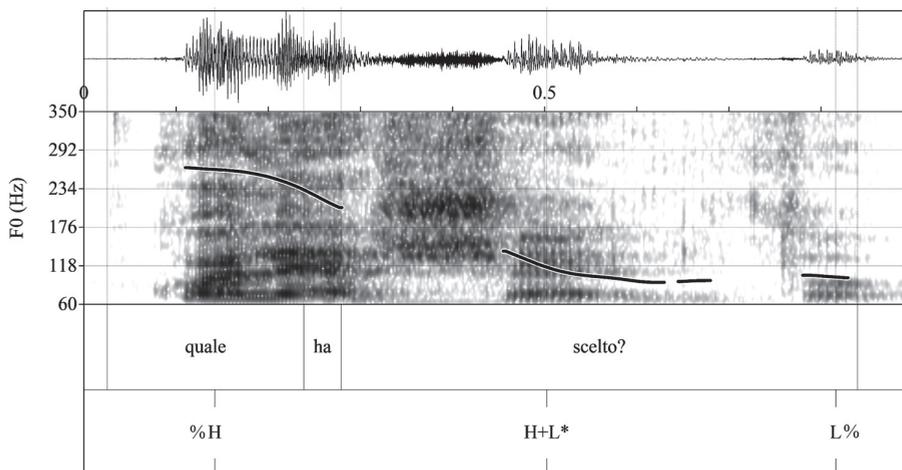


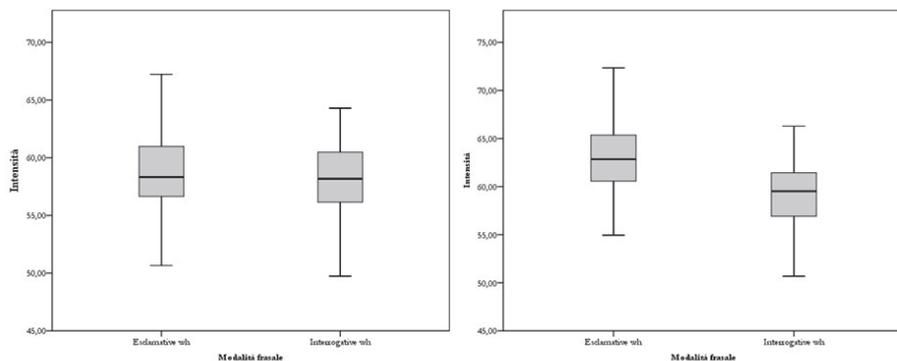
Figura 2 - Forma d'onda, spettrogramma a banda larga, curva di F0, trascrizione ortografica e ToBI della frase interrogativa wh "Quale ha scelto?" pronunciata dal locutore GR



7.1.2 Intensità

I risultati raccolti hanno suggerito che anche il parametro dell'intensità sembrerebbe subire l'influenza della modalità frasale: le esclamative con introduttore wh sono più intense della loro controparte interrogativa; tale tendenza si riscontra, tuttavia, principalmente nei locutori di sesso maschile. Come riportato nella fig. 3, mentre nelle donne la differenza fra le due modalità non è significativa [+0.5 dB ($\tau=1.299$; $df=99$; $p>.05$)], negli uomini la discrepanza fra le due modalità è più ampia [2.1 dB ($\tau=6.339$; $df=99$; $p<.001$)]. Questa tendenza non era disattesa: la frase esclamativa è un atto espressivo e, in quanto tale, il parlante tende a realizzarlo impiegando maggiore energia e, quindi, un volume più alto.

Figura 3 - *Intensità degli enunciati esclamativi e interrogativi realizzati da locutori femmine (sx) e maschi (dx)*



7.1.3 Durata e velocità di eloquio

Al fine di osservare le dinamiche temporali che interessano le frasi esclamative, sono state misurate sia la durata totale dell'intero enunciato che dell'ultima vocale tonica: la ricerca ha rivelato che le esclamative wh sono mediamente più lunghe delle interrogative. Nelle donne la differenza è di +232 ms ($t=13.02$; $df=99$; $p<.001$), negli uomini di +350 ms ($t=15.526$; $df=99$; $p<.001$). L'allungamento temporale sembra interessare principalmente l'ultima vocale tonica: nelle donne la differenza è di +32.6 ms ($t=7.165$; $df=99$; $p<.001$), negli uomini di ben +71.4 ms ($t=8.637$; $df=99$; $p<.001$). I parlanti sembrano allungare strategicamente la vocale nucleare, probabilmente al fine di rinforzare l'espressività convogliata dalla frase esclamativa e quindi enfatizzare il senso di sorpresa. Conseguentemente le esclamative sono più lente delle interrogative: nelle donne le prime sono pronunciate con una velocità di 5.1 sill/s (dev. st. 0.9), mentre le seconde presentano una media di 6.3 sill/s (dev.st. 0.86), ($t=-10.530$; $df=99$; $p<.001$); negli uomini la velocità elocutiva media nelle esclamative è di 4.8 sill/s (0.8), mentre nelle interrogative è di 6.6 sill/s (dev.st. 0.90), ($t=-12.600$; $df=499$; $p<.001$).

7.2 Esclamative positive vs Esclamative negative

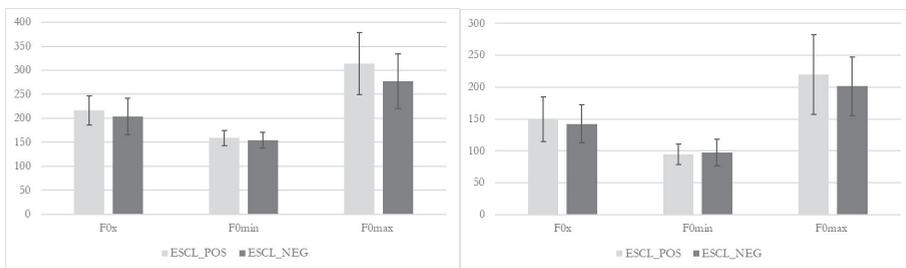
Per quanto riguarda la seconda domanda di ricerca, il nostro studio ha dimostrato che la connotazione affettiva influenza alcuni parametri prosodici dell'enunciato esclamativo; specificamente, gli indici fonetici coinvolti nella differenziazione delle esclamative con connotazione positiva e negativa (da qui in poi ESCL_POS ed ESCL_NEG) sono la F0max, l'estensione tonale, l'Onset di F0, l'intensità e la durata dell'ultima vocale tonica. Tuttavia, i nostri dati rivelano che le differenziazioni evidenziate interessano in misura diversa le femmine e i maschi.

7.2.1 Frequenza fondamentale

I dati relativi alla F0x rivelano che ESCL_POS ed ESCL_NEG sono comparabili in riferimento ai valori medi e minimi di F0, tuttavia si rinviene una discrepanza

statisticamente significativa a livello della F0max. Le tendenze frequenziali sono riportate nella fig. 4. In riferimento alla F0x si evidenzia una differenza non significativa fra le due modalità nelle donne [+13.3 Hz, +1.1 ST²⁸ ($t=1.851$; $df=49$; $p>0.5$)] e negli uomini [+7.7 Hz, +0.8 ST ($t=1.327$; $df=49$; $p>0.5$)]; i valori di F0min sono comparabili in entrambe le tipologie di affettività [+4.5 Hz, +0.5 ST per le donne ($t=1.847$; $df=49$; $p>0.5$), -2,5 Hz, -0.3 ST per gli uomini ($t=-.700$; $df=49$; $p>0.5$)]. Per quanto concerne la F0max, i risultati rivelano una discrepanza significativa fra le due attitudini a favore delle ESCL_POS per le donne [+37.1 Hz, +2.1 ST ($t=2.903$; $df=49$; $p<.05$)] ma più debole per gli uomini [+18.3, +1.2 ST ($t=2.141$; $df=49$; $p<.05$)]. Anche per quanto concerne l'estensione tonale le differenze fra ESCL_POS ed ESCL_NEG sono statisticamente significative: si rileva infatti una distanza media fra le due attitudini di +1.6 ST sia nelle donne ($t=2.424$; $df=49$; $p<.05$) che negli uomini ($t=2,141$; $df=49$; $p<.05$).

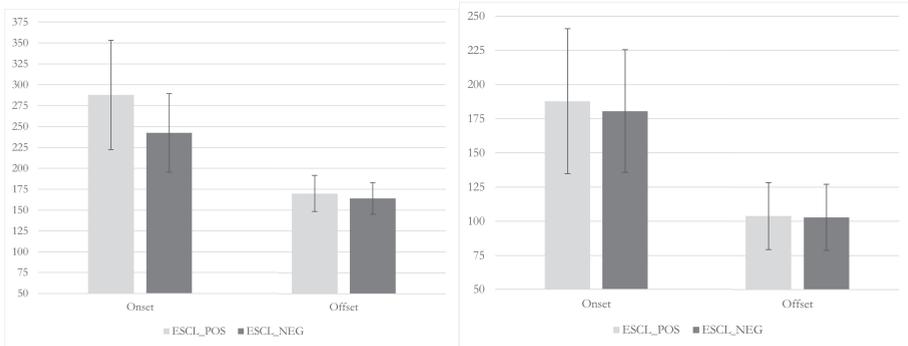
Figura 4 - Valori medi dei 3 parametri di F0 (in Hz) per femmine (sx) e maschi (dx) nelle ESCL_POS e nelle ESCL_NEG



Passiamo ora ad analizzare i valori iniziali e finali del contorno di F0; in riferimento a tali parametri, il quadro è poco trasparente. Come riportato nella fig. 5, i dati relativi all'Onset di F0 dimostrano che nelle donne le ESCL_POS presentano valori medi significativamente più alti rispetto alle ESCL_NEG [+45.5 Hz, +2.8 ST ($t=3.981$; $df=49$; $p<.05$)]; diversamente negli uomini non si rileva significatività statistica [+7.2 Hz, 0.4 ST ($t=.921$; $df=49$; $p>.05$)]. I valori relativi all'Offset di F0 sono comparabili nelle due attitudini, sia per gli uomini [+1 Hz, +0.1 ST ($t=.921$; $df=49$; $p>.05$)] che per le donne [+0.5 ST ($t=2.267$; $df=49$; $t<.05$)].

²⁸ Nei nostri risultati la differenza fra le due modalità frasali è espressa sempre in base al rapporto ESCL_POS/ESCL_NEG

Figura 5 - Valori medi di Onset e Offset di F0 (in Hz) per femmine (sx) e maschi (dx) nelle ESCL_POS e ESCL_NEG



I dati raccolti mostrano che le donne tendono ad attivare i parametri frequenziali in misura maggiore nelle ESCL_POS che nelle ESCL_NEG.

7.2.2 Intensità

Anche nel caso dei valori dell'intensità media, i dati differiscono in base al genere. Nelle donne le ESCL_POS rivelano una intensità più alta in media di 2.3 dB rispetto a quanto riscontrato per le ESCL_NEG ($t=4.833$; $df=49$; $p<.000$). La stessa tendenza non si rinviene negli uomini, per i quali le esclamative connotate da diversa affettività presentano risultati comparabili [$+0.5$ dB ($t=1.031$; $df=49$; $p>.05$)]. I dati relativi alle donne sembrerebbero suggerire che le locutrici coinvolte nella nostra ricerca tendono a impiegare maggior energia nella realizzazione degli enunciati che convogliano un senso di sorpresa positivo; diversamente, gli enunciati con connotazione emotiva negativa rispondono a una strategia espressiva meno marcata dal volume della voce. Tuttavia, il fatto che la medesima tendenza non si rinvenga nell'eloquio maschile mette in dubbio questa prospettiva.

7.2.3 Durata e velocità elocutiva

Per quanto concerne le dinamiche temporali, il confronto fra ESCL_POS ed ESCL_NEG potrebbe risentire della non corrispondenza sintattica e morfologica fra i due gruppi; tuttavia, nella composizione delle frasi target si è provveduto a scegliere enunciati che fossero formati quanto più possibile dallo stesso numero di costituenti. In riferimento alla durata dell'intero enunciato non si rilevano differenze significative fra le due attitudini in dipendenza del genere. Le donne, però, mostrano in media una vocale tonica significativamente più estesa nelle ESCL_POS che nelle ESCL_NEG [$+18$ ms ($t=2.157$; $df=49$; $p<.05$)]; diversamente, i valori raccolti per gli uomini non restituiscono significatività statistica [$+14$ ms ($t=1.065$; $df=49$; $p>.05$)].

Infine, i risultati hanno rivelato che ESCL_POS ed ESCL_NEG sono realizzate dalle donne con una velocità di eloquio comparabile (rispettivamente 5 e 5.3 sill/s; $t=-1.492$; $df=49$; $p>.05$); negli uomini, la differenza rilevata è sempre di 0.3

sill/s (tuttavia in questo caso il confronto è significativo: ESCL_POS=4.8 sill/s, ESCL_NEG=5.1 sill/s; $t=-2.128$; $df=49$; $p<.05$).

8. *Discussione e conclusioni*

Nello studio del 1989, Bolinger ha definito ampio e profondo il rapporto fra esclamativa e intonazione, evidentemente poiché quest'ultima contribuisce a convogliare il senso di sorpresa e di inaspettato che distingue l'enunciato esclamativo. Lo stesso autore sostiene che la continua ricerca dell'estremo determina nell'esclamativa un'intonazione fuori controllo, che eccede la norma. Non è, infatti, facile identificare un'unica intonazione esclamativa o comunque associare a essa dei contorni prevalenti (Bolinger, 1989). L'esclamativa potrebbe essere considerata imprevedibile dal punto di vista intonativo, in quanto dipendente da una serie infinita di variabili contestuali, emotive o linguo-specifiche. Tuttavia, nel 2008 Castroviejo Mirò sostiene che, in definitiva, l'unico aspetto che le strutture esclamative hanno in comune è un'intonazione enfatica.

Il presente studio si pone, quindi, l'obiettivo di provare a raggiungere quella profondità scorta da Bolinger nella relazione fra l'esclamativa e l'intonazione. La nostra ricerca ha rilevato che l'esclamativa wh nell'italiano di Bari è, nella maggior parte dei casi, più intensa, più lunga e quindi più lenta della frase interrogativa wh, che mostra, però, valori di F0 più alti. Le esclamative presentano, difatti, un range tonale più compresso: tuttavia, il pitch range è positivamente correlato al livello di espressività dell'enunciato, quindi la presenza di un grado di sorpresa maggiore dovrebbe provocare un innalzamento dei valori massimi di F0 (Lieberman, Pierrehumbert, 1984), tendenza non rilevata nel nostro studio. Anche il parametro dell'intensità media è generalmente influenzato dal grado di espressività della frase: nei nostri risultati i valori di intensità sono più alti nelle esclamative, ma tale dato è statisticamente significativo soltanto nei parlanti di sesso maschile.

Secondo la letteratura disponibile (in particolare Sorianello, 2010, 2012) la struttura maggiormente marcata a livello prosodico è quella non wh: in questo caso la disambiguazione rispetto alla controparte assertiva, in assenza di un elemento introduttore, sarebbe convogliata esclusivamente dalla prosodia. Si ipotizza, dunque, che le esclamative wh, seppure connotate sul piano soprasegmentale rispetto a un enunciato non marcato, come quello assertivo, non necessitano dell'attivazione di specifiche risorse prosodiche per distinguersi dalle interrogative. Dinamiche frequenziali simili a quelle rilevate nella nostra ricerca sono state riscontrate nelle esclamative ungheresi, che sono caratterizzate da una F0max più bassa e livelli di Onset inferiori rispetto alle interrogative (Gyuris *et al.*, 2013). Nel nostro studio, i parlanti sembrano però affidare l'esplicitazione del senso di sorpresa a livello soprasegmentale al parametro della durata, l'indice fonetico che più di tutti sembra distinguere l'esclamativa dall'interrogativa; la durata dell'intero enunciato è maggiore nella prima tipologia frasale, l'allungamento temporale si rileva in particolare in corrispondenza dell'ultima vocale tonica. Questo dato si allinea sia ai risultati raccolti per il porto-

ghese brasiliano (Morales, 2008), che per l'italiano (Soriano, 2010). Tuttavia, gli esiti finora raccolti richiedono di essere opportunamente avvalorati da un test di identificazione percettiva che possa gettare luce sul ruolo svolto da ciascun parametro nella caratterizzazione della forza illocutiva dell'enunciato esclamativo²⁹.

La nostra ricerca ha inoltre confermato l'assunto secondo cui la frase esclamativa presenta un innalzamento del contorno iniziale dell'enunciato e una discesa finale. Tuttavia, il confronto con la modalità interrogativa ha mostrato che anche quest'ultima rivela valori medi di Onset più alti rispetto alla F0x; ciò suggerisce che un picco iniziale alto non sia da considerarsi un elemento esclusivo e caratterizzante dell'enunciato esclamativo *wh*, in quanto condiviso anche dalle domande *wh*. I dati finora raccolti necessitano, dunque, di essere sottoposti ad analisi fonologica al fine di osservare in modo accurato il contorno intonativo delle frasi indagate e, in particolare, la distribuzione delle prominenze³⁰.

La possibilità che una diversa connotazione emotiva dell'enunciato possa influenzarne il livello soprasegmentale ha rappresentato il secondo focus di questo studio. La sorpresa è un'emozione poco indagata rispetto alle altre, a causa del fatto che, secondo alcuni studiosi (cfr. § 4), viene espressa principalmente attraverso espressioni facciali; inoltre, spesso la sorpresa può sfociare in altre emozioni come la gioia o la paura. Si tratta quindi di un'emozione base dai confini sfumati che in questa sede abbiamo voluto mettere sotto esame, analizzando la modalità frasale che per definizione dovrebbe veicolarla. Generalmente, in quanto emozione ad alta attivazione, la sorpresa dovrebbe essere intensa e connotata da una F0x alta. Tuttavia, la possibilità che la sorpresa possa avere diversi gradi di affettività complica il quadro. Quanto emerge dalla verifica dei dati è, infatti, poco chiaro: le esclamative che convogliano sorpresa positiva mostrano, rispettando le nostre attese, una maggiore attivazione dei parametri frequenziali e dell'intensità rispetto a quelle con connotazione negativa; pur tuttavia, tale tendenza non può essere generalizzata, visto che coinvolge solo i soggetti femminili. Questi risultati hanno messo in luce che uomini e donne presentano una realizzazione prosodica differente non solo in riferimento alla modalità frasale, ma anche in relazione al tipo di affettività convogliato dall'enunciato. Uno studio futuro specificamente incentrato sulla comunicazione del senso di sorpresa, positivo o negativo, mediante atti frasali esclamativi, potrà gettare ulteriore luce sulla misura in cui i diversi livelli di questa emozione influenzano la prosodia dell'enunciato, modificandone il carattere gradiente. È probabile che la connotazione emotiva venga trasmessa anche dalla qualità della voce, un indice non contemplato in questa ricerca.

Il nostro studio ha, quindi, contribuito all'osservazione degli indici fonetici di una tipologia frasale che merita di essere indagata approfonditamente, non soltanto per l'imprevedibilità e la complessità che la contraddistingue, ma anche per il suo stretto legame con la dimensione emotiva del parlante.

²⁹ A tal proposito, alcuni studi si sono interrogati sul ruolo svolto dal parametro della durata nella corretta identificazione dei tipi frasali (di grande interesse, fra gli altri, lo studio di Cangemi & D'Imperio, 2013).

³⁰ È attualmente in corso un'analisi più approfondita sulle prominenze, corredata da annotazione ToBI.

Riferimenti bibliografici

- AVESANI, C. (1995). ToBI: un sistema di trascrizione per l'intonazione italiana. In LAZZARI, G. (Ed.), *Metodologia di analisi e di descrizione delle caratteristiche prosodiche e intonative dell'italiano*, Atti delle V Giornate di Studio del Gruppo di Fonetica Sperimentale. Roma: Esagrafica, 85-98.
- BANSE, R., SCHERER, K.R. (1996). Acoustic profiles in vocal emotion expression. In *Journal of Personality and Social Psychology*, 70(3), 614-636.
- BENINCÀ, P. (1995). Il tipo esclamativo. In RENZI, L., SALVI, G. & CARDINALETTI A. (Eds.), *Grande Grammatica Italiana di Consultazione*, vol. III. Bologna: Il Mulino, 127-152.
- BENINCÀ, P. (1996). La struttura della frase esclamativa alla luce del dialetto padovano. In BENINCÀ, P., CINQUE, G., DE MAURO, T. & VINCENT, N. (Eds.), *Italiano e dialetti nel tempo. Saggi di grammatica per Giulio Lepschy*. Roma: Bulzoni, 23-43.
- BOLINGER, D. (1989). *Intonation and Its Uses*. Palo Alto: Stanford University Press.
- CANEPARI, L. (1992). *Manuale di pronuncia italiana*. Bologna: Zanichelli.
- CANGEMI, F., D'IMPERIO, M. (2013). Tempo and the perception of sentence modality. In *Laboratory Phonology*, vol. 4 (1), 191-220.
- CASTROVIEJO, E. (2008). Deconstructing Exclamations. In *Catalan Journal of Linguistics*, 7, 41-90.
- D'EUGENIO, A. (1976). The intonation systems of Italian and English. In *Rassegna Italiana di Linguistica Applicata*, 8, 57-85.
- DARDANO, M., TRIFONE, P. (1995). *Grammatica italiana con nozioni di linguistica*. Bologna: Zanichelli.
- DEL MAR VANRELL, M., BALLONE, F., SCHIRRU, C. & PRIETO, P. (2015). Sardinian intonational phonology: logudorese and campidanese varieties. In FROTA, S., PRIETO, P. (eds), *Intonation in Romance*. Oxford: Oxford University Press, 317-349.
- DELAIS-ROUSSARIE, E., POST, B., AVANZI, M., BUTHKE, C., DI CRISTO, A., FELDHAUSEN, I., JUN, S., MARTIN, P., MEISENBURG, T., RIALLAND, A., SICHEL-BAZIN, R. & YOO, H. (2015). Intonational phonology of French: developing a ToBI system for French. In FROTA, S., PRIETO, P. (Eds.), *Intonation in Romance*. Oxford: Oxford University Press, 63-100.
- DELATRE, D. (1966). Les dix intonations de base du français. In *French Review*, 40, 1-14.
- DELANCEY, S. (1997). Mirativity: the grammatical marking of unexpected information. In *Linguistic Typology*, 1, 33-52.
- DE MARCO, A., PAONE, E. (2014). L'espressione e la percezione delle emozioni vocali in apprendenti di Italiano L2 uno studio cross-linguistico. In *Educazione linguistica. Language education*, 3, 483-500.
- EKMAN, P. (2003). *Emotions Revealed. Understanding faces and feelings*. London: Weidenfeld and Nicolson.
- ELLIOTT, D. (1971). The Grammar of Emotive and Exclamatory Sentences in English. In *Ohio State Working Papers in Linguistics*, 8, viii-110.
- ELLIOTT, D. (1974). Toward a grammar of exclamations. In *Foundation of language*, 11, 231-246.

- ESTEBAS-VILAPLANA, E., PRIETO, P. (2010). Castilian Spanish intonation. In PRIETO, P., ROSEANO, P. (Eds.), *Transcription of intonation of the Spanish language*. Munich: LINCOM Europa, 17-48.
- GILI FIVELA, B., AVESANI, C., BARONE, M., BOCCI, G., CROCCO, C., D'IMPERIO, M., GIORDANO, R., MAROTTA, G., SAVINO, M. & SORIANELLO, P. (2015). Intonational phonology of the regional varieties of Italian. In FROTA, S., PRIETO, P. (Eds.), *Intonation in Romance*. Oxford: Oxford University Press, 140-197.
- GRICE, M., D'IMPERIO, M.P., SAVINO, M. & AVESANI, C. (2005). Strategies for intonation labelling across varieties of Italian. In SUN-AH JUN (Ed.), *Prosodic Typology: the Phonology of Intonation and Phrasing*. New York: Oxford University Press, 362-389.
- GRIMSHAW, J. (1979). Complement selection and the lexicon. In *Linguistic Inquiry*, 10, 279-326.
- GYURIS, B., MÁDY, K. & SZALONTAI, Á. (2013). *Experimental investigations on the prosody of Hungarian exclamatives*. Budapest: Research Institute for Linguistics, Hungarian Academy of Sciences.
- GUTIÉRREZ-REXACH, J. (2008). Spanish Root Exclamatives at the Syntax/Semantics Interface. In *Catalan Journal of Linguistics*, 7, 117-133.
- HUDDLESTON, R., PULLUM, G. (2002). *The Cambridge grammar of the English language*. Cambridge: Cambridge University Press.
- IZARD, C.E. (1994). Innate and Universal Facial Expression: Evidence From Developmental and Cross-Cultural Research. In *Psychological Bulletin*, vol. 115, no. 2, 288-299.
- JITCA, D., APOPEI, V. PADURARU, O. & MARUSCA, S. (2015). Transcription of Romanian intonation. In Frota, S., Prieto, P. (Eds.), *Intonation in Romance*. Oxford: Oxford University Press, 284-316.
- JOHNSTONE, T., SCHERER, K.R. (2000). Vocal communication of emotions. In LEWIS, M., HAVILAND, J. (Eds.), *The handbook of emotions*. New York: Guilford, 226-235.
- JUSLIN, P.N., LAUKKA, P. (2003). Communication of emotions in vocal expression and music performance: Different channels, same code?. In *Psychological Bulletin*, 129, 770-814.
- KELLERT O., PETRONE, C. & PANIZZA, D. (2015). Psycholinguistic investigation of the prosody of wh-exclamatives and wh-questions. In *Methods in empirical prosody research*, Jul 2015. Germany: Mannheim.
- KOTZ, S.A., PAULMANN, S. (2011). Emotion, Language, and the Brain. In *Language and Linguistics Compass*, 5/3, pp. 108-125.
- KÖNIG, E., SIEMUND, P. (2007). Speech act distinctions in grammar. In SHOPEN, T. (Ed.) *Language typology and syntactic description*, 2nd ed., Vol. 1, *Clause structure*. Cambridge: Cambridge University Press, 276-324.
- LIBERMAN, M., PIERREHUMBERT, J.B. (1984). Intonational invariance under changes in pitch and length. In ARONOFF, M., OEHRLE, R.T. (Eds.), *Language sound structure*, Cambridge (Ma.): MIT Press, 157-233.
- MAFFIA, M., PELLEGRINO, E. & PETTORINO, M. (2014). Labelling expressive speech in L2 Italian: The role of prosody in auto-and external annotation. In CAMPBELL, A.W., GIBBON, D., HIRST, D. (Eds.). In *Proceedings of the 7th International Conference on Speech Prosody*,

Berlin, Germany, 20-23 May 2014, Speech Prosody Special Interest Group (SProSIG). Urbana: Illinois, 81-84.

MICHAELIS, L. (2001). Exclamative constructions. In HASPELMATH, M., KÖNIG, E., ÖSTERREICHER, W. & RAIBLE, W. (Eds.), *Language Universals and Language Typology: An International Handbook*. Berlin: Walter de Gruyter, 1038-1050.

MORAES, J. (2008). The pitch accent in Brazilian Portuguese: analysis by synthesis. In BARBOSA, P.A., MADUREIRA, S. & REIS, C. (Eds.), *Speech Prosody 2008*, 389-397.

O'CONNOR, J.D., ARNOLD, G.F. (1961). *Intonation of colloquial English*. London: Longmans.

POGGI, I., MAGNO CALDOGNETTO, E. (2004). Il parlato emotivo, Aspetti cognitivi, linguistici e fonetici. In ALBANO LEONI, F., CUTUGNO, F., PETTORINO, M. & SAVY, R. (Eds.), *Atti del Convegno nazionale Il Parlato Italiano*. Napoli: D'Auria Editore.

RETT, J. (2008). A degree account of exclamatives. In FRIEDMAN T., ITO, S. (Eds.), *Proceedings of SALT XVIII*, Ithaca, Greece. New York: CLC Publications, 601-608.

RETT, J., STURMAN, B. (in stampa). *Prosodically marked mirativity*. Los Angeles: University of California.

ROSEANO, P., del MAR VANRELL, M., PRIETO P. (2015). Intonational Phonology of Friulan and its dialects. In FROTA, S., PRIETO, P. (Eds.), *Intonation in Romance*. Oxford: Oxford University Press, 101-139.

SADOCK, J., ZWICKY, A. (1985). Speech act distinctions in syntax. In SHOPEN, T. (Ed.), *Language typology and syntactic description*. Cambridge: Cambridge University Press, 155-196.

SCHERER, K.R. (2003). Vocal communication of emotion: a review of research paradigms. In *Speech communication*, 40 (1-2), 227-256.

SCHERER, K.R. (1986). Vocal affect expression: A review and a model for future research. In *Psychological Bulletin*, 99 (2), 143-165.

SERIANNI, L. (1997). *Italiano*. Garzanti: Milano.

SICHEL-BAZIN, R., MESENBURG, T & PRIETO, P (2015). Intonational phonology of Occitan: towards a prosodic transcription system. In FROTA, S., PRIETO, P. (Eds.), *Intonation in Romance*. Oxford: Oxford University Press, 198-233.

SORIANELLO, P. (2011). Aspetti pragmatici e prosodici dell'atto esclamativo. In *Studi Linguistici e Filologici Online (SLIFO)*, 9.1, 287-332.

SORIANELLO, P. (2012). A prosodic account of Italian exclamative sentences: a gating test. In MA, Q., DING, H. & HIRST, D. (Eds.), *The Proceedings of the 6th International Conference on Speech Prosody*, Shanghai (China), May 22-25, 2012, vol. I. Tongji University Press, 298-301.

SORIANELLO, P. (2010). Il tipo esclamativo. Analisi e percezione delle risorse prosodiche. In CUTUGNO, F., MATURI, P., SAVY, R., ABETE, G. & ALFANO, I. (Eds.), *Parlare con le persone, parlare con le macchine. La dimensione interazionale della comunicazione verbale. Proceedings of 6^o National Congress of AISV*. Torriana: EDK, 85-104.

TOMKINS, S. (1962). *Affect Imagery Consciousness: Volume I, The Positive Affects*. London: Tavistock.

TOMKINS, S. (1963), *Affect Imagery Consciousness: Volume II, The Negative Affects*. London: Tavistock.

VILLALBA, X. (2003). An exceptional exclamative sentence type in Romance. In *Lingua*, 13, pp. 713-745.

ZANUTTINI, R., PORTNER, P. (2003). Exclamative Clauses: At the Syntax-Semantics Interface. In *Language*, 79, 39-81.

PATRIZIA SORIANELLO

Sul riconoscimento percettivo delle domande retoriche

On the perceptual recognition of rhetorical questions

This study investigates the perception of Italian minimal pair sentences formed by information-seeking questions and rhetorical questions. Two perception experiments were performed, the first using natural stimuli, the second using three different series of stimuli generated by manipulating respectively: 1) the duration of the nuclear vowel, 2) the final intonation contour, 3) both duration and final contour. The two types of questions behaved differently. The results showed that temporal manipulations contributed to the rhetorical interpretation more often than the manipulations of the final intonation contour.

Keywords: yes-no question prosody, rhetorical questions, perception test.

1. *La domanda retorica*

A prescindere dalla struttura sintattica, gli enunciati interrogativi possono essere catalogati a seconda della funzione e del significato che assumono nella dinamica conversazionale. Con riferimento allo scopo pragmatico che intendono raggiungere, si distinguono domande tese a richiedere informazioni e domande volte a ottenere conferma. Le prime, idealmente neutre e dal valore letterale, notoriamente denominate *information-seeking-questions* (da ora ISQ) sono contrassegnate da un valore epistemico di ignoranza e dunque di ‘sincerità’, chi le produce non ha aspettative sulla risposta che riceverà, (1) *Dove sei stato?* Le seconde invece sono formulate con il fine di ricevere una conferma, sono domande orientate aventi uno stato epistemico di probabilità soggettiva; il parlante presume cioè di conoscere la risposta; ne sono un esempio le domande eco, le domande coda o quelle che esplicano una funzione fatica, (2) *Sei stato al mare, vero?* Le domande retoriche (da ora DR) rappresentano lo stadio ultimo di questo secondo raggruppamento in quanto non sono realizzate per richiedere informazioni. Sul versante epistemico, le DR si collocano sul polo opposto rispetto alle ISQ incorporando uno stato di conoscenza totale e la trasmissione di un significato già noto, (3) *Dove vuoi che sia stato?*

Studiata da diverse angolature e spesso al centro di dibattiti controversi, la DR è una pseudo-domanda, ossia un enunciato che, pur mantenendo la forma interrogativa, realizza funzioni ‘altre’. A differenza di una domanda ‘vera’ non chiede informazioni e diversamente da un’assertiva non le fornisce. La sua natura fittizia è confermata anche dal fatto che la DR il più delle volte non è seguita da una risposta esplicita. Si tratta, dunque, di un enunciato indiretto in cui la struttura locutoria non coincide con la sua funzione comunicativa. Il ‘senso’ è presupposto e ovvio e

si pone oltre l'interpretazione letterale della frase (cfr. Sadock, 1971; Ilie, 1994, Gutiérrez-Rexach, 1998; Han, 2002).

Secondo alcuni autori le DR sarebbero delle asserzioni enfatiche (Anzilotti, 1982; Stati, 1982; Han, 2002); per altri studiosi sarebbero invece da includere nella classe delle interrogative (Gutiérrez-Rexach, 1998; Rooy, 2003; Rohde, 2006). Il novero delle funzioni attribuite alla DR è vario e dipende anche dal dominio linguistico di osservazione e dalla tipologia testuale di riferimento¹. A questo proposito, Stati (1982) riconosce alla DR le funzioni di critica, autocritica, ma anche di rimprovero, sfida e ironia. Più in generale, tale domanda sarebbe tesa a mitigare (Brown, Levinson, 1978) o anche a rafforzare la forza illocutoria dell'enunciato (Frank, 1990; Ilie, 1994); in un'ottica più vasta, assumendo come riferimento l'intera dinamica conversazionale, la DR sarebbe invece impiegata con lo scopo di risincronizzare il dialogo e rinnovare l'impegno cooperativo tra gli interlocutori (Ilie, 1994)².

Sul piano sintattico, in più lingue, queste domande contengono dei *Negative Polarity Items* (NPI), specifici indicatori lessicali che inducono una polarità negativa dell'enunciato, dirigendo l'ascoltatore verso un'interpretazione retorica della domanda, tra questi: *forse mica, mai, dopotutto, vuoi che*, ad es. (4) *Ti ho forse(mai\ mica) detto che sarei venuto?* (Han, 2002)³. Altre volte, le DR sfruttano le medesime strutture delle interrogative con valore letterale; si rinvengono così DR polari (5) *La smetti di dire sciocchezze?* e wh- (6) *Perché avrei dovuto farlo?* le quali, identiche per composizione morfo-sintattica e lessicale alle rispettive ISQ, contrastano per intenzione comunicativa formando delle coppie minime frasali.

In questi casi, si assume che la struttura prosodica di DR e ISQ contenga dei tratti divergenti, tali da rendere possibile la disambiguazione del loro contenuto illocutivo. Questo punto, però, non è condivisivo da Jung e Schrott (2003: 360) le quali ritengono che «whether a question is rhetorical or not depends on the context». Occorre dunque postulare che, in caso di potenziale ambiguità, il discrimine sia dato dal contesto enunciativo e dall'insieme delle conoscenze condivise tra locutore e interlocutore. Quest'ultimo aspetto si prefigura come un tassello di grande rilevanza ai fini della nostra ricerca, tesa a saggiare, per mezzo di un confronto tra coppie minime costituite da DR e ISQ, il peso percettivo di alcuni indici prosodici.

2. *Gli studi prosodici*

Le informazioni sulla prosodia della DR, per molti anni frammentarie e affidate a poche righe contenute all'interno di alcuni studi a carattere semantico e pragma-

¹ Ad esempio, nella prosa giornalistica, la DR favorisce il cambio di argomento, l'apertura e la chiusura argomentativa (Anzilotti, 1982).

² Per un approfondimento sul rapporto tra funzione della DR e intonazione si rinvia a Soriano (2019).

³ Le DR sono considerate enunciati aventi la forza illocutiva di un'asserzione di polarità opposta a quella che appare nella loro forma interrogativa, ad es. la domanda *Come potevo immaginarlo?* implica l'asserzione negativa *Io non potevo immaginarlo*, laddove la domanda *Non potevi pensarci prima?* corrisponde all'asserzione positiva *Tu potevi pensarci prima*.

tico, indicano con una certa continuità la presenza di un contorno melodico discendente finale (Anzilotti, 1982; Stati, 1982; Frank, 1990; Ilie, 1994; Gutiérrez-Rexach, 1998; Han, 2002; Koshik, 2005; Rohde, 2006; Caponigro, Sprouse, 2007). Il ruolo dell'intonazione è però tutt'altro che secondario nella costituenza della DR. Recenti ricerche svolte in più lingue in prospettiva sperimentale dimostrano infatti che la prosodia della DR contiene sempre degli indizi fonetici e fonologici utili al riconoscimento del suo valore illocutivo (Wochner *et al.*, 2015. Dehé *et al.*, 2018, Braun *et al.*, 2020, Zahner *et al.*, 2020).

A un confronto tra studi, l'intonazione della DR sembra altresì possedere un certo grado di condizionamento linguo-specifico. In questa direzione si pongono i dati ricavati per alcune lingue dal confronto puntuale tra DR e ISQ di pari struttura, pur con qualche cautela dovuta all'adozione di procedure solo parzialmente sovrapponibili. Da queste indagini si desume come l'azione differenziante sia spesso svolta dal contorno intonativo terminale, altre volte dall'accento intonativo nucleare, altre ancora dalla durata o da una diversa distribuzione degli accenti intonativi. La forma sintattica della domanda, polare o wh-, concorre inoltre a definire la sua realizzazione prosodico-intonativa. Le ricerche condotte hanno mostrato che in giapponese il fattore differenziante tra DR e ISQ è la durata (Miura, Hara, 1995), una tendenza di recente confermata anche per l'estone (Asu *et al.*, 2020), mentre in tedesco (Wochner *et al.*, 2015) e in islandese (Dehé *et al.*, 2018) sarebbe il tono nucleare a distinguere i due atti interrogativi⁴. Per il tedesco, Neitsch, Niebuhr (2019), in uno studio teso a cogliere le divergenze tra ISQ e DR in due diversi contesti attitudinali, la derisione e il disgusto, mettono in luce come le DR siano sempre dipendenti dal contesto discostandosi dalle ISQ per durata, andamento intonativo e qualità della voce. Secondo gli autori, l'analisi della DR può essere inquadrata in una prospettiva multidimensionale in cui i parametri fonetici svolgono un ruolo interdipendente con il contesto comunicativo e reciprocamente condizionante.

Con riferimento alla lingua italiana, da un'analisi sperimentale (Soriano, 2018) condotta su 20 coppie di DR e ISQ con struttura polare e wh- prodotte seguendo un protocollo di elicitazione controllato da parlanti baresi, si deduce come uno dei parametri più incisivi ai fini della distinzione sia la durata. L'incremento temporale coinvolge la DR e si estrinseca in primo luogo sull'ultima vocale tonica dell'enunciato, ovvero la vocale su cui si realizza l'accento intonativo nucleare. Nel dettaglio, la durata media normalizzata della VN è maggiore nelle retoriche sia nel tipo wh- (+ 15%) che in quello polare (+ 25%). Anche il contorno finale concorre a tale distinzione, la DR polare è più spesso (62%) dotata di un *pattern* discendente rispetto alla ISQ (42%), una percentuale che si riduce nelle wh- (56%)⁵.

⁴ Precisamente, in tedesco DR e ISQ sono entrambe caratterizzate da un contorno intonativo discendente (L%), la presenza dell'accento nucleare L*+H sarebbe però un tratto tipico della DR. Analogamente, in islandese DR e ISQ polari sono discendenti finali, ma si oppongono per la presenza di un diverso *pitch accent*: L+H* nella DR e L*+H nell'ISQ. Nella lingua inglese invece DR e ISQ mostrano sia *pattern* finali ascendenti (H%) che discendenti (L%), (Banuazizi, Creswell, 1999; Bartels, 2013).

⁵ Le DR presentano inoltre una più ampia escursione tonale rispetto alle ISQ, ma solo nel tipo wh-.

A fronte degli studi incentrati sulla produzione della DR, le ricerche improntate alla definizione delle dinamiche che agiscono in percezione sono davvero poche. Per il giapponese, si dispone degli esiti raccolti a seguito della somministrazione a un gruppo di uditori di due diverse batterie di stimoli sintetici ottenuti modificando la durata finale e, rispettivamente, l' f_0 iniziale o finale di uno stimolo di tipo ISQ (Miura, Hara, 1995). I risultati provano che le risposte a favore dell'interpretazione retorica si innalzano con l'aumentare della durata vocalica e con l'abbassamento dell' f_0 iniziale. Tale effetto percettivo sembra cogliere le peculiarità della DR sebbene, in assenza di ulteriori confronti cross-linguistici, deve essere assunto con una certa cautela, in quanto verosimilmente condizionato anche dalle caratteristiche metrico-accentuali del giapponese.

Alla luce di questa premessa, lo svolgimento anche per altre lingue di esperimenti percettivi potrebbe rivelarsi particolarmente utile per l'individuazione del peso distintivo esercitato dai singoli parametri prosodici, come pure degli effetti linguo-specifici della DR.

3. *La ricerca*

Questo studio ha inteso verificare il ruolo di alcuni indici fonetici nella categorizzazione percettiva di ISQ e DR. A tale scopo, sono state selezionate coppie minime frasali per la costruzione di due diversi test percettivi i quali sono stati somministrati *online* in due diversi momenti temporali.

Le coppie di frasi che formano il corpus hanno identità segmentale, ma diversa forza illocutiva, essendo interpretabili, nel contesto di riferimento da cui sono state estratte, rispettivamente come DR o ISQ. I quesiti che hanno mosso questa indagine sono stati essenzialmente due. Da un lato abbiamo voluto accertare se alla DR potesse essere riconosciuta, anche a livello percettivo, una propria veste prosodica. Per rispondere a tale domanda di ricerca abbiamo condotto un test di identificazione percettiva (Test 1) costituito da coppie di stimoli naturali (DR vs. ISQ). L'ipotesi che abbiamo assunto è che la DR, in virtù dei tratti che la caratterizzano, abbia una realizzazione prosodica autonoma e dunque riconoscibile nel suo valore pragmatico anche quando estratta dal suo intorno enunciativo. Con riferimento alla produzione, alcuni dati raccolti per l'italiano indicano infatti come la DR, almeno per certi aspetti, sia, nella varietà di Bari distinta dalla ISQ (Soriano, 2018). Ciò che non sappiamo al momento è se tale base differenziante sia sufficiente anche ai fini di una sua categorizzazione sul fronte percettivo. L'ipotesi sarà confermata se la DR, isolata dal proprio contesto dialogico, sarà ancora in fase di ascolto correttamente identificata in un numero rilevante di casi. La confutazione di tale ipotesi porterà, viceversa, ad assumere che la sola tessitura prosodica della DR non sia bastevole di per sé a indurre nell'ascoltatore il suo riconoscimento, ragione per cui la DR, non più riconosciuta nella sua *facies* naturale, sarà processata come ISQ. Ciò vorrà dire, pertanto, che l'attuazione del significato e della forza illocutoria di una DR andrà cercata altrove, verosimilmente nel proprio contesto situazionale.

In secondo luogo, ci siamo chiesti quale fosse il ruolo di due parametri chiave, precisamente durata della vocale tonica nucleare (da ora VN) e contorno intonativo finale (da ora CF), ai fini della classificazione della DR su base percettiva. Questo aspetto è stato affidato all'esecuzione di un secondo test sperimentale (Test 2) contenente, diversamente dal primo, stimoli manipolati tramite sintesi. Gli indici selezionati non sono stati casuali, bensì scelti tra quelli che hanno mostrato più potere differenziante a livello di produzione (Soriano, 2018). L'ipotesi assunta è che la durata della VN sia nella DR un elemento robusto capace di guidare il giudizio degli ascoltatori: una riduzione della VN potrebbe dunque incrinare la probabilità che la domanda sia interpretata come retorica. L'ipotesi iniziale è che gli stimoli con incremento temporale sintetico della VN siano assegnati più frequentemente alla categoria delle domande retoriche, in tal caso, alla durata potrà essere effettivamente riconosciuto un ruolo differenziante.

Con riferimento all'intonazione, invece, ci siamo chiesti in che misura il contorno intonativo finale delle frasi bersaglio riuscisse a modificare il giudizio pragmatico degli uditori. L'azione svolta dal CF mostra alcune criticità. Se, infatti, la DR è caratterizzata in modo prevalente, ma beninteso non esclusivo, da un CF discendente, l'ISQ presenta, invece, in modo pressoché alternato sia *pattern* di f_0 discendenti che ascendenti (Soriano, 2018). Di conseguenza, l'assenza di un andamento omogeneo da assegnare alle due tipologie interrogative potrebbe, realisticamente, influire sull'azione condizionante di tale parametro in fase di sintesi. In merito a questo aspetto, ci aspettiamo che una variazione della direzione del contorno terminale di f_0 possa dirottare, sia nella DR che nella ISQ, il giudizio degli ascoltatori, o quanto meno confonderlo. Ad ogni modo, sarà interessante verificare quali saranno gli effetti percettivi indotti dall'ascolto di domande (DR e ISQ) dotate di un CF opposto rispetto a quello dello stimolo naturale⁶.

Le procedure impiegate nei due test sperimentali, così come l'analisi delle risposte fornite dagli uditori, essendo parzialmente diverse, ma soprattutto legate da un rapporto sequenziale, saranno analizzate in dettaglio e discusse separatamente nei paragrafi che seguono.

3.1 Test 1

Il primo stadio è stato quello di accertare se le DR del campione fossero correttamente identificate se ascoltate in isolamento contestuale. Tale obiettivo, tutt'altro che irrilevante, ha permesso infatti di verificare la peculiarità percettiva della DR rispetto alla sua versione non retorica. Questo stadio può essere legittimamente considerato una precondizione dell'intero paradigma sperimentale. Per verificare tale ipotesi, sono state selezionate 12 coppie di stimoli naturali, per un totale di 24 interrogative polari di cui 12 DR e 12 ISQ⁷. Gli enunciati sono stati tratti da brevi scenari dialogici prodotti da 4 locutori di Bari. Ogni coppia è formata da domande aventi identica struttura segmentale e sintat-

⁶ Si è scelto di manipolare il contorno intonativo finale delle domande e non l'andamento della vocale nucleare, in quanto nel barese domande polari sincere e retoriche condividono con elevate percentuali lo stesso *pitch accent* nucleare, vale a dire L+H* (Savino, 2012; Soriano 2018).

⁷ Gli enunciati interrogativi fanno parte di un corpus più ampio la cui analisi fonetica e fonologica è stata discussa in Soriano (2018).

tica, ma diverso valore illocutivo: retorico o non retorico. Tali enunciati sono stati posti in ordine randomizzato e poi sottoposti al giudizio percettivo da parte di ascoltatori; nel test sono state inserite anche 6 interrogative *filler*. Il test è costituito da tre diverse sezioni: la prima raccoglie le informazioni personali degli ascoltatori (età, genere, provenienza e grado di istruzione), la seconda contiene le istruzioni per l'ascolto e infine la terza la sequenza degli stimoli. Agli uditori è stato chiesto di ascoltare gli stimoli audio e di esprimere un giudizio sul loro valore pragmatico; tre le opzioni previste: (1) *domanda con valore retorico*, (2) *domanda con valore non retorico*, (3) *non so rispondere*. Il test non prevedeva un limite di tempo per la sua esecuzione e consentiva l'ascolto reiterato dello stimolo proposto⁸.

All'esperimento hanno partecipato 38 uditori di area barese⁹, prevalentemente femmine (75%), di età compresa tra 20 e 48 anni, aventi grado di istruzione universitario (94%). Complessivamente sono state raccolte ed elaborate 912 risposte. I dati sono stati testati con un'Anova a una via (SPSS 20) ponendo le risposte ottenute come parametro dipendente e la categoria della domanda come fattore indipendente.

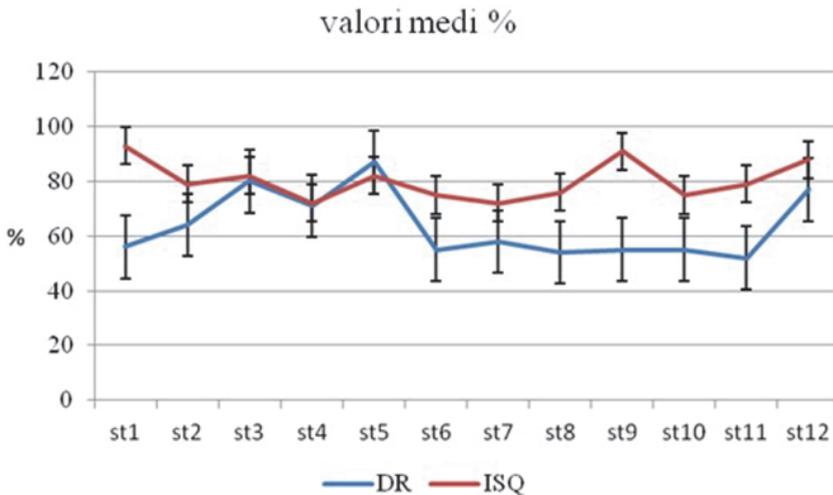
3.1.1 Risultati: Test 1

I dati forniscono un quadro interessante. Un primo aspetto meritevole di attenzione è che gli stimoli naturali delle due categorie frasali, sebbene classificati con percentuali piuttosto alte, non raggiungono valori di identificazione assoluti. Nella Fig. 1 sono riportati i valori medi e l'indice di errore calcolati per ogni stimolo nella sua duplice natura: DR o ISQ. Come si evince dal grafico, alcune domande, corrispondenti agli stimoli 3, 5 e 12 per le DR e agli stimoli 1, 9 e 12 per le ISQ, hanno ottenuto le più alte percentuali di riconoscimento. In generale, la variabilità con cui gli uditori hanno classificato gli stimoli sembra non dipendere da un unico tratto, l'esperimento però lascia intravedere un punto rilevante, la DR sembra profilarsi come una tipologia interrogativa più debole ai fini della sua identificazione percettiva (*ultra*, § 4). Nell'insieme, gli *scores* relativi alle DR sono infatti più bassi rispetto a quelli delle ISQ. Precisamente, prescindendo dai singoli stimoli, le ISQ sono state riconosciute con un *range* compreso tra il 72% e il 93%, laddove il tasso di identificazione delle DR mostra una flessione, oscillando tra il 55% e l'87%; la differenza è statisticamente significativa ($p=0,0024$). In questa prima fase di indagine, nessun ascoltatore ha selezionato l'opzione *Non so rispondere*. Il fatto che le DR siano state riconosciute con percentuali spesso al di sotto del 60% mette in evidenza come tali interrogative, estrapolate dal loro contesto testuale, per quanto artificiale, perdano parte della loro identità illocutiva.

⁸ La metodologia adottata non ci ha consentito purtroppo l'acquisizione dei tempi di risposta, un parametro che avrebbe di certo fornito informazioni aggiuntive.

⁹ Per la precisione, il 36,3% proviene da Bari, il 30,3% dalla provincia barese, i rimanenti dalla provincia BAT.

Figura 1 - Valori percentuali medi e indice di errore rilevati per le 12 coppie di stimoli naturali (DR e ISQ)



3.2 Test 2

Al secondo esperimento percettivo hanno preso parte 24 uditori¹⁰ di età compresa tra 23 e 47 anni di cui il 78% costituito da femmine¹¹. La procedura metodologica è fondamentalmente la stessa di quella adottata nel primo esperimento; in questo caso, però, il test somministrato si compone di 3 coppie di stimoli audio manipolati tramite sintesi con il *software PSOLA* implementato in *PRAAT* (Boersma, Weenink, 2013). Le coppie di enunciati prescelti, la prima realizzata da una locutrice femmina e le rimanenti da un parlante maschio, sono le seguenti:

DR1\ISQ1) *Lo sapete leggere quel segnale?*

DR2\ISQ2) *Sei capace di guidare?*

DR3\ISQ3) *La vedi da qualche parte?*

Al fine di assicurare la loro adeguatezza pragmatica, le domande sono state selezionate all'interno del campione utilizzato per il Test 1 tra quelle aventi il più alto *score* di riconoscimento. Analogamente a quanto effettuato nel primo esperimento, gli uditori sono stati invitati ad ascoltare gli enunciati e a classificarli in base al loro significato retorico o non retorico.

Per ogni enunciato bersaglio sono state costruite tre diverse batterie per un totale di 36 stimoli, a cui sono stati aggiunti anche 6 distrattori; in questo caso le risposte esaminate sono state nel complesso 864 (36 x 24 locutori). Nel dettaglio, la prima serie è stata creata modificando tramite sintesi la durata della VN, la seconda

¹⁰ Tutti gli ascoltatori hanno partecipato anche al Test 1.

¹¹ Il 50% di essi proviene dal capoluogo, il 29,1% dalla provincia barese, la restante parte dalla provincia BAT. Il grado di istruzione prevalente è la laurea di primo o di secondo livello (95,8%).

modificando il CF di f_0 , la terza, infine, ha agito simultaneamente sulla durata della VN e sul CF. Per evitare l'introduzione nel disegno sperimentale di componenti 'altre', difficilmente controllabili in fase di analisi, le coppie di enunciati rispettano le seguenti condizioni: (1) gli stimoli relativi alle DR mostrano un contorno intonativo discendente (L%), laddove in modo opposto, gli stimoli relativi alle ISQ hanno un *pattern* finale ascendente (H%)¹²; (2) sia le DR che le ISQ condividono il medesimo accento intonativo nucleare, ovvero L+H*, il *pitch accent* più rappresentativo di queste domande¹³. La sintesi ha agito in una duplice direzione, nel primo caso l'enunciato *source* è rappresentato dalla DR, nel secondo dalla ISQ. Le manipolazioni sono state speculari: nelle ISQ è stata incrementata la durata della VN e abbassato il valore di f_0 finale, nelle DR, invece, in modo diametralmente opposto, è stata ridotta la durata temporale della VN e innalzato il valore del *boundary tone*. In ambedue i casi, la durata della vocale è stata rispettivamente aumentata o ridotta di 40 ms in due stadi di 20 ms¹⁴, mentre l' f_0 finale è stata innalzata nelle DR e abbassata nelle ISQ in due stadi di 1,5 semitoni (da ora ST) per un totale complessivo di 3 ST¹⁵. Le soglie prescelte non sono causali, in quanto incorporano piuttosto bene la differenza media rinvenuta durante l'indagine acustica per i due tipi frasali posti a confronto.

Le modifiche effettuate per le tre coppie di domande sono riportate nella Tabella 1 (sintesi durata VN) e nella Tabella 2 (sintesi f_0 CF). Dai valori riassunti si osserva da un lato come le DR abbiano una durata originale maggiore delle ISQ, anche se lo scarto temporale è variabile¹⁶, dall'altro come l'esito della manipolazione effettuata determini, al secondo stadio di sintesi, durate comparabili a quelle rinvenute negli omologhi stimoli naturali di tipo ISQ; il *trend* è confermato anche per le VN delle ISQ le quali, a sintesi avvenuta, si allineano per durata a quella delle rispettive DR originali.

I dati sono stati testati con un'Anova a due vie (SPSS 20) ponendo le risposte ottenute come parametro dipendente e il tipo di sintesi e la categoria della domanda come parametri indipendenti.

¹² Nel barese la presenza di un andamento terminale ascendente è tipico del parlato letto e controllato (Savino, 2012, Gili Fivela *et al.*, 2015).

¹³ Precisamente, tale *pitch accent* è stato rinvenuto nell'89% delle ISQ e nel 93% delle DR [$p=.553$], (Soriano, 2018).

¹⁴ Nel dettaglio, la durata è stata modificata in corrispondenza della porzione formantica stazionaria della vocale nucleare, in modo da non alterare l'effetto uditivo degli attacchi transizionali. L'attacco consonantico della sillaba nucleare non ha invece subito alcuna variazione temporale.

¹⁵ La manipolazione del tono di confine non ha determinato una concomitante variazione dell'accento intonativo. La sintesi ha avuto inizio in corrispondenza del *turning point* del contorno finale, collocato di solito sulla consonante d'attacco della sillaba finale ed è proseguita in modo lineare fino al raggiungimento del valore frequenziale stabilito. In questo processo di sintesi si è reso necessario talvolta arrotondare il valore per avvicinarsi quanto più possibile alle soglie di 1,5 e 3 semitoni.

¹⁶ Si tenga conto che la VN della coppia DR3\ISQ3, diversamente dalle altre, ricorre in sillaba chiusa.

Tabella 1 - *Valori in ms, originali e sintetici, delle 3 coppie di stimoli sottoposte a manipolazione della durata VN*

stimolo	VN	VN	VN
	originale	sintesi 1° stadio	sintesi 2° stadio
DR1	208	188	168
DR2	255	235	215
DR3	175	155	135
ISQ1	170	190	210
ISQ2	204	224	244
ISQ3	140	160	180

Analogamente, anche per quanto concerne il CF (Tab. 2), dal confronto incrociato tra valori originali e manipolati e tipo di domanda, si evince come la sintesi fornisca frequenze comparabili.

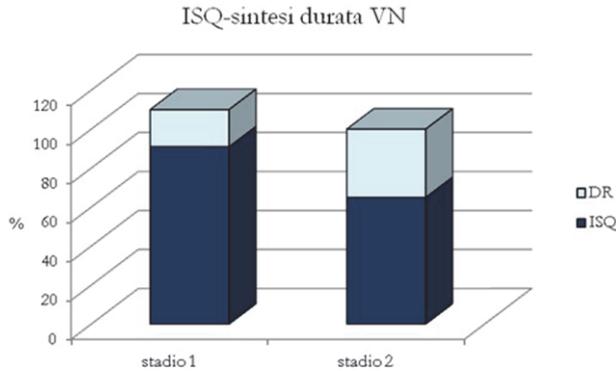
Tabella 2 - *Valori in Hz, originali e sintetici, delle 3 coppie di stimoli sottoposte a manipolazione del CF*

stimolo	f0	f0	f0
	originale	sintesi 1° stadio	sintesi 2° stadio
DR1	181	198	216
DR2	114	124	136
DR3	109	119	130
ISQ1	222	204	187
ISQ2	130	119	109
ISQ3	137	126	115

3.2.1 Test 2: la sintesi della durata

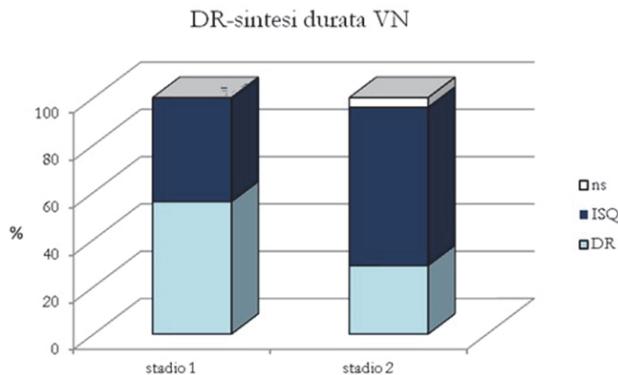
Preliminarmente, occorre dire che la durata media della VN rinvenuta in produzione nei due tipi interrogativi è diversa (Tab. 1). Il peso esercitato da tale parametro trova conferma in percezione nelle risposte fornite dagli ascoltatori. Il primo aspetto che merita di essere commentato riguarda la diversa efficacia esercitata dai due stadi di sintesi: la prima manipolazione ha intaccato le risposte degli uditori solo in un numero limitato di casi, probabilmente a causa del ridotto tasso di variazione introdotto, compreso, a seconda dello stimolo, tra l'8% e il 14% rispetto alla durata originale della vocale. Il secondo elemento da considerare concerne il fatto che la manipolazione della VN ha prodotto un effetto differente a seconda dell'enunciato di partenza, come riassunto nella Figura 2. Il confronto tra il primo e il secondo stadio di sintesi operato su DR e ISQ è significativo ($p=0,0044$). Con riferimento alla ISQ, la prima manipolazione non incide sulla domanda nel 91% delle risposte. Tale percentuale si abbassa nel secondo stadio di sintesi ($p=0,066$): nel 65% dei casi la ISQ non subisce un cambio di percezione; ciò vuol dire che l'incremento della durata vocalica non è stato sufficiente a sviare la valutazione degli uditori verso una classificazione retorica.

Figura 2 - Valori percentuali rinvenuti per la ISQ



Diverso è però il comportamento che si osserva quando si agisce sulla VN di una DR (Fig. 3); in questo caso, la riduzione della durata vocalica produce una variazione di giudizio in circa la metà delle risposte fornite (44%), una percentuale che si innalza nel secondo stadio di sintesi ($p=0,036$): nel 67% dei casi la DR originale viene infatti percepita come ISQ; solo il 35% degli uditori, nonostante la modifica apportata, continua ad assegnare allo stimolo un senso retorico, mentre il 4% propende per l'opzione 3.

Figura 3 - Valori percentuali rinvenuti per la DR



Questo risultato è significativo poiché mette in evidenza come la presenza di una vocale nucleare allungata sia un correlato positivo ai fini della decodifica di una domanda come retorica. È dunque sufficiente ridurre la durata della VN di circa il 20% rispetto alla sua lunghezza naturale per orientare gli ascoltatori verso una diversa categorizzazione dell'interrogativa.

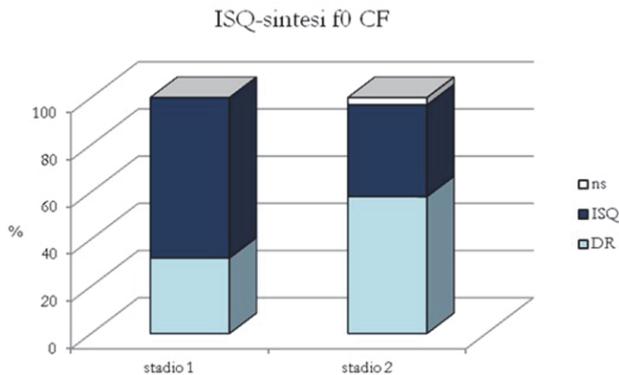
Le tre coppie di stimoli presentano un comportamento omogeneo, con una sola eccezione: lo stimolo ISQ3 nel 70% dei casi è interpretato come DR, una valuta-

zione in controtendenza rispetto alle altre ISQ; è verosimile supporre che l'allungamento sintetico della VN, avvenuto in sillaba chiusa¹⁷, abbia conferito alla domanda un senso amplificato di ovvietà, tale da far propendere gli ascoltatori verso una classificazione retorica.

3.2.2 Test 2: la sintesi del contorno finale

Meno trasparente è la lettura delle risposte raccolte per gli stimoli ottenuti tramite la manipolazione dell'andamento intonativo finale. Nuovamente, i dati sono diversi a seconda dell'enunciato *source* su cui è stata effettuata la sintesi e del grado di manipolazione. Il confronto tra primo e secondo stadio di identificazione tra DR e ISQ raggiunge la significatività statistica ($p=0,010$). Consideriamo innanzitutto ciò che avviene nelle ISQ (Fig. 4). Negli enunciati bersaglio del nostro campione, con riferimento alla variazione del *boundary tone*, si osserva come il primo stadio di sintesi è poco incisivo ed influenza solo il 32% delle risposte, mentre il secondo abbassa sensibilmente la percentuale di riconoscimento di queste domande le quali continuano a essere identificate solo nel 39% dei casi; nel 58% l'ISQ è invece percepita come DR, mentre nel 3% rimanente l'uditore sceglie l'opzione *non so rispondere*.

Figura 4 - Valori percentuali rinvenuti per l'ISQ

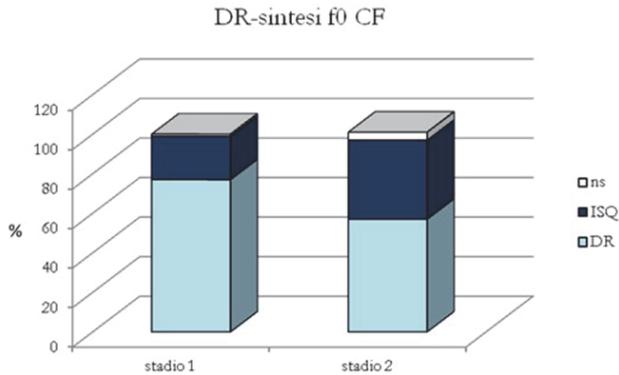


Il quadro descrittivo relativo alle DR presenta diverse problematiche (Fig. 5). La manipolazione del contorno finale operato sulla DR, da discendente (L%) ad ascendente (H%), produce risultati ambigui, il confronto interno tra primo e secondo stadio di sintesi non è significativo ($p=0,483$). Un aspetto che merita di essere messo in luce riguarda la scarsa efficacia percettiva operata dal primo stadio di manipolazione sintetica: nel 77% dei casi la DR continua a essere classificata tale. Nello stadio di sintesi successivo la situazione è lievemente diversa, in oltre la metà del-

¹⁷ La presenza in sillaba chiusa di una vocale allungata, percepita come marcata, si è dunque caricata di un significato interrogativo non prototipico. Quest'aspetto richiama l'annosa questione sul condizionamento esistente, in una lingua come l'italiano a isocronia sillabica, tra durata vocalica e struttura sillabica, come messo in luce da un revisore, che ringraziamo.

le risposte fornite dagli uditori (57%) il giudizio espresso risulta immutato, anche quando associato, per effetto della manipolazione, a un *boundary tone* ascendente. In questa condizione sperimentale, il cambio di valutazione il più delle volte non avviene, mettendo in dubbio la portata distintiva di questo parametro ai fini dell'identificazione percettiva di una DR. Il 4% degli ascoltatori sospende il giudizio selezionando l'opzione *Non so rispondere*. Ne consegue, pertanto, che l'associazione di un *pattern* ascendente finale, in buona parte degli ascoltatori, non è sentito come estraneo alla natura della DR, probabilmente per il fatto che tale profilo è comunque presente, seppure con percentuali minori, tra le possibili realizzazioni intonative della domanda.

Figura 5 - Valori percentuali rinvenuti per la DR



3.2.3 Test 2: la sintesi della durata VN e del CF

La manipolazione congiunta dei due parametri induce su ambedue le categorie frasali uno spostamento di giudizio. Ma ancora una volta è il secondo stadio a produrre risultati più incisivi (Figure 6 e 7). Nel complesso, le percentuali di riconoscimento della DR si innalzano, quelle computate per la ISQ si abbassano; allo stesso tempo crescono, per quest'ultima categoria, anche i valori relativi alla terza opzione (*non so rispondere*, 8%), un dato che sembra testimoniare come negli enunciati bersaglio di tipo ISQ la sintesi di entrambi i parametri abbia trasmesso, durante la procedura di ascolto, anche un certo grado di innaturalità. Ciò non è sorprendente se consideriamo che la manipolazione ha agito contemporaneamente sulla realizzazione acustica di due indici fonetici rilevanti ai fini della decodifica pragmatica dell'enunciato, cioè la durata della VN e l'andamento del tono di confine, determinando combinazioni fonetiche in un certo senso atipiche, talora anche forzate.

Figura 6 - Valori percentuali rinvenuti per la ISQ

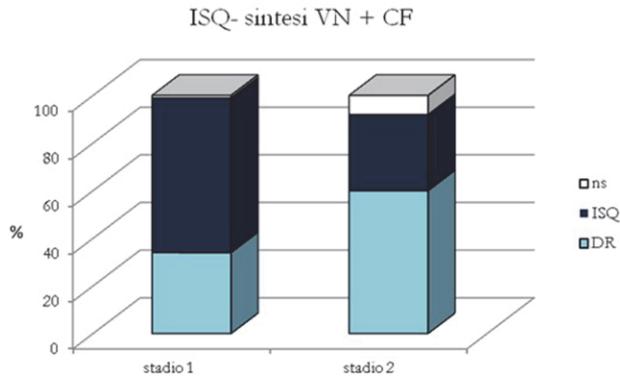
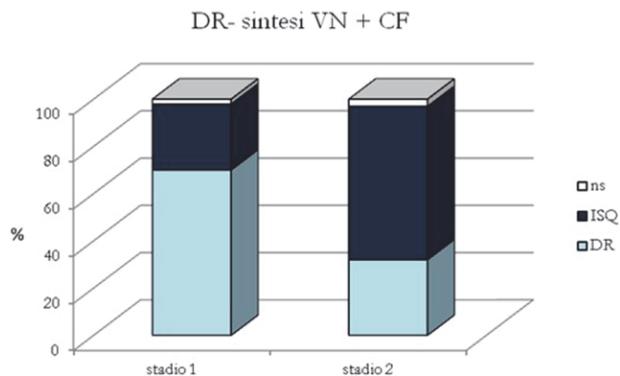
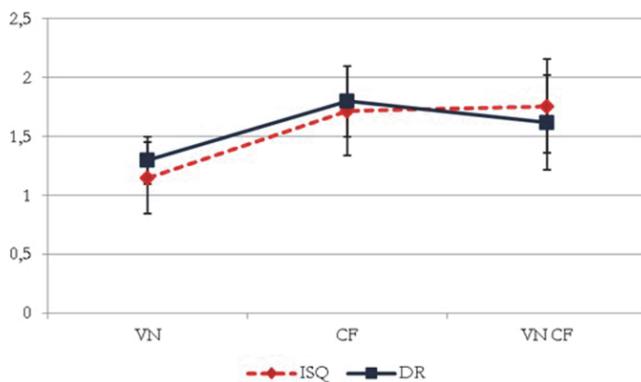


Figura 7 - Valori percentuali rinvenuti per la DR



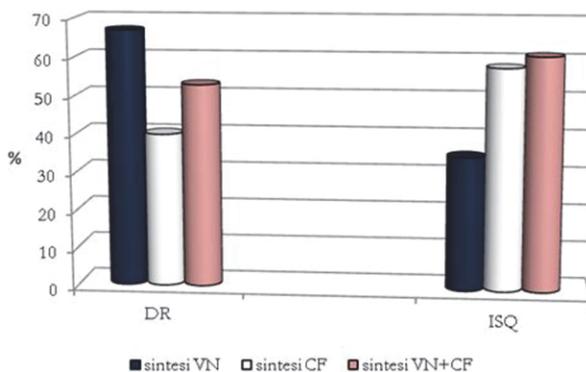
Il computo della media delle risposte fornite dagli uditori per il secondo stadio di manipolazione mostra alcuni comportamenti peculiari (Figura 8): nella ISQ la sintesi di VN non sposta il giudizio degli ascoltatori ($p=0,82$) che pertanto rimane prossimo al valore 1, cioè la codifica assegnata alla domanda ISQ, mentre spinge verso il basso il giudizio delle DR ($p= 0,022$). Un comportamento speculare si osserva per la sintesi del CF, in questo caso mentre le risposte ottenute per le DR, la cui codifica è pari a 2, mantengono sostanzialmente un'interpretazione retorica, poco variata, quelle delle ISQ si allontanano dal loro valore originale, spostandosi verso l'area occupata dai significati retorici; le risposte sono significative a un confronto sia tra ISQ e DR ($p= 0,011$) sia tra le due condizioni di sintesi (ISQ_VN vs ISQ_CF: $p= 0,001$; DR-VN vs DR_CF: $p= 0,029$). La tendenza è ancora più evidente nella terza condizione di sintesi. In tale set sperimentale le risposte ottenute per DR e ISQ mostrano però una certa vischiosità e anche una maggiore dispersione dei valori; il confronto globale tra tutte le risposte è significativo [$F=(2,538)=14,730$; $p=0,000$], ma non il confronto interno tra CF e VN+CF ($p=0,607$).

Figura 8 - Valori medi delle risposte ottenute e indice di errore per ISQ e DR nelle tre condizioni di sintesi (stadio 2)



In termini percentuali, se l'enunciato di partenza è una DR il 65% degli ascoltatori ritiene che lo stimolo ascoltato sia una ISQ e non più una retorica, il 32% di essi però continua a percepire una DR, mentre il 3% non sa rispondere. Viceversa, quando la domanda di origine è una ISQ, nel 32% il cambio di giudizio è sospeso, la domanda continua infatti a essere decodificata nella sua veste pragmatica originale. I due tipi di domanda si comportano diversamente. Nella Figura 9 è riportato in un quadro d'insieme il *trend* rilevato per le tre batterie di stimoli sintetici. L'immagine riporta il tasso percentuale di variazione di giudizio categoriale emerso nelle tre condizioni sperimentali al secondo stadio di sintesi.

Figura 9 - Valori percentuali relativi al cambio di giudizio rinvenuti per ISQ e DR nelle tre condizioni di sintesi (stadio 2)



Come è evidente, nelle ISQ il grado percentuale di deviazione percettiva è linearmente ascendente, cioè basso nella prima condizione sperimentale (35%) che rappresenta la sintesi della durata della VN, ma progressivamente maggiore nelle altre

due, rispettivamente sintesi del CF e sintesi congiunta di VN e CF. Questi ultimi due processi di sintesi alterano in modo significativo la struttura percettiva della ISQ, tanto da non renderla più identificabile nella sua funzione interrogativa e confondendola con la tipologia retorica, come emerge dalla maggior parte delle risposte espresse dagli uditori.

Dagli esiti ottenuti si evince che le tre condizioni di sintesi sono inoltre suscettibili di essere ordinate gerarchicamente: se la manipolazione della durata sembra incidere piuttosto poco sull'identificazione della domanda polare epistemicamente sincera, diverso è il discorso per la manipolazione del CF che induce negli ascoltatori un significativo salto di giudizio. Questo vuol dire che nel tipo ISQ il contorno finale sia un parametro più robusto nella trasmissione degli indici illocutivi, laddove la durata si prefigura come un indice debolmente incisivo.

Lo scenario che si delinea per le DR è diverso: la durata sembra rappresentare l'indice più saliente ai fini dell'interpretazione pragmatica dell'interrogativa, in quanto riesce a fuorviare, deviandola (67%), la valutazione percettiva dell'ascoltatore. La modifica attuata sul CF produce un minore impatto sulla sua resa percettiva, mentre la sintesi aggregata di durata e f_0 , rialza le percentuali di deviazione, sebbene rimangono più basse rispetto alla manipolazione della sola VN. Nel complesso, i confronti tra ISQ e DR e tipo di sintesi raggiungono la soglia di validità statistica, rispettivamente ISQ [$F=(2,302)=7,310$; $p=0,001$]; DR [$F=(2,304)=3,632$; $p=0,028$]; ad eccezione delle seguenti condizioni: tipo ISQ=CF\VN+CF ($t=0,232$); tipo DR=VN\VN+CF ($t=0,569$).

4. *Discussione*

I risultati raccolti mostrano tendenze diversificate la cui interpretazione deve essere ricondotta a piani diversi. I fattori coinvolti nel disegno sperimentale sono essenzialmente di tre ordini: 1) il tipo di interrogativa polare coinvolta: DR o ISQ, 2) il tipo di stimolo audio sottoposto all'attenzione degli uditori: naturale o sintetico, 3) il tipo di sintesi acustica effettuata e, di conseguenza, il tipo di parametro fonetico testato: durata VN o CF.

In merito al primo punto, c'è un aspetto che non può passare inosservato: DR e ISQ presentano delle tendenze specifiche non sovrapponibili. Questa forbice, che si delinea già nella prima condizione sperimentale (Test 1), assume proporzioni maggiori nella seconda sperimentazione (Test 2) quando precipui indici acustici, opportunamente selezionati e manipolati sinteticamente, entrano in gioco nella dinamica percettiva degli stimoli, inducendo negli ascoltatori frequenti salti di giudizio categoriale, non sempre prevedibili, almeno per ciò che riguarda la DR.

Il diverso comportamento di DR e ISQ si ripercuote a cascata sugli altri aspetti determinando continui condizionamenti.

Un elemento differenziante riguarda il *range* di corretta identificazione degli stimoli naturali (Test 1), un indice che varia, in modo statisticamente significativo, a seconda del significato pragmatico della domanda, retorica o non retorica. Il

fatto che le ISQ siano riconosciute con percentuali più alte rispetto alle DR apre la strada ad alcune riflessioni. Vale la pena ricordare che gli stimoli naturali sottoposti all'attenzione degli ascoltatori, estrapolati da contesti dialogici elicitati in modo sorvegliato, costituiscono delle coppie minime frasali, diverse solo per valore comunicativo, ma identiche per composizione lessicale e morfosintattica. Estratti dalla loro cornice testuale, gli stimoli perdono quell'indispensabile aggancio contestuale che, sebbene costruito artificialmente, è comunque adeguato alla trasmissione del loro contenuto pragmatico. Ciò che resta è, dunque, un enunciato provvisto della sola veste prosodica. Stando ai dati ottenuti nel corso del Test 1, gli stimoli somministrati riescono comunque, con percentuali piuttosto alte, a proiettare durante l'ascolto i propri tratti specifici, favorendo la corretta identificazione delle domande. Questo vuol dire che la realizzazione sonora di DR e ISQ, anche quando privata da ogni informazione contestuale, conserva sufficienti specificità prosodiche tali da influenzare positivamente, in percezione, il riconoscimento e il recupero del loro significato pragmatico. L'ipotesi da noi formulata e sottesa a questa prima fase sperimentale è in larga parte confermata, visto che DR e ISQ conservano il più delle volte una propria identità sul piano prosodico. Pur nella validità di questo assunto, va tuttavia osservato come le ISQ siano state nel complesso meglio identificate rispetto alle DR. È ragionevole sostenere, anche alla luce dei numerosi tratti caratterizzanti (§ 1), che la DR sia una tipologia interrogativa debole, maggiormente dipendente dal contesto. Ciò è vero sul piano informativo: la DR non è prodotta con l'intenzione di aggiungere o di ricevere informazioni nuove, è infatti una domanda fittizia. Totalmente priva di peso informativo, è per lo più volta alla risincronizzazione del dialogo, anche attraverso l'esternazione di alcune sue funzioni tipiche, come il rafforzamento o la mitigazione della forza illocutoria, rese possibili anche attraverso l'adozione di una *facies* ironica. Sul versante pragmatico, la DR è, invece, un enunciato indiretto dal significato presupposto che si estrinseca all'interno di un *common ground*, un'area di conoscenze condivise a cui i partecipanti alla comunicazione fanno continuo riferimento tramite inferenza. Sulla base di questa premessa, è dunque lecito assumere che nella concezione cognitiva del parlante la DR non abbia uno statuto prototipico, diversamente dalla ISQ la quale, nelle sue possibili declinazioni strutturali e nelle sue numerose varianti sociolinguistiche, è ampiamente riconosciuta come atto rappresentativo della categoria interrogativa. Se la ISQ si colloca idealmente come elemento focale e saliente della classe degli atti interrogativi, la DR, essendo fondamentalmente una non-domanda che trae sostanza dal contesto e da quelle specifiche e continue relazioni che intrattengono emittente e ricevente nella dinamica comunicativa, è ragionevolmente da considerarsi un enunciato cognitivamente atipico che occupa uno spazio categoriale periferico, una condizione congruente con la sua natura indiretta.

Un ulteriore aspetto degno di osservazione riguarda poi il comportamento degli stimoli di DR e ISQ sottoposti a manipolazione acustica. In merito a questo punto, una preliminare precisazione è doverosa: nei test eseguiti la sintesi non era mirata a creare equivalenza percettiva tra enunciati. Di conseguenza, altri parametri hanno

continuato ad agire sulle domande, sovrapponendosi a quelli modificati. Pur con questo distinguo, va rilevato come, anche in questo caso, le DR presentano tendenze parzialmente diverse dalle ISQ. Non solo la sintesi dei due punti chiave selezionati, durata VN e CF, produce comportamenti differenti, ma cambiano anche le percentuali ottenute in fase di giudizio percettivo. I due parametri investigati sembrano avere un ruolo diverso nel processo di decodifica uditiva delle domande. Di conseguenza, la nostra seconda ipotesi è dimostrata solo parzialmente; trattandosi di un atto linguistico multidimensionale è altresì possibile che l'esplorazione della DR debba avvenire mediante protocolli sperimentali più potenti rispetto a quello da noi utilizzato capaci di cogliere, in modo più fine, la complessa interazione dei fattori fonetici. Fatta questa precisazione, dalla ricerca emerge che nelle DR la modifica del CF non si è rivelata molto incisiva, l'ascolto di una DR con tono di confine H% non è stato infatti giudicato, non sempre almeno, innaturale; ne è prova il fatto che molti uditori hanno continuato a percepire una DR anche quando dotata di un contorno sintetico. Diversamente, in queste domande la durata ha agito con più forza, in quanto capace di virare in modo significativo il giudizio verso un'interpretazione letterale della domanda. Questo vuol dire che nella produzione naturale l'allungamento temporale di alcuni punti chiave strategici, come per l'appunto la vocale nucleare, sia effettivamente un elemento caratterizzante una DR, come d'altra parte rinvenuto anche in altre lingue (Miura, Hara, 1995; Nietsch, Niebuhr, 2019). La durata si rivela ritenersi un efficace predittore di illocuzione retorica; sarà utile osservare come la trasmissione di molti sensi paralinguistici, non solo l'ovvietà, ma anche il tono canzonatorio o il sarcasmo sono affidati a precipui, quanto mirati, allungamenti temporali.

Nelle ISQ è invece il CF l'indice più sensibile ai fini della sua categorizzazione percettiva, lo dimostra da un lato il fatto che la modifica tramite sintesi della durata vocalica ha prodotto un effetto piuttosto blando, dall'altro il rilevante mutamento di giudizio indotto dalla variazione del tono di confine. L'interpretazione di questo risultato rimane però cruciale. Nell'italiano barese, l'ISQ è infatti associata sia a un CF discendente, è quanto accade soprattutto nell'eloquio spontaneo, sia a un CF ascendente; ciò avviene tipicamente nel parlato controllato, ma anche ogni volta che la domanda intende sollevare un maggiore coinvolgimento personale. C'è da aggiungere che la modifica del CF ha determinato una concomitante variazione dell'escursione tonale delle domande; precisamente, nelle DR il *pitch range* ha subito un incremento poiché si è passati dal tono L% a uno H%, nelle ISQ, invece, il passaggio dal tono H% a quello L% ha causato una riduzione del *range* melodico. L'effetto percettivo è stato però diverso: nelle ISQ la presenza, operata tramite sintesi, del *boundary tone* discendente ha cambiato il giudizio degli uditori che hanno spesso associato tale resa acustica a una DR; ciò non è avvenuto, in modo speculare, nelle DR per le quali il salto di giudizio è stato più ridotto. D'altra parte, tale tendenza è congruente con il significato pragmatico veicolato rispettivamente dai toni di confine L% e H%. Come più volte messo in luce, la presenza di un andamento finale basso o discendente, in trascrizione ToBI L%, è legato a un significato di assertività e di conclusione, e quindi

di ovvietà e presupposizione, laddove l'andamento alto o ascendente, H%, è notoriamente collegato alla trasmissione di un senso di interrogazione e di continuazione (Cruttenden, 1981; Pierrehumbert, Hirschberg, 1990).

La ricerca, la prima incentrata sulla percezione delle DR della lingua italiana, incrementa le nostre conoscenze di questo particolare atto linguistico. Pur tuttavia, il quadro descrittivo, tutt'altro che esaustivo e a tratti problematico, necessita di essere integrato in altre direzioni. Innanzitutto, coglie l'obbligo precisare che si tratta di un esperimento pilota che ha indagato solo due dei numerosi indici fonetici che entrano in gioco nella complessa dinamica della categorizzazione percettiva di un significato retorico. La manipolazione degli indici selezionati (VN e CF) è stata effettuata in modo lineare e non ha previsto il controllo degli effetti concomitanti generati sugli altri parametri; la sintesi ha inoltre coinvolto solo la porzione finale delle domande, lasciando pressoché inalterato il resto della frase. Allo stadio attuale non è quindi possibile escludere l'azione di altri fattori nel processo di categorizzazione percettiva delle domande, come ad esempio il contorno prenucleare o l'influenza svolta dalla qualità della voce. Oltre a queste doverose cautele di ordine metodologico, perseguendo una prospettiva ecolinguistica si avverte da un lato l'esigenza di estendere l'analisi ad altre varietà di italiano, al fine di cogliere la diversità linguistica del territorio, evitando una visione locale, dall'altro la necessità di considerare anche le tendenze riscontrabili dall'inevitabile interazione tra parlanti, lingua e ambiente, accogliendo quindi forme di eloquio spontanee.

Bibliografia

- ANZILOTTI, G.I. (1982). The rhetorical question as an indirect speech device in English and Italian. In *Canadian Modern Language Review*, 38, 290-302.
- ASU, E.L., SAHKAI, H., LIPPUS, P. (2020). The prosody of rhetorical and information-seeking questions in Estonian: preliminary results. In *Proceeding of the 10th International Conference on Speech Prosody*, 25-28 May 2020, Tokyo, Japan, 381-384.
- BANUAZIZI, A., CRESWELL, C. (1999). Is that a real question?: Final rises, final falls and discourse function in yes-no question intonation. In *Proceedings of the 35th Regional Meeting of the Chicago Linguistics Society* (CLS 35), Chicago Linguistic Society, 1-14.
- BARTELS, S.C. (2013). *The intonation of English statements and questions. A compositional interpretation*. Routledge, New York & London: Garland Publishing, 1 ediz. 1999.
- BOERSMA, P., WEENINK, D. (2013). *Praat: doing phonetics by computer*, version 5.3.51, <http://www.praat.org>, accessed 2 June 2017.
- BRAUN, B., EINFELDT, M., ESPOSITO, G., DEHÉ, N. (2020). The prosodic realization of rhetorical and information-seeking questions in German spontaneous speech. In *Proceeding of the 10th International Conference on Speech Prosody*, 25-28 May 2020, Tokyo, Japan, 342-346.
- BROWN, P., LEVINSON, S. (1978). Universals in language usage: Politeness phenomena. In GOODY, E. (Ed.), *Questions and politeness: strategies in social interaction*, Cambridge: Cambridge University Press, 56-311.

- CAPONIGRO, I., SPROUSE, J. (2007). Rhetorical questions as questions. In PUIG-WALDMULLER, E. (Ed.), *Proceedings of Sinn und Bedeutung*, 11, Barcelona, Spain: Universitat Pompeu Fabra, 121-133.
- CRUTTENDEN, A. (1981). Falls and rises: meaning and universals. In *Journal of Linguistics*, 17, 77-91.
- DEHÉ, N., BRAUN, B., WOCHNER, D. (2018). The prosody of rhetorical vs. information-seeking questions in Icelandic. In KLESSA, K., BACHAN, J., WAGNER, A., KARPIŃSKI, M., ŚLEDZIŃSKI, D. (Eds.), *Proceedings of the 9th International Conference on Speech Prosody*, Poznań (Poland), 403-407.
- FRANK, J. (1990). You call that a rhetorical question? Forms and functions of rhetorical questions in conversation. In *Journal of Pragmatics*, 14 (5), 723-738.
- GILI FIVELA, B., AVESANI, C., BARONE, M., BOCCI, G., CROCCO, C., D'IMPERIO M.P., GIORDANO, R., MAROTTA, G., SAVINO, M., SORIANELLO, P. (2015). Intonational phonology of the regional varieties of Italian. In FROTA, S., PRIETO, P. (Eds.), *Intonation in Romance*, Oxford: Oxford University Press, 140-197.
- GUTIÉRREZ-REXACH, J. (1998). Rhetorical question, relevance and scales. In *Revista Alicantina de Estudios Ingleses*, 11, 139-155.
- JUNG, V., SCHROTT, A. (2003). A question of time? Question types and speech act shifts from a historical-contrastive perspective: some examples from Old Spanish and Middle English. In JASCZOLT, K.M., TURNER, K. (Eds.) *Meaning Through Language Contrast*. Amsterdam: John Benjamins, 345-371.
- HAN, C. (2002). Interpreting interrogatives as rhetorical questions. In *Lingua*, 112, 112-229.
- KOSHIK, I. (2005). *Beyond rhetorical question. Assertive questions in everyday conversation*, Amsterdam: John Benjamins.
- ILIE, C. (1994). What else can I tell you? A pragmatic study of English rhetorical questions as discursive and argumentative acts. In *Stockholm Studies in English*, vol. 72. Stockholm: Almqvist & Wiksell International.
- MIURA, I., HARA, N. (1995). Production and perception of rhetorical questions in Osaka Japanese. In *Journal of Phonetics*, 23, 291-303.
- NEITSCH, J., NIEBUHR, O. (2019). Questions as prosodic configurations: how prosody and context shape the multiparametric acoustic nature of rhetorical questions in German. In *Proceedings of the 19th International Congress of Phonetic Sciences*, Melbourne, Australia 2019, 2425-2429.
- PIERREHUMBERT, J.B., HIRSCHBERG, J. (1990). The meaning of intonational contours in the interpretation of discourse. In COHEN, P.R., MORGAN, J., POLLACK, M.E. (Eds.), *Intentions in communication*, Cambridge (MA): MIT Press, 271-311.
- ROHDE, H. (2006). Rhetorical questions as redundant interrogatives. In *San Diego Linguistics Paper*, 2, San Diego: University of California, 134-168.
- ROOY, R. VAN (2003). Negative polarity items in questions: Strength as relevance. In *Journal of Semantics*, 20(3), 239-273.
- SADOCK, J.M. (1971). Queclaratives. In *Papers from the Seventh Regional Meeting of the Chicago Linguistics Society*, (CLS 7), Chicago: Linguistics Society, 223-231.

- SAVINO, M. (2012). The intonation of polar questions in Italian: where is the rise?. In *Journal of the International Phonetic Association*, 42(1): 23-48.
- SORIANELLO, P. (2018). Tra prosodia e pragmatica: il caso delle domande retoriche. In *Studi e Saggi Linguistici*, LVI (2), 39-71.
- SORIANELLO, P. (2019). 'A che serve saperlo?' Funzioni pragmatiche e variazioni intonative della domanda retorica. In Nuzzo, E., Vedder, I. (Eds.), *Lingue in contesto: la prospettiva pragmatica*, Studi AItLA 8, Milano: Officinaventuno, 89-108.
- STATI, S. (1982). Le frasi interrogative retoriche. In *Lingua e Stile*, XVII(2), 195-207.
- WOCHNER, D., SCHLEGEL, J., DEHÉ, N., BRAUN, B. (2015). The prosodic marking of rhetorical questions in German. In *Interspeech 2015, 16th Annual Conference of the international speech communication association, Dresden, Germany, September 6-10, 2015*, 987-991.
- ZAHNER, K., MANLUOLAN, X., CHEN, Y., DEHÉ, N., BRAUN, B. (2020). The prosodic markers of rhetorical questions in Standard Chinese. In *Proceeding of the 10th International Conference on Speech Prosody*, 25-28 May 2020, Tokyo, Japan, 389-393.

RICCARDO ORRICO, VIOLETTA CATALDO, MARIAPAOLA D'IMPERIO

The effect of early exposure in the production of Salerno Italian question intonation¹

The study explores the effect of early language exposure in the production of question intonation in Salerno Italian. In particular, we tested the hypothesis that exposure to other varieties represents a possible source of variability in nuclear tune choice. Ten speakers from Salerno were selected and divided into two groups according to whether or not they had at least a non-native parent. Information-seeking yes-no questions and information-seeking wh-questions were elicited using a Reading Task. Results show that part of the variability found in the use of intonation can be attributed to the presence vs absence of early exposure. Specifically, such results concern differences in the distribution of tunes mirroring that of the variety to which speakers were exposed. Finally, the study highlights the need of taking into account different sources of variation when investigating a language's intonational grammar.

Keywords: question intonation, Salerno Italian, dialect contact, intonational variation.

1. Introduction

Traditionally, phonology has been conceived in a completely abstract manner, as a way to provide generalization over variation (e.g., Halle, 1985; Lahiri, Marslen-Wilson, 1991). According to this view, variation within the speech signal is considered to be useless noise which is systematically filtered out by listeners. Nevertheless, a long line of research initiated by Goldinger (1998; 2000) has shown that not only does such variation exist, but it is also stored in memory and affects the way speakers produce and perceive future utterances. Indeed, several studies show that an individual's system can adapt to the variability that is present in the input (Pierrehumbert, 2016). Goldinger (1998), for example, using a shadowing task, showed a systematic shifting of speaker productions towards the stimulus, in terms of general acoustic properties. Later research showed that such mechanisms are much more automatic and unintentional than shadow tasks seem to indicate. Delvaux and Soquet (2007), for example, found that the presence of a non-native dialect as ambient noise during a recording session affected the way vowels were produced by participants, and that, crucially, the effect lasted for several minutes after the exposure.

¹ This paper is the result of a close collaboration of the three authors. Specifically, Riccardo Orrico wrote sections 2 and 4, with relative subsections; Violetta Cataldo wrote sections 1 and 3, with relative subsections. Discussion and conclusion (sections 5 and 6) were written by all three authors. Mariapaola D'Imperio also supervised and provided corrections for the writing of all the parts of the paper.

General patterns of variation in the production and perception of intonation have also been attested (Mennen, 2004; Cangemi, Krüger & Grice, 2015; Orrico, D'Imperio, 2020a). Additionally, a number of recent studies have shown that speakers are able to accommodate their intonational contours to other varieties of their native languages in a great deal of detail, both in direct imitation experiments (Cole, Shattuck-Hufnagel, 2011; D'Imperio, German, 2015, and German, 2012 for different varieties of English; D'Imperio, Cavone & Petrone, 2014 for varieties of Italian) and in semi-spontaneous interactions (Savino, 2017). Research has also shown the impact of two major factors affecting the way people use language, i.e. cognitive traits (Kidd, Donnelly & Christiansen, 2018) and linguistic environment. As for the role of cognitive traits, recent studies have investigated how they shape the way language is processed by an individual. Among cognitive factors, we find working memory capacity (Yu, Grove, Martinovic & Sonderegger, 2011), attentional abilities (Kim, Hazan, 2010; Kong, Lee, 2018), autistic-like traits in neurotypical population (Yu et al., 2011; Jun, Bishop, 2015) and, more recently, empathy (Esteve-Gibert, Schafer, Hemforth, Portes, Pozniak & D'Imperio, 2020; Orrico, D'Imperio, 2020a).

Moreover, the linguistic environment, which refers to the linguistic input that an individual is exposed to during her life, shapes the phonological system of individuals. The present study is concerned with this source of variability and, more specifically, with early exposure to a non-native variety.

The effects of a certain degree of exposure to a non-native input are well known as far as L2 acquisition is concerned. Second Language Acquisition (SLA) models such as the Speech Learning Model (SLM; Flege, 1995) and the Perceptual Assimilation Model-L2 (PAM-L2; Best, Tyler, 2007) claim that non-native sounds are perceived, and consequently more or less likely to be acquired, with reference to the similarity or dissimilarity with the speaker's native sounds. Crucially, a highly variable L2 input, in terms of multiple speakers, accents, and phonetic contexts (Best, 2015) is responsible for achieving phonological distinctiveness and constancy (Best, Tyler, Gooding, Orlando & Quann, 2009). Such phonetic variation in the L2 input is considered as responsible for learning advantages in both beginners and advanced learners (Iverson, Pinet & Evans, 2012).

These models succeed to predict the likelihood of acquiring L2 phonetic categories making reference to an acquisition situation which typically involves a condition of L2 immersion in L2-dominant environments. Similar effects are pointed out within the Foreign Language Learning (FLL) framework as well. Indeed, even in a classroom setting, which typically takes place in an L1-dominant environment, a key factor appears to be the exposure to a rich and phonetically variable L2 input (Tyler, 2019).

Clearly, as noticed by Bohn and Bundgaard-Nielsen (2009), learners may vary greatly in their degree of previous exposure to the foreign language (FL); indeed, exposure appears to be intertwined with other individual factors, namely the linguistic environment within the family, prior FL classroom instruction, exposure to film or TV, periods of study abroad.

On the other hand, SLA and FLL studies normally set the native monolingual norm as the point of reference learners may achieve as improving their L2/FL proficiency. The target for learners in the Common European Framework of Reference for Languages (CEFR; Council of Europe, 2001) is represented by a proficiency level which can be compared to a native-speaker or near native-speaker competence. But this native standard or the consequent notion of “nativeness” are rarely unproblematic for both L2 and L1 studies. Indeed, several factors need to be considered as potential sources of variability even when dealing with monolingual speakers. Among these, consistent exposure to non-native diatopic varieties should also be taken into account as a possible source of variability. In this regard, Romera and Elordieta (2013) identify a process of accommodation to an L2 variety of Spanish (Majorcan Catalan) by adult monolingual speakers of Peninsular Spanish; in particular, transfer phenomena at the intonational level from the L2 variety to the native one lead to such accommodation. The authors highlight that, different from between-language settings, interference phenomena are likely to take place in situations of contact between language varieties, since the latter are structurally more similar and might only differ from each other in terms of prosodic and/or intonational features.

2. *Intonational variation in Italian*

Starting from the 1990s, a growing interest in the intonational phonology of Italian has spurred, and, to date, detailed information is available for a number of varieties. A great deal of studies has been conducted on Southern varieties: Neapolitan Italian (*inter alia*, D’Imperio, 1999; 2000; 2002), Bari Italian (Grice, Savino, 1997; Gili Fivela, Savino, 2003; Savino, Grice, 2011), Palermo Italian (Grice, 1991; 1995; Grice, D’Imperio, Savino & Avesani, 2005 and, more recently, Gili Fivela, Iraci, 2017) and Salerno Italian (Orrico, 2020, see §2.1). Furthermore, a large number of studies has been published also for the varieties spoken in Tuscany, especially those spoken in Pisa, Florence, and Siena (e.g., Marotta Sorianello, 2001; Avesani Vayra, 2003; Gili Fivela, 2008). Several studies have also been dedicated to the comparison of different varieties. Within the AM approach, Grice et al. (2005) includes four regional varieties (Naples, Bari, Florence, and Palermo) by covering commonalities and differences in intonational inventories. A greater number of varieties have been reviewed in Savino (2012) and Crocco (2013), in which an analysis of question intonation across Italy was reported. Finally, Gili Fivela and colleagues (Gili Fivela, Avesani, Barone, Bocci, Crocco, D’Imperio, Giordano, Marotta, Savino & Sorianello, 2015), offer a description including thirteen varieties, underlying the strong variation across Italian intonational systems.

While earlier studies (e.g., Grice et al., 2005) have been conducted with the aim of clustering together different varieties by areal features, more recent ones have failed to find support for a general geographical classification. This is the case of Savino (2012), which puts into question the traditional point of view classifying

Northern varieties as expressing yes-no questions with a final rise, while Southern ones would make use of nuclear pitch accent alone. This represents a major starting point for understanding the source of variation within Italian varieties and how to account for it. However, Savino's treatment of the final rise in questions in some Italian varieties as a mere stylistic choice (e.g., read speech, without marked pragmatic context), appears too simplistic and experimental support for this argument is needed.

Additionally, a crucial aspect that recent studies on Italian intonation have brought to the attention of the research community is the intra-varietal variability. As an example, within the data analyzed by Gili Fivela and colleagues, consistent intra-varietal contours have been found only for some very specific functions (e.g., counter-expectational wh-questions), while others present high levels of variety internal variability, especially in the case of yes-no questions. Gili Fivela (2008) in her description of the intonational system of Pisa Italian had already proposed that a one-to-one mapping between a specific pragmatic category and a specific tune cannot be found. Extending the scope of the research to a much higher number of varieties, similar results have been obtained. Additionally, a point that might be inferred from Gili Fivela et al.'s (2015) description, even if not explicitly addressed in that study, is the impact of speaker-specific use of intonation.

What is more, a point that is particularly stressed in Gili Fivela et al. (2015) is that while some phonological features might be shared by geographically distant varieties, at the same time, differences might be found across varieties spoken within the same region. This points out to the fact that behind the uncontroversial differences across varieties in both their phonological nature of intonational events and their phonetic implementation, there is a phenomenon of 'mixing of patterns' throughout the varieties. The authors speculate that this might be due, at least to a certain extent, to a process of interference among varieties. In this respect, the study of Gili Fivela and Nicora (2018) on Liguria and Tuscany varieties, takes a first step in the investigation of varieties of spoken Italian in contact conditions. The relation among the varieties, in terms of phonetic similarity, is explored comparing the La Spezia variety with other geographically neighboring varieties, namely Imperia, Genoa, Florence and Pisa Italian. The area under investigation is considered as being divided by an isogloss detected for vernaculars. Particular attention is paid to contact among the varieties and to the process by which those varieties may have influenced each other in terms of intonational features. The authors identify a possible common origin for information-seeking yes-no questions on the phonetic level, which leads them to suggest the need to abandon the idea of intonational isoglosses.

2.1 Salerno Italian (SI)

With specific reference to the variety spoken in Salerno, recent investigations have unveiled a strong intra-varietal variability in the way specific tunes are used to express and identify specific meanings (Orrico, Savy & D'Imperio, 2019a; Orrico, 2020; Orrico, D'Imperio, 2020a; Orrico, D'Imperio, 2020b). In fact, the picture yielded by analyses

focusing on Salerno Italian does not differ much from the general picture of Italian varieties. Specifically, these studies show that, on the one side, there are functions (such as broad and narrow focus statements) that are produced by Salerno Italian speakers using specific tunes in a consistent way, with very little variation reported. On the other side, *wh*- and *yes-no* questions, for instance, appear to show significant variation so that, in some specific cases, a general predictable pattern is hard to define. Higher levels of variation, however, have been attested for SI *yes-no* as opposed to *wh*- questions, hence different pictures should be drawn for the two modalities.

In the case of *wh*-questions, a clearer pattern emerges from the tune-function mapping. Specifically, as reported in Orrico and D'Imperio (2020b), the use of pitch accents in nuclear position is a strong predictor of the expression of speaker commitment to salient propositions evoked by the *wh*-questions, specifically with falling ($H+L^*$) accents indicating the presence of speaker commitment, while rising (L^*+H) accents would indicate that the speaker is rejecting such commitment. A different scenario is however registered for boundary tones, which do not appear to be dependent on specific pragmatic meanings, but on speaker-specific choice.

Moreover, results for SI *yes-no* questions show the absence of a one-to-one form-meaning mapping, confirming what has been argued by Gili Fivela et al. (2015) for several Italian varieties. Specifically, different tunes appear to be used to express different pragmatic functions in SI questions (i.e., information-seeking, confirmation-seeking, echo, and counter-expectational questions), with the most frequent nuclear tune being a rise-fall-rise, notationally expressed as a $L+H^* HL-H\%$. Significantly, this specific tune is the most frequent one in all conditions, except for counter-expectational instances. The function expressed, therefore, does not allow to establish a stable correspondence between tune and meaning, rather it appears that variability depends on individual speakers.

In fact, the analysis of individual speaker behavior shows that some of the variation can be explained by allowing intonational meaning models to take into account speaker heterogeneity within a language community. Even if speakers share their native language, they are characterized by different socio-indexical features and have been exposed to different linguistic input interacting with their common phonological system. Much more clarity regarding this point has been reached in perception. In fact, perception experiments conducted by Orrico and colleagues (Orrico, Savy, & D'Imperio, 2019b; Orrico, D'Imperio, 2020a) have been specifically designed to test for group-based differences in the identification of the epistemic bias in question tunes². Despite the difficulties of measuring the degree of exposure of an individual to other systems (see §1 for a discussion about this), the participants were divided into two groups according to whether or not they experienced exposure by living in other cities for prolonged periods of time or by being raised by non-SI

² Biased questions are non-canonical questions, i.e. they depart from the norm by not only asking for information, but they also convey an epistemic bias (i.e., an expectation) about the truth of the proposition p expressed in the question. The bias conveyed in a *yes-no* question can be positive or negative, depending on whether the speaker is expressing an expectation towards the truth of p or the truth of $\neg p$.

families. Results show that listeners without those specific experience of exposure classify different types of question bias on the basis of boundary tone type (H% or L%), while listeners who underwent those types of exposure base their judgments on pitch accent type (L+H* or L*+H).

While perception results allow us to classify tune-meaning mapping according to specific grouping of listeners, the production studies reported above failed to find specific generalizable traits which would account for the way different speakers within a language community behave. This means that the question of where production variation comes from and how it could be integrated in a generalizable model is still wide open.

2.2 Variability in questions

This study investigates SI question intonation production (both yes-no and wh-questions) and, more specifically, the impact that early exposure to non-SI varieties has on their realization. This section will review some of the literature on yes-no and wh-questions in Italian varieties, by highlighting differences and commonalities between SI and such varieties.

The literature reviewed in the previous section reports that the highest levels of variability across Italian varieties are encountered for yes-no questions, while for wh-questions, much more stability has been registered. With specific reference to yes-no questions, the main issue in previous studies has been the status of the rising/high boundary tone: since the first investigations, it was made clear that for some varieties, especially Southern ones, the prototypical yes-no question tune is marked by a L% boundary (Grice et al., 2005). However, Savino (2012) and Gili Fivela et al. (2015), both conducting studies on a high number of varieties, attested both versions of terminals in the majority of the varieties under investigation. Despite this observation, quantitative data from Savino (2012) show that many varieties, regardless of geographical position, make a large use of L%. This is the case of Turin, Venice, Parma, Bari, Naples, Catanzaro, and Palermo, all of which show an accentual rise (L+H*/L*+H) ending in a low (L%) boundary tone, as opposed to others, e.g. Bergamo, Milan, Lecce, Perugia, and Cagliari, which prefer rising H% terminals.

Data from both Gili Fivela et al. (2015) and Orrico et al. (2019a) show that the picture depicted for SI is much more variable: both rising and falling boundaries are attested in this variety, with very similar distributions. The same appears to be the case for Genoa, Rome, and Florence. Variability in yes-no question tunes appears to be also related to the nuclear pitch accent, though a greater intra-variety consistency is revealed, as opposed to the variability within boundary realization. Both rising and falling accents have generally been attested in nuclear position: falling accents (H*+L/H+L*) characterize, for example, the varieties spoken in Milan, Bergamo,

Pisa, and Lecce, while accentual rises appear much more frequently in Naples, Bari, Palermo, Venice, Turin, and Salerno³.

Among varieties using a rising accent in nuclear position, an additional divide can be drawn between those using L*+H and those using L+H*. The late-peak variant appears to be typical of Naples, Turin, Venice, and Palermo, while the early-peak variant is mostly attested in Bari. Despite the different phonological notation might be due to variety-internal reasons (see Grice et al., 2005), a review of the literature shows that there is, everything else being equal, (at least) a phonetic difference between the two rising accents, with the Bari Italian one being realized with a peak around the middle of the stressed vowel, while the late-peak variant, used in Venice or Neapolitan Italian, is realized with a peak at the offset of the vowel. Again, as reported above, both rising accents are attested in SI and, despite the distributional differences discussed above, none of them appears to be the most frequent, hence representative of the variety.

As for *wh*-questions, a more consistent use of intonation has been found, both within and across varieties. Gili Fivela et al. (2015) report that the main pitch accent used in this condition is H+L*, which can be followed by either terminal rises or falls. For some varieties, mainly Southern ones, it appears to be possible to have rising pitch accents too, which are generally followed by a rising boundary tone⁴. The same, as shown above, is registered within SI. Rising nuclear tunes have often been linked to specific pragmatic meaning of *wh*-questions and, specifically, with reference to echo and incredulous questions (Gili Fivela et al., 2015; Crocco, Badan, 2016; Orrico, D'Imperio, 2020).

In the present study we aim at exploring the impact of the presence versus absence of early exposure to a non-SI variety in adult SI speakers. The specific hypothesis tested is that a certain degree of non-SI input would affect question intonation variability (both nuclear pitch accent and boundary tone) for both *yes-no* and *wh*-questions. The non-SI varieties targeted in this study are Venice, Foggia, Bari, and Neapolitan Italian.

3. *Method*

The following sections provide information about the dataset used for the study, the participants and the methodology used for the analysis.

³ Both for nuclear pitch accents and boundary tones, difference might be found in the literature with reference to specific varieties. Investigations for Bari Italian, for example, report that also falling accents might be used in questions, which are more commonly used to express a positive bias of the speaker (Savino, Grice, 2011) or high boundary tones in the same variety (Gili Fivela et al., 2015). The same is true for SI: some of the instances of nuclear pitch accents in *yes-no* questions are considered as falling H*+L accents by Gili Fivela et al. (2015). This, however, might be also due to labeller-specific differences.

⁴ Falling terminals after a rising accent have also been attested for southern varieties, as for example in Cosenza and Pescara Italian (Soriano, 2001; Gili Fivela et al., 2015). These varieties, however, show both rising and falling tunes as well.

3.1 Corpus and dataset

For the present production study, a limited dataset coming from two larger corpora has been used. Specifically, these corpora were collected and used in previous studies aimed at investigating the learning process of prosodic features by Italian learners of Spanish and English as Foreign Languages (Savy, Luque Moya, 2014; Orrico, Cataldo, Savy & Barone 2016; Luque Moya, Savy, 2017; Cataldo, Orrico & Savy, 2017). Both corpora include data in Salerno Italian L1.

The dataset used in the present study consists of Italian neutral, information-seeking yes-no questions, with SVO syntactic order (18 items), as shown in (1) below, and information-seeking wh-questions with fronted wh- (9 items), as shown in (2).

- (1) *La rondine ama volare?* [Does the swallow love to fly?]
- (2) *Dove vive la rondine?* [Where does the swallow live?]

Productions were elicited using a Reading Task. Each question type for each target word was inserted in a situational context in order to both ease the naturalness of productions and control the pragmatic modality of utterances. Speakers were asked to read silently the contexts and then to read aloud the target question. Further information about the corpus building and the recording session are provided in Savy, Luque Moya (2014) and Orrico et al. (2016).

3.2 Speakers

Participants had to complete a sociolinguistic questionnaire. We selected 10 subjects, all female students at the University of Salerno, aged from 19 to 25 (mean: 21.5). They were all born and raised in Salerno and had never lived abroad or in other Italian cities. We later divided them into two equally balanced groups according to whether or not they had experienced early exposure to other systems (non-SI varieties of Italian). Specifically, we labelled as *Early-Exposed* (EE) those speakers who had at least one parent that was a non-native Salerno Italian (SI) speaker and as *Non-Early-Exposed* (NEE) all the others. Among the 10 speakers, half belonged to the NEE group. Of those non-native parents, two came from Naples, one from Venice, one from Bari and one from Foggia.

Clearly, during the lifetime, speakers can be explicitly and/or implicitly exposed to or come in contact with different non-native varieties and languages. The degree of such exposure and contact is not easy to quantify. Accordingly, in this study we limit our considerations to the effects of early exposure, i.e. the exposure to a non-SI variety spoken by one of subjects' parents (see §1).

3.3 Analysis

The intonational analysis was carried according to the Autosegmental-Metrical approach (Pierrehumbert, 1980; Ladd, 2008) and annotations were performed using a ToBI-like system (Grice et al., 2005; Gili Fivela et al., 2015, Orrico, 2020). We

specifically labelled nuclear pitch accents and boundary tones. The total amount of productions is 270 (18 items for information-seeking yes-no questions and 9 items for information-seeking wh-questions * 10 speakers).

Given the theoretical framework outlined above, we predict that question tune variability in production is a function of having been exposed to (at least) a non-native variety of Italian. Given the varieties EE speakers were exposed to, we also predict effects for variety-specific exposure. Specifically, for yes-no questions we predict: i) higher percentage of L% boundaries in speakers exposed to Venice, Neapolitan and Bari Italian as opposed to NEE; ii) higher percentage of L*+H pitch accents in speakers exposed to Neapolitan and Venice Italian; iii) higher percentage of L+H* in speakers exposed to Bari Italian. For wh-questions we predict: i) lower overall levels of variability and ii) higher percentage of falling tunes in speakers exposed to Bari and Neapolitan varieties⁵.

4. Results

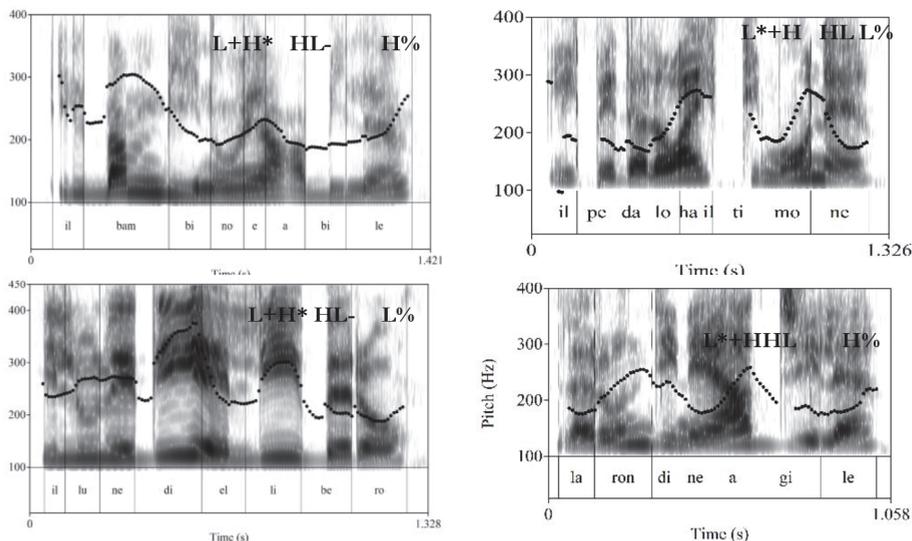
Intonational analyses of the productions of the 10 Salerno Italian speakers showed that the predictions formulated above were largely confirmed. Specifically despite keeping the pragmatic function of the utterances elicited the same, we found high levels of variability in the question tunes, which can be partly explained as a function of exposure to non-SI phonological systems. The next sections report results for the two question types separately.

4.1 Yes-no questions

In line with previous investigations of the intonation of SI questions, the analyses of the productions reported here uncovered several intonational contours that can be used by SI speakers to express an information-seeking yes-no question. Focusing on the nuclear configuration alone, we found four different nuclear tunes within the dataset analyzed. Specifically, we found that yes-no questions can be expressed by means of two phonologically different rising nuclear pitch accents, i.e. an early L+H* and a later L*+H, which can combine with either a falling (HL-L%) or a rising (HL-H%) edge tone configuration. Figure 1 below reports an example for each of the four yes-no question tunes attested.

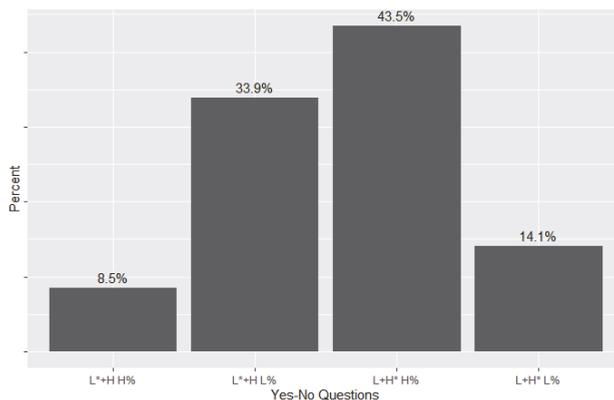
⁵ Predictions about varieties for which no information was available in the literature are missing.

Figure 1 - *Tunes attested for yes-no questions: Il bambino è abile? [Is the kid skilled?] uttered with a L+H* HL-H% nuclear tune (top left); Il lunedì è libero? [Is Monday free?] uttered with a L+H* HL-L% nuclear tune (bottom left); Il pedalò ha il timone? [Does the pedalò has the wheel?] uttered with a L*+H HL-L% (top right), La rondine è agile? [Is the swallow agile?] uttered with a L*+H HL-H% nuclear tune (bottom right)*



The frequency of occurrence of each of the four tunes was also measured. Figure 2 shows that the most frequent tune is the rise-fall-rise with an early peak pitch accent (L+H* HL-H%), which alone accounts for almost half of the total occurrences (43.5%). The second most frequent tune is the rise-fall with a late peak pitch accent (L*+H HL-L%). Finally, the other two tunes are by far less represented within the corpus analyzed.

Figure 2 - *Percentage of occurrence of yes-no question tunes⁶*

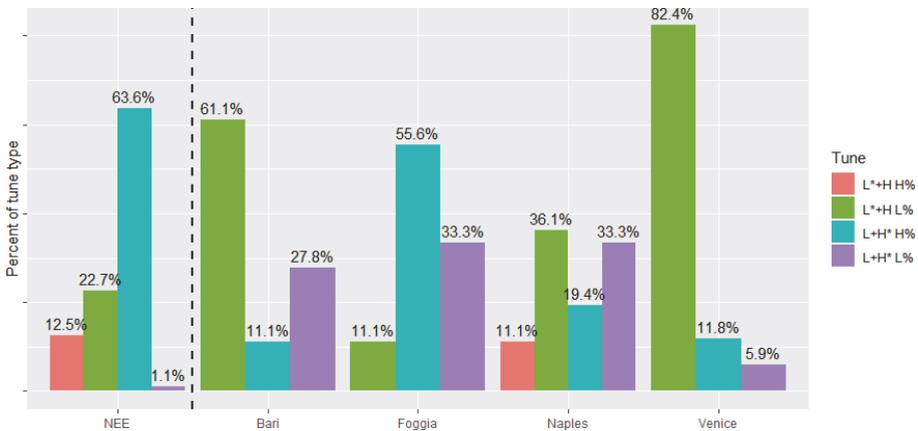


⁶ Here and in the following plots, label for tunes within plots have been shortened for convenience: the HL-phrase accent has been omitted since it stays the same in the four tunes.

These data also appear to be very consistent with previous investigations, both in terms of relative frequency of occurrence of the four yes-no question tunes and in terms of probabilistic distributions of the boundary type as a function of the nuclear pitch accent preceding it. In fact, Orrico et al. (2019a) reported that a falling edge tone configuration is much more likely to be found in combination with L*+H accents, while a rising edge correlates with the presence of a L+H*. Here, though we report data from a larger number of speakers, the relative distribution appears to be the same.

In order to test the hypothesis that the variability found in these data could partly be explained by early language contact, we observed the distribution of the yes-no tunes as a function of speakers' exposure to a non-SI variety⁷. The plot in Figure 3 shows the distribution of the four tunes according to the specific variety EE speakers were exposed to.

Figure 3 - Distribution of yes-no question tunes as a function of exposure to non-SI varieties



As a general result, the NEE group shows a clear preference for the use of the rise-fall-rise L+H* HL-H% tune, while the other tunes are much less represented within this group, with only the L*+H HL-L% going beyond 20%. As for the EE groups, different distributions are found according to the specific variety they were exposed to. Specifically, speakers with exposure to Bari and Venice Italian show a clear preference for the rise-fall with a late-peak accent, i.e. L*+H HL-L%, and they rarely use rise-fall-rise tunes. Additionally, the speaker exposed to Bari Italian makes a higher use of L+H* HL-L% tune, as opposed to NEE speakers. As for the speaker exposed to Foggia Italian, she shows a consistent use of the L+H* pitch accent,

⁷ We also tried to fit a generalized linear mixed model with the data that we collected, which showed some statistically significant effects. The strongest effect found was that of boundary tone as a function of exposure, showing that tunes with H% boundaries are used more frequently by NEE speakers. Nevertheless, due to the fact that the dataset under investigation is limited and several convergence errors were reported also for models with an extremely reduced random effect structure, we rather opted for a qualitative analysis instead.

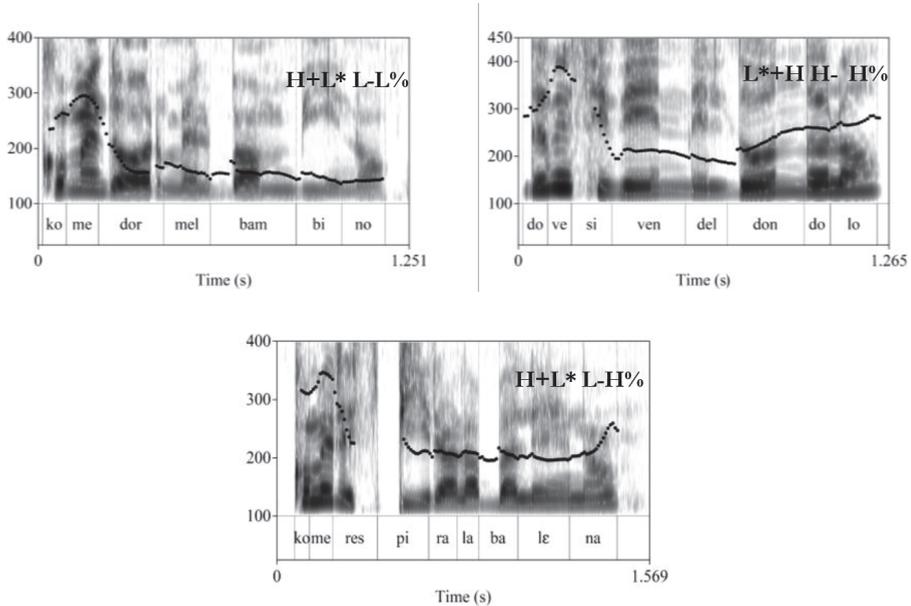
which is more frequently combined with a H% boundary. Her distributions are very similar to the NEE group, and the main difference lies in the use of the L+H* HL-L%, far less frequent in NEE speakers. Finally, the EE speakers with exposure to Neapolitan Italian show a generally variable distribution: as opposed to the NEE speakers, this group shows a sensible reduction of the L+H* HL-H% tune while the two rising-falling tunes are the most frequent ones.

The general picture drawn here mirrors the predictions outlined in 3.3. When we separate the NEE group from the EE one, the distribution of tunes changes. If we take the speakers exposed to the Naples, Venice, and Bari varieties, we see that the use of L% boundary tones becomes more frequent, mirroring the fact that rising-falling tunes are the most attested in the varieties they were exposed to. Also, the higher occurrence of L+H* HL-L% in the Bari Italian EE as opposed to NEE was expected owing to the fact that this tune has been largely attested in that variety. Finally, we do not have specific information about question intonation for Foggia Italian, though the fact that we registered a high percentage of the L+H* HL-L% with respect to NEE speakers might be interpreted as an effect of interference with the variety they were exposed to.

4.2 Wh-questions

Similar to what was found for yes-no questions, different nuclear patterns were also attested for wh-questions and, again, the tunes that were found are consistent with previous investigations. Specifically, three different patterns were attested. As reported in Figure 4 below, we found that wh-questions are most frequently realized with a falling pattern, analyzed as a H+L* nuclear pitch accent and an L-L% edge configuration. Other patterns attested are a fall-rise, analyzed as a H+L* falling accent and a rising edge configuration (L-H%), and a nuclear rise, labeled as L*+H H-H%.

Figure 4 - *Tunes attested for wh-questions: Come dorme il bambino? [How does the kid sleep?] uttered with a H+L* L-L% nuclear tune (top left), Dove si vende il dondolo? [Where do they sell the porch swing?] uttered with a L*+H H-H% nuclear tune (top right), and Come respira la balena? [How does the whale breathe?] uttered with a H+L* L-H% nuclear tune (bottom)*



As for the relative frequencies, shown in the plot below (Figure 5), the most frequent pattern was the H+L* L-L% fall, accounting for 60% of the total realizations, while the fall-rise and the rise were respectively found in the 23.3% and 16.7% of the cases. Hence, different from yes-no questions, here we found that the great majority of wh-question realizations are produced with a falling H+L*. This result is also consistent with previous investigations of intonation in SI. Specifically, Orrico and D'Imperio (2020b) found a clear separation of intonational cues between information-seeking and echo wh-questions, arguing that while pitch accent type in nuclear position determines the pragmatic interpretation of the question, boundary tones, on the other hand, seem to be much more dependent on a speaker-specific choice for which no account can be offered at this point.

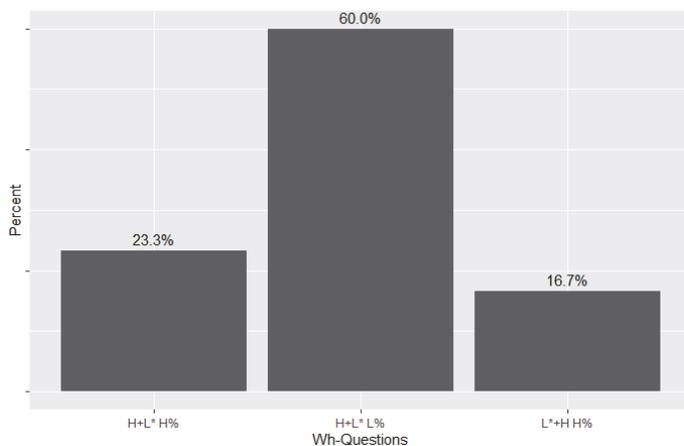
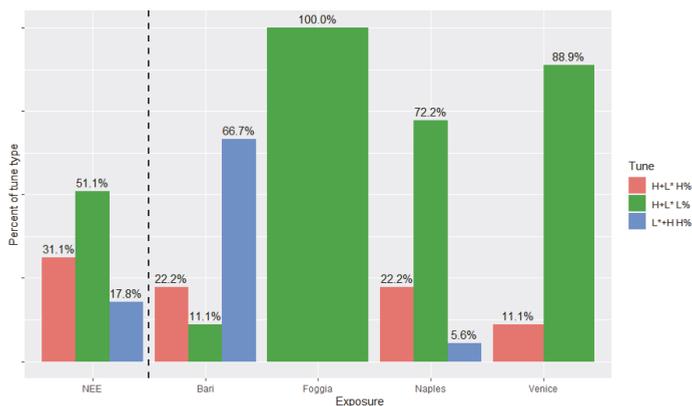
Figure 5 - Percentage of occurrence of *wh*-question tunes⁸

Figure 6 below shows the relative frequencies of tunes according to the EE variable, with productions divided according to the variety speakers were exposed to.

Figure 6 - Distribution of *wh* question tunes as a function of Exposure

The distribution for the NEE group does not change much with respect of the general distribution shown in Figure 5 above. The most frequent contour is the falling H+L* L-L% tune, followed by the falling-rise tune. The low frequency of L*+H H-H% was expected since it has been shown to be linked to specific pragmatic meanings in SI that were not elicited here (Orrico, D'Imperio, 2020b, but see also section 2.1 above). Similar distributions were found for speakers exposed to Venice and Neapolitan Italian, while the speaker exposed to Bari Italian showed high frequency for the rising tune. Finally, the speaker exposed to Foggia Italian

⁸ Here, too, phrase accents have been omitted for convenience. Falling accents are always followed by L-, while rising accents by H-.

only produced *wh*-questions with a falling tune. While data from EE to Neapolitan Italian was expected since the falling tune was the only attested tune in the variety the speakers were exposed to, results for the EE to Bari are a bit surprising: the rising tune in *wh*-questions has been attested for Salerno and Pescara Italian in Gili Fivela et al. (2015), therefore the behavior of this speaker is not easily accounted for by hypothesizing an effect of exposure.

5. Discussion

Our aim in this study was to report variability in the use of intonation to encode a question (both *yes/no* and *wh*-questions) in Salerno Italian. We build on previous research suggesting that the high degree of variation found in the use of intonation, especially in Italian varieties, might be due to varietal contact. Specifically, we tested the hypothesis that early exposure to different Italian varieties might induce such variation. We hence asked 10 Salerno Italian female speakers to perform a production task (read speech) created with the aim of eliciting both information-seeking *yes-no* and *wh*-questions. The 10 speakers were divided into two groups on the basis of having experienced early exposure to another variety of Italian, besides the one spoken in Salerno. More specifically, while all the speakers were born and brought up in Salerno, without ever living in other cities for prolonged periods of time, a group of them (5 speakers) had one non-SI parent. Results show that while high levels of variability were found in the data analyzed, this variation can be partly explained as an effect of experiencing early-exposure to a different phonological system. In particular, NEE speakers show a preference for specific tunes (L+H* HL-H% for *yes-no* and H+L* L-L% for *wh*-questions), while productions by the exposed speakers can be partially explained by looking at the tunes attested for the varieties they were exposed to.

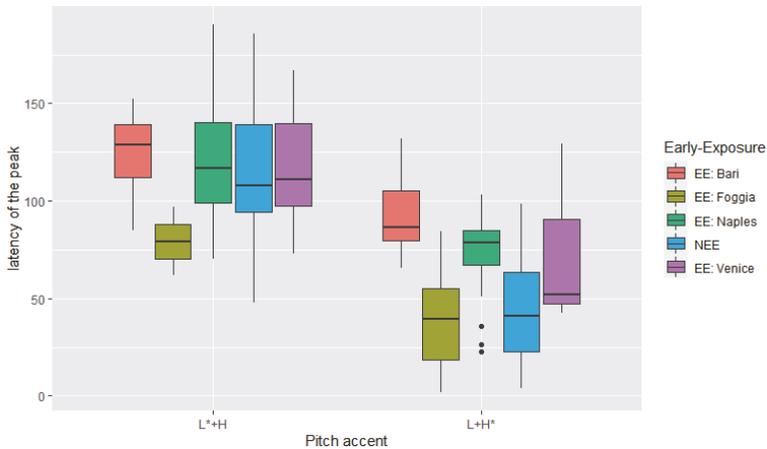
This is particularly visible in *yes-no* questions. While the four patterns found in the present study were also attested in other studies focusing on the variety spoken in Salerno, we found that their distribution changes when we consider early exposure as a possible factor of variation. The distributional differences might therefore be attributed to the different frequency of distribution of these tunes in the input the speakers received their whole life, which, as mentioned in the introduction above, is a key factor in shaping the phonological system of an individual. Therefore, if we suppose a different input as a function of having a non-SI parent, this might explain why in *yes-no* questions we registered a higher percentage of L*+H HL-L% in Venice- and Neapolitan-EE and of L+H* HL-L% in Bari-EE as opposed to the percentages registered for the NEE, and why, in the case of *wh*-questions, Neapolitan-EE speakers show a higher percentage of falling tunes as opposed to NEE.

One problem that would arise if we follow this path is represented by how the different phonological categories are shaped within an individual being exposed to different phonological systems. This might be extremely problematic in the case of pitch accents, since the same phonological category might be realized differently

(in terms of peak alignment, shape of the f_0 curve, and so on) in different regional varieties. For example, what has been labelled as $L+H^*$ in Neapolitan or Salerno Italian, two varieties that have in their phonological inventories two different accentual rises, is different from the $L+H^*$ attested in Bari, having a peak aligned later than SI and the Neapolitan $L+H^*$. What is more, the same Bari Italian accent appears to be aligned earlier with respect to an L^*+H in the Salerno and Neapolitan varieties. Therefore, if we posit that the higher frequency of the $L+H^*$ in Bari-EE yes-no questions derives from the high presence of this accent in the input, we also have to suppose that that accent is realized differently in terms of peak alignment in that speaker.

Figure 7 shows the distribution of peak alignment (distance of the peak from the onset of the stressed vowel) in the different groups.

Figure 7 - *Distribution of peak alignment in information-seeking yes-no questions according to early exposure and pitch accent type*



We are aware of the fact that the data shown in the plot hardly allow for any generalization: firstly because they consider only one of the possible dimensions that characterize category membership of tonal events and, more importantly, these refer to a very small dataset (we had only one speaker by exposure to either Venice, Bari, or Foggia). Nevertheless, the plot allows us to focus our attention on an interesting trend. While all speakers appear to keep the two categories apart by means of (at least) the timing of the accentual peak, differences can still be noticed. Specifically, if we look at data of EE to Bari, as opposed to NEE, it is possible to see that both $L+H^*$ and L^*+H are aligned later.

We can speculate that the Bari-EE speaker has learned, at an early stage of life, the later-peaked version of $L+H^*$, typical of Bari Italian; then, the earlier-peaked SI version, present in the SI input she was later exposed to, has been assimilated to that category. This process is accounted by Flege's (1995) SLM, in which it is reported that the formation of a new category is prevented in case the sounds of two

competing phonological systems are perceived as similar. The same model can also explain why the L*+H within the same speaker is also aligned later, as an effect of category dissimilation: in that case, the phonetic distance between two categories might be exaggerated in order to keep the two categories apart. In other words, since the L+H* is realized with a later-than-expected peak, the speaker does the same for the L*+H, to make sure to maintain the phonological contrast between the two pitch accents; see also D'Imperio et al. (2014) for similar results concerning the adjustment of peak alignment within categories in imitation.

A similar behavior is observed for the realization of L+H* by Neapolitan-EE, which is realized with a later peak than NEE speakers. This might be explained in a similar fashion as we did with the Bari Italian case: the speakers exposed to Neapolitan Italian might have learned both categories, though they might avoid realizing it with an extremely early peak to keep the interrogative early peak accent apart from the L+H* used for narrow focus statement. These arguments are only speculative for now and future research should be carried out with an appropriate number of speakers to test the validity of this theory.

In addition to a perspective which takes into account a condition of language contact at the individual level, some effects of exposure can also be highlighted looking at language contact among the phonological systems of such varieties. Clearly, unlike Gili Fivela and Nicora (2018), the present study does not strictly deal with geographical adjacency among the varieties. Nevertheless, being Neapolitan Italian one of the non-SI varieties under investigation, some observations about the relation between these two geographically neighboring varieties need to be made. Salerno and Naples share the same regional area and are traditionally classified as belonging to the same dialectal area. In the light of this, we looked at the distribution of the patterns attested for information-seeking yes-no questions⁹ in both varieties.

As far as Neapolitan Italian is concerned, for polar questions, the rise-fall tune – L*+H HL-L% – is largely attested (*inter alia* D'Imperio, 2002). Similarly, in our data, this tune is the most frequent one in the productions of speakers who had been early exposed to Neapolitan Italian. Conversely, in previous investigations on SI this tune is attested with a lower frequency of occurrence (about 20% of cases; Orrico, 2020); our data confirm such frequency (22.7% in the NEE group). As for SI, both in this study (63.6%) and in previous ones (more than 50% of cases; Orrico, 2020), the most frequent attested pattern is the rise-fall-rise – L+H* HL-H% – tune. Note that, for Neapolitan Italian, rising terminals have only been attested in Canepari (1986) and Cangemi and Grice (2016), but information on their frequency of occurrence is not available. Our EE speakers to Neapolitan Italian do make use of this tune, even though in less than 20% of productions. In short, the most frequent attested tunes in the two varieties of Neapolitan and Salerno Italian, i.e. the rise-fall

⁹ We focus on information-seeking yes-no questions because in this condition the highest levels of intra and inter-variety variability have been found both in the literature (Gili Fivela, Nicora, 2018) and in our data.

and the rise-fall-rise respectively, are possible tunes in both varieties; crucially, the main difference lies in their distribution.

A close look at the Neapolitan-Salerno pair of varieties suggests on the one hand that both NEE and Neapolitan-EE speakers' productions mirror the distribution of patterns attested for the two varieties: both tunes are present in their native phonological systems, though the frequency of occurrence is reversed, at least as far as yes-no questions are concerned. On the other hand, in line with Gili Fivela and Nicora's (2018) investigation, a certain degree of continuity in terms of possible tunes can be ascribed to vicinity at the diatopic level. Indeed, such a geographical proximity between Naples and Salerno implicates that Neapolitan and Salerno Italian speakers are subject to continuous exposure to each other. Nevertheless, in order to assess such a contact between the varieties, further analyses appear to be necessary.

The results reported here have strong implications for linguistic theories of intonation, spanning from improving our knowledge of the phonological organization of a language and its varieties to the criteria used to select speakers and listeners. To the best of our knowledge, only few studies have investigated language contact in case of geographically close varieties. Indeed, several studies are, to date, still building on the assumption that individuals sharing the same geographical space are, in all respects, homogeneous from a phonological point of view. What is important, however, is to point out that lack of speech variability within the same linguistic community should not necessarily be expected. Chang (2019) indeed advises caution when considering a speaker as monolingual (e.g., in our case, the non-exposed ones), since a linguistic system keeps evolving during the entire life span as an effect of all kinds of linguistic input a speaker is exposed to during her life (e.g., learning an L2). Several factors, either linked to linguistic input or other aspects, such as cognitive differences (e.g., musical abilities or empathy skills), play a strong role in the definition of an individual's phonological system (see Cason, Marmursztejn, D'Imperio & Schön, 2019; Esteve-Gibert et al., 2020, Orrico, D'Imperio, 2020a).

Our study underlines the extreme importance of taking into account speaker-specific factors in the study of intonation, in order to reach a fairly good understanding of the way this system contributes to the communicative process.

6. Conclusion

A production study was designed to investigate intonation variation in Salerno Italian yes-no and wh-questions as a consequence of the presence vs absence of early exposure to a non-native variety. Results point to differences in the use of intonational contours, both in terms of pitch accent and boundary tone specification, across the different exposed speakers. Moreover, the patterns shown by the exposed speakers were similar to those typical of the variety they were exposed to. The study adds to the literature by showing that the type of phonological input an individual is exposed to is an important predictor of the way she uses intonation to encode specific pragmatic meanings. Furthermore, our study emphasizes the importance of

taking into account several potential sources of speech variability, in order to reach a good understanding of the intonation-meaning mapping.

Bibliography

- AVESANI, C., VAYRA, M. (2003). Broad, narrow and contrastive focus in Florentine Italian. In *Proceedings of the 15th International Congress of Phonetic Sciences*, Vol. 2, 1803-1806.
- BEST, C.T., TYLER, M.D. (2007). Nonnative and second-language speech perception: Commonalities and complementarities. In BOHN, O.-S., MUNRO, M.J. (Eds.), *Language experience in second language speech learning*. Amsterdam: John Benjamins, 13-34.
- BEST, C.T. (2015). Devil or angel in the details? Perceiving phonetic variation as information about phonological structure. In ROMERO, J., RIERA, M. (Eds.), *The Phonetics-Phonology Interface: Representations and methodologies*, Amsterdam: John Benjamins, 3-31.
- BEST, C.T., TYLER, M.D., GOODING, T.N., ORLANDO, C.B., & QUANN, C.A. (2009). Development of phonological constancy: Toddlers' perception of native – and Jamaican – accented words. *Psychological science*, 20(5), 539-542.
- BOHN, O.-S., & BUNDGAARD-NIELSEN, R.L. (2009). Second language speech learning with diverse inputs. In T. PISKE & M. YOUNG-SCHOLTEN (Eds.), *Input matters in SLA* (pp. 207-218). Clevedon, UK: Multilingual matters.
- CANGEMI, F., KRÜGER, M., & GRICE, M. (2015). Listener-specific perception of speaker-specific production in intonation. In S. FUCHS, D. PAPE, C. PETRONE, & P. PERRIER (Eds.), *Individual Differences in speech production and perception*. Frankfurt am Main: Peter Lang, 123-145.
- CASON, N., MARMURSZTEJN, M., D'IMPERIO, M., & SCHÖN, D. (2019). Rhythmic Abilities Correlate with L2 Prosody Imitation Abilities in Typologically Different Languages. *Language and Speech*, 63(1), 149-165.
- CATALDO, V., ORRICO, R., SAVY, R. (2017). Phonetic variations of f0 range in L1 and L2: a comparison between Italian, English and Spanish native and non-native speakers. In BERTINI, C., CELATA, C., LENOCI, G., MELUZZI, C., RICCI, I., (Eds.). *Fattori sociali e biologici nella variazione fonetica*, Studi AISV 3. Milan: Officinaventuno, 227-256.
- CHANG, C.B. (2019). Language change and linguistic inquiry in a world of multicompetence: Sustained phonetic drift and its implications for behavioral linguistic research. In *Journal of Phonetics*, 74, 96-113.
- COLE, J., SHATTUCK-HUGNAGEL, S. (2011). The phonology and phonetics of perceived prosody: What do listeners imitate?. In *Interspeech-2011*, Florence, Italy, 27-31 August 2011, 969-972.
- COUNCIL OF EUROPE (2001). *Common European framework of reference for languages: learning, teaching and assessment*. Cambridge University Press.
- CROCCO, C. (2013). Is Italian Clitic Right Dislocation grammaticalised? A prosodic analysis of yes/no questions and statements. In *Lingua*, 133, 30-52.
- CROCCO, C., & BADAN, L. (2016). 'L'hai messo dove il focus?' Un'analisi prosodica delle domande eco wh. In SAVY, R., ALFANO, I. (Eds.) *La fonetica nell'apprendimento delle lingue*. Studi AISV 2. Milano: Officinaventuno, 191-207).

- D'IMPERIO, M. (1999). Tonal structure and pitch targets in Italian focus constituents. In Ohala, J. (Ed.), *Proceedings of the 14th International Congress of Phonetic Sciences*, 3, San Francisco, USA, 1757–1760.
- D'IMPERIO, M. (2000). The role of perception in defining tonal targets and their alignment. PhD dissertation, Ohio State University.
- D'IMPERIO, M. (2002). Italian intonation: An overview and some questions. In *Probus*, 14(1), 37–69.
- D'IMPERIO, M., CAVONE, R., & PETRONE, C. (2014). Phonetic and phonological imitation of intonation in two varieties of Italian. In *Frontiers in psychology*, 5, 1226.
- D'IMPERIO, M., GERMAN, J.S. (2015). Phonetic detail and the role of exposure in dialect imitation. In *Proceedings of the 18th International Congress of Phonetic Sciences (ICPhS 2015)*, Glasgow, 10–14 August 2015.
- DELVAUX, V., SOQUET, A. (2007). The influence of ambient speech on adult speech productions through unintentional imitation. In *Phonetica*, 64(2-3), 145–173.
- ESTEVE-GIBERT, N., SCHAFER, A.J., HEMFORTH, B., PORTES, C., POZNIAK, C., & D'IMPERIO, M. (2020). Empathy influences how listeners interpret intonation and meaning when words are ambiguous. *Memory & Cognition*, 1–15.
- FLEGE, J.E. (1995). Second language speech learning: Theory, findings, and problems. In STRANGE, W. (Ed.), *Speech perception and linguistic experience: Issues in cross-language research*. Baltimore: York Press, 233–277.
- GERMAN, J. (2012). Dialect adaptation and two dimensions of tune. In *Proceedings of the 6th International Conference of Speech Prosody*, Shanghai, China, 22–25 May 2012, 430–433.
- GILI FIVELA B., NICORA F. (2018). Intonation in Liguria and Tuscany: checking for similarities across a traditional isogloss boundary. In VIETTI, A., SPREAFICO, L., MEREU, D., GALATÀ, V. (Eds.). *Il parlato nel contesto naturale*. Studi AISV 4, Milano, Officinaventuno. 131–156.
- GILI FIVELA, B., AVESANI, C., BARONE, M., BOCCI, G., CROCCO, C., D'IMPERIO, M., GIORDANO, R., MAROTTA, G., SAVINO, M. & SORIANELLO, P. (2015). Intonational phonology of the regional varieties of Italian. In FROTA, S., PRIETO, P. (Eds.). *Intonation in Romance*. Oxford: Oxford University Press, 140–197.
- GILI FIVELA, B., IRACI, M.M. (2017). Variation in intonation across Italy: The case of Palermo Italian. In BERTINI, C., CELATA, C., LENOCI, G., MELUZZI, C. & RICCI, I. (Eds.), *Fattori sociali e biologici nella variazione fonetica*. Studi AISV 3. Milano: Officinaventuno, 167–188.
- GILI FIVELA, B., SAVINO, M. (2003). Segments, syllables and tonal alignment: A study on two varieties of Italian. In *Proceedings of the XVth international congress of phonetic sciences*, 2933–2936.
- GILI FIVELA, B. (2008). *Intonation in production and perception: The case of Pisa Italian*. Alessandria: Edizioni dell'Orso.
- GOLDINGER, S.D. (1998). Echoes of echoes? An episodic theory of lexical access. In *Psychological review*, 105(2), 251.
- GOLDINGER, S.D. (2000). The Role of Perceptual Episodes in Lexical Processing. In *SWAP-2000*, Nijmegen, Netherlands, 29–31 May 2000, 155–158.

- GRICE, M. (1991). The intonation of interrogation in two varieties of Sicilian Italian. In *Proceedings of the XIIth International Congress of Phonetic Sciences*, vol. 5, 210-213.
- GRICE, M. (1995). *The intonation of interrogation in Palermo Italian*. Tübingen: Niemeyer
- GRICE, M., D'IMPERIO, M., SAVINO, M., & AVESANI, C. (2005). Strategies for intonation labelling across varieties of Italian. In JUN, S.A. (Ed.). *Prosodic typology: The phonology of intonation and phrasing (Vol. 1)*. Oxford: Oxford University Press, 362-389.
- GRICE, M., SAVINO, M. (1997). Can pitch accent type convey information status in yes-no questions? In *Proceedings of a Workshop Sponsored by the Association for Computational Linguistics*, Madrid, Spain, pp. 29-38.
- HALLE, M. (1985). Speculations about the representation of words in memory. In *Phonetic linguistics*, 101-114.
- IVERSON, P., PINET, M., & EVANS, B.G. (2011). Auditory training for experienced and inexperienced second-language learners: Native French speakers learning English vowels. *Applied Psycholinguistics*, 33(1), 145.
- JUN, S.-A., BISHOP, J. (2015). Priming implicit prosody: Prosodic boundaries and individual differences. In *Language and Speech*, 58(4), 459-473.
- KIDD, E., DONNELLY, S., & CHRISTIANSEN, M.H. (2018). Individual differences in language acquisition and processing. In *Trends in cognitive sciences*, 22(2), 154-169.
- KIM, Y.H., HAZAN, V. (2010). Individual variability in the perceptual learning of L2 speech sounds and its cognitive correlates. In DZIUBALSKA-KOLACZYK, K., WREMBEL, M. & KUL, M. (Eds.). *New Sounds 2010: Proceedings of the Sixth International Symposium on the Acquisition of Second Language Speech*. Poznań: School of English, Adam Mickiewicz University.
- KONG, E.J., LEE, H. (2018). Attentional modulation and individual differences in explaining the changing role of fundamental frequency in Korean laryngeal stop perception. In *Language and speech*, 61(3), 384-408.
- LADD, D.R. (2008). *Intonational phonology*. Cambridge: Cambridge University Press.
- LAHIRI, A., MARSLÉN-WILSON, W. (1991). The mental representation of lexical form: A phonological approach to the recognition lexicon. In *Cognition*, 38(3), 245-294.
- LUQUE MOYA, J.A., SAVY, R. (2017). Un metodo para la enseñanza asistida de la prosodia en aprendientes italofonos de ELE. In *Sistema, codificación e interpretación: aproximaciones al análisis de la lengua y a su didáctica en una perspectiva metaoperacional*. MarcoELE: *Revista de Didáctica Español Lengua Extranjera*, (24), 8, 192-217.
- MAROTTA, G., SORIANELLO, P. (2001). La teoria autosegmentale nell'analisi dell'intonazione interrogativa in due varietà di italiano toscano (Lucca e Siena). In *Atti del XXXIII Congresso della Società di Linguistica Italiana*. Roma: Bulzoni Editore, 177-204.
- MENNEN, I. (2004). Bi-directional interference in the intonation of Dutch speakers of Greek. *Journal of phonetics*, 32(4), 543-563.
- ORRICO, R. (2020) Individual variability in intonational meaning identification: The role of cognitive and sociolinguistic variables. *Ph.D. Thesis*. University of Salerno & University of Aix Marseille.
- ORRICO, R., & D'IMPERIO, M. (2020a) Individual empathy levels affect gradual intonation-meaning mapping: The case of biased questions in Salerno Italian. *Laboratory Phonology*:

Journal of the Association for Laboratory Phonology 11(1): 12, pp. 1–39. DOI: <https://doi.org/10.5334/labphon.238>

ORRICO, R., CATALDO, V., SAVY, R. & BARONE, L. (2016). Transfer, Fossilization and Prosodic Drift in Foreign Language Learning. In SAVY, R., ALFANO, I. (Eds.). *La fonetica nell'apprendimento delle lingue*, Studi AISV 2. Milan: Officinaventuno, 117-132.

ORRICO, R., D'IMPERIO, M. (2020b) Tonal specification of speaker commitment in Salerno Italian wh-questions. Proc. 10th International Conference on Speech Prosody 2020, 361-365, DOI: 10.21437/SpeechProsody.2020-74.

ORRICO, R., SAVY, R., D'IMPERIO, M. (2019a). Salerno Italian: intonational phonology and dimensions of variation. In PICCARDI, D., ARDOLINO, F., CALAMAI, S. (Eds.) *Gli archivi sonori al crocevia tra scienze fonetiche, informatica umanistica e patrimonio digitale* Studi AISV 6. Milan: Officinaventuno, 309-328.

ORRICO, R., SAVY, R., D'IMPERIO, M. (2019b). The perception of speaker certainty in Salerno Italian intonation. In CALHOUN, S., ESCUDERO, P., TABAIN, M., & WARREN, P. (Eds.). *Proceedings of the 19th International Congress of Phonetic Sciences*, Melbourne, Australia, 5-9 August 2019, 2946-2950.

PIERREHUMBERT, J.B. (1980). The phonology and phonetics of English intonation. *Ph.D. thesis*, Massachusetts Institute of Technology, Cambridge, MA.

PIERREHUMBERT, J.B. (2016). Phonological representation: Beyond abstract versus episodic. In *Annual Review of Linguistics* (2), 33-52.

ROMERA, M., & ELORDIETA, G. (2013). Prosodic accommodation in language contact: Spanish intonation in Majorca. *International journal of the sociology of language*, 2013(221), 127-151.

SAVINO, M. (2012). The intonation of polar questions in Italian: Where is the rise?. In *Journal of the International Phonetic Association*, 42(1), 23-48.

SAVINO, M. (2017). The dynamics of prosodic adaptation between Italian conversational partners. In *ExLing 2017: Proceedings of 8th Tutorial and Research Workshop on Experimental Linguistics*, Heraklion, Crete, Greece, 19-22 June 2017, 93-96.

SAVINO, M., GRICE, M. (2011). The perception of negative bias in Bari Italian questions. In FROTA, S., PRIETO, P., ELORDIETA, G. (Eds.). *Prosodic categories: Production, perception and comprehension*. London, UK: Springer Verlag, 187–206.

SAVY, R., LUQUE MOYA, J.A. (2014). Aspectos prosódicos de las interrogativas en aprendientes italianos de ELE. In *Anejos n° 7 de Normas- Revista de Estudios Linguísticos Hispánicos*, 283-296.

TYLER, M.D. (2019). PAM-L2 and Phonological Category Acquisition in the Foreign Language Classroom. A Sound Approach to Language Matters: In Honor of Ocke-Schwen Bohn, 607-630.

YU, A.C., GROVE, J., MARTINOVIC, M., & SONDEREGGER, M. (2011). Effects of working memory capacity and “autistic” traits on phonotactic effects in speech perception. In *Proceedings of the International Congress of the Phonetic Sciences XVII*, Hong Kong, 17-21 August 2011, 2236-2239.

SIMONA SBRANNA, FRANCESCO CANGEMI, MARTINE GRICE

Quantifying L2 interactional competence

The Common European Framework of Reference for languages (CEFR) defines foreign language competence as communicative competence, emphasizing its interactional aspect. Nonetheless, L2 (second language) assessment often focuses on the quantification of grammar and lexical competence, neglecting interactional aspects, which are only subject to an impressionistic evaluation. With this exploratory study, we test a method for quantification and visualization of interaction management on L2 data. Our corpus includes 40 conversations both in L1 (first language) and L2 produced by Italian learners of German as L2. Results suggest that low levels of proficiency negatively affect the smoothness of the interactional flow, whereas the difference between L1 and L2 interactional patterns reduces with increasing L2 proficiency. Extracting reliable and testable metric, this method could represent a valid starting point to develop an instrument for a quantifiable assessment of interactional competence.

Keywords: L2 acquisition, communicative competence, interactional competence, assessment of L2 proficiency.

1. Introduction

The Common European Framework of Reference for languages describes learners as “social agents” (Council of Europe, 2001:9) who, as members of the society, constantly need to accomplish communicative tasks in various contexts and under different circumstances. To carry out the tasks they are faced with, they need to strategically combine their general and communicative language competence in the most appropriate way according to the context.

Figure 1 shows a schematic description of the communicative language competence extracted from the CEFR (Figueras, North, Takala, Van Avermaet, & Verhelst, 2009:32). Here communicative language competence is depicted as a compound competence formed by several components: linguistic competence, intended as lexical and grammatical knowledge, together with its cognitive organisation and accessibility; sociocultural competence, including the knowledge of register appropriateness, degree of formality, rules of politeness and the knowledge of linguistic rituals specific to a community; pragmatic and strategic competence, referring to the functional skills necessary to arrange the message and manage the conversation according to interactional schemata. The table also includes four different aspects of language activities learners can perform: production, reception, mediation and interaction.

Figure 1 - *Communicative Language Competence in the CEFR (Figueras et al., 2009:32)*

	RECEPTION		INTERACTION		PRODUCTION		MEDIATION	
	Listening	Reading	Spoken Interaction	Written Interaction	Spoken Production	Written Production	Spoken Mediation	Written Mediation
Linguistic Competence								
▪ General Linguistic Range	✓	✓	✓	✓	✓	✓	✓	✓
▪ Vocabulary Range	✓	✓	✓	✓	✓	✓	✓	✓
▪ Vocabulary Control			✓	✓	✓	✓	✓	✓
▪ Grammatical Accuracy			✓	✓	✓	✓	✓	✓
▪ Phonological Control			✓		✓		✓	
▪ Orthographic Control				✓		✓		✓
Socio-linguistic Competence								
▪ Socio-linguistic Appropriateness	✓	✓	✓	✓	✓	✓	✓	✓
Pragmatic Competence								
▪ Flexibility			✓	✓			✓	✓
▪ Turntaking			✓					
▪ Thematic Development	✓	✓		✓	✓	✓	✓	✓
▪ Cohesion and Coherence	✓	✓			✓	✓	✓	✓
▪ Spoken Fluency			✓		✓		✓	
▪ Propositional Precision	✓	✓			✓	✓	✓	✓
Strategic Competence								
▪ Identifying cues/infering	✓	✓					✓	✓
▪ Turntaking (repeated)			✓					
▪ Cooperating			✓	✓				
▪ Asking for clarification			✓	✓				
▪ Planning					✓	✓		✓
▪ Compensating			✓	✓	✓	✓	✓	✓
▪ Monitoring and Repair			✓	✓	✓	✓	✓	✓

Production and reception are primary activities considered firstly in isolation since they are preliminary to mediation and interaction. Mediation consists in making communication possible among parties who are not able to directly communicate with each other, generally through the reformulation of an oral or written source message. Interaction is described as a complex process in which:

at least two individuals participate in an oral and/or written exchange in which production and reception alternate and may in fact overlap in oral communication. Not only may two interlocutors be speaking and yet listening to each other simultaneously. Even where turn-taking is strictly respected, the listener is generally already forecasting the remainder of the speaker's message and preparing a response. Learning to interact thus involves more than learning to receive and to produce utterances (Council of Europe, 2001:14).

In its face-to-face form, interaction involves productive and receptive skills at the same time, as well as additional abilities which allow speakers to monitor the development of this process and constantly adjust to it in real-time.

As described into the CEFR, different abilities come into play at different moments of the interaction. In planning, speakers try to anticipate the forthcoming linguistic exchange, identify the common background information and opinion of the interlocutor to plan which conversational move best suits. In order to take the floor, speakers use turn-taking strategies and during the exchange, they co-operate

and build the conversation together. Speakers also need to be able to cope with unexpected misunderstandings and ask for clarification, or repair communication breakdowns, thus interlocutors constantly evaluate the on-going process to be able to appropriately react.

The description offered by the CEFR mirrors the definitions of interactional competence and interactional abilities provided by the existing literature. Jacoby and Ochs (1995) have described interaction as a form of co-construction, a joint creation of discourse between interlocutors. Hall (1993, 1995 as reported in He, Young, 1998) has argued that interactional competence is, therefore, context-specific since it emerges in varied interactive practices, to which participants contribute with the appropriate linguistic and pragmatic resources.

Moreover, especially in its oral form, interaction is the most common and direct way of communication in our daily life and for this reason, it should be ensured a central role in language teaching and testing. Indeed, the Framework states:

High importance is generally attributed to interaction in language use and learning in view of its central role in communication (Council of Europe, 2001:14).

However, assessment of learners' proficiency often tends to focus mainly on grammar and lexicon, neglecting or even excluding interactional aspects. Many language test formats only involve a written form, such as the cloze format, where some words of a text are replaced with gaps to be filled in by learners. These kinds of test are often used with the explanation that they show correlations to all receptive and productive abilities (reading, listening, writing and speaking). Nevertheless, they mainly put to test grammar and vocabulary knowledge, leaving out the full range of pragmatic and strategic resources required in oral interaction.

Bachmann and Palmer in "Language Testing Practice" (1996) mention interactivity as one of the fundamental characteristics a good quality language proficiency test should have: reliability, construct validity, authenticity and interactivity. Reliability refers to the consistency of the measurement, thereby of the results given by the test. Construct validity indicates the possibility to interpret the score of the test as a valid indication of global language proficiency. Authenticity defines how correspondent the task given to learners in test circumstances is to real-life tasks they would perform using the L2. Finally, interactivity refers to the degree of involvement of learners' different abilities in accomplishing the task, i.e. the extent to which a test involves various learners' skills, which include general language knowledge, metacognitive strategies and strategic competence for planning and dealing with unexpected difficulties, topical knowledge and affective schemata, which refers to learners' emotional response to the task. According to the authors, tasks with a high level of interactivity are role play and long conversations, as they require learners to draw on all these abilities.

Some possible reasons for neglecting highly interactive tasks in language proficiency testing may be practical. He and Young (1998) point out that having learners interviewed by native or highly proficient speakers can create certain difficulties. First, such interviewers have to be available; secondly, the interviews need to be

carried out for a reasonable length of time to allow the interviewer to elicit enough linguistic data from the learner so that these data can be considered representative of the learner's global knowledge. Hence, such testing would require more assessors at the same time, be time-consuming and consequently more expensive than a test format, such as the cloze test, which can optimize time for testing and correction. Furthermore, the quantification of the skills involved in interaction may turn out to be extremely time-consuming and complex to synthesise. As a result, interactional competence is often subject only to a qualitative evaluation based on illustrative scales¹, the interpretation of which may include a certain degree of subjectivity.

Intending to propose quantification methods for L2 speaking ability, research has mainly focussed on the measurement of fluency. Indeed, fluency is consistently mentioned as a fundamental component of learners' oral performance in various assessment traditions and its correlation with general L2 proficiency has been demonstrated by several studies (De Jong, Steinel, Florijn, Schoonen & Hulstijn, 2013; De Jong, Groenhout, Schoonen & Hulstijn, 2015; Segalowitz, Freed, 2004).

In the following paragraphs we will shortly review theories and findings on fluency as a measure of L2 speaking proficiency.

2. *Background*

Fluency has been identified as one of the main aspects ensuring the success of the speaking performance (De Jong, 2016). One of its first definitions can be traced back to Fillmore (1979). He defines fluency as a measure of how well a language is spoken, in other words the skill to use L2 knowledge efficiently, and enumerates four dimensions of fluency including both quantitative and qualitative aspects: the ability to speak at length with few breaks; the ability to speak in a coherent, reasoned, and semantically dense way; the ability to talk appropriately according to the context; the ability to be creative and imaginative in speech production.

Starkweather (1987) suggests instead four dimensions of fluency mainly related to physical aspects of speech: continuity, rate, rhythm, effort. In other words, fluent speech should present few discontinuities, have a regular rhythm and a fast rate, and not require too much cognitive and physical effort (Zmarich, 2017).

In his model of fluency, Logan (2015) adds to the described eight dimensions two additional ones: naturalness, i.e. how much speech resembles that uttered by a typical speaker with regard to continuity, rate, rhythm and effort; and stability, i.e. how similar are speaker's performances over time if subject to repeated measurements.

Segalowitz (2010) focuses on L2 fluency from a dynamical system perspective. He argues that fluency is strongly linked to the social context in which the speech

¹ An example from the illustrative scale for overall spoken interaction (C1 level): "Can express him/herself fluently and spontaneously, almost effortlessly. Has a good command of a broad lexical repertoire allowing gaps to be readily overcome with circumlocutions. There is little obvious searching for expressions or avoidance strategies; only a conceptually difficult subject can hinder a natural, smooth flow of language" (Council of Europe, 2001:74).

performance takes place and distinguishes three aspects: utterance fluency, cognitive fluency and perceived fluency.

L2 utterance fluency refers to the fluidity observable in a speech sample and quantifiable by temporal measures, among which the author mentions syllable rate, duration and rate of hesitations, filled and silent pauses, breakdown fluency (pausing phenomena) and repair fluency (false starts, corrections, repetitions). Indeed, most studies calculating fluency temporal measures follow the classification in the sub-components breakdown, repair and speed fluency (Tavakkoli and Skehan, 2005).

L2 cognitive fluency refers to the fluidity of the cognitive processes underlying speech production, such as processing skills (declarative and procedural knowledge), efficiency and speed of semantic retrieval, and cognitive load in working memory. Some measures of cognitive fluency have been found to correlate with L2 proficiency, e.g. reaction time and switch cost measures. Reaction time speed and its coefficient of variability have been used to operationalise the efficiency of semantic retrieval (Segalowitz, Freed, 2004), while switch cost measures has been used as an indicator of linguistic attention, which refers to attention shifting guided by connections among grammatical elements within utterances (Duncan, Segalowitz & Phillips, 2014).

Such systemic understanding of fluency is assumed also by Kormos' psycholinguistic model (2006) in which different cognitive processes underlie the three above-mentioned sub-components of fluency. In particular, breakdown measures are related to learners' effort, such as final-clause pauses, which reflect learners' conceptualization and planning of the message, and mid-clause pauses, which represent the time taken by learners to encode and formulate the linguistic information; repair measures signal the monitoring of the speech output and consequently the amount of attention required for speaking in L2; and finally, speed-related measures inform about the degree of automatization of all these processes.

Another important aspect to take into account is the fact that both utterance and cognitive fluency are specific to each person. Still, individual variability in L1 can only partially explain individual variability in L2 (De Jong *et al.*, 2013) since disfluency is also characterised by L2-specific features, e.g. a higher cognitive load. Therefore, it may be a good scientific practice to consider L1 fluency measures as a baseline for each individual's fluency characteristics (as in De Jong *et al.*, 2015; Saito *et al.*, 2019) to get a clearer picture of L2-specific fluency measures by partialling out the variables that are not especially related to L2 disfluency phenomena (Segalowitz, 2010).

Finally, L2 perceived fluency indicates subjective listeners' ratings on how fluent a speaker is. One disadvantage is that being subjective, perceived fluency is only moderately informative about utterance fluency and cannot explain all the variance of its objective measures. However, it is helpful to get an understanding of what cues are relevant to native listeners when judging L2 speech fluency in relation to L2 proficiency. Moreover, a listener's judgment of their interlocutor's fluency can affect the interaction and influence both speakers' fluency.

2.1 Fluency measures and operationalisations

Research on L2 fluency has focused on individuating which objective measures can better explain L2 fluency judgments and has mainly concentrated on temporal features.

A categorisation of aspects of fluency comparable to the more recent triad “breakdown, repair and speed” (Tavakkoli, Skehan, 2005) was already proposed in one pioneering study. Riggensbach (1991) classifies the features which can characterise a judgement of fluent or non-fluent in non-native speech into hesitation and repair phenomena, and rate and amount of speech. The study also includes an analysis of interactive features contributing to the turn-taking alternation, such as overlaps, pauses between turns and collaborative completions. Hesitation phenomena and speech rate were found to be significantly correlated to ratings of L2 fluency, with hesitations placement and the resulting discourse chunking playing a central role. On the contrary, results related to repair phenomena appeared to be less clear, probably due to the small set of data. The same holds for interactive features, which revealed a high variability because of the idiosyncratic nature of interactions that vary according to many linguistic and non-linguistic factors. Furthermore, this pragmatic-oriented analysis is referred to be extremely time-consuming and, indeed, studies on turn-taking fluency are relatively scarce.

Later on, many studies confirmed that perceived fluency by both native listeners (Kormos, Dénes, 2004; Derwing, Rossiter, Munro, & Thomson, 2004; Rossiter, 2009; Bosker, Pinget, Quené, Sanders & De Jong, 2013; Préfontaine, Kormos, & Johnson, 2016; Saito, Ilkan, Magne, Tran, & Suzuki, 2018; Suzuki, Kormos, 2019) and L2 listeners (Magne, Suzuki, Suzukida, Ilkan, Tran, & Saito, 2019) is closely related to the speed of delivery and pausing phenomena.

As reported in Suzuki and Kormos (2019), a more recent approach in research on fluency differentiates three independent dimensions of breakdown fluency – frequency, duration and location of pauses – and all of them have been demonstrated to independently contribute to fluency. Moreover, concerning pause location, mid-clause pauses have been found to have a more distinctive role than final-clause pauses (Saito *et al.*, 2018; Magne *et al.*, 2019; Suzuki, Kormos, 2019). One possible reason, following Kormos’ model (2006), would be that mid-clause pauses, being associated with the time required for linguistic encoding, are more representative of proficiency than final-clause pauses associated with content planning.

However, these studies present several differences in methodology, in particular regarding sample size, rating methods, task used for data collection and operationalisation of measures. The following table summarises the most listed temporal measures in literature reviews on fluency and some interactional measures.

Table 1 - *Most used temporal and interactional measures in research on fluency*²

Speed of speech measures	Formula / Description	Reference
speech rate	n° of syllables / total time	Kormos (2006) in: De Jong (2016)*
pruned speech rate	n° of syllables - n° of disfluent syllables / total time	De Jong (2016)
phonation time ratio	speaking time / total time	Kormos (2006) in: De Jong (2016)*
articulation rate	n° of syllable / speaking time	Kormos (2006) in: De Jong (2016)
mean length of syllables	speaking time / n° of syllable	De Jong (2016)*
mean length of run	n° of silent pauses / n° of syllables	Kormos (2006)*
Breakdown fluency measures		
<i>Frequency</i>		
number of pauses	n° of pauses / total time or speaking time	as in Saito et al. (2019)*
number of silent pauses	n° of silent pauses / total time or speaking time	Kormos (2006) in: De Jong (2016)*
mean length of utterance	total speaking time / n° of utterances	De Jong (2016)
number of filled pauses	n° of filled pauses / total time or speaking time	Kormos (2006) in: De Jong (2016)
<i>Duration</i>		
duration of silent pauses	pausing time / n° of silent pauses	Kormos (2006) in: De Jong (2016)
<i>Location</i>		
mid-clause pauses	interrupting the clause	Saito et al. (2019)*
final-clause pauses	junction pauses (clause and phrase boundaries)	Saito et al. (2019)*
Repair fluency measures		
n° of repetitions	n° of repetition / total time or speaking time	De Jong (2016)*
n° of repairs	n° of corrections and restarts / total time or speaking time	De Jong (2016)*
Turn-taking fluency measures		
pause	within speaker silence	Heldner & Edlund (2010)
gap	between speakers silence	Heldner & Edlund (2010)
overlap	turn-changing and non-turn-changing	Heldner & Edlund (2010)
backchannel	not taking the turn	Riggenbach (1991)
collaborative completion	attempt to complete a sentence or a phrase of the other	Riggenbach (1991)

Moreover, tasks used in experimental and assessment settings can have a large impact on L2 fluency. Derwing *et al.* (2004) argue that the task used to collect data can exert an influence on learners' fluency depending on its degree of freedom, since picture narrative and description impose a given range of lexicon and syntactic structures, while a monologue, or a conversation with free choice of topic allow learners to have much more control on the content and the expressions used to deliver it.

In non-experimental environments, assessment of L2 oral proficiency is often conducted through oral proficiency interviews, where a native speaker tries to elicit linguistic information from learners using a script representative of real-life language use settings. However, although these interviews try to simulate ordinary conversation, they present various constraints that can affect fluency (He, Young, 1998), e.g. interviews take place in an institutional setting; speech activities are predetermined; participants have different statuses, two different L1s and cultural backgrounds, as well as two different proficiency levels of the language used during the interview. As there are numerous factors which can negatively affect learners' fluency performance, a wider range of speaking tasks, in particular more open and interactive ones, should be employed.

² For the references of the studies in which the measures signalled with a star have been found to be significant see Saito *et al.* (2019).

For these reasons, especially in experimental settings, assuming L1 fluency measures as a baseline for each speakers' L2 fluency can help to explain some idiosyncratic differences by controlling for non-linguistic factors possibly affecting learners' performance (Segalowitz, 2016), such as contextual factors (e.g. attitude to the task and the interlocutor), but also learner-specific ones (e.g. personality and motivation).

2.2 Turn-taking fluency visualisation methods

Most studies mentioned in the previous paragraphs have focussed mainly on the concept of fluency as an individual phenomenon. Their main concern was to define if and how much a speaker is fluent, its perception from listeners' perspective and implications for L2 assessment, considering also that in many experiments monologic tasks were used for data collection.

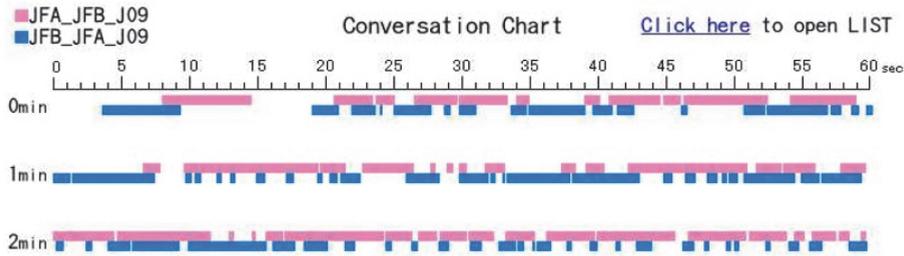
However, researchers agree that the circumstances in which learners' performance takes place is fundamental to make assumptions about the performance itself. Therefore, the measures above proposed only provide incomplete information about L2 fluency if considered out of context. For example, it has been suggested that speed of delivery and pausing behaviour are accommodated to the interlocutor during the interaction (Kourisidis, Dorran, 2009) so that a jointly achieved harmonisation of tempo occurs. This phenomenon has been depicted through the metaphor of interactional "flow" (McCarthy, 2009).

Moreover, being ruled by the turn-taking system (Sacks, Schegloff, & Jefferson, 1974), dyadic interactions better put to test learners' degree of automaticity. Indeed, turn-boundaries (also called transition relevant places – TRP – in the field of conversational analysis) are the places in which smooth or disfluent transition of turns can take place and which require to appropriately anticipate the end of the interlocutor's turn to be able to quickly react (Levinson, 2015; Bögels, Torreira, 2015).

For these reasons, judgements of fluency based on a single speaker and ignoring the interlocutors' contribution to the conversation would lack the interactive perspective and miss important information about learners' ability to co-create fluency (remember that interaction is also described as a co-creation process in the CEFR).

Some studies focussing on conversational speech rhythm have developed visualisation tools to represent the speech activity performed in dyadic exchanges. In his study on speech activity patterns in telephone conversations, Campbell (2007) created a plot to visualise the timing organisation of the interaction using four classes of activities: both interlocutors silent, both interlocutors talking simultaneously, only interlocutor A speaking, only interlocutor B speaking (Fig. 3). On the horizontal axis of the plot, a time window of one minute is displayed, whereas the vertical axis reproduces time passing throughout the interaction. The speaking activity was then represented on this graphic scaffold by colour-coded bars, whose length represents the duration of each speakers' turn. As a consequence, when the two different bars are upon each other, speakers are speaking at the same time causing an overlap, whilst when bars interrupt and a white space follows, speakers are being silent.

Figure 2 - A section of Campbell's conversation chart (2007:344)



The same type of plot has been used to display laughter in dyadic conversations (Trouvain, Truong, 2013) to investigate the interplay of laughing and speaking activities (Fig. 3). In this plot the four classes of activities are of different nature: filled and empty red bars signal respectively overlapping and non-overlapping laughs, whereas filled dark and grey bars represent speakers' speech (blue rectangles mark instead cases of interest to the study).

Figure 3 - Visualisation of speech and laugh activity in dyadic conversation (Trouvain et al., 2013:2)



This kind of charts displaying the duration and timing of specific classes of interest has two main benefits. It serves as an “eye-opener” (Trouvain, Truong, 2013:4) helping researchers to evaluate their intuitions by means of a visual exploration and comparison of data. Additionally, the data on duration and timing extracted to create the plot provide the material for acoustic analysis and statistics for hypothesis testing.

In this study, we test this type of visualisation plot on L2 interactional data to assess its representative power and usefulness in capturing differences in oral interaction management by learners with different levels of L2 proficiency. Due to its preliminary nature, we do not conduct any hypothesis testing, but instead concentrate on the application of the method, i.e. the extraction of metrics, the quantification of L1 and L2 interactions and the discussion of analysis results. To fill the gap in quantification methods for interactional competence, we suggest these visualisation and quantification tools as possible groundwork for developing a more complete instrument for a standardised assessment of L2 interactional competence.

3. Method

As a proxy for interactional competence, we explore the degree of fluency of the interaction co-created by participants through the turn-taking system. Following Campbell (2007), we operationalise the interactional flow by quantifying four classes of activities: the percentage of time each speaker takes the floor (1, 2), how long the speech of the participants overlaps (3) and the amount of total silence (4).

Subjects are 40 Italian students of German as L2 with different L2 proficiency levels assessed by the German language courses they were attending. Speakers were aged between 19 and 65, 33 were female and 7 were male³. Some learners were students at the “Goethe Institut” of Naples and others at the “Università degli Studi di Napoli L’Orientale”, with German as foreign language as one of the main subjects⁴. Speakers were matched according to their L2 proficiency level to perform the recording in pairs. In table 2 the dyads of students who performed the task together and their corresponding level of German are listed. However, because of constraints due to subjects’ time availability, dyads from 11 to 14 were mixed and are thus composed by an interlocutor with B1 level and the other with B2 level.

Table 2 - *Subjects matched to perform the task for data collection according to their level of L2 German. Dyads from 11 to 14 are mixed and composed by interlocutors with B1 and B2 levels respectively*

DYAD	LEVEL	DYAD	LEVEL
1	A1	11	B1-B2
2	A2	12	B1-B2
3	A2	13	B1-B2
4	B1	14	B1-B2
5	B1	15	B2
6	B1	16	B2
7	B1	17	B2
8	B1	18	B2
9	B1	19	C1
10	B1	20	C1

The corpus consists of forty dialogues, twenty in the native language – Italian, Neapolitan variety – and twenty in German as L2, following the suggestion to use

³ Apart from one student of the Goethe Institut aged 65, all other participants were aged between 19 and 38. Median age of participants = 21; standard deviation = 7,94.

⁴ 24 learners had benefited from a stay in German-speaking countries for a variable length of time (from one to ten months) either for a short language course, or an exchange period at a partner university. However, the effect of a period abroad is neither accurately quantifiable, nor equal for everyone. It varies according to the amount of input and use of the foreign language (consider, for example, exchange students who do not manage to establish regular contact with the locals, or decide not to attend a German language course). That said, since a period of immersion in the foreign language contributes to language proficiency, we did not consider this variable separately.

learners' L1 as a baseline against which to assess their L2 interactional patterns (Segalowitz, 2016). Indeed, we supposed that their way of interacting would not be similar to a native German one, since our subjects are living in Italy, studying in Italy and currently talking to an Italian interlocutor with whom they share the same L1 and culture. There is no foreign exposure which could favour them in approaching a German native style. Moreover, using the L1 as a baseline was informative in a previous exploration phase (Sbranna, Cangemi, Grice, 2019) conducted on two dyads, one with beginners and the other with advanced learners. The pilot study showed that with a higher L2 proficiency the interaction in L2 approached the same interactional pattern learners used in their own L1. Recordings were made at the "Goethe Institut" in Naples using headset microphones (AKG C 544 L) connected through an audio interface (Alesis iO2 Express) to a computer running *Praat* (Boersma, 2001).

The spontaneous speech data were elicited using the Map Task (Anderson *et al.*, 1991; Grice, Savino, 2003 for set up, map layout and instructions), a task-oriented dialogue which matches a kind of task described in the CEFR, the goal-oriented co-operation task. Participants sit opposite each other and have no eye-contact. They are provided with two maps; one speaker receives a map with a route drawn across landmarks – instruction giver – and has to describe the route to the other participant – instruction follower –, whose map only has landmarks. The goal is to co-operate so that the instruction follower can reproduce the route on their map thanks to the instructions given by the partner. Some landmarks are however different and speakers only discover it during the task, which creates unexpected problem-solving situations.

This task was chosen for two reasons. It can be performed at every proficiency level since learners should address the topic of grammar and vocabulary knowledge related to road indications at a beginner level according to the CEFR. In addition, following the suggestions of previous studies, it presents a fair degree of openness thanks to the unforeseen unmatched landmarks increasing the degree of spontaneity in interaction.

Participants first read an Italian version of the game instructions and carried out the task in their native language. Afterwards, before performing the task in L2, they watched a video with a German native speaker explaining the instructions again in German to help them get into the language and reduce the L1 bias. They kept the same role (either instruction giver, or follower) in both languages to prevent cross-language differences in their fluency from being attributable to the factors regarding their role in the task.

At the end of the recording session, learners were provided with an online test for lexical competence, the German version of Lextale (Lemhöfer, Broersma, 2012), originally designed to test English as L2 in an experimental setting and found to be a reliable indicator of general L2 proficiency.

3.1 Procedure

The annotation and extraction procedure is composed of three steps. After having extracted the two channels from the stereo recordings, a first step consisted in the

automatic labelling of interpausal units in *Praat* using the function of silent interval detection. Secondly, boundaries were manually checked and corrected to make sure that interpausal units were rightly identified, since some voiceless consonants were automatically labelled as silence. Finally, the Text Grids resulting from the two channels of each dialogue were used as input files for a *Praat* script, which generated a figure depicting each speaker's contribution to the interaction as it develops over time (Fig. 4, 5, 6, 7).

Each horizontal bar corresponds to an interpausal unit uttered by one of the two speakers involved in the task. Speakers are colour-coded with red being the instruction giver and black the instruction follower. Time unfolds from top to bottom – minutes –, and from left to right – seconds – so that the interaction can be followed as on a written page in a left-to-right writing system.

Figure 4 and 5 display a low proficiency dyad performing the task in L1 and L2. The first striking difference is the total length; in L2 these speakers need roughly half of the time more than in L1 to conclude the task. Moreover, differently from the L1 smooth pattern, the flow of interaction in L2 appears much more fragmented, with shorter turns and more frequent and longer pauses, especially during the turn of an individual speaker.

This observation is in line with the systemic perspective of fluency mentioned in the background, according to which utterance fluency measures mirror learners' cognitive fluency. Indeed, in the case of this dyad, who has a low proficiency in German (A1 level, beginner), it is not surprising to find lengthened within-speaker pauses in L2 as compared to their interactional behaviour in L1, since this reflects a high cognitive effort necessary for the retrieval of linguistic information and formulation of the linguistic message they want to convey to their interlocutor.

Figure 4 - *Visualisation of interactional flow (L1 Italian by dyad 1 – low L2 proficiency)*

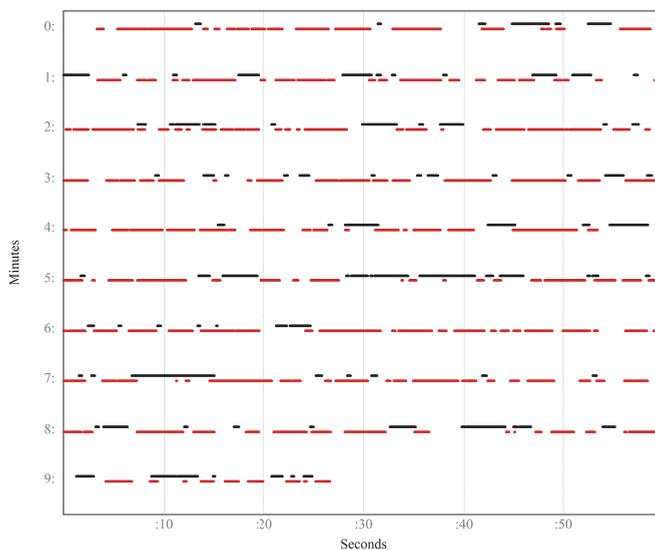
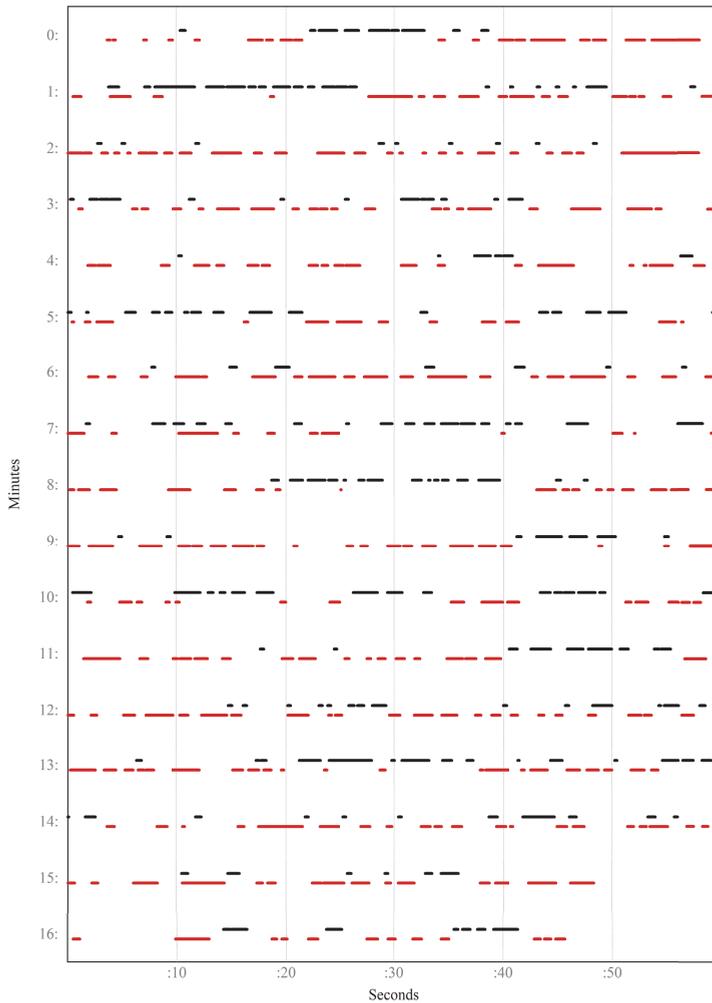


Figure 5 - *Visualisation of interactional flow (L2 German by dyad 1 – low L2 proficiency)*



Figures 6 and 7 show the interactional pattern of a dyad with high proficiency in L2 German (C1 level, advanced). In this case, it is difficult to identify at first sight which dialogue was carried out in the foreign language since the two interactional patterns look very similar. We can still notice a that in L2 the two highly proficient speakers need a little more time to complete the task, yet the difference is extremely slight and may be due to other factors generating variability in total duration of the interaction.

Figure 6 - Visualisation of interactional flow (L1 Italian by dyad 19 – low L2 proficiency)

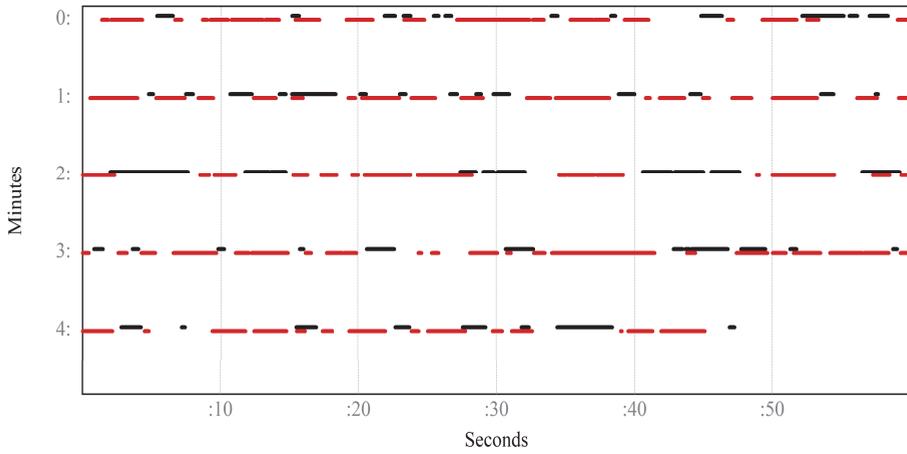
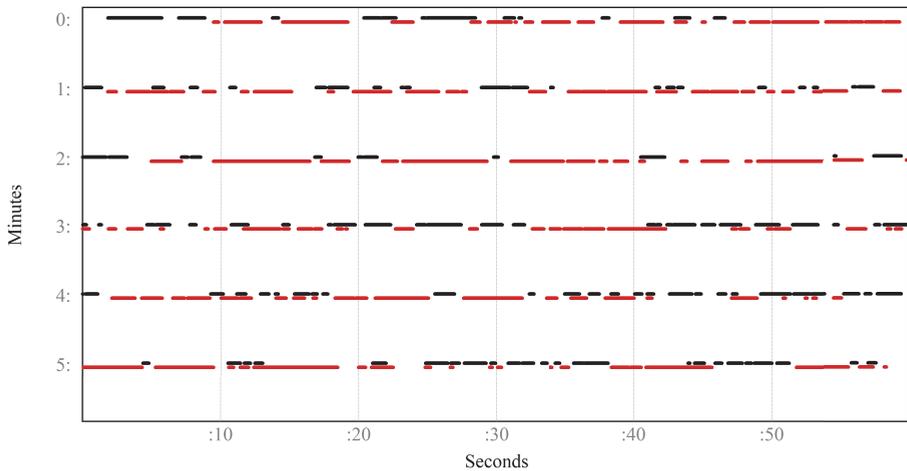


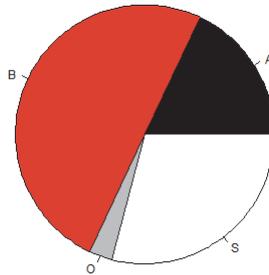
Figure 7 - Visualisation of interactional flow (L2 German by dyad 19 – low L2 proficiency)



The first exploration of data using these visualisation plots suggests that a higher proficiency level in L2, thanks to an enhanced automatization of the cognitive processes required to speak a foreign language, enables a degree of smoothness in managing the interactional flow closer to the one learners have in their native language.

In addition to this figure, the *Praat* script derives from the extracted data a table used to generate pie plots (example in Fig. 8) in *R* (R Core Team, 2013). The four sections of the pie plots use the same colour-coding of the visualisation plot to show the percentages of speech uttered by each speaker (A, instruction follower in black; B, instruction giver in red), the total amount of silence (S, in white) and overlap between speakers (O, in grey). The radius of the circle represents the total duration of the interaction, so that the bigger the pie is, the longer the interaction.

Figure 8 - Example of pie plot summarising the interactional flow



While the visualisation plot is useful to observe the time-aligned development of the interaction, this pie plot is helpful to summarize and quantify the partition of the interactional pattern into the four classes of activities and its total duration.

4. Results

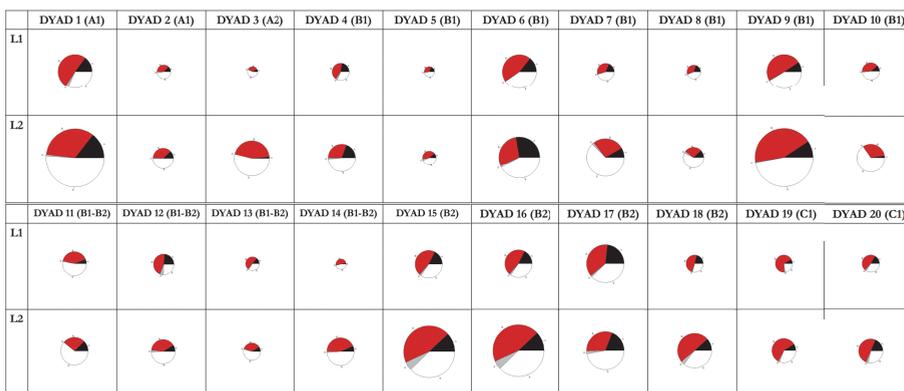
The pie plots described above capture the changing of the interactional patterns with increasing command of L2. As an example of their informativeness, we show four pie plots representing four dialogues performed by two dyads in L1 and L2 (Fig. 9): on the right there are two learners with low L2 proficiency and on the left two learners with high L2 proficiency.

Figure 9 - Pie plots for dyad 1 (low proficiency – A1, on the left) and dyad 20 (high proficiency – C1, on the right). Plots in the upper line display dialogues in L1 and plots below dialogues in L2

	DYAD 1 (A1)	DYAD 20 (C1)
L1		
L2		

The dyad with high proficiency of German (C1) presents two very similar patterns of interaction across languages. The ratio of time speaking between the giver and the follower remains 3:1 when they repeat the task in L2. Differently, the dyad with low proficiency of German (A1) features two very different interactional patterns. In L1 the ratio of time speaking between the giver and the follower is 3:1, whereas in L2 the ratio changes to 2:1, with the giver speaking less in L2 than in L1. Further, more than half of the conversation is in silence. The following figure depicts the pie plots for all interactions of the corpus (Fig. 10).

Figure 10 - Pie plots for the whole corpus ordered by increasing proficiency level of L2 German

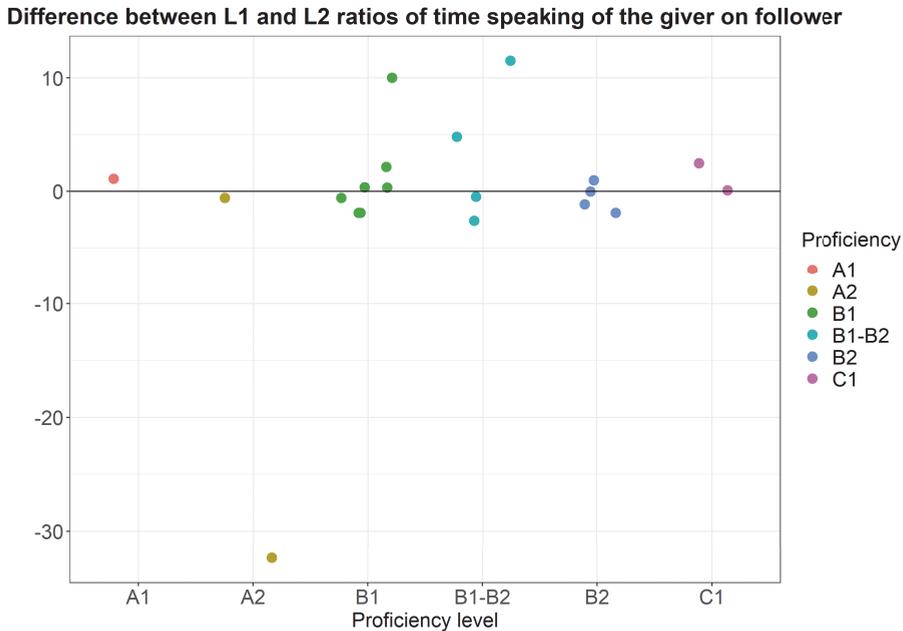


One commonality is that all interactions in L2 have a longer duration compared to those produced by the same dyad in L1. However, L2 interactional patterns start consistently resembling those in L1 from the dyads at B2 level of competence. Below this level interactional patterns in L1 and L2 reveal high variability across the two languages (except for dyads 2, 6 and 9).

To summarize data contained in all pie plots, we derived a graph calculating the difference between L1 and L2 ratios of time speaking between the giver and follower (Fig.11), so that each point in the graph represent a dyad and not a single speaker for the above-stated reason that interaction performances should not be analysed or evaluated in isolation as they are a form of co-creation of discourse by both – or, in case of multi-party conversations, all – interlocutors.

As in a first exploration, we noticed that the difference between the L1 and L2 ratios of time speaking between the giver and the follower is around 0 for high proficient learners (B2 and C1) and different from 0 for less proficient learners, we expect that with increasing proficiency, points will approach to 0 as indicating less difference in the interactional behaviour learners have across languages.

Figure 11 - *Difference between L1 and L2 ratios of time speaking of the giver on follower. The x-axis displays the proficiency level of L2 German, the y-axis shows the values resulting from the formula. The more points approach to 0, the less difference there is between learners' interactional behaviour in L1 and L2*



The graph shows high idiosyncratic variability. To start with the beginner group (A1 and A2 levels of proficiency), we can notice that two dyads are near to 0 and, therefore, present few differences in learners' interactional behaviour between L1 and L2. Indeed, grammatical and lexical resources are not the only factor contributing to conversational rhythm, since many other linguistic and extralinguistic factors play a role, e.g. personality, engagement in the task, relationship between speakers, not to forget the skill to strategically draw on the few resources beginners have to reach the goal of the interaction.

The golden outlier represents a dyad with particular behaviour – the follower only utters a few sentences towards the end of the dialogues in both L1 and L2 –. Especially in this case, having used the L1 as a baseline for learners' interactional behaviour has revealed to be useful, because this piece of data would have generally been associated with poor command of the L2, while we could observe that this speaker behaves exactly the same in L1, possibly due to lack of motivation and engagement in the task. Nevertheless, not disposing of many samples for the beginner level of proficiency, these dyads cannot be considered informative about a general trend.

On the other hand, the high variability displayed in the intermediate groups (B1 and mixed B1-B2 levels of proficiency) tends to reduce in the more advanced B2 and C1 levels. This observation would confirm the hypothesis emerged from the exploratory phase, i.e. learners with a higher proficiency of the L2 tend to repro-

duce in L2 the same interactional pattern they produce in their native language. In other words, less cognitive load, time for information retrieval and formulation, and attention required in L2 due to a higher degree of automatization allow learners to gain in smoothness of the interactional flow. Nonetheless, having only two dyads for the C1 level, a more conspicuous and homogenously distributed number of samples across proficiency groups would be required to confidently test this pattern.

Finally, to check an eventual relation of total duration of the dialogue and of lexical competence with the interactional pattern, we created two additional graphs and reduced the range of values on the y-axis cutting out the golden outlier to better visualise the remaining points (Fig. 12, 13).

In figure 12, the dimension of the circles represents the score obtained in the test for vocabulary knowledge by the giver, who is generally the one leading the task, so that the larger the circle is, the higher is the score they received. In figure 13, the dimension of the circles results from the difference in dialogue duration between L1 and L2, so that the bigger the point is, the longer is the time learners needed to accomplish the task in L2.

Figure 12 - *Relation between lexical competence and interactional pattern. The x-axis displays the proficiency level of L2 German, the y-axis shows the difference between L1 and L2 ratios of time speaking of the giver on follower. The dimension of the circles represents the score for givers' vocabulary knowledge*

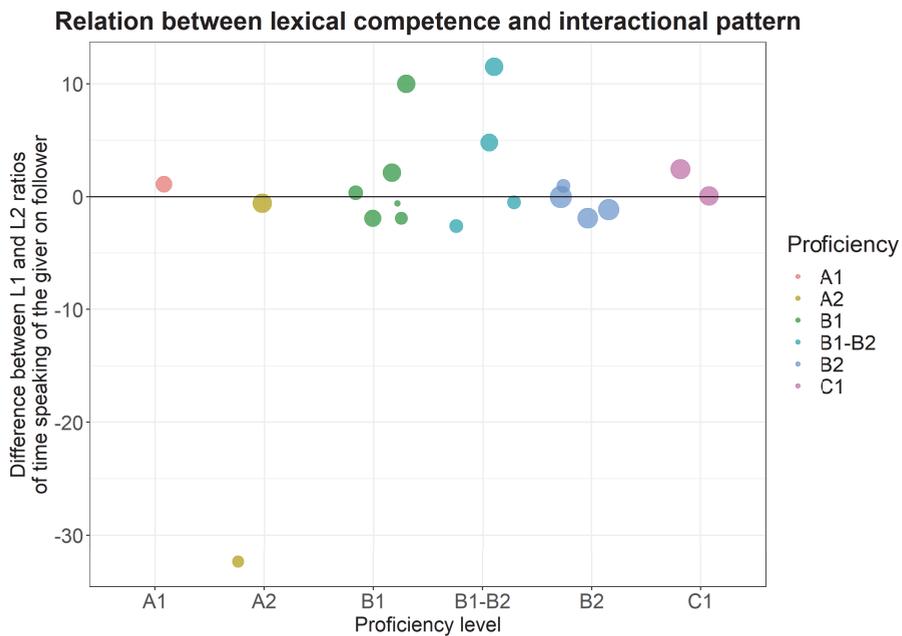
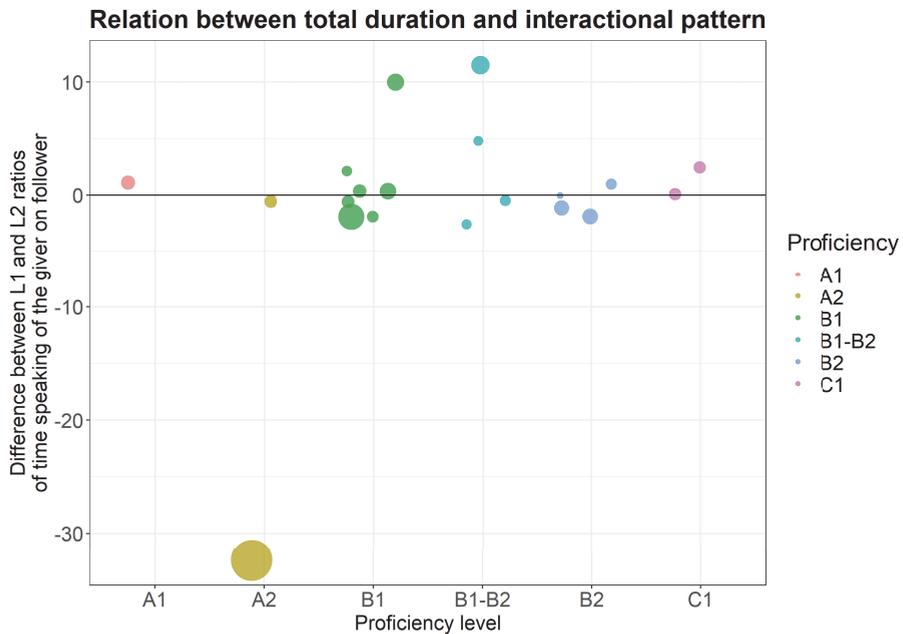


Figure 13 - Relation between total duration of the interaction and interactional pattern. The x-axis displays the proficiency level of L2 German, the y-axis shows the difference between L1 and L2 ratios of time speaking of the giver on follower. The dimension of the circles represents the difference in dialogue duration between L1 and L2



In neither case, it is possible to observe a clear trend. Especially in the graph displaying vocabulary scores (Fig. 12), in B1, B1-B2 and B2 groups both very low and high scores are near to zero, suggesting that lexical competence does not seem to be a factor determining how different the interactional patterns are in L1 and L2. Neither the second graph (Fig. 13) shows a clear pattern, but as the total duration of interaction may be at least partly dependent on a higher number of breakdown phenomena and slower speed of speech, which are directly related to L2 proficiency, it could be worth testing the relation between the difference in dialogue duration between L1 and L2 and interactional patterns on a larger corpus, which might provide a better understanding of the relationship between these two factors.

5. Conclusion

In this contribution, we problematised the absence of a standardised instrument for the quantification of interactional competence in L2. To open up new perspectives for L2 assessment, we focussed on the development of visualisation tools and a quantification method that can extract reliable and testable metrics for interactional aspects of communication.

We tested its informativeness on a corpus of L1 and L2 interactions including different levels of L2 proficiency. We found that more proficient learners are able

to maintain the natural interactional rhythm they have in L1 in their L2, as demonstrated by the interactional patterns displayed. We also observed that lexical competence does not seem to influence learners' interactional behaviour, which would suggest that mastering the lexicon does not automatically ensure a higher degree of success in oral interactions. Moreover, lexical scores do not seem to lead to corresponding levels of general competence. Indeed, the learning process is not linear and there is no discrete order in L2 knowledge acquisition (Nava, 2010), so that different skills can improve at different speeds. Since open interactional tasks put to test learners' L2 abilities in a more comprehensive way, an enhancement of interactional tasks in L2 experimental and testing settings can help to obtain a clearer picture of learners' L2 general proficiency and possibly shed light on the interplay among the different skills.

However, our corpus only included a few samples for beginner and advanced groups, in contrast to the more conspicuous intermediate group. Using the proposed metrics to draw statistical inferences on a corpus with a homogeneous number of samples across all L2 proficiency levels would make it possible to test these observations and collect additional evidence.

Despite this limitation in the sample, the method proposed was shown to be highly beneficial for the exploration and analysis of L2 oral interactions. It permits an immediate comparison of learners' interactional behaviour in L1 with their performance in L2 and can be used to observe the developmental trajectory of L2 interactional competence across different stages of the learning process. For these reasons, this method of visualisation and quantification of oral interactions could represent a starting point for quantifying L2 interactional competence in a standardised way.

In this study, we based our analysis on temporal measures only and did not provide an analysis of the transitions, i.e. sequences in which speakers change turns. A future development of this tool contributing to a more fine-grained analysis would be the integration of measurements for pragmatic and strategic competence from the CEFR scale for interaction (e.g. asking for clarification, compensating, cooperating, monitoring and repair). One option may be an analysis of conversational moves to assess the strategies learners use to coordinate the interaction throughout the learning process. However, such analysis presents some disadvantages inasmuch as it is time-consuming and may provide extremely variable results due to the highly idiosyncratic nature of interaction. Therefore, a pilot study evaluating the time cost and the informativeness of different pragmatic measurements about L2 interactional competence needs to be carried out to identify which ones can best and most effectively contribute to a fully comprehensive instrument for L2 assessment purposes.

Acknowledgement

We would especially like to thank the Goethe Institut of Naples for generously providing rooms for the recording sessions. Financial support for carrying out the

research is gratefully acknowledged from a.r.t.e.s Graduate School and CRC-1252 “Prominence in Language”, funded by the German Research Foundation.

Bibliography

- ANDERSON, A.H., BADER, M., BARD, E.G., BOYLE, E.H., DOHERTY, G.M., GARROD, S.C., (...) & WEINERT, R. (1991). The HCRC map task corpus. *Language and speech*, 34(4), 351-366.
- BACHMAN, L.F., PALMER, A.S. (1996). *Language testing in practice*. Oxford: Oxford University Press.
- BOERSMA, P. (2001). Praat, a system for doing phonetics by computer. *Glott International* 5: 9/10, 341-345.
- BÖGELS, S., TORREIRA, F. (2015). Listeners use intonational phrase boundaries to project turn ends in spoken interaction. *Journal of Phonetics*, 52, 46-57.
- BOSKER, H.R., PINGET, A.F., QUENÉ, H., SANDERS, T. & DE JONG, N.H. (2013). What makes speech sound fluent? The contributions of pauses, speed and repairs. *Language Testing*, 30(2), 159-175.
- CAMPBELL, N. (2007). Approaches to conversational speech rhythm: speech activity in two-person telephone dialogues. *Proceedings of XVIth International Congress of the Phonetic Sciences*, Saarbrücken, Germany, 6-10 August 2007.
- COUNCIL OF EUROPE (2001). *Common European framework of reference for languages: Learning, teaching, assessment*. Cambridge, U.K: Press Syndicate of the University of Cambridge.
- DE JONG, N.H. (2016). Fluency in second language assessment. In TSAGARI, D., BANERJEE, J. (Eds.), *Handbook of second language assessment*. Mouton de Gruyter, 203-218.
- DE JONG, N.H., GROENHOUT, R., SCHOONEN, R. & HULSTIJN, J.H. (2015). Second language fluency: speaking style or proficiency? Correcting measures of second language fluency for first language behavior. *Applied Psycholinguistics*, 36(2), 223-243. DOI: 10.1017/S0142716413000210.
- DE JONG, N.H., STEINEL, M.P., FLORIJN, A.F., SCHOONEN, R. & HULSTIJN, J.H. (2013). Linguistic skills and speaking fluency in a second language. *Applied Psycholinguistics*, 34(5), 893-916.
- DERWING, T.M., ROSSITER, M.J., MUNRO, M.J. & THOMSON, R.I. (2004). Second language fluency: Judgments on different tasks. *Language Learning*, 54, 655-679.
- DUNCAN, H., SEGALOWITZ, N. & PHILLIPS, N. 2014. Differences in L1 linguistic attention control between monolinguals and bilinguals. *Bilingualism: Language and Cognition*, 19(1), 106-121. doi:10.1017/S136672891400025X.
- FIGUERAS, N., NORTH, B., TAKALA, S., VAN AVERMAET, P., & VERHELST, N. (2009). *Relating language examinations to the common European framework of reference for languages: learning, teaching, assessment (CEFR): a manual*. Strasbourg, France: Council of Europe, Language policy division.
- FILLMORE, C.J. (1979). On fluency. In KEMPLER, D., WANG, W.S.Y. (Eds.), *Individual differences in language ability and language behavior*. New York: Academic Press, 85-102.

GRICE, M., SAVINO, M. (2003). Map Tasks in Bari Italian: Asking Questions about Given, Accessible and New Information. Invited contribution to special issue *Intonation in Romance*. *Catalan Journal of Linguistics* 2.2, 153-180.

HALL, J.K. (1993). The Role of Oral Practices in the Accomplishment of Our Everyday Lives: The Sociocultural Dimension of Interaction with Implications for the Learning of Another Language. In *Applied Linguistics*, 14, 145-166.

HALL, J.K. (1995). (Re)creating Our Worlds with Words: A Sociohistorical Perspective of Face-to-face Interaction. In *Applied Linguistics*, 16, 206-232.

HE, A.W., YOUNG, R. (1998). Language proficiency interviews: A discourse approach. In YOUNG, R. & HE, A.W. (Eds.), *Talking and testing: Discourse approaches to the assessment of oral proficiency*. Amsterdam: John Benjamins, 1-24.

JACOBY, S., OCHS, E. (1995). Co-Construction: An Introduction. In *Research on Language and Social Interaction* 28, 171-183.

KORMOS, J. (2006). *Speech production and second language acquisition*. Mahwah, N.J.: Lawrence Erlbaum Associates.

KORMOS, J., & DÉNES, M. (2004). Exploring measures and perceptions of fluency in the speech of second language learners. *System*, 32(2), 145- 164.

KOUSIDIS, S., & DORRAN, D. (2009). Monitoring convergence of temporal features in spontaneous dialogue speech. Dublin Institute of Technology: *Digital media Centre Conference Papers*. <http://arrow.dit.ie/cgi/viewcontent.cgi?article=1003&context=dmcon>.

LEMHÖFER, K., & BROERSMA, M. (2012). Introducing LexTALE: A quick and valid Lexical Test for Advanced Learners of English. In *Behavior Research Methods*, 44, 325-343. URL <http://www.lextale.com/>.

LEVINSON, S. (2015). Turn-taking in Human Communication – Origins and Implications for Language Processing. *Trends in Cognitive Sciences*. 20. 10.1016/j.tics.2015.10.010.

LOGAN, K.J. (2015). *Fluency disorders*. San Diego, CA: Plural.

MCCARTHY, M. (2009). Rethinking spoken fluency. *Estudios de lingüística inglesa aplicada*, 9, 11-29.

MAGNE, V., SUZUKI, S., SUZUKIDA, Y, ILKAN, M., TRAN, M.N. & SAITO, K. (2019) *Exploring the Dynamic Nature of Second Language Listeners' Perceived Fluency: A Mixed Methods Approach*. TESOL Quarterly. ISSN 0039-8322. <https://doi.org/10.1002/tesq.528>

NAVA, E.A. (2010). *Connecting phrasal and rhythmic events: evidence from second language speech*. PhD dissertation. University of Southern California.

PRÉFONTAINE, Y., KORMOS, J. & JOHNSON, D.E. (2016). How do utterance measures predict raters' perceptions of fluency in French as a second language?. *Language Testing*, 33(1), 53-73.

R CORE TEAM (2013). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. URL <http://www.R-project.org/>.

RIGGENBACH, H. (1991). Toward an understanding of fluency: A microanalysis of non-native speaker conversations. *Discourse Processes*, 14(4), 423-441. <https://doi.org/10.1080/01638539109544795>

- ROSSITER, M.J. (2009). Perceptions of L2 fluency by native and non-native speakers of English. *Canadian Modern Language Review*, 65(3), 395-412. <http://doi.org/10.3138/cmlr.65.3.395>
- SACKS, H., SCHEGLOFF, E.A. & JEFFERSON, G. (1974). A simplest systematics for the organization of turn-taking for conversation. *Language*, 50(4), 696-735.
- SAITO, K., ILKAN, M., MAGNE, V., TRAN, M.N. & SUZUKI, S. (2018). Acoustic characteristics and learner profiles of low-, mid- and high-level second language fluency. *Applied Psycholinguistics*, 39(3), 593-617.
- SBRANNA, S., CANGEMI, F., GRICE, M. (2019). Towards quantifying conversational competence in second language learning. Poster presented at *Phonetik und Phonologie Tagung*, Düsseldorf, Germany, 25-29 September 2019.
- SEGALOWITZ, N. (2010). *Cognitive bases of second language fluency*. New York: Routledge.
- SEGALOWITZ, N. (2016). Second language fluency and its underlying cognitive and social determinants. *International Review of Applied Linguistics in Language Teaching*, 54, 79-95.
- SEGALOWITZ, N., FREED, B.F. (2004). Context, contact and cognition in oral fluency acquisition: Learning Spanish in at home and study abroad contexts. *Studies in Second Language Acquisition*, 26(2), 173-99.
- STARKWEATHER, C.W. (1987). *Fluency and stuttering*. Englewood Cliffs, NJ: Prentice-Hall.
- SUZUKI, S. & KORMOS, J. (2019). Linguistic Dimensions of Comprehensibility and Perceived Fluency: An Investigation of Complexity, Accuracy, and Fluency in Second Language Argumentative Speech. *Studies in Second Language Acquisition*. 10.1017/S0272263119000421.
- TAVAKOLI, P. & SKEHAN, P. (2005). Strategic planning, task structure, and performance testing. In ELLIS, R. (ed.), *Planning and task performance in a second language*. Amsterdam: John Benjamins, 239-273.
- TROUVAIN, J. & TRUONG, K.P. (2013). Exploring sequences of speech and laughter activity using visualisations of conversations. *Proceedings of Workshop on Affective Social Speech Signals* (WASSS 2013), Grenoble, 22-23 August, 2013.
- ZMARICH, C. (2017). Stuttering and phonetic theory: An introduction. In *Social and biological factors in speech variation*, studi AISV – Collana peer reviewed, 3, 357-372.

PARTE V

RELAZIONI
A TEMA LIBERO

SONIA CENCESCHI, CHIARA MELUZZI, NICHOLAS NESE

Speaker's identification across recording modalities: a preliminary phonetic experiment¹

This work investigates how WhatsApp audio messages could be compared to high quality professional recordings and low quality ones in a forensic framework. A controlled experiment with 12 Italian students (6F, 6M) was performed in order to ascertain whether formants' values of the three cardinal vowels /a/-/i/-/u/ will help in distinguishing the same speaker across three different recording modalities. Both unnormalized data in Hertz and normalized values (Lobanov and Bark) were compared across the male and female subsets. Results indicate that unnormalized data performed better than normalized ones, and that a qualitative investigation has to be combined with a quantitative one. This preliminary work opens the way to further investigations on the possibilities of WhatsApp audio messages for forensic purposes at the crossroads between linguistics and engineering.

Keywords: forensic phonetics, WhatsApp, forensics linguistics, environmental recordings, vowels' formants, intra-speaker variation.

1. Introduction

This work addresses a central issue in forensic phonetics: the possibilities of identifying a speaker's voice through different recording modalities. With the growth of environmental recordings (Orletti & Mariottini, 2017; Ministero della Giustizia, 2018), and the spread of WhatsApp technology a new challenge opens for both linguists and engineers working in the forensic field. To put it simply, how is it possible to ascertain the identity of a speaker across different recordings, especially when those recordings are highly deteriorated? The growing interest in the possibilities of automatic or semi-automatic comparisons have usually worked on a suprasegmental or acoustic level, by leaving aside linguistic variation as expressed (Drygajlo, Jessen, Gfroer, Wagner, Vermeulen & Niemi, 2016; Tirumala, Shahamiri, Garhwal & Wang, 2017; Jagdale, Shinde & Chitode, 2020). At the same time, sociophonetics researches have demonstrated the great heuristic power of vowels' formants in explaining variabilities across and within speakers (e.g., Quené, 2008).

¹ This work has been conceived and written jointly by the three authors. However, for the Italian evaluation system, author 1 is responsible for section 2, 3, and 4.2 (with subsections); author 2 is responsible for sections 1, 4, 4.1, 5 and 6; author 3 is responsible for sections 4.3 and 7. Authors 1 and 2 have also conceived the experimental design of the work and analyzed the data, whereas author 3 performed the data collection and the annotation of the whole corpus.

In this respect, this work proposes a preliminary phonetic experiment focused on vowels' variation as produced by 12 speakers (6F, 6M) recorded in three different settings: a professional high quality recording, a WhatsApp audio message, and an environmental low quality recording.

The paper is organized as follows: section 2 presents the theoretical premises of the work, and section 3 the research questions we aimed at answering in this paper. The experimental design, the research protocol, and the corpus with the associated technical issues will be presented in section 4 (and subsections).

The fifth section presents the results of both a qualitative (5.1) and a quantitative (5.2) analysis on F1 and F2 of the three cardinal vowels /a/-/i/-/u/, with a summary of results (5.3) for the individual variation. Finally, section 6 discusses the results in light of their possible applications in forensic phonetics, and section 7 presents our first conclusions and further perspectives.

2. Some theoretical premises

Audio forensics daily deals with spontaneous speech, but the specialized research suffers from a lack of spontaneous speech corpora, mainly due to privacy reasons. As already introduced in Cenceschi, Trivilini, Sbattella & Tedesco (2019), we think that social media apps could represent a very large digital pool from which to draw (Kaplan, 2015). They constitute a fundamental part of modern human communication, increasing enormously in recent years among users of all ages. However, for the purposes of both linguistic and forensic analysis, it is necessary to investigate similarities and variations with respect to phone-calls and live speech, because the audio message as a category belongs to a new speech communication style (Nencioni, 1983; Cenceschi, Sbattella & Tedesco, 2018). As a consequence, audio messages have introduced new interaction behaviors such as different speakers' expectations, rhythm, pauses, etc. From a technical point of view, audio messages also provide data with different qualities, if compared to laboratory recordings.

Through the years, a large variety of spontaneous Italian speech corpora have been collected by scholars (Cresti, Moneglia, do Nascimento, Moreno-Sandoval, Véronis, Martin & Blum, 2002; Albano Leoni, 2006; Cresti & Panunzi, 2013). However, none of them focused on social media speech style. The inter-device speech features variability has also been investigated by various works, in order to highlight any differences between the audio in the various compressed formats (Nolan, Grigoras, 2005; Khan, Wiil & Memon, 2010; Gold, French, & Harrison, 2013; van Braak & Heeren, 2015). It remains up to investigate whether this two dimensions co-occur in shaping the variability of speech data, that is to say how speech varies both across style (and, in particular, social-media styles) and recording modality. In this respect, some explorative studies have addressed the issue of intra-speaker and inter-devices variations for a limited number of speakers. For instance, Cenceschi et al. (2018) consider only 2 speakers, and compare WhatsApp

data with phone-calls, in order to verify the hypothesis that these data will be equivalent for most of the parameters despite the diatechnical variations.

The forensic consequence will be that WhatsApp data can be used as a parallel or even alternative source in forensic investigations.

However, the main difference between recording modality was of prosodic and temporal nature, thus leaving an open hypothesis on what happens on the segmental level.

3. *Research questions and main objectives*

Given the lack of studies considering speech variation between recording modalities from a linguistic perspective, we decided to start with a first explorative study to investigate the intra-speaker and inter-device variations for a limited number of speakers. The main purpose of this paper is to understand whether and to what extent is possible to compare WhatsApp data with high quality and low quality recordings. In particular we will consider if different audio typologies affect the main phonetic characteristics for voice identification in forensic settings. Therefore, our specific research questions for this first study are:

1. Are recordings made with different devices directly comparable among them?
2. What is the impact of common normalization procedures on comparability between different recording formats?
3. To what extent is, then, possible to compare/identify the voice of the same speaker from different recording devices in forensic analysis?

Indeed, when working in forensics it frequently happens to compare two audio samples recorded in different modalities. The usual request from the law forces to the expert (i.e., the phonetician) is to ascertain whether the two samples could belong to the same speaker.

Conversely, in the present experiment, we are certain that the voices belong to the same speakers recorded in different modalities. The main aim of our work, thus, is to verify if a semi-automatic investigation of some phonetic features will confirm that the samples belong to the same voice. If the results will lead to the emergence of a difference in speakers' identification according to the recording devices, this will tell something important for what it concerns the possibilities of vocal comparison in phonetic forensics.

4. *The WAsp Corpus*

Our preliminary study consisted in the creation of a small corpus (called *Wasp*) based on the productions of 12 speakers (6M, 6F) with the same sociolinguistic characteristics. They were all students enrolled at the University of Pavia in various courses, with a preference for non-linguists. All students were born and living in the north-west of Italy, and they were Italian L1; no bilingual students have been included, although all participants have knowledge of various foreign languages

(e.g., English). All speakers were consciously and freely taken part in the experiment as volunteers, receiving no compensation for their participation in the project. Speakers personal data were anonymized and protected according to the current ethical and privacy dispositions. Ethical and privacy agreement was signed by both the researchers and the participants.

Each speaker in the Wasp corpus performed two different tasks in three different recording conditions. The two tasks consist in a sentence-list reading of 30 sentences, with a pause of about 3" between each sentence, and a description task, thus producing short monologues about the furniture of their room and the explanation of the cooking of their favorite food.

Each speaker has asked record with three different modalities (audio formats are deepened in the related paragraph):

1. Recorded by an expert, simulating a high quality comparative forensic registration in a sound-proof environment, by means of a Tascam DR-05.
2. Auto-recording of a WhatsApp audio message.
3. Through a phone call made by the researcher and recorded through the App Voice Recorder 2.81 in *.mp3*.

During each recording session, the speaker was asked to repeat the sentence-list reading and the description tasks twice, with a short pause between the two recordings, in partial accordance with the Protocol for the collection of databases of forensic recordings (Morrison, Rose & Zhang, 2012).

Speakers were recorded by the third author in a soundproof room at the boarding school 'Giasone del Maino' in Pavia, in the afternoon of 4th December 2019. The boarding school was chosen because all participants were hosted there, and the soundproof room offered the ideal environment for high quality recordings.

After this first session, speakers were asked to reach their private room and, when they felt comfortable, send a WhatsApp message to the second author by following the same protocol (i.e., reading list and short description, each one repeated twice, possibly in two different messages). When receiving the messages, the researcher controlled them for their completion (e.g., presence of both the tasks, and of the two repetitions), eventually soliciting the repetition of a task.

Finally, the second researcher called each participant and recorded them repeating the two tasks twice though the App Voice Recorder installed on her phone. The speakers were asked to receive the phone calls in a possibly silent environment like their personal bedrooms, and the researcher made the calls from her office, which was not a soundproof room.

The data were, thus, acquired in three different moments, albeit very close to each other. It was judged difficult to record the same speaker at a single moment in the three modalities, because the mobiles would have created interferences with the microphone. We do not believe that speakers will change much in style while performing a reading task, and this was also one reason for opting for a reading task vs. a real dialogical task: as a preliminary research, our main interest was only devoted to recording modalities, without other variables involved. The data acquired

according to these three different modalities were stored according to task, and by identifying each speaker through an alphanumeric label in order to later compare their speech through the different recording modality.

4.1 Task and acoustic features

The sentence-reading task contains 12 target words, balanced by target stressed vowels /a/, /i/ and /u/ followed by a singleton or geminate alveolar or bilabial consonant (e.g., *Papa, pappa, Tita, Titti*). An equal number of fillers was also added in each list. The list was presented to the speaker in a randomized order.

Target words and fillers appear in sentences with similar prosodic contour (e.g., *La tata guarda i bimbi al parco* “The nanny looks after the kids in the park”).

The similarity of prosodic profile was aimed at reducing the involved variables as much as possible, to ensure repeatability, and to lay the ground for future studies gradually introducing further parameters (e.g., enlarge the dataset to sentences containing a pragmatic accent to understand how its characterization changes according to the recording methodology). According to the proposal of Cenceschi *et al.* (2018) and van Braak & Heeren (2015), we limited the study to read speech and short spontaneous descriptions, without pragmatic accents, emotions, dialog interaction, and by maintaining the same talking speed (as far as possible) in the three settings.

For the same reasons, the recording must be realized in a silent room, but without specific details regarding the environmental soundproofing. The aim is to simulate a recording in normal everyday life, without any particular noisy conditions that will be introduced in future studies.

4.2 Technical equipment and digital formats

The three recording modalities produce different audio outputs whose digital quality mainly depend on: hardware equipment (Microphone, CPU, chipsets, etc.), and compression formats. The different characteristics are shown below with those of the professional microphone.

4.2.1 Hardware smartphone equipment

The 12 speakers have phones of different brands, and specifically: 5 iPhone (3 iPhone-7, 2 iPhone-8, 1 iPhone-11), 2 Huawei (1 P10 and 1 P20), and one unit for Xiaomi (Redme note7), Asus (Z00ED), Google (pixel 2 XL), Honor (version 10), and Samsung (J330FN). Although the technology of the different smartphone models can cause quality variations, albeit minimal to the human ear, this variable will not be investigated here as the sample is too limited. Then, we consider the possible discrepancies momentarily irrelevant in the context of a macroscopic analysis. A future enlargement of the sample will allow further analyzes concerning the variability between the various categories.

4.2.2 Professional microphone

The high quality recordings were performed with a Tascam DR-05, without any external microphones connected to the recorder. Its sound quality guarantees over 92dB signal to noise ratio, under 0.05% total harmonic distortion and 20Hz to 40kHz response (-1/+3dB) at 96kHz/24-bit resolution. Its output has been setted as *wav* at 44.100 Hz - 16 bit in order to simulate the higher quality recordings typical of forensic investigations.

4.2.3 Format compression and conversion

The difference between low and high quality file formats depends on the modality used to encoding or decoding a digital data stream. The corpus comprises four different data formats: *wav*, *mp3*, *ogg*, and *m4a* depending on the recording modality:

- Professional recordings: uncompressed *wav* files 44.100 Hz - 16 bit.
- Voice Recorder 2.81 *mp3*, 16.000 Hz - 128 kbps, codec Lame..
- WhatsApp: *ogg*, 64 kbps, codec Vorbis.
- WhatsApp: *m4a* 64 kbps, codec AAC.

WhatsApp is based on the SILK VoIP (Voice over Internet Protocol) codec developed by Skype and now licensed out, being available as open-source freeware: the voice (and other media) are delivered over an IP packet switched network. It is a foundation, with CELT, of the hybrid codec Opus. WhatsApp exports now Opus files with pseudo file extension *m4a* (AAC codec) or as *ogg* (Vorbis codec) because from about 2018, the Opus format is not recognized by many apps. The VoIP technique has a major impact on the spectro-acoustic properties of the signal (as already addressed for automatic speaker recognition in Khan, Baig & Youssef (2010) because it was introduced to preserve the network bandwidth to the detriment of signal quality (Singh & Mian, 2016). Voice Recorder version 2.99 by Splend App works for Android 4.1+. It allows a variable bitrate from 32 up to 320 kbps, and sampling rate from phone quality (8 kHz) to CD quality (44 kHz). It has been setted with *mp3* 16 kHz - 128 kbps. The app has been installed on one author's smartphone and started just before the phone calls with the speaker. It exploits the microphone of the mobile phone on which it is installed: as a consequence the two voices have unbalanced intensity, and the quality of the speaker's speech is very limited with respect to other recording modalities. This phenomenon well approximates the worst forensic recordings, as realized in real context scenarios. Furthermore, it should be remembered that private dialogues are often recorded with similar modalities, without adequate technological knowledge, and through applications found on the web. The *ogg* and *m4a* recordings have been converted to *wav* in order to allow the Praat analysis as usually performed in linguistics (De Decker, Nycz, 2011; Styler, 2013). As underlined in (Wang et al. 2018), the up-sampled recording does not modify speech features because acting a PCM linear conversion from a compressed format to a higher quality one.

4.3 The corpus

We base our present analysis on 864, that is 12 target words, twice repeated in three different recording settings by 12 different speakers. The corpus is balanced for sex of the speaker, recording condition, target vowels (/a/, /i/, /u/), phonotactic environment (singleton, geminate), and surrounding consonants (either bilabials or alveodentals, in both cases voiceless). Each target word was manually annotated on three different tiers in PRAAT: a first tier included the whole sentence, whereas on the second tier the target word was isolated, and on the third tier we segmented the target vowel and the following consonant.

For setting the vowel's left and right boundaries we base on the beginning and end of the second formant (F2); the occlusive consonants include the whole silence phase and the following VOT, and it ends when the F2 of the following vowel appears.

After the annotation, we automatically extracted the following acoustic parameters: pitch, F0, F1 and F2 of the target vowels, jitter and shimmer, duration of both the target vowel and the following consonant. It should be mentioned that we extracted the formants' values twice: at the midpoint, at five different timepoints through the whole segment. Although in this paper we will focus only on formants' variations in a static approach (i.e., by looking at midpoint values), further research will include a dynamic approach on formants' variation across the segment, together with jitter and shimmer analysis. The data have been inserted in a matrix on the software IBM SPSS 20, and also visually inspected through the web application Visible Vowels (Heeringa & Van de Velde, 2018).

5. *Analysis*

Since the preliminary nature of this work, the dataset was balanced but limited in the amount of samples, so that a detailed statistical analysis with all the possible variables (e.g., phone label) can't be performed. The analysis will thus consist of a first qualitative analysis, performed through the inspection of formants' variations and vowel space variation through Visible Vowles, and a second quantitative analysis on formants' variation by performing different Anovas on IBM SPSS 20.

5.1 Qualitative analysis

A first visual inspection of formants' variation in the three recording modalities highlights some differences with respect also to vowel quality. In the graphs, as well as in the following statistical analysis, we maintain as separated the values for males and females in our corpus, because of the notorious biological differences affecting formant values (especially in non-normalized data). The three different recording modalities will be indicated as MIC for the high quality recordings made by the expert, WA for the WhatsApp audio messages sent by the participants, and VR for the phone call recordings made through the Voice Recorder Application.

Figure 1 - *Graphic representations of mean values variation of F1 (above) and F2 (below), with unnormalized data, in the three recording modalities (MIC = microphone, VR = Voice Recorder, WA = Whats.App) divided by speakers' sex. The visualization method indicated as TL is based on Fox & Jacewicz (2009)*

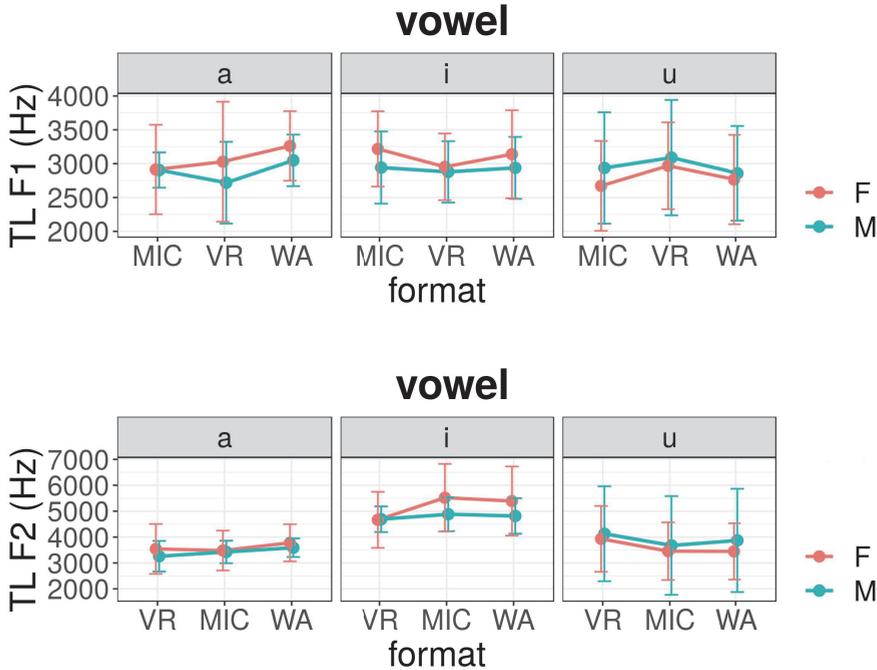
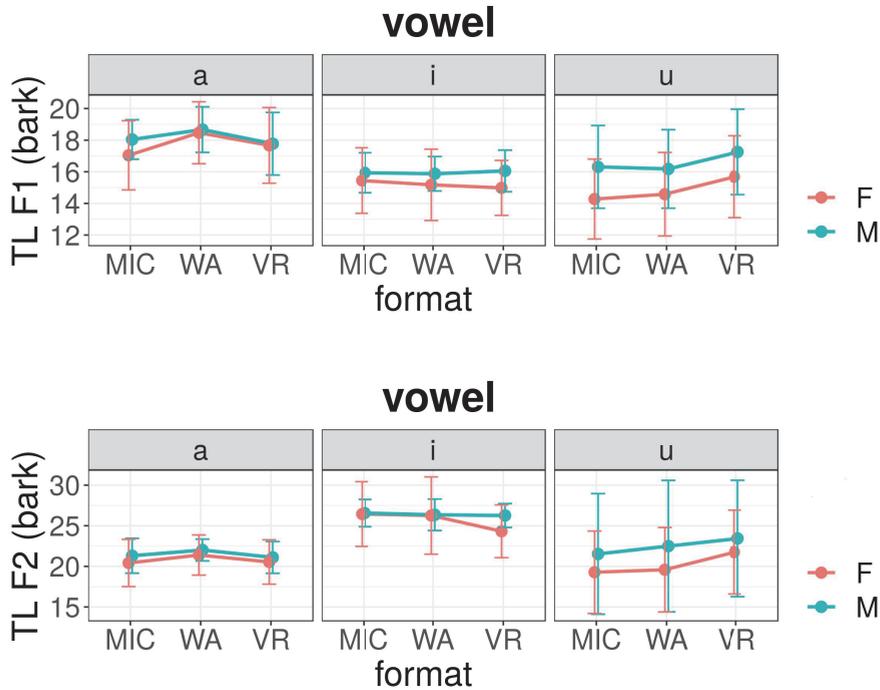


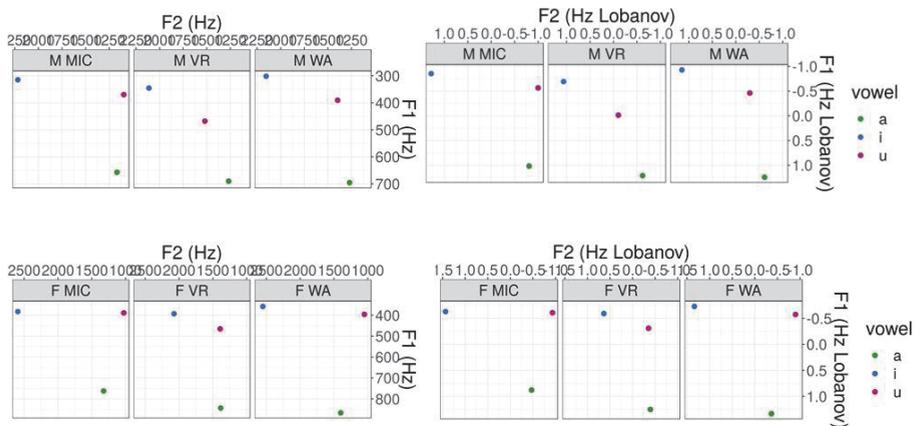
Figure 1 presents the variation of the mean values in the female and male subgroups according to the three recording modalities. It appears that the variation is, in general, quite minimal, with a similar behavior between males and females, but with some differences with respect to the formant and the vowel quality. For instance, as it concerns F1, it is possible to notice how values are higher for /u/ in the low quality recordings (VR) than in the two other modalities. Conversely, for /i/, F1 values are lower in the VR modality than in either high quality recordings and WhatsApp message. As for F2, an overall similarity is noticeable for /a/, whereas for the two high values there is again a difference between the low quality recordings (VR), on the one side, and the high quality recordings and WhatsApp messages, on the other side. Thus, from this preliminary investigation it appears that the low quality recordings as performed through phone call registration present the most different values, whereas the mean values emerging from WhatsApp audio messages are similar to the ones recorded in a professional environment.

Figure 2 - Graphic representations of mean values variation of F1 (above) and F2 (below), with data normalized in Bark (Traunmüller, 1990), in the three recording modalities (MIC = microphone, VR = Voice Recorder, WA = WhatsApp) divided by speakers' sex. The visualization method indicated as TL is based on Fox & Jacewicz (2009)



In Fig. 2 the same data are presented, but the values of F1 and F2 have been normalized in Bark through Traunmüller's (1990) formula. As it appears from the graphs above, for both formants the variation is minimal for the vowel /a/, although there is a greater dispersion in F2 for male subjects. As noticed for non-normalized data, both high quality recordings (MIC) and WhatsApp messages (WA) show similar values if compared to low quality recordings (VR).

Figure 3 - Vowel space representation of variation of the three cardinal vowels /a/-/i/-/u/ in the three recording modality (MIC = microphone, VR = Voice Recorder, WA = WhatsApp), with respect to speaker's sex (M = males, above; F = females, below). Unnormalized data are presented on the left, normalized data with Lobanov's (1971) formula are presented on the right



We also compare the position of the three cardinal vowels accordingly to their mean values of F1 and F2, with data unnormalized and normalized through Lobanov's (1971) formula (Fig. 3). Indeed, many studies in sociophonetics have highlighted how Lobanov's (1971) normalization procedure should be preferred because it preserves more information on the socio-indexical dimension (cf. Van der Harst, 2011; Adank, 2003). However, this normalization procedure usually works better with huge corpora: this is not the case either in our work nor, more generally, in forensics.

Data presented in Figure 3 show little or no difference between unnormalized mean values and normalized ones, both in the male subgroup (above) and in the female one (below). Moreover, once again it seems possible to highlight a major similarity between high quality recordings (MIC) and WhatsApp audio messages (WA). However, the values for /a/ represent an exception since WA modality shows a lower F2 and a higher F1 than in both MIC and VR. Conversely, VR setting seems to play a major role on the posterior vowel /u/, which appears to be more centralized than in MIC and WA settings. In the female subgroup, there seems also to be an influence on the values of /i/, which appears to be more centralized in VR recordings than in the other two settings.

To sum up, this first qualitative inspection of our data has pointed out some important points. Firstly, the variation between recording modalities doesn't appear to be a huge one, in particular between professional recordings and WhatsApp audio messages. In this respect, the phone call recordings realized with the app Voice Recorder are more dissimilar from the other two settings, in particular for what it concerns the two extreme vowels /i/ and /u/. Furthermore, it has been observed that Lobanov's normalization does not differ much from unnormalized data; for this reason, it has been decided to exclude this procedure from the next quantitative analysis.

5.2 Quantitative analysis

After a qualitative analysis of the different formant values across the three recording modalities, the question arises on whether these differences are only descriptive or not. In other words, we would like to investigate if a statistical analysis on formants' values would certify that the speaker is the same across the different recording modalities. Indeed, if the test will result significantly (i.e., $p < 0.05$), it will mean that the difference between the two (or three) settings is so huge to be attributed to different speakers. Conversely, if no significance will be found, it will mean that, albeit some differences due to the quality of the recordings, it is still possible to recognize that the different vowels belong to the same speakers.

We run Anovas on both unnormalized values and values normalized in Bark. Since male and female formant values are very different due to biological reasons, in both cases we maintain the two subgroups separated. A post hoc Tukey test was also performed on both F1 and F2 values, in Hertz and in Bark, in order to verify if differences between the recording modalities are statistically significant.

From the data it emerges that in both male and female subgroups the comparison of formant values in the three recording modalities is always statistically significant ($p < 0.05$). An exception is represented by the F2 of the vowel /a/, but only in the female subgroup ($p = 0.063$).

By looking at the post hoc Tukey test results, the following picture emerges for the two subgroups. For the female speakers, F1 unnormalized values are not significant for /a/ between WhatsApp and Voice Recorder settings, for /i/ between high quality recordings and the other two modalities, and for /u/ only between high quality recordings and WhatsApp messages. In this subgroup, F2 values are significantly different among all devices for /a/, but only in comparing high quality recordings and WhatsApp messages for /i/ and /u/. Conversely, no significance has been found in the cross-modalities comparison with Bark values, with the exception of the MIC-WA comparison for the F2 of /i/ and /u/.

For the male group, F1 values for /a/ were statistically significant in comparing WhatsApp messages and Voice Recorder, both with unnormalized and normalized Bark data. However, for /i/ and /u/ only unnormalized data show a difference between high quality recordings and WhatsApp audio messages. The same could be said for the F2 of /i/ and /u/, both for unnormalized and normalized Bark data. With Hertz value a difference also has emerged between WhatsApp and the Voice Recorder App for the F2 of /i/. Finally, with Bark data, there was always statistical significance in comparing the three recording modalities when considering the F1 of both /i/ and /u/.

As for individual variation, unnormalized data predicted better the coincidence of the speaker across recording modalities, whereas normalized Bark data always resulted in a statistical significant difference, with but 1 exception (cf. Appendix). Indeed, the ANOVA also shown that for some speakers unnormalized data could predict their identity across the three recording modalities (e.g., speaker CM), and

that the back vowel /u/ preserved better this individual difference, with respect to both /i/ and /a/.

To sum up, it appears that normalized data performed worse than unnormalized data in recognizing the same voices across recording modalities. Among vowels, /a/ seems to create major confusion, especially for what it concerns F1 values. Conversely, the vowel /i/ appears to better perform in recognizing the three recording modalities as belonging to the same speakers. Moreover, the two low quality recording modalities (i.e., WhatsApp and Voice Recorder) generate more confusion than the comparison between a low quality recording (especially WA) and high quality one (MIC).

6. Discussion

This work has focused on a common problem in forensic phonetics: the comparability and recognizability of speakers across different speech samples, recorded with different modalities. This problem has usually been addressed from an engineering point of view (cf. 4.2.3): for instance, Khan et al. (2010) and Singh et al. (2016) proposed a semi-automatic speech recognition system by considering the loss of spectral information. In this work, we intended to address the issue from a linguistic point of view, with a phonetic analysis of vowels' formants variation across recording modalities.

The results of both our qualitative and quantitative analysis both point at a major reliability of direct values, without normalization. As expected, the low quality recordings realized with the app Voice Recorder badly performed because of the high formants variability. It is important to stress that this represents the typical forensic case, when an environmental interception (similar, for quality, to our VR setting) has to be compared with professional recordings (our MIC setting). Therefore, our results suggest that this comparison should be addressed with a semi-automatic analysis only with extreme caution. A combination of quality and quantity analysis seems to be preferable, especially when working with a small dataset, as it frequently happens in phonetic forensics. For what it concerns the different vowels, the central vowel /a/ seems to be more indicated for comparing low-quality audio files (such as VR and WA). Conversely, extreme vowels /i/ and /u/ seem more suitable for forensic comparison, especially between low quality recordings and high quality ones. However, this could be a language-specific difference that should be tested on inter-linguistically with a similar research protocol.

Finally, it has been repeatedly noted how normalization prevents recognition of different samples as belonging to the same voices across different recording modalities. Bark normalization presents some exceptions, but with a considerable variability not only among the cardinal vowels considered but also between the male and female subgroups. The Lobanov normalization was completely ineffective or counter-productive in our case, but this is probably due to the small dimension of our corpus, since Lobanov's formula squeezes the values too much. However, this

lack of informativity of normalized data could also depend on the target approach chosen for this analysis, with values extracted on the midpoint of the stressed vowels. Although this is the most common practice in forensic phonetics, it is true that (socio)phonetic analysis nowadays relies more on dynamic approaches for vowel analysis (e.g., Farrington et al. 2018, van der Harst et al. 2014).

7. Conclusions and further perspectives

Forensic linguists are always asked to ascertain the identity of a speaker across short speech samples frequently recorded with different sound qualities. In this paper, we proposed a laboratory experiment aimed at exploring the possibilities of (semi) automatic comparison of formant values. We recorded 12 speakers in three different modalities (i.e., professional high quality audio, WhatsApp messages, low quality environmental recordings). The values of F1 and F2 of the three cardinal vowels /a/-/i/-/u/ were compared both qualitatively and quantitatively, and with and without normalizing the original Hertz data. The analysis allows us to answer our research questions, by also opening the fields for further discussions and experiments on this topic.

Indeed, we show that audio files made with different recording devices, and especially low and high quality ones, suggest to combine a qualitative inspection of data distribution through graphic representations (e.g., with the online software Visible Vowels) with a statistical comparison of formants' values. Normalization procedures usually adopted in (socio)phonetic analysis (i.e., Bark and Lobanov) do not work well with small subsets like the ones commonly available for forensic comparisons. In particular, Lobanov squeezes the values too much for allowing a comparison, whereas Bark normalized data are randomly significant in recognizing the same speakers across recording modalities. In particular when working with extremely compromised audios, a substantial precautions in linguistic and phonetic analysis is needed. Comparing the voices of a possible same speaker from different recording devices for forensic purposes is possible, but a qualitative analysis has to be combined with a quantitative one.

Furthermore, WhatsApp audio messages turn out to be a good compromise between good quality (professional) recordings and low quality (environmental) ones. This leads to hypothesize that their use in forensic phonetics will increase in the future, also because of their availability. From a linguistic point of view, WhatsApp messages could also be said to represent a new form of expressive modality (Nencioni, 1983), conceivable halfway between spontaneous and recited speech.

Obviously, further experiments should be conducted on more spontaneous samples, since a word list reading task is quite different from real speech from a stylistic and a phonological point of view, as it has been pointed out by previous scholars (cf. 2). However, a preliminary investigation like the one proposed here was necessary to ascertain the difference between recording modalities without the 'noise' generated by variability in spontaneous speech. Further studies will, thus, address

other phonetic and phonological variables (e.g., prosodic contour), and widen the analysis to spontaneous social-media speech. It will also be desirable to confirm production analysis with perceptive tests, in order to verify whether and to what extent recording modalities affect our capability to recognize the speakers, especially with extremely deteriorated recordings.

All these issues are extremely important for the practical application of linguistic analysis to forensics. Although until now they have been scarcely addressed, as far as we know, from a linguistic point of view, much work has been done by engineers. Therefore, an interdisciplinary approach will benefit the investigation and strengthen the results on speakers' semi-automatic comparability.

Bibliography

- ADANK, P.M. (2003). *Vowel Normalization. A Perceptual acoustic study of Dutch Vowels*. Netherlands Graduate School of Linguistics: LOT.
- ALBANO LEONI, F. (2006). *Il corpus CLIPS*, presentazione del progetto. Dostupno na: [http://www.clips.unina.it/it/\[25.10.2012\]](http://www.clips.unina.it/it/[25.10.2012]).
- CENCESCHI, S., SBATELLA, L. & TEDESCO, R. (2018). Verso il riconoscimento automatico della prosodia. In *STUDI AISV*, 433-440.
- CENCESCHI S., TRIVILINI A., SBATELLA L. & TEDESCO R. (2019) *Collecting Italian spontaneous social media speech: the WAsp2 project*, AISV conference 2019, Arezzo (Italy).
- CRESTI, E., MONEGLIA, M., DO NASCIMENTO, F.B., MORENO-SANDOVAL, A., VÉRONIS, J., MARTIN, P. & BLUM, C. (2002). The C-ORAL-ROM Project. *New methods for spoken language archives in a multilingual romance corpus*. In LREC.
- CRESTI, E., PANUNZI A. (2013). *Introduzione ai corpora italiani*. Bologna: Il Mulino.
- DE DECKER, P., NYCZ, J. (2011). For the record: Which digital media can be used for sociophonetic analysis?. *University of Pennsylvania Working Papers in Linguistics*, 17(2), 7.
- DRYGAJLO, A., JESSEN, M., GFROERER, S., WAGNER, I., VERMEULEN, J., & NIEMI, T. (2016). *Methodological guidelines for best practice in forensic semiautomatic and automatic speaker recognition*. Verlag für Polizeiwissenschaft.
- FARRINGTON, C., KENDALL, T., & FRIDLAND, V. (2018). Vowel Dynamics in the Southern Vowel Shift. In *American Speech: A Quarterly of Linguistic Usage*, 93(2), 186-222.
- FOX, R.A. & JACEWICZ, E. (2009). Cross-dialectal variation in formant dynamics of American English vowels. In *The Journal of the Acoustical Society of America*, 126(5): 2603-18.
- GOLD, E., FRENCH, P. & HARRISON, P. (2013). Examining long-term formant distributions as a discriminant in forensic speaker comparisons under a likelihood ratio framework. In *Proceedings of Meetings on Acoustics*, 19, 1-7.
- HEERINGA, W., VAN DE VELDE, H. (2018). Visible Vowels: a Tool for the Visualization of Vowel Variation. In *Proceedings CLARIN Annual Conference 2018, 8 - 10 October, Pisa, Italy*. CLARIN ERIC.
- JAGDALE, S.M., SHINDE, A.A., & CHITODE, J.S. (2020). Robust Speaker Recognition Based on Low-Level-and Prosodic-Level-Features. In J. Vanita, G. Chaudhary, M.C.

- Taplamacioglu & M.S. Agarwal (eds.) *Advances in Data Sciences, Security and Applications*, Singapore: Springer, pp. 267-274.
- KAPLAN, A.M. (2015). Social Media, the Digital Revolution, and the Business of Media. In *International Journal on Media Management*, 17(4), 197-199.
- KHAN, A., WILL, U.K., & MEMON, N. (2010). Digital forensics and crime investigation: Legal issues in prosecution at national level. In *2010 Fifth IEEE International Workshop on Systematic Approaches to Digital Forensic Engineering* (pp. 133-140). IEEE.
- KHAN, L.A., BAIG, M.S., & YOUSSEF, A.M. (2010). Speaker recognition from encrypted VoIP communications. In *Digital investigation*, 7(1-2), 65-73.
- MINISTERO DELLA GIUSTIZIA (2018). *Relazione del Ministero sull'amministrazione della giustizia*, p. 25.
- MORRISON, G.S., ROSE, P., & ZHANG, C. (2012). Protocol for the collection of databases of recordings for forensic-voice-comparison research and practice. In *Australian Journal of Forensic Sciences*, 44(2), 155-167.
- NENCIONI, G. (1983). *Di scritto e di parlato*. Bologna: Zanichelli.
- NOLAN, F. & GRIGORAS, C. (2005). A case for formant analysis in forensic speaker identification. In *Journal of Speech, Language and the Law*, 12, 143-173.
- ORLETTI F. & MARIOTTINI L. (2017). *Forensic Communication in Theory and Practice: A Study of Discourse Analysis and Transcription*, Cambridge: Cambridge Scholars Publishing.
- QUENÉ, H. (2008). Multilevel modeling of between-speaker and within-speaker variation in spontaneous speech tempo. In *The Journal of the Acoustical Society of America*, 123(2), 1104-1113.
- SINGH, H., MIAN, M. (2016). Comparative Study and Analysis of various VoIP coding Algorithms. In *International Journal of Computer Applications*, 141(2).
- STYLER, W. (2013). Using Praat for linguistic research. *University of Colorado at Boulder Phonetics Lab*.
- TIRUMALA, S.S., SHAHAMIRI, S.R., GARHWAL, A.S., & WANG, R. (2017). Speaker identification features extraction methods: A systematic review. In *Expert Systems with Applications*, 90, 250-271.
- TRAUNMÜLLER, H. (1990) Analytical expressions for the tonotopic sensory scale. In *Journal of the Acoustical Society of America*, 88 (1), 97-100.
- VAN BRAAK, P., HEEREN, W.F.L. (2015). "Who's calling, please?" Is there speaker-specific information in twins' vowels?, Bachelor's thesis, Utrecht University.
- VAN DER HARST, S. (2011). *The vowel space paradox: A sociophonetic study on Dutch*. Netherlands Graduate School of Linguistics: LOT.
- VAN DER HARST, S., VAN DE VELDE, H., & VAN HOUT, R. (2014). Variation in Standard Dutch vowels: The impact of formant measurement methods on identifying the speaker's regional origin. In *Language Variation and Change*, 26(2), 247-272.
- WANG, Z., YAN, D., WANG, R., XIANG, L., & WU, T. (2018). Speech resampling detection based on inconsistency of band energy. *CMC-Comput., Mater. Continua*, 56(2), 247-259.

*Appendix**Anova on individual variation across recording modalities*

Speaker	Vowel	Formant	Normalizatin	Anova F(2,21)	p value
AB_F	a	F1	None (Hertz)	45.855	0.0001
			Bark	80.179	0.0001
		F2	None (Hertz)	2.217	0.134*
			Bark	86.47	0.0001
AM_F	a	F1	None (Hertz)	24.774	0.0001
			Bark	36.219	0.0001
		F2	None (Hertz)	16.997	0.0001
			Bark	1.194	0.323*
CM_F	a	F1	None (Hertz)	30.793	0.0001
			Bark	137.379	0.0001
		F2	None (Hertz)	5.179	0.15
			Bark	132.519	0.0001
DS_M	a	F1	None (Hertz)	6.712	0.06
			Bark	77.347	0.0001
		F2	None (Hertz)	1.895	0.175*
			Bark	21.32	0.0001
ER_F	a	F1	None (Hertz)	9.829	0.0001
			Bark	100.417	0.0001
		F2	None (Hertz)	0.624	0.545*
			Bark	152.526	0.0001
JF_M	a	F1	None (Hertz)	1.603	0.203*
			Bark	188.648	0.0001
		F2	None (Hertz)	1.465	0.254*
			Bark	16.416	0.0001
LS_M	a	F1	None (Hertz)	311.404	0.091*
			Bark	104.406	0.0001
		F2	None (Hertz)	6.142	0.0001
			Bark	41.649	0.0001
OQ_M	a	F1	None (Hertz)	13.553	0.0001
			Bark	299.105	0.0001
		F2	None (Hertz)	16.76	0.0001
			Bark	119.552	0.0001

Speaker	Vowel	Formant	Normalizatin	Anova F(2,21)	p value
SB_M	a	F1	None (Hertz)	2.18	0.138*
			Bark	12.429	0.0001
		F2	None (Hertz)	1.475	0.252
			Bark	8.211	0.002
SR_M	a	F1	None (Hertz)	0.122	0.088*
			Bark	545.137	0.0001
		F2	None (Hertz)	0.542	0.589*
			Bark	64.511	0.0001
VG_F	a	F1	None (Hertz)	6.041	0.0001
			Bark	313.315	0.0001
		F2	None (Hertz)	2.272	0.128*
			Bark	10.898	0.001
VL_L	a	F1	None (Hertz)	16.09	0.0001
			Bark	409.9	0.001
		F2	None (Hertz)	3.177	0.062*
			Bark	212.769	0.0001
AB_F	i	F1	None (Hertz)	12.82	0.0001
			Bark	340.616	0.0001
		F2	None (Hertz)	340.616	0.0001
			Bark	26.257	0.0001
AM_F	i	F1	None (Hertz)	31.617	0.0001
			Bark	207.533	0.0001
		F2	None (Hertz)	45.659	0.0001
			Bark	23.322	0.0001
CM_F	i	F1	None (Hertz)	0.105	0.901*
			Bark	208.323	0.0001
		F2	None (Hertz)	2.7	0.09*
			Bark	194.302	0.0001
DS_M	i	F1	None (Hertz)	1.057	0.365*
			Bark	61.709	0.001
		F2	None (Hertz)	5.191	0.15
			Bark	10.351	0.001
ER_F	i	F1	None (Hertz)	13.096	0.0001
			Bark	362.458	0.001
		F2	None (Hertz)	5.117	0.015
			Bark	73.722	0.0001

Speaker	Vowel	Formant	Normalizatin	Anova F(2,21)	p value
JF_M	i	F1	None (Hertz)	17.193	0.0001
			Bark	215.032	0.0001
		F2	None (Hertz)	2.866	0.079*
			Bark	13.935	0.0001
LS_M	i	F1	None (Hertz)	1.387	0.272*
			Bark	80.012	0.0001
		F2	None (Hertz)	1.791	0.191*
			Bark	7.915	0.003
OQ_M	i	F1	None (Hertz)	6.917	0.005
			Bark	100.186	0.0001
		F2	None (Hertz)	1.875	0.178*
			Bark	10.374	0.001
SB_M	i	F1	None (Hertz)	16.002	0.001
			Bark	101.575	0.0001
		F2	None (Hertz)	8.377	0.002
			Bark	10.674	0.001
SR_M	i	F1	None (Hertz)	45.603	0.0001
			Bark	140.523	0.0001
		F2	None (Hertz)	13.499	0.001
			Bark	10.888	0.01
VG_F	i	F1	None (Hertz)	5.688	0.011
			Bark	82.799	0.0001
		F2	None (Hertz)	5.879	0.009
			Bark	34.256	0.0001
VL_L	i	F1	None (Hertz)	2.486	0.107*
			Bark	74.345	0.0001
		F2	None (Hertz)	51.672	0.0001
			Bark	16.389	0.001
AB_F	u	F1	None (Hertz)	3.132	0.064*
			Bark	552.616	0.0001
		F2	None (Hertz)	8.594	0.002
			Bark	163.334	0.0001
AM_F	u	F1	None (Hertz)	21.158	0.0001
			Bark	121.192	0.001
		F2	None (Hertz)	1.839	0.184*
			Bark	8.544	0.002

Speaker	Vowel	Formant	Normalizatin	Anova F(2,21)	p value
CM_F	u	F1	None (Hertz)	0.122	0.886*
			Bark	169.6	0.0001
		F2	None (Hertz)	0.13	0.879*
			Bark	57.544	0.001
DS_M	u	F1	None (Hertz)	4.572	0.022
			Bark	214.468	0.0001
		F2	None (Hertz)	8.558	0.002
			Bark	37.126	0.0001
ER_F	u	F1	None (Hertz)	9.184	0.001
			Bark	162.399	0.0001
		F2	None (Hertz)	0.904	0.42*
			Bark	10.452	0.001
JF_M	u	F1	None (Hertz)	8.228	0.002
			Bark	56.516	0.001
		F2	None (Hertz)	19.969	0.001
			Bark	8.173	0.02
LS_M	u	F1	None (Hertz)	2.434	0.112*
			Bark	211.774	0.001
		F2	None (Hertz)	0.074	0.929*
			Bark	4.327	0.027
OQ_M	u	F1	None (Hertz)	11.791	0.0001
			Bark	137.459	0.0001
		F2	None (Hertz)	48.9	0.63*
			Bark	56.019	0.0001
SB_M	u	F1	None (Hertz)	14.302	0.0001
			Bark	421.666	0.0001
		F2	None (Hertz)	2.377	0.117*
			Bark	11.651	0.0001
SR_M	u	F1	None (Hertz)	0.656	0.529*
			Bark	86.624	0.0001
		F2	None (Hertz)	2.475	0.108*
			Bark	28.585	0.0001
VG_F	u	F1	None (Hertz)	3.323	0.056*
			Bark	196.262	0.0001
		F2	None (Hertz)	3.242	0.059*
			Bark	10.899	0.001

Speaker	Vowel	Formant	Normalizatin	Anova F(2,21)	p value
VL_L	u	F1	None (Hertz)	3.191	0.062*
			Bark	67.891	0.0001
		F2	None (Hertz)	2.565	0.101*
			Bark	10.493	0.001

BARBARA GILI FIVELA, SONIA I. D'APOLITO, GIORGIA DI PRIZIO

Controllo motorio e disartria nella malattia di Parkinson: uno studio pilota sulla labializzazione

Motor control in Parkinson's Disease: An investigation on lip rounding

Parkinson's disease is often associated with the hypokinetic type of dysarthria, which is characterized by movements that are reduced in both amplitude and speed. As far as dysarthric speech is concerned, these features of dysarthria have been correlated with a reduction in the phonetic space used by speakers, e.g. in vowel articulation. The aim of the study described in this paper is to investigate the production of Italian high vowels by Parkinsonian dysarthric subjects, with specific attention to both the acoustic and the articulatory characteristics of their speech, by looking at formant values, amplitude of the lingual and labial gestures as well as duration and articulation rate. In particular, the discussion focuses on the acoustic and articulatory measurements (electromagnetic articulography – AG501) obtained by recording three dysarthric subjects and three healthy controls from Lecce. Results show that, in addition to the expected tendency to reduce articulatory gestures as for both their amplitude and duration, an increase concerning measures obtained for some articulators and along some axes may also be observed. Data collected in relation to tongue and lip movement suggest that gestures are modulated to carry out compensatory strategies aimed at preserving relevant linguistic features, such as the acoustic characteristics of rounded in comparison to unrounded vowels.

Keywords: Pathological speech, Parkinson's Disease, dysarthria, speech production, labialization.

Introduzione

La maggior parte degli individui affetti dalla malattia di Parkinson sviluppa una forma di disartria, un disturbo motorio che compromette la realizzazione articolatoria dei suoni linguistici e che non implica necessariamente una compromissione sul piano cognitivo. Tra le disartrie, quella che si sviluppa più di frequente come disturbo secondario tipico della malattia di Parkinson è la disartria ipocinetica che, per quanto è di interesse in questa sede, causa una riduzione dell'ampiezza di movimento degli arti con conseguente *target undershooting* (Ackermann & Ziegler, 1991); inoltre, i movimenti sono caratterizzati da particolare lentezza, o bradicinesia (Duffy, 2005). Sebbene alcuni studi sulle vocali (Skodda et al. 2011, 2012) abbiano confermato che nel Parkinson la riduzione dell'ampiezza del gesto ha effetti sullo spazio acustico utilizzato per le vocali e sull'articolazione dei suoni

linguistici (nell'analisi dei quali il gesto articolatorio può essere l'unità fonologica di riferimento – Browman & Goldstein 1992), altri studi effettuati con articolografia elettromagnetica (Wong et al., 2010, 2011) hanno riportato risultati diametralmente opposti, ossia che nei Parkinson l'ampiezza del gesto è maggiore rispetto a quanto osservato nei soggetti di controllo; altri ancora mostrano che l'ampiezza dei gesti è ridotta o meno a seconda dell'articolatore e addirittura dell'asse lungo il quale avviene il movimento (Gili Fivela et al., 2014, dove si osservano peculiarità relativamente all'asse antero-posteriore), benché, considerando i dati intra-speaker, le modifiche non interferiscano con una distinzione fonologicamente rilevante (Iraci, 2017). Altra caratteristica tipica della disartria ipocinetica conseguente al Parkinson è la forte variabilità nella velocità di eloquio, caratterizzata da accelerazioni esagerate e pause inattese (Canter, 1963, Monrad-Krohn 1947, Teston, Viallet 2001, Duez 2006), benché questo si osservi soprattutto nel parlato semispontaneo, più che nel parlato letto (De Looze et al., 2012), dove però si può riscontrare spesso un'elevata velocità di lettura a discapito dell'accuratezza del parlato (Wisniecki et al., 2006).

Per cercare di chiarire ciò che accade ai suoni linguistici, soprattutto quelli che implicano la sincronizzazione di più articolatori e il loro movimento sia sull'asse verticale che sull'asse orizzontale antero-posteriore, in questa sede analizziamo i principali correlati acustici e articolatori delle vocali alte in italiano, la vocale anteriore non arrotondata /i/ e la posteriore arrotondata /u/.

2. Studio pilota sulle vocali alte italiane: obiettivi e ipotesi

In questo studio, ci proponiamo di verificare se i principali correlati acustici e articolatori delle vocali alte in italiano, la vocale anteriore non arrotondata /i/ e la posteriore arrotondata /u/, ci permettano di comprendere meglio qual è l'impatto della disartria ipocinetica sull'articolazione del parlato e, in particolare sulla realizzazione dei gesti articolatori lungo i diversi assi interessati nella produzione dei suoni. In particolare, intendiamo verificare se 1) lo spazio acustico utilizzato dai parlanti disartrici differisca da quello usato dai parlanti di controllo e 2) si osservino differenze relative ai gesti articolatori di lingua e labbra che possano essere correlati a eventuali differenze acustiche; inoltre, siamo interessate a verificare se 3) se ci siano differenze che riguardano anche la durata e la velocità del movimento. Le nostre ipotesi sono che a) lo spazio acustico sia ridotto nel parlato disartrico, in linea con quanto osservato per altre lingue (Skodda et al. 2010, 2011) e che b) coerentemente, i gesti articolatori siano in genere ridotti, come atteso in caso di disartria ipocinetica. Tuttavia, in linea con risultati ottenuti precedentemente (Gili Fivela et al., 2014), ci aspettiamo b1) di riscontrare una maggior ampiezza del gesto linguale antero-posteriore e, in linea con l'idea che la compensazione giochi un ruolo importante nel preservare le principali caratteristiche linguistiche (Iraci, 2017), di poter correlare la maggior ampiezza del gesto linguale con l'ampiezza del gesto labiale, sempre sull'asse antero-posteriore. Infine, ipotizziamo di riscontrare c) differenze nella durata e velocità degli articolatori dei parlanti disartrici rispetto a quelli dei controlli, ma non siamo in

grado di ipotizzare la direzione della variabilità, per via dei frequenti rallentamenti ed accelerazioni che caratterizzano il parlato disartrico (e.g., Duez 2006).

2.1 Corpus e metodo

Il corpus usato per questa indagine è parte di un corpus già utilizzato per analisi sulle consonanti dell'italiano (Gili Fivela et al., 2014; Iraci, 2017), ottenuto selezionando le pseudoparole del tipo 'CVCV in cui le V sono /i/-/u/ o /u/-/i/ e le C sono sempre bilabiali, sorde o sonore (/p/, /b/). Le pseudoparole sono state inserite nella frase cornice "Lu/i X blu" (con "Lu/i" a seconda delle V in X, anche grazie al fatto che nel dialetto/varietà della zona esiste l'articolo maschile singolare "lu"). Soggetti affetti dalla malattia di Parkinson che hanno sviluppato una disartria di tipo ipocinetico (3, età compresa tra 65 e 80 anni) e soggetti di controllo (3, della stessa età dei soggetti disartrici, ma non affetti da problemi neurologici e/o articolatori), tutti appartenenti all'area linguistica di Lecce, sono stati registrati sia dal punto di vista acustico che articolatorio per mezzo di un sistema EMA-AG501 (CarstensGmbH), mentre leggevano per 7 volte le frasi bersaglio proposte in ordine casuale. I soggetti disartrici al momento della registrazione avevano raggiunto lo stesso livello di compromissione, in base alla valutazione clinica, mentre i soggetti di controllo non avevano riportato alcun rilevante problema di salute. Per effettuare la registrazione articolatoria, sono stati applicati 2 sensori sul piano medio sagittale della lingua (per dorso e punta), 2 sulle labbra (labbro superiore e inferiore) e 3 dietro alle orecchie e sul naso (per la normalizzazione dei dati).

I materiali sono stati verificati percettivamente durante la fase di analisi uditivo-acustica. Si è verificato che le produzioni includessero le sequenze attese (benché in alcuni casi le occlusive potessero essere prodotte più come fricative o approssimanti), mentre si segmentavano manualmente, in PRAAT (Boersma, Weenink, 2019), le vocali e le consonanti, individuando tutti i confini segmentali. In seguito, sono state ricavate in modo semiautomatico le durate di tutti i segmenti presi in considerazione e le prime due formanti di entrambe le vocali, in sillaba tonica e posttonica. La durata dei segmenti è stata poi usata per calcolare la durata delle sillabe e dell'intero enunciato, al fine di ottenere anche la velocità di articolazione, espressa come la durata media delle sillabe e ottenuta grazie al rapporto tra la durata dell'enunciato e il numero di sillabe realizzate.

I dati cinematici sono stati ricavati grazie a MAYDAY (Sigona et al. 2015). Il segnale è stato etichettato in relazione alla traiettoria e alla velocità sia del dorso della lingua (TD) che del labbro inferiore (LL), lungo gli assi verticale e orizzontale (anteriore-posteriore). Il bersaglio articolatorio è stato individuato in corrispondenza dello zero di velocità e, per ogni segmento considerato, è stato anche etichettato il massimo di velocità. Per quanto riguarda la Consonante e la Vocale tonica (nella sequenza 'CVCV), sono state poi misurate ampiezza e durata dei gesti di apertura, chiusura di lingua/labbro e protrusione del labbro inferiore. Ad esempio, per ogni gesto di chiusura, è stata calcolata la durata come intervallo di tempo tra l'apertura massima e la massima chiusura degli articolatori; l'ampiezza è stata calcolata come

componente verticale/orizzontale dello spostamento durante il gesto. Tutti i gesti presi in considerazione sono stati analizzati sia sul piano verticale che su quello orizzontale. Si noti che per la durata di sillabe ed enunciato, si è fatto riferimento all'onset/offset dei gesti (identificato come il punto nel tempo in cui l'articolatore lascia/raggiunge la posizione stabile) calcolato automaticamente come il 20% del picco di velocità in modulo nei pressi dello zero di velocità¹. Per la velocità di articolazione, anche nel caso dei dati articolatori è stata calcolata la durata media della sillaba, ottenuta grazie al rapporto tra la durata articolatoria dell'enunciato (ma si veda nota 1) e il numero di sillabe realizzate².

I dati sono stati analizzati statisticamente in R per mezzo del pacchetto *lme4* (Bates et al., 2014; R Core Team, 2019). Sono stati considerati come effetti fissi, con interazione, *Sonorità* (2 livelli, *sordo* vs. *sonoro*), *Ciclo vocalico* (2 livelli, *i_u_i* vs. *u_i_u*), *Popolazione* (PAT vs. CONTR) e *Ripetizione* (7 livelli) sono stati usati per informare il modello circa la variabilità attesa tra le due popolazioni, PAT e CONTR, anche in relazione alla stanchezza dovuta alle ripetizioni; *intercept* e *random slope* sono stati inseriti per rendere conto della variabilità inter-soggettiva. L'ispezione visiva dei *residual plot* non ha mostrato evidenti deviazioni dalla normalità e omoschedasticità. Per verificare la significatività degli effetti fissi è stato utilizzato il *Likelihood Ratio Test*.

2.2 Risultati

2.2.1 Risultati acustici

Il fattore *Ciclo vocalico* qui considerato (IU, ossia il contesto vocalico /i-u-i/ vs. /u-i-u/ in UI), influenza la produzione del parlato, e, almeno nella dimensione orizzontale, i soggetti patologici sembrano usare uno spazio vocale ridotto rispetto ai parlanti di controllo.

La prima formante (F1) misurata nella vocale tonica (V0, /i/ in UI e /u/ in IU) diminuisce di circa 34 Hz \pm 4 (S.E.) nella vocale /i/rispetto alla vocale /u/ (ossia in UI rispetto a IU; $\chi^2(1)=60.884$ $p<0.001$), mentre la seconda formante (F2) ovviamente aumenta in /i/ di circa 1193 Hz \pm 23 (S.E.) rispetto ad /u/ (ossia nel ciclo UI rispetto a IU; $\chi^2(1)=441.08$, $p<0.001$). Ma ciò che è rilevante per il nostro studio è che si riscontrano interazioni con il fattore *Popolazione*. In particolare, per quanto riguarda F1 rileviamo l'interazione tra *Popolazione* e *Sonorità*, con la F1 che diminuisce di circa 19 Hz \pm 8 (S.E.) nei soggetti patologici, quando la consonante è sorda ($\chi^2(1)=6.1123$, $p<0.05$).

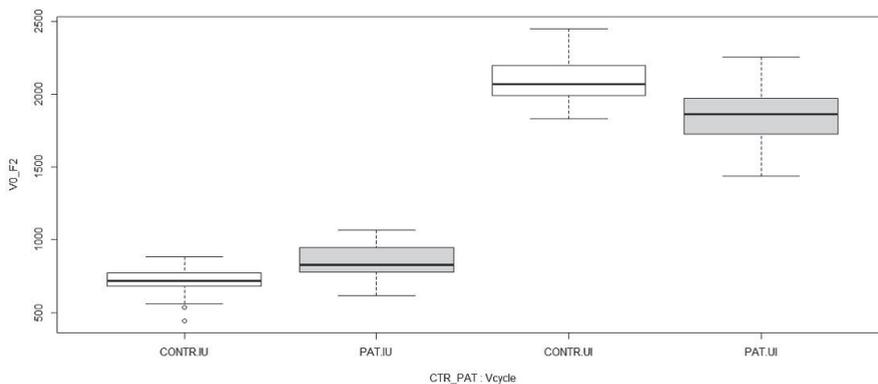
¹ Per l'enunciato, dal punto di vista articolatorio non disponiamo di una misura veramente completa, perché la nostra finestra di osservazione corrisponde piuttosto a quella della parola prosodica interna alla frase target.

² Data la struttura molto controllata delle frasi bersaglio e la modalità di elicitazione dei materiali, il calcolo della velocità di articolazione non è sostanzialmente più informativo rispetto a quello relativo alla durata dell'enunciato, per quanto riguarda l'indicazione circa la velocità. Abbiamo comunque riportato i risultati inerenti al calcolo suddetto per dare un'idea della durata media della sillaba nel compito effettuato dai soggetti.

Relativamente a F2, invece, è interessante l'interazione *Popolazione e Ciclo vocalico*, con la F2 che diminuisce di circa $375 \text{ Hz} \pm 35$ (S.E.) nei soggetti patologici rispetto al valore riscontrato nei soggetti di controllo, nel ciclo vocalico UI, quindi in /i/ ($\chi^2(1)=84.294$, $p<0.001$) – (Figura 1).

Nel caso dei patologici, quindi, lo spazio vocalico risulta essere ridotto sull'asse antero-posteriore (e i dati acustici suggeriscono un minor avanzamento della lingua nel passaggio da /u/ a /i/ tonica) e più alto quando le consonanti in contesto sono sorde (suggerendo che in questi casi la posizione della lingua sia leggermente abbassata).

Figura 1 - Valori della seconda formante (F2) nella vocale tonica, al variare della popolazione (CONTR vs. PAT) e del ciclo vocalico (IU = /u/tonica; UI = /i/ tonica)



Per quanto riguarda le durate della sillaba accentata e dell'intero enunciato, risultano fattori significativi la sonorità, il ciclo vocalico e la loro interazione con la popolazione.

La durata della sillaba tonica diminuisce di circa $14 \text{ ms} \pm 4$ (S.E.) nel ciclo vocalico UI (quindi quando il nucleo vocalico è /i/; $\chi^2(1)=12.185$, $p<0.001$), e di circa $12 \text{ ms} \pm 4$ (S.E.) quando la consonante in contesto è sorda ($\chi^2(1)=9.3478$, $p<0.001$). L'interazione tra i fattori *Popolazione e Sonorità* è significativa ($\chi^2(1)=6.9974$, $p<0.01$), con la durata della sillaba tonica che diminuisce nei patologici di circa $20 \text{ ms} \pm 7$ (S.E.) quando la consonante è sorda; anche l'interazione tra *Popolazione e Ciclo vocalico* è significativa, con la durata che aumenta di circa $13 \text{ ms} \pm 7$ (S.E.) nei patologici e in UI ($\chi^2(1)=2.7816$, $p<0.1$).

Coerentemente, la durata dell'enunciato diminuisce di circa $40 \text{ ms} \pm 10$ (S.E.) quando la consonante è sorda ($\chi^2(1)=14.145$, $p<0.001$). Nel caso di ciclo vocalico UI, quindi quando /i/ è tonica e la maggior parte delle vocali nella frase sono /u/, la durata dell'enunciato, contrariamente a quanto osservato per la sillaba, aumenta di circa $33 \text{ ms} \pm 10$ (S.E.) ($\chi^2(1)=9.68$, $p<0.01$). Anche la durata dell'enunciato, così come quella della sillaba, è influenzata dall'interazione tra *Popolazione e Sonorità*, con una diminuzione di circa $52 \text{ ms} \pm 20$ (S.E.) per i patologici e quando le consonanti sono sorde ($\chi^2(1)=6.4024$, $p<0.05$), e dall'interazione tra *Popolazione e Ciclo vocalico* ($\chi^2(1)=4.665$, $p<0.05$), con un aumento di circa $45 \text{ ms} \pm 20$ (S.E.) nei soggetti patologici e nel ciclo vocalico UI.

Anche l'*Articulation rate* (durata media della sillaba) è influenzata dai fattori *Sonorità* e *Ciclo vocalico* ($\chi^2(1)=17.427$, $p<0.001$), con la durata media della sillaba che diminuisce di circa $10 \text{ ms} \pm 2$ (S.E.) quando la consonante è sorda (con un incremento quindi dell'*articulation rate*), mentre aumenta di circa $8 \text{ ms} \pm 2$ (S.E.) nel caso di ciclo vocalico UI ($\chi^2(1)=9.6819$, $p<0.01$) (con una diminuzione di *articulation rate*, quindi, quando sono presenti più /u/ nella frase). L'*articulation rate* è influenzata anche dall'interazione tra *Popolazione* e *Sonorità* ($\chi^2(1)=6.4017$, $p<0.05$), con una diminuzione della durata media delle sillabe di circa $13 \text{ ms} \pm 5$ (S.E.) per i patologici e quando la consonante è sorda; è influenzata dall'interazione tra *Popolazione* e *Ciclo vocalico* ($\chi^2(1)=4.6659$, $p<0.05$), con un aumento della durata media delle sillabe di circa $11 \text{ ms} \pm 5$ (S.E.) per i patologici e nel ciclo vocalico UI, ossia nel caso in cui siano presenti più /u/ all'interno della frase.

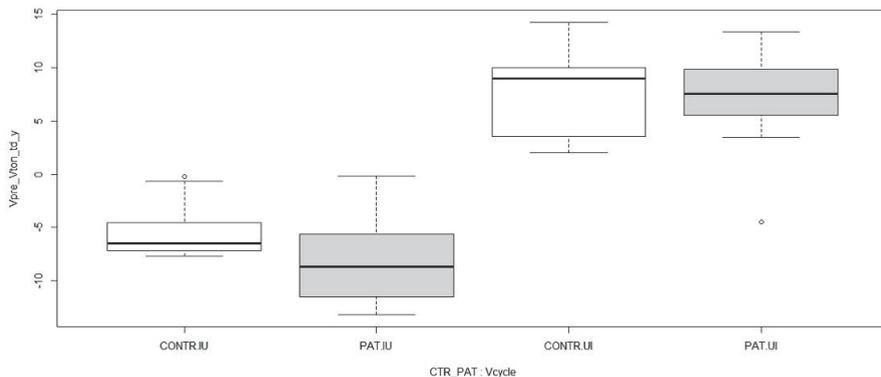
2.2.2 Risultati articolatori

2.2.2.1 Gesto linguale

Per quanto riguarda i gesti della lingua lungo l'asse verticale, sia verso che dalla vocale accentata, si osserva un'ovvia influenza del fattore *Ciclo vocalico*, ma anche la sua interazione con la patologia, quindi con il fattore *Popolazione*. In particolare, i gesti della lingua verso la vocale accentata sono caratterizzati da un incremento dell'ampiezza di circa 15 mm nel ciclo vocalico UI, quindi nel gesto verso /i/, rispetto al ciclo IU, ossia nel gesto verso /u/ ($\chi^2(1)=317.52$, $p<0.001$); i risultati inerenti all'interazione tra *Ciclo vocalico* e *Popolazione* ($\chi^2(1)=9.2983$, $p<0.01$) mostrano che l'incremento è maggiore nei patologici e di circa 3 mm (0.8 S.E.) in UI. Di fatto, complessivamente, il gesto risulta essere più ampio nei soggetti patologici (Figura 2).

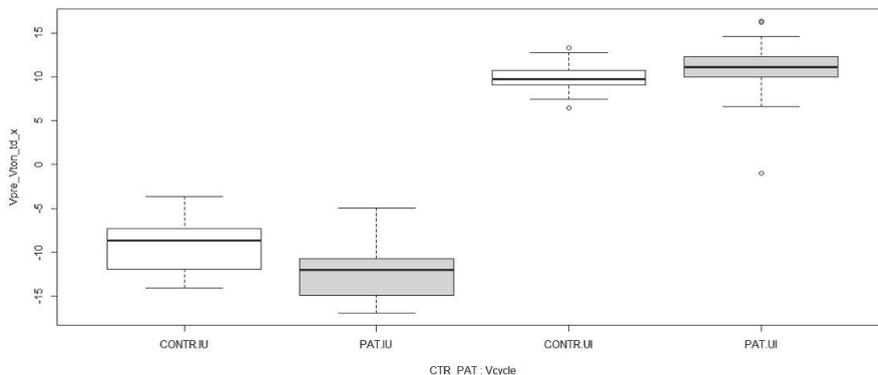
Anche nel caso del movimento della lingua dalla vocale accentata, l'ampiezza aumenta nel ciclo vocalico UI, di circa 15 mm (0.4 S.E.) ($\chi^2(1)=322.55$, $p<0.001$); i risultati inerenti all'interazione tra *Ciclo vocalico* e *Popolazione* ($\chi^2(1)=32.73$, $p<0.001$) mostrano che anche in questo caso l'incremento è maggiore nei patologici, per i quali è di circa 4 mm nel caso di UI.

Figura 2 - Ampiezza (mm) del gesto verticale della lingua verso la vocale target, al variare della popolazione (CONTR vs. PAT) e del ciclo vocalico (IU vs. UI = /u/ vs. /i/ tonica)



Anche sull'asse orizzontale, i gesti della lingua da e verso la vocale accentata risultano influenzati dal ciclo vocalico e dalla sua interazione con il fattore *Popolazione*. In particolare, il movimento della lingua verso la vocale accentata risulta ovviamente più ampio nel ciclo vocalico UI ($\chi^2(1)=467.13$, $p<0.001$), con un aumento di circa 21 mm (0.4 S.E.), e l'interazione tra *Ciclo vocalico* e *Popolazione* ($\chi^2(1)=33.192$, $p<0.001$) mostra che l'aumento è maggiore nei patologici (di circa 4 mm - 0.6 S.E., Figura 3) in UI. Peraltro, l'ampiezza del movimento della lingua verso la vocale accentata è influenzata anche dalla sonorità ($\chi^2(1)=4.4646$, $p<0.05$) e risulta più ampia di circa 1 mm (0.3 S.E.) quando la consonante è sorda. Nel movimento della lingua dalla vocale accentata, l'ampiezza aumenta di circa 25 mm (0.5 S.E.) nel ciclo vocalico UI ($\chi^2(1)=414.68$, $p<0.001$) e di circa 3.7 mm (1 S.E.) quando c'è l'interazione tra ciclo vocalico UI e popolazione $\chi^2(1)=12.519$, $p<0.001$). L'ampiezza, dunque, è maggiore in tutti i casi nel ciclo vocalico UI e nei soggetti patologici.

Figura 3 - Ampiezza (mm) del gesto antero-posteriore della lingua verso la vocale target, al variare della popolazione (CONTR vs. PAT) e del ciclo vocalico (IU vs. UI = /u/ vs. /i/ tonica)



2.2.2.2 Gesto labiale

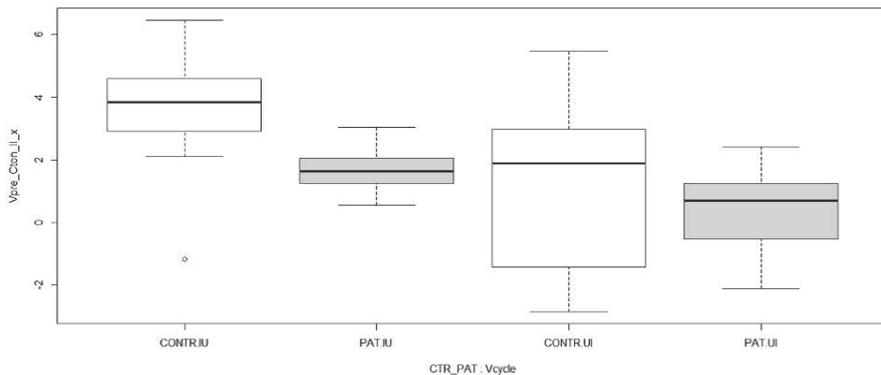
I dati relativi al gesto verticale del labbro inferiore risultano statisticamente significativi, ma le differenze sono talmente ridotte da non poter essere considerate veramente indicative di eventi rilevanti (in alcuni casi, inoltre, ci si avvicina al limite di accuratezza garantito dall'azienda costruttrice dell'AG501). Riteniamo quindi inutile riportare i risultati nel dettaglio e li considereremo solo se otterremo significatività statistica anche in ulteriori, future indagini.

Nel caso del gesto labiale sull'asse orizzontale, antero-posteriore, invece, i risultati ottenuti sono più interessanti. In particolare, il gesto del labbro inferiore verso la consonante target (l'attacco della sillaba accentata), è influenzato dalla *Sonorità* ($\chi^2(1)=6.9879$, $p<0.01$), dal *Ciclo vocalico* ($\chi^2(1)=53.195$, $p<0.001$) e dalla sua interazione con il fattore *Popolazione* ($\chi^2(1)=4.1641$, $p<0.05$), e in tutti i casi il movimento è ridotto nei patologici rispetto a quanto riscontrato nei controlli. Quando la consonante è sorda l'ampiezza è ridotta di circa 0,5 mm (0.2 S.E.), nel ciclo vocalico UI è ridotta di circa 2 mm (0.2 S.E.), e nel caso dell'interazione tra *Ciclo vocalico* e

Popolazione (Figura 4) l'ampiezza è ridotta di circa 1.2 mm (0.4 S.E.) nei soggetti patologici rispetto ai soggetti di controllo.

Al contrario, non si riscontra alcuna influenza del fattore *Popolazione* sul gesto del labbro inferiore dalla consonante verso la vocale target. Quest'ultimo risulta influenzato solo dal ciclo vocalico UI ($\chi^2(1)=27.723$, $p<0.001$) e, in particolare, aumenta di circa 1 mm (0.2 S.E.) in UI rispetto a IU.

Figura 4 - Ampiezza (mm) del gesto antero-posteriore del labbro inferiore verso la consonante target, al variare della popolazione (CONTR vs. PAT) e del ciclo vocalico (IU vs. UI = /u/ vs. /i/ tonica)



2.2.2.3 Durate articolatorie

Per quanto riguarda le misurazioni della durata articolatoria della sillaba non abbiamo riscontrato l'influenza significativa dei fattori da noi considerati, mentre risulta significativa l'interazione tra *Ciclo vocalico* UI e *Popolazione* per le durate "dell'enunciato" (si veda la nota 1) ($\chi^2(1)=3.3148$, $p<0.05$) e per l'*articulation rate* ($\chi^2(1)=5.5581$, $p<0.05$) con una diminuzione rispettivamente di circa 54 ms (29.4 S.E.) e di 36 ms (15.14 S.E.) per i patologici nel ciclo vocalico UI, quindi una diminuzione delle durate e un incremento dell'*articulation rate* nel caso di numerose /u/ nella produzione.

2.2.3 Discussione

I risultati delle misurazioni della prima e della seconda formante mostrano che lo spazio acustico utilizzato dai parlanti disartrici per le vocali è ridotto sull'asse antero-posteriore (e i dati acustici suggeriscono un minor avanzamento della lingua nel passaggio da /u/ a /i/ tonica); inoltre gli stessi parlanti sembrano usare uno spazio acustico più alto, in termini di frequenza, quando le consonanti in contesto sono sorde (suggerendo che in questi casi la posizione della lingua sia leggermente abbassata).

Benché i risultati acustici siano in linea con le aspettative inerenti alla disartria ipocinetica, le analisi articolatorie forniscono una visione molto più articolata della produzione delle vocali nel caso del parlato patologico, che non è compatibile con l'idea di una riduzione generalizzata dei gesti articolatori. Per quanto riguarda i gesti della lingua lungo l'asse verticale, sia verso che dalla vocale accentata, si os-

serva un'ovvia influenza del fattore *Ciclo vocalico*, ma anche la sua interazione con la patologia, quindi con il fattore *Popolazione*, con valori maggiori nel caso dei patologici. Inoltre, anche sull'asse orizzontale i gesti della lingua, da e verso la vocale accentata, risultano influenzati dal *Ciclo vocalico* e dalla sua interazione con il fattore *Popolazione*. In particolare, l'ampiezza del movimento orizzontale è ovviamente maggiore in tutti i casi nel ciclo vocalico UI e, aspetto più interessante, nei soggetti patologici. I risultati articolatori che riguardano i gesti linguali, quindi, non sono coerenti con quelli ricavati sul piano acustico, perché lo spazio acustico ridotto, sia per quanto riguarda F1 che per quanto riguarda F2, corrisponde invece a gesti linguali di maggior ampiezza dal punto di vista articolatorio.

Tuttavia, l'analisi dei gesti labiali sembra chiarire i risultati. Diversi fattori influenzano il gesto labiale in senso antero-posteriore, ma, a parte il ruolo della *Sonorità* e del *Ciclo vocalico*, è l'interazione con il fattore *Popolazione* e la riduzione della protrusione nel gesto verso la consonante che può motivare la presenza di gesti linguali di maggior ampiezza (con una differenza di circa 2 mm tra patologici e controlli). La minor protrusione potrebbe essere quindi compensata dalla maggior estensione dei movimenti della lingua. D'altronde, la protrusione labiale normalmente partecipa nel garantire l'abbassamento dei valori formantici corrispondenti alle vocali posteriori e, quindi, la differenza acustica rispetto alle vocali anteriori. La minor protrusione potrebbe, quindi, richiedere una compensazione operata grazie a un articolatore molto mobile come la lingua, che, con gesti di maggior ampiezza, permetterebbe di garantire la sufficiente differenziazione delle caratteristiche acustiche di vocali anteriori e posteriori. Si deve comunque considerare che l'analisi presentata in questa sede non prende in considerazione la variazione tra soggetti (il fattore è stato solo considerato come *intercept* e *random slope* nel modello misto), che potrebbe poi rivelare la presenza di variazione inter-soggettiva non coerente con la compensazione ipotizzata in questa sede.

I dati sulla durata acustica di sillaba ed enunciato, così come quelli ottenuti sull'*articulation rate*, mettono in evidenza che le durate, al di là del variare in base alla sonorità delle consonanti (che causano una diminuzione di durata e incremento di *articulation rate* se sono sorde) e del ciclo vocalico (con un aumento delle durate e diminuzione del *rate* quando siano presenti (più) vocali posteriori nel contesto misurato), variano anche per effetto della disartria. In particolare, le variazioni in *articulation rate* e durata dell'enunciato sono significativamente più marcate nei patologici e nella direzione sopra indicata, e la durata della sillaba nei patologici addirittura aumenta, invece che diminuire, nel ciclo UI. Dal punto di vista articolatorio, i fattori considerati non risultano significativi se presi singolarmente, ma l'effetto della patologia si manifesta tramite la significatività dell'interazione tra *Ciclo vocalico* UI e *Popolazione* per le durate "dell'enunciato" e, in particolare, emerge una diminuzione delle durate articolatorie e un incremento dell'*articulation rate* nel caso di numerose /u/ nella produzione, risultato stranamente opposto rispetto a quello ottenuto in acustica e per il quale al momento non abbiamo una spiegazione se non facendo riferimento alla leggera discrepanza nell'intervallo misurato (v. nota 1).

L'effetto della disartria risulta quindi evidente sia dai dati acustici che dai dati articolatori, ma i risultati mostrano soprattutto che i parlanti sembrano mettere in atto delle strategie di compensazione che portano a gesti più ampi della lingua per ovviare alla minor mobilità delle labbra, qui rappresentate dal labbro inferiore, il più mobile. L'analisi della variazione inter-soggettiva permetterà di confermare questa interpretazione o di mettere in evidenza altre strategie.

Un'ultima considerazione riguarda lo studio del fattore *Ripetizione* per tener conto dell'affaticamento del parlante durante il compito sperimentale: il fattore non ha influito sui risultati, probabilmente perché le frasi erano molto brevi e, inoltre, i soggetti si potevano riposare sufficientemente tra una frase e l'altra e tra un blocco e il successivo.

3. Conclusioni

Le analisi acustiche e articolatorie (articulografia elettromagnetica – AG501) delle produzioni orali di tre soggetti disartrici affetti da malattia di Parkinson, provenienti dall'area di Lecce, e di altrettanti soggetti di controllo, della stessa età e provenienti dalla stessa area geolinguistica, ci hanno permesso di studiare nel dettaglio le caratteristiche dell'articolazione di vocali alte italiane da parte di soggetti disartrici. Particolare attenzione è stata dedicata alle caratteristiche acustiche, sia all'ampiezza, durata e sincronizzazione dei gesti linguali e labiali.

Oltre alla prevista tendenza alla riduzione dello spazio acustico e dell'ampiezza e durata dei gesti articolatori, attesa per via della forma ipocinetica di disartria che colpisce i soggetti in esame, i risultati mostrano anche un aumento delle misure ricavate sia per i movimenti lingua, non solo lungo l'asse antero-posteriore, come già osservato in passato (Gili Fivela et al., 2014), ma anche lungo l'asse verticale. I gesti più ampi osservati per la lingua accompagnano, però, una minor mobilità dei gesti labiali sull'asse antero-posteriore. Questo suggerisce che il gruppo di parlanti possa mettere in atto strategie di compensazione volte a preservare le caratteristiche acustiche distintive delle vocali, compensando con la maggior mobilità linguale, il ridotto spostamento labiale. Tuttavia, solo un'analisi più approfondita, relativa alla variazione tra soggetti (in questo caso, infatti, il fattore *Soggetto* è stato solo considerato come *intercept* e *random slope* nel modello misto) potrà confermare il tipo di compensazione o rivelare la presenza di altre strategie, magari differenti a seconda del soggetto.

Anche le durate e la velocità di articolazione sembrano risentire della disartria. In particolare, le variazioni in *articulation rate* e durata dell'enunciato sono significativamente più marcate nei patologici che nei controlli (es., nel caso della diminuzione di durata e incremento di *articulation rate* in contesto di consonanti sorde e dell'aumento delle durate e diminuzione del *rate* quando siano presenti (più) vocali posteriori nel contesto misurato); inoltre la durata della sillaba nei patologici addirittura aumenta, invece che diminuire come nei controlli, nel ciclo UI. I risultati articolatori circa le durate offrono un quadro diverso (ossia suggeriscono una dimi-

nuzione delle durate dell' "enunciato" e un incremento dell'*articulation rate* nel caso di numerose /u/ nella produzione", che al momento non riusciamo a spiegare se non facendo riferimento alla leggera discrepanza nell'intervallo misurato (v. nota 1).

Infine, è importante notare che il fattore *Ripetizione*, considerato per rendere conto della stanchezza accumulata durante il compito sperimentale, non ha influito sulle misure considerate. Questo risultato non viene considerato come una prova dell'assenza di affaticamento e di un possibile peggioramento dell'accuratezza del parlato dovuto ad esso, ma del fatto che le ripetizioni o le frasi brevi possano non essere il modo ideale per indagare l'influenza della fatica sull'accuratezza del parlato disartrico. Future indagini potranno permettere di verificare se la lunghezza dell'enunciato e/o la complessità della frase, piuttosto che la lunghezza del compito, possano rappresentare fattori migliori per rendere conto delle modifiche di accuratezza dovute all'affaticamento.

Ringraziamenti

Questo lavoro è stato parzialmente finanziato dal progetto PRIN 2017 2017JNKCYZ. Ringraziamo tutti i soggetti che hanno partecipato all'esperimento, il dott. M. Iraci che ha aiutato nell'acquisizione dei materiali, e il dott. F. Sigona che ha offerto supporto tecnico per le misurazioni.

Bibliografia

- ACKERMANN, H., ZIEGLER, W. (1991). Articulatory deficits in Parkinsonian dysarthria: an acoustic analysis. In *Journal of Neurology, Neurosurgery, and Psychiatry*, 54, 1093–8.
- BATES, D.S, MAECHLER, M., BOLKER, B. & WALKER, S. (2014). *lme4: Linear mixed-effects models using Eigen and S4*. R package version 1.1-5. <http://CRAN.R-project.org/package=lme4>
- BOERSMA, P., WEENINK, D. (2019). Praat: doing phonetics by computer [Computer program]. Version 6.1.04 from <http://www.praat.org/>.
- BROWMAN, C.P. & L. GOLDSTEIN (1992). Articulatory phonology: an overview. In *Phonetica*, 49, 155–180.
- CANTER, G.J. (1963). Speech characteristics of patients with Parkinson's disease: Intensity, pitch and duration. In *Journal of Speech Hearing Disorders*, 28, 217–224.
- DE LOOZE, C., GHIO, A., SCHERER, S., POUCHOULIN, G. & VIALLET, F. (2012). Automatic analysis of the prosodic variations in Parkinsonian read and semi-spontaneous speech. In *Proceedings of Speech Prosody 2012*, 71–74.
- DUEZ, D. (2006). Syllable structure, syllable duration and final lengthening in parkinsonian French speech. In *Journal of Multilingual Communication Disorders*, 4, 1, 45–57.
- DUFFY, J.R. (2005). *Motor Speech Disorders: Substrates, Differential Diagnosis, and Management*. 2° ed., Elsevier Mosby.

- GILI FIVELA B., IRACI M., SALLUSTIO V., GRIMALDI M., ZMARICH C. & PATROCINIO D. (2014). Italian Vowel and Consonant (co)articulation in Parkinson's Disease: extreme or reduced articulatory variability?, In S. Fuchs, M. Grice, A. Hermes, L. Lancia, D. Mücke (eds), *Proc. of the 10th International Seminar on Speech Production (ISSP)*, 5-8 May, Cologne, Germany, 146-149.
- IRACI, M. (2017). Vowels, consonants and co-articulation in Parkinson's Disease. *Unpublished PhD Dissertation*, University of Salento.
- MONRAD-KROHN, G.H. (1947). Dysprosody or altered melody of language, In *Brain*, 70, 405-415.
- R CORE TEAM (2019). R: A language and environment for statistical computing. *R Foundation for Statistical Computing*, Vienna, Austria. URL: <http://www.Rproject.org/>
- SIGONA, F., STELLA A., GRIMALDI M. & GILI FIVELA B. (2015). MAYDAY: A Software for Multimodal Articulatory Data Analysis. In A. Romano, M. Rivoira, and I. Meandri (eds.), *Aspetti prosodici e testuali del raccontare: dalla letteratura orale al parlato dei media*, *Atti del X Convegno Nazionale AISV*, Edizioni dell'Orso, 173-184.
- SKODDA, S., W. GRONHEIT & U. SCHLEGEL (2012). Impairment of Vowel Articulation as a Possible Marker of Disease Progression in Parkinson's Disease. In *PLoS ONE*, 7 (2), 1-8.
- SKODDA S., W. VISSER & U. SCHLEGEL (2011). Vowel Articulation in Parkinson's Disease. In *Journal of Voice*, 25, 4, 467-72.
- TESTON, B., VIALLET, F. (2001). L'évaluation objective de la prosodie in Les dysarthries. In Auzou et al. (ed.), *Problèmes en médecine de rééducation*, Paris: Masson, 109-121.
- WISNIECKI, A., CANNIZZARO, M., COHEN, H. & SNYDER, P.J. (2006). Speech Impairment in Neurodegenerative Diseases/Psychiatric Illnesses. In H.A. Whitaker (ed.) *Concise Encyclopedia of Brain and Language*, Amsterdam, NL: Elsevier Science and Academic Press, 477-482.
- WONG, M.N., MURDOCH, B.E. & WHELAN, B.M. (2010). Kinematic analysis of lingual function in dysarthric speakers with Parkinson's disease: An electromagnetic articulograph study. In *Int. JS-L Pathology*, 12, 414-25.
- WONG, M.N., MURDOCH, B.E. & WHELAN, B.M. (2011). Lingual Kinematics in Dysarthric and Nondysarthric Speakers with Parkinson's Disease. In *Parkinson's Disease*, 1-8.

Autori

FABIO APREA – U. Wien (Austria) & U.G. Marconi, Roma
fabio.aprea@univie.ac.at

FABIO ARDOLINO – Dipartimento di Filologia, Letteratura e Linguistica,
Università di Pisa; Laboratoire Parole et Langage, Aix-Marseille Université
fabio.ardolino@fileli.unipi.it

ANCUTA BUDEANU – Dipartimento di Studi Letterari, Università di Napoli
L'Orientale
anca_0506@yahoo.com

FRANCESCO CANGEMI – University of Cologne, Cologne, Germany
fcangemi@uni-koeln.de

TERESA CARBUTTI – Centro Internazionale di Dialettologia, Università degli
Studi della Basilicata
nisia100@hotmail.it

MARINA CASTAGNETO – Dipartimento di Studi Umanistici, Università del
Piemonte Orientale
marina.castagneto@uniupo.it

VIOLETTA CATALDO – Department of Humanities Studies, University of Salerno
& Department of Linguistics, Ghent University, Belgium
vcataldo@unisa.it

SONIA CENCESCHI – Digital Forensic Service, Department of Innovative
Technologies, University of Applied Sciences and Arts of Southern Switzerland,
Via Cantonale 2C, 6928 Manno, Switzerland
sonia.cenceschi@supsi.ch

SONIA I. D'APOLITO – Università del Salento, CRIL – DREAM, Lecce – Italia
sonia.dapolito@gmail.com

MAYARA DA SILVA NETO – Department of Modern Languages, University of São Paulo/Brazil

mayara.neto@usp.br

ANNA DE MEO – Dipartimento di Studi Letterari, Università di Napoli L'Orientale
ademeo@unior.it

MARIAPAOLA D'IMPERIO – Rutgers University, Department of Linguistics and
The Center for Cognitive Science (RuCCS), USA

mdimperio2@gmail.com

DALILA DIPINO – Romanisches Seminar, Università di Zurigo, Svizzera
dalila.dipino@uzh.ch

GIORGIA DI PRIZIO – Università del Salento, CRIL – DREAM, Lecce – Italia
giorgiadiprizio@libero.it

MARGHERITA DI SALVO – Dipartimento di Studi Umanistici, Università Federico
II, Napoli

margherita.disalvo@unina.it

FLAVIANE ROMANI FERNANDES SVARTMAN – Department of Classical and
Vernacular Languages and Literatures, University of São Paulo/Brazil

flavianesvartman@usp.br

MANUELA FRONTERA – Università degli Studi Internazionali di Roma

manuela.frontera@unint.eu

BARBARA GILI FIVELA – Department of Humanities, University of Salento

barbara.gilifivela@unisalento.it

PIA GRECA – Institute of Phonetics and Speech Processing, Ludwig-Maximilians
Universität München, Germany

greca@phonetik.uni-munchen.de

MARTINE GRICE – University of Cologne, Cologne, Germany

martine.grice@uni-koeln.de

GLENDA GURRADO – Dip.to Lelia, Università degli Studi di Bari

glenda.gurrado@uniba.it

JONATHAN HARRINGTON – Institute of Phonetics and Speech Processing,
Ludwig-Maximilians Universität München, Germany
jmh@phonetik.uni-muenchen.de

CHIARA MELUZZI – Department of Humanities, Section of Linguistics, University
of Pavia, corso strada Nuova 65, 27100, Pavia, Italy
chiara.meluzzi@unipv.it

DANIELA MEREU – Libera Università di Bolzano
daniela.mereu@unibz.it

NICHOLAS NESE – Department of Humanities, Section of Linguistics, University
of Pavia, corso strada Nuova 65, 27100, Pavia, Italy
nicholas.nese01@universitadipavia.it

RICCARDO ORRICO – Department of Humanities Studies, University of Salerno
& Aix Marseille University, France
rorrico@unisa.it.

MASSIMO PETTORINO – Dipartimento di Studi Letterari, Università di Napoli
L'Orientale
mpettorino@gmail.com

FRANCESCA PEZZELLA – Dipartimento di Filologia, Letteratura e Linguistica –
Università di Pisa
francesca.pezzella@phd.unipi.it

KRISTIJAN RAJIC – Brandstrasse 21, CH-8952 Schlieren, Svizzera
kristijan.rajic@uzh.ch

LUCIANO ROMITO – Dipartimento di Culture, Educazione e Società, Università
della Calabria
luciano.romito@unical.it

MICHELA RUSSO – U. Lyon & UMR 7023 SFL/U. Paris 8 (Francia)
mrusso@univ-paris8.fr

ELISABETTA SANTORO - Department of Modern Languages, University of São
Paulo/Brazil
esantoro@usp.br

SIMONA SBRANNA – University of Cologne, Cologne, Germany
s.sbranna@outlook.com

STEPHAN SCHMID – Phonetisches Laboratorium, Universität Zürich, Rämistrasse 71, CH-8006 Zürich, Svizzera
stephan.schmid@uzh.ch

PATRIZIA SORIANELLO – Dip.to LeLiA, Università degli Studi di Bari 'Aldo Moro'
patrizia.sorianello@uniba.it

IRINA STAN – Dipartimento di Scienze della Mediazione Linguistica e di Studi Interculturali, Università degli Studi di Milano
irina.stan@unimi.it

ANDREA TARASI – Dipartimento di Culture, Educazione e Società, Università della Calabria
tarasiandrea17@gmail.com

ALESSANDRO VIETTI – Libera Università di Bolzano
alessandro.vietti@unibz.it

Studi AISV è una collana di volumi collettanei e monografie dedicati alla dimensione sonora del linguaggio e alle diverse interfacce con le altre componenti della grammatica e col discorso. La collana, programmaticamente interdisciplinare, è aperta a molteplici punti di vista e argomenti sul linguaggio: dall'attenzione per la struttura sonora alla variazione sociofonetica e al mutamento storico, dai disturbi della parola alle basi cognitive e neurobiologiche delle rappresentazione fonologiche alle applicazioni tecnologiche. I testi sono selezionati attraverso un processo di revisione anonima fra pari e vengono pubblicati nel sito dell'Associazione Italiana di Scienze della Voce con accesso libero a tutti gli interessati.

Luciano Romito è Professore Ordinario di Glottologia e Linguistica presso l'Università della Calabria. Attualmente è membro del Senato Accademico, direttore e responsabile scientifico del Laboratorio di Fonetica dell'Unical e dell'Universidad National de Rosario (UNR-Argentina), coordinatore dell'Osservatorio sulla Linguistica Forense (OLF), direttore del Dipartimento di Scienze Forensi dell'Osservatorio Nazionale per le Investigazioni, la Sicurezza e le Scienze Forensi (ONISSF), Responsabile scientifico dei moduli di *Diritti, Democrazia e Sicurezza e Intelligence, Nuove Tecnologie* nel Master di II livello in "Intelligence", Consulente per il Ministero di Grazia e Giustizia e per il Ministero degli Interni per l'identificazione del parlatore in casi di sequestri, terrorismo e criminalità organizzata e Consulente Linguista per la Commissione Europea Eulex per le stragi di guerra in Kosovo.

AISV - Associazione Italiana Scienze della Voce

sito: www.aisv.it

email: aisv@aisv.it | redazione@aisv.it

ISBN: 978-88-9765-749-1

Edizione realizzata da

Officinaventuno

info@officinaventuno.com | sito: www.officinaventuno.com

via F.lli Bazzaro, 18 - 20128 Milano - Italy