



AISV
Associazione Italiana di Scienze della Voce



AISV 2009

5° Convegno Nazionale

AISV - Associazione Italiana di Scienze della Voce

“LA DIMENSIONE TEMPORALE DEL PARLATO”

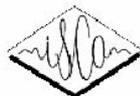
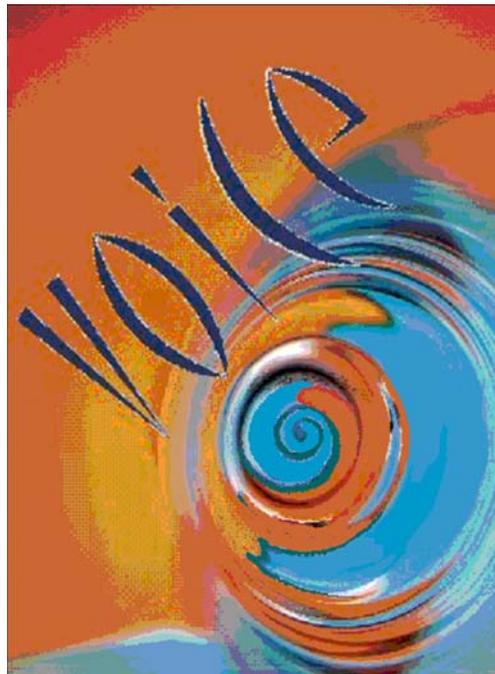
a cura di
Stephan Schmid
Michael Schwarzenbach
Dieter Studer



CD-rom incluso



Università di Zurigo
Kollegiengebäude
4-6 Febbraio 2009



international speech
communication association

Universität Zürich
Phonetisches Laboratorium

“LA DIMENSIONE TEMPORALE DEL PARLATO”

Atti del 5° Convegno Nazionale AISV 2009

a cura di
Stephan Schmid
Michael Schwarzenbach
Dieter Studer



Università di Zurigo
Kollegiengebäude
4-6 Febbraio 2009



international speech
communication association



Universität Zürich
Phonetisches Laboratorium

“LA DIMENSIONE TEMPORALE DEL PARLATO”

Atti del 5° Convegno Nazionale AISV 2009
Università di Zurigo, *Kollegengebäude*
4-6 Febbraio 2009

a cura di
Stephan Schmid
Michael Schwarzenbach
Dieter Studer

ISBN: 978-88-6368-087-4

Copyright © 2010 by EDK Editore srl
Via Santarcangiolese, 6
47825 Torriana (RN)

INDICE

INDICE	i
ORGANIZZAZIONE CONVEGNO	v
ORGANIZZAZIONE GENERALE	v
COMITATO SCIENTIFICO	v
SEGRETERIA DEL CONVEGNO	v
WEBMASTER	v
ORGANIZZAZIONE LOCALE	vi
STAFF	vi
SPONSOR	vi
“LA DIMENSIONE TEMPORALE DEL PARLATO”	vii
PREMESSA	vii
PREMIO FRANCO FERRERO	viii
RINGRAZIAMENTI	ix
ROUND TABLE: “DIFFERENT WAYS OF ANALYZING SPEECH RHYTHM”	1
RHYTHM MEASURES IN RETROSPECT. REFLECTIONS ON THE NATURE OF SPOKEN-LANGUAGE RHYTHM	3
<i>William J. Barry</i>	
CHOOSING THE RIGHT RATE NORMALIZATION METHOD FOR MEASUREMENTS OF SPEECH RHYTHM	5
<i>Volker Dellwo</i>	
SPEECH RHYTHM AND WORD SEGMENTATION: A PROMINENCE-BASED ACCOUNT OF SOME CROSSLINGUISTIC DIFFERENCES	7
<i>Christopher S. Lee</i>	
SPEECH RHYTHM AND TIMING: STRUCTURAL PROPERTIES AND ACOUSTIC CORRELATES	9
<i>Antonio Romano</i>	

LINGUISTICA, FONETICA E FONOLOGIA	11
UN CONFRONTO TRA DIVERSE METRICHE RITMICHE USANDO CORRELATORE	13
<i>Paolo Mairano, Antonio Romano</i>	
VARIABILITÀ RITMICA DI VARIETÀ DIALETTALI DEL PIEMONTE	15
<i>Antonio Romano, Paolo Mairano, Barbara Pollifrone</i>	
TEMPI E MODI DI CONSERVAZIONE DELLE R ITALIANE NEI <i>FRIGORIFERI</i> CLIPS	17
<i>Alessandro Vietti, Lorenzo Spreafico, Antonio Romano</i>	
NOTE SULLE OPPOSIZIONI DI QUANTITÀ VOCALICA	19
<i>Arianna Uguzzoni</i>	
FENOMENI D'ARMONIA VOCALICA IN AREA FRIULANA E IBERICA E LE SORTI DI -A FINALE LATINA	21
<i>Renzo Miotti</i>	
COARTICOLAZIONE E MUTAMENTO. UNA RICERCA SUL VOCALISMO ATONO FINALE NELL'ENTROTERRA MACERATESE	23
<i>Tania Paciaroni</i>	
DURATA E STRUTTURE FORMANTICHE NEL PARLATO TOSCANO: INDAGINI PRELIMINARI SU UN CAMPIONE DI DIALOGHI SEMISPONTANEI	25
<i>Nadia Nocchi, Silvia Calamai</i>	
DIAGNOSTICA FONOLOGICA E DIAGNOSI FONETICA. OSSITONI LUNGI IN SILLABA LIBERA NEI DIALETTI DI SAMBUCA PISTOIESE (PT)	27
<i>Lorenzo Filippino, Nadia Nocchi</i>	
ELISIONE OBBLIGATORIA, VARIABILE E POCO FREQUENTE NEL FIORENTINO: UN CASO DI ALLOMORFIA FRASALE PRECOMPILATA CON FORME PREFERENZIALI	29
<i>Luigia Garrapa</i>	
CONTINUUM DIAFASICO E DINAMICHE DIAGENERAZIONALI NEL BASSO E ALTO CASERTANO ORIENTALE	31
<i>Edoardo Mastantuoni</i>	
CONFINI PROSODICI E VARIAZIONE SEGMENTALE. ANALISI ACUSTICA DELL'ALTERNANZA MONOTTONGO/DITTONGO IN ALCUNI DIALETTI DELL'ITALIA MERIDIONALE	33
<i>Giovanni Abete, Adrian P. Simpson</i>	
PHONETIC DETAIL IN INTONATION CONTOUR DYNAMICS	35
<i>Francesco Cangemi</i>	
INTERROGATIVE E ASSERTIVE IN UN CORPUS DIALETTALE RECUPERATO (BOMARZO)	37
<i>Amedeo De Dominicis</i>	
BALBUZIE E COARTICOLAZIONE	39
<i>Caterina Pisciotto, Massimiliano Marchiori, Claudio Zmarich</i>	

CANTO	41
OSSERVAZIONI PRELIMINARI SUGLI ASSETTI INTERVALLARI NEL CANTO A <i>MUTETUS</i> DELLA SARDEGNA MERIDIONALE	43
<i>Paolo Bravi</i>	
PERCEZIONE E APPRENDIMENTO	45
FUNCTIONS OF THE LEFT AND RIGHT POSTERIOR TEMPORAL LOBE DURING SEGMENTAL AND SUPRASEGMENTAL SPEECH PERCEPTION	47
<i>Cyrill Ott, Martin Meyer</i>	
PHONETIC CONTRASTS IN FOREIGN LANGUAGE PERCEPTION: A NEUROPSYCHOLOGICAL STUDY ON SERBIAN AFFRICATES	49
<i>Nuria Kaufmann, Martin Meyer, Stephan Schmid</i>	
DOES A TALKER'S OWN RATE OF SPEECH AFFECT HIS/HER PERCEPTION OF OTHERS' SPEECH RATE?	51
<i>Sandra Schwab</i>	
CROSS-LANGUAGE SPEECH PERCEPTION: LEXICAL STRESS IN SPANISH WITH ITALIAN AND FRANCOPHONE SUBJECTS	53
<i>Ioland Alfano, Sandra Schwab, Renata Savy, Joaquim Llisterri</i>	
PERSISTENZA DELL'ACCENTO STRANIERO. UNO STUDIO PERCETTIVO SULL'ITALIANO L2	55
<i>Giovana Marotta, Philippe Boula de Mareuil</i>	
PERCEZIONE E PRODUZIONE DEI FONEMI DELL'INGLESE AMERICANO IN PARLANTI CON UN SISTEMA PENTAVOCALICO	57
<i>Bianca Sisinni, Mirko Grimaldi</i>	
LA DIMENSIONE TEMPORALE IN TRE TIPI DI PARLATO: UN CONFRONTO TRA ARABO E ITALIANO	59
<i>Dalia Gamal</i>	
TECNOLOGIE DEL PARLATO	61
ALCUNE CONSIDERAZIONI SULL'IMPORTANZA DEGLI ASPETTI DINAMICI NELLA PERCEZIONE, PRODUZIONE ED ELABORAZIONE DEL PARLATO	63
<i>Piero Cosi</i>	
RECENTI SVILUPPI DI 'SONIC' PER L'ITALIANO: RICONOSCIMENTO AUTOMATICO DEL PARLATO INFANTILE	65
<i>Piero Cosi</i>	
TEST FONETICO DELLA PRIMA INFANZIA PER BAMBINI DAI 18 AI 36 MESI: ANALISI CON 'PHON' DEI PRIMI DATI RACCOLTI	67
<i>Claudio Zmarich, Maria Pia Bardozzetti, Caterina Pisciotto, Serena Bonifacio</i>	
ENFASI E CONFINI PROSODICI IN DUE STILI DI ELOQUIO EMOZIONALE	69
<i>Pier Luigi Salza, Enrico Zovato, Morena Danieli</i>	

UN CORPUS SPERIMENTALE PER LO STUDIO CROSS-LINGUISTICO EUROPEO DELLE EMOZIONI VOCALI	71
<i>Vincenzo Galatà, Luciano Romito</i>	
STABILITÀ DEI PARAMETRI NELLO <i>SPEAKER RECOGNITION</i> . LA VARIABILITÀ INTRA E INTER PARLATORE: F0, DURATA E <i>ARTICULATION</i> <i>RATE</i>	73
<i>Luciano Romito, Rosita Lio, Pier Francesco Ferri, Sabrina Giordano</i>	
LOUDNESS E ‘LIVELLO DEL DIALOGO’ NELLE TRASMISSIONI RADIOTELEVISIVE	75
<i>Mauro Falcone, Antonino Barone, Alessandro Bonomi, Alessandro Balestri, Anna Grazia Santoro, Maria Dell’Osso</i>	
SONORITY BASED SYLLABLE SEGMENTATION	77
<i>Bogdan Ludusan, Serena Soldo</i>	
STATICO VS. DINAMICO. UN POSSIBILE RUOLO DELLA SILLABA NEL RICONOSCIMENTO AUTOMATICO DEL PARLATO	79
<i>Serena Soldo, Bogdan Ludusan</i>	

ORGANIZZAZIONE CONVEGNO

ORGANIZZAZIONE GENERALE

Stephan Schmid

COMITATO SCIENTIFICO

Cinzia Avesani (ISTC – CNR, Padova)
Pier Marco Bertinetto (Scuola Normale Superiore, Pisa)
Silvia Calamai (Università degli Studi di Siena)
Piero Così (ISTC – CNR, Padova)
Francesco Cutugno (Università Federico II, Napoli)
Amedeo De Dominicis (Università della Tuscia, Viterbo)
Mauro Falcone (Fondazione Ugo Bordonì, Roma)
Barbara Gili-Fivela (Università degli Studi di Lecce)
Michele Loporcaro (Università di Zurigo)
Giovanna Marotta (Università degli Studi di Pisa)
Pietro Maturi (Università Federico II, Napoli)
Maurizio Omologo (FBK – IRST, Trento)
Andrea Paoloni (Fondazione Ugo Bordonì, Roma)
Antonio Romano (Università di Zurigo)
Luciano Romito (Università della Calabria, Arcavata di Rende)
Pier Luigi Salza (Loquendo S.p.A., Torino)
Renata Savy (Università degli Studi di Salerno)
Carlo Schirru (Università di Sassari)
Stephan Schmid (Università di Zurigo)
Mario Vayra (Università di Bologna)
Claudio Zmarich (ISTC – CNR, Padova)

SEGRETERIA DEL CONVEGNO

AISV 2009
Phonetisches Laboratorium der Universität Zürich
Rämistrasse 71
CH-8006 Zurigo
Tel. 0041 44 634 3001 - Fax 0041 44 634 6948
E-mail: aisv2009@pholab.uzh.ch
URL: <http://www.pholab.uzh.ch/aisv2009.html>

WEBMASTER

Stephan Schmid
Dieter Studer

ORGANIZZAZIONE LOCALE

Vincenzo Faraoni
Lorenzo Filipponio
Michele Loporcaro
Nadia Nocchi
Susanne Oberholzer
Tania Paciaroni
Dieter Studer

STAFF

Marina Albertini
Camilla Bernardasci
Francesca Beyeler
Francesco Cangemi
Sarah Heim
Fritz Herrmann
Endri Llanaj
Luana Massaro
Lucia Picuccio
Michael Schwarzenbach

SPONSOR

Hochschulstiftung der Universität Zürich
Zürcher Universitätsverein (ZUNIV)
Phonetisches Laboratorium der Universität Zürich
Phonogrammarchiv der Universität Zürich
Harman/Becker Automotive Systems GmbH
Förderverein "Amici del Liceo Artistico"

“LA DIMENSIONE TEMPORALE DEL PARLATO”

PREMESSA

Il 5° Convegno Nazionale dell’Associazione Italiana di Scienze della Voce si è svolto, per la prima volta nella breve storia dell’AISV, all’estero, e più precisamente all’Università di Zurigo. Dal 4 al 6 febbraio si sono riuniti ricercatori italiani, svizzeri, inglesi, tedeschi e francesi per discutere del tema della ‘dimensione temporale del parlato’ nei suoi più svariati aspetti.

Com’è noto, la dimensione temporale è un elemento costitutivo della comunicazione orale, che interviene sia nella produzione che nella percezione della parola. A livello segmentale la dimensione temporale determina non solo fenomeni come durata e quantità (particolarmente interessanti in ambito italo-romanzo), ma più in generale la pianificazione e il controllo dei gesti articolatori. A livello suprasegmentale la dimensione temporale caratterizza tra l’altro l’allineamento dei contorni intonativi con le parti dell’enunciato. Inoltre, la velocità di eloquio costituisce un importante parametro in ambito forense, ma essa può essere analizzata anche da un punto di vista ‘diacronico’ (attraverso l’evoluzione del parlato dei mass media). Infine, i vari aspetti del *timing* sono di notevole rilevanza per la tecnologia del linguaggio, in particolare nel campo della sintesi della voce. Questi sono stati soltanto alcuni degli aspetti trattati nei numerosi contributi presentati durante le tre fittissime giornate di lavoro, sia sotto forma di relazione orale sia come poster.

Per volontà degli organizzatori del convegno, particolare attenzione è stata rivolta alla fenomenologia del ‘ritmo’. La tavola rotonda del primo giorno su *Different ways of analyzing speech rhythm*, introdotta e moderata magistralmente dal prof. William Barry, ha permesso un confronto e uno scambio di opinioni tra alcuni specialisti in questo ambito di ricerca. Ad esempio, Volker Dellwo ha messo in evidenza come qualsiasi studio del ritmo debba tener conto anche della velocità di eloquio. Chris Lee ha invece proposto un approccio in termini di ‘prominenza’ che prende in considerazione non solo durata, ma anche F0 e intensità. Infine, Antonio Romano ha presentato alcuni studi recenti sul ritmo che sono stati condotti da ricercatori italiani. I quattro contributi alla tavola rotonda aprono il presente volume.

La tematica del ritmo è stata ripresa all’ultimo giorno nella relazione plenaria di Eric Keller dal titolo *From sound to rhythm expectancy*. Lo studioso svizzero ha da un lato sviluppato alcuni argomenti discussi durante il *workshop* da lui organizzato per il congresso ICPhS di Saarbrücken nel 2007; dall’altro lato le sue riflessioni hanno testimoniato la ricca esperienza di una vita dedicata alla ricerca sulla prosodia. Della conferenza di Eric Keller si trova la presentazione .ppt nel CD-ROM.

La relazione plenaria di Martin Meyer, in apertura dei lavori scientifici del primo giorno, verteva invece sui meccanismi neuronali coinvolti nella percezione del parlato, focalizzando soprattutto l’aspetto temporale del segnale acustico. Questa impressionante revisione dello stato dell’arte nelle neuroscienze è presente negli Atti come articolo co-

firmato da Cyrill Ott e si contraddistingue anche per la ricchissima bibliografia.

Come di consueto, anche in questo convegno non si è parlato solo di un tema specifico, ma ai partecipanti è stata offerta anche una variegata rassegna delle ricerche attualmente in corso nei settori disciplinari coltivati dai membri dell’AISV – dalla linguistica, fonetica e fonologia all’ambito forense alle varie tecnologie del parlato, ecc. All’interno di questa vocazione interdisciplinare (che è parte integrante dell’identità della nostra associazione), dall’incontro zurighese sono comunque emerse due aree di ricerca che attualmente godono di un notevole interesse. La prima di queste aree di ricerca rientra in una delle tematiche da sempre presenti nei convegni AISV e si colloca all’interfaccia tra dialettologia e fonetica sperimentale, a riprova del fatto che il patrimonio linguistico dell’Italia costituisce tuttora un oggetto di ricerca di grandissima rilevanza scientifica. La seconda area di ricerca, di tradizione più recente, sembra invece nascere all’insegna della globalizzazione e ha come oggetto la percezione e l’acquisizione delle strutture sonore in una seconda lingua; dallo sviluppo che questi studi hanno sperimentato negli ultimi anni anche in Italia si evince che la loro ragione non si esaurisce nelle eventuali applicazioni pratiche, ma che questa linea di ricerca ha dei risvolti teorici considerevoli per la comprensione della facoltà umana del linguaggio.

Un convegno non è fatto solo di scienza. Le tre giornate zurighesi hanno offerto un’occasione non solo per incontrare amici di vecchia data, ma anche di fare nuove conoscenze durante gli eventi sociali – quali il rinfresco al *Romanisches Seminar*, la cena sociale oppure durante una delle pause caffè nel *Lichthof*. Un momento particolarmente emozionante è stato sicuramente il bellissimo concerto del ‘Trio Fontane’ nell’Aula Magna.

A differenza del convegno precedente, AISV 2009 non conteneva un’apposita sezione riservata ai dottorandi. Ciononostante è stata numerosa la partecipazione sia attiva che passiva di giovani provenienti da varie Università – sicuramente un buon auspicio per il futuro delle scienze della voce.

PREMIO FRANCO FERRERO

Come nelle due edizioni precedenti, anche quest’anno è stato assegnato il ‘Premio Franco Ferrero’ all’autore (studente o dottorando) del miglior articolo pubblicato negli Atti del 3° e del 4° Convegno AISV. Il ‘Premio Ferrero 2009’ è stato consegnato durante l’apertura del convegno dal prof. Andreas Fischer, Magnifico Rettore dell’Università di Zurigo. Sono stati premiati

– nella categoria ‘Linguistica, Fonetica, Fonologia’:

PAOLO MAIRANO

per l’articolo “Lingue isosillabiche e isoaccentuali: misurazioni strumentali su campioni di italiano, francese, inglese e tedesco”, pubblicato da Paolo Mairano & Antonio Romano negli Atti del 3° Convegno AISV (2006) tenutosi a Povo (Trento);

– nella categoria ‘Tecnologie del Parlato’:

GIACOMO SOMMAVILLA

per l’articolo “SMS-Festival: un nuovo ambiente di lavoro per la sintesi vocale da testo scritto”, pubblicato da Giacomo Sommovilla, Carlo Drioli, Piero Cosi & Graziano Tisato negli Atti del 3° Convegno AISV (2006) tenutosi a Povo (Trento).

RINGRAZIAMENTI

Un ringraziamento particolare va ai membri del comitato scientifico per la valutazione degli Abstract e la revisione dei lavori che in questo volume sono contenuti. Grazie anche a Piero Così e a Luciano Romito, interlocutori dell' AISV sempre disponibili per discutere qualsiasi dettaglio organizzativo.

Mi preme ringraziare tutte le persone dell'Università di Zurigo che in un modo o l'altro hanno contribuito alla riuscita del convegno, a cominciare dal Magnifico Rettore, Prof. Andreas Fischer, per le gentili parole di benvenuto.

Ringrazio inoltre

- il dott. Maximilian Jaeger, delegato del Rettore, per la sua disponibilità riguardo all'organizzazione logistica,
- il personale del *Veranstaltungsdienst Zentrum* per l'assistenza tecnica,
- i musicisti del 'Trio Fontane',
- la dott.ssa Katharina Maier-Troxler e Vera Ziswiler per l'aiuto nell'organizzazione del rinfresco al *Romanisches Seminar*,
- tutti gli studenti e assistenti di linguistica italiana per le ore di servizio passate nella segreteria del convegno,
- gli Sponsor che con il loro contributo finanziario hanno agevolato la realizzazione delle tre giornate zurighesi.

Infine, ringrazio di cuore Michele Loporcaro per aver sostenuto dall'inizio alla fine l'avventura del convegno zurighese – nonché Didi e Michi, cari compagni di lavoro.

Stephan Schmid

TAVOLA ROTONDA:

**“DIFFERENT WAYS
OF ANALYZING
SPEECH RHYTHM”**

RHYTHM MEASURES IN RETROSPECT. REFLECTIONS ON THE NATURE OF SPOKEN-LANGUAGE RHYTHM

William J. Barry
Institut für Phonetik, Universität des Saarlandes
wbarry@CoLi.Uni-SB.DE

ABSTRACT

Against the background query of “what spoken-language rhythm actually is”, an account is offered of the starting-point and the developments which have led to the present state of research into quantitative rhythm analysis. The emergence of a language-rhythm-typology concept from the observation of speech performance in a language-teaching context is shown to run parallel to a search for some objective performance measure of the regularity of the phonological units – morae, syllables and feet – which the typology suggests. A picture is drawn of the paradigm shift from isochrony-based to variability-based research, and the two most influential rhythm metrics of the past decade – Ramus’ global variability and vowel-percentage measures and the Grabe-Low Pairwise Variability Indices (PVI) – are explained (cf. Ramus *et al.*, 1999; Grabe & Low, 2002). What these measures achieve and fail to achieve are discussed, and some of the modified metrics of the same ilk, which have arisen as a reaction to their weaknesses are described. Returning to the background query as to the essence of spoken-language rhythm, the potential and the theoretical limitations of rhythm metrics to reflect perceived rhythmic differences are discussed. Finally, the needs for future rhythm research are considered.

REFERENCES

- Grabe, E. & Low, E.L. (2002), Durational Variability in Speech and the Rhythm Class Hypothesis, in *Papers in Laboratory Phonology 7* (C. Gussenhoven & N. Warner, editors), Berlin: Mouton de Gruyter, 515-546.
- Ramus, F., Nespor, M. & Mehler, J. (1999), Correlates of linguistic rhythm in the speech signal, *Cognition*, 73, 262-292.

CHOOSING THE RIGHT RATE NORMALIZATION METHOD FOR MEASUREMENTS OF SPEECH RHYTHM

Volker Dellwo

Division of Psychology and Language Studies, University College London

v.dellwo@ucl.ac.uk

ABSTRACT

Some acoustic correlates of language rhythm are durational characteristics of consonants and vowels. The present study investigates the influence of speech rate on these acoustic correlates. In experiment I four widely applied correlates of speech rhythm (%V, ΔC , nPVI and rPVI) were correlated with the rate of consonantal and vocalic intervals using speech from five different languages (Czech, English, French, German, Italian) that was characterized by high tempo variability within each language (very slow to very fast produced speech). It was found that rhythm measures based on consonantal interval durations (ΔC , rPVI) correlate negatively with rate measures and that rhythm measures based on vocalic intervals (%V, nPVI) are not influenced by rate. In experiment II the effectiveness of rate normalization procedures on the rate dependent measures, ΔC and rPVI, was tested by correlating these measures with speech rate before and after normalization using the same speech data as in Experiment 1. ΔC was normalized by logarithmically transforming the consonantal interval durations and rPVI was normalized by previously proposed ways for the normalization of nPVI. It was found that rate effects on ΔC and rPVI could be normalized for effectively using the suggested rate normalization procedures. In Experiment III it was tested whether rate normalized measures of speech rhythm support the impression that some languages can be categorized according to their auditory rhythmic characteristics (e.g. stress- and syllable-timing). Strong support for this was only found for the rate normalized rPVI why the normalized ΔC revealed mixed results. It was concluded that ΔC is less appropriate for rhythmic measurements that aim to separate languages of different rhythmic classes.

SPEECH RHYTHM AND WORD SEGMENTATION: A PROMINENCE-BASED ACCOUNT OF SOME CROSSLINGUISTIC DIFFERENCES

Christopher S. Lee
Goldsmiths, University of London
chrisslee@ntlworld.com

SUMMARY

Most work on crosslinguistic rhythmic differences has focused exclusively on temporal factors: the temporal distribution of phonological units at various levels in the classical timing-based accounts, or more recently the durational variability of phonetic units. Lee & Todd (2004) present a prominence-based account of speech rhythm, according to which a crucial determinant of rhythmic organisation is the variability in the auditory prominence of phonetic events (in particular vowels), as primarily determined by their duration, intensity and F0. According to this view, the key difference between so-called ‘stress-timed’ languages (e.g. English and Dutch) and other types is the greater variability in the prominence of their syllabic nuclei. They describe an auditory model developed by Todd and his associates and propose several prominence measures, all of which yield results consistent with their claim on the two multi-language corpora investigated.

Can such an account offer a possible explanation of why infant learners of languages such as English, as opposed to those of languages such as French, adopt a word-segmentation strategy based on the principle that a stressed syllable marks the likely boundary of a content word (Nazzi *et al.*, 2006)? The following claim is advanced here: perceptible prominence distinctions between neighbouring syllables (marking probable transitions between stressed and unstressed syllables) are sufficiently frequent in languages such as English (and Dutch) to delimit large numbers of learnable proto-words (i.e. no more than 2-3 syllables in length), whereas in languages such as French (and Italian), they are too sparsely distributed to serve a similar useful function. The results of analyses of the two multi-language corpora investigated in Lee & Todd (2004) are presented in support of the claim: they yield large and robust differences between the putative rhythm classes.

REFERENCES

- Lee, C.S. & Todd, N.P.M. (2004), Towards an auditory account of speech rhythm: application of a model of the auditory ‘primal sketch’ to two multi-language corpora, *Cognition*, 93, 225-254.
- Nazzi, T., Iakimova, G., Bertoncini, J., Fredonie, S., & Alcantara, C. (2006), Early segmentation of fluent speech by infants acquiring French: emerging evidence for crosslinguistic differences, *Journal of Memory and Language*, 54, 283-299.

SPEECH RHYTHM AND TIMING: STRUCTURAL PROPERTIES AND ACOUSTIC CORRELATES

Antonio Romano

Laboratorio di Fonetica Sperimentale 'Arturo Genre'

Dipartimento di Scienze del Linguaggio – Università degli Studi di Torino

antonio.romano@unito.it

ABSTRACT

In my intervention to the round table I summarised results from a selection of recent contributions to the research on rhythm and speech timing coming from two Italian laboratories: the *Laboratorio di Linguistica* of the *Scuola Normale Superiore di Pisa* and the *Laboratorio di Fonetica Sperimentale 'Arturo Genre'* of the University of Turin.

In my short presentation I emphasised reference to papers by Pier Marco Bertinetto and Chiara Bertini (Bertinetto & Bertini, 2008, forthcoming; Bertini & Bertinetto, 2009) and Paolo Mairano and Antonio Romano (Mairano & Romano, 2007, 2010), with an introduction explaining the reasons of my own interests in it.

REFERENCES

Bertinetto, P.M. & Bertini, C. (2008), On modeling the rhythm of natural languages, in *Proceedings of the 4th International Conference on Speech Prosody*, Campinas, Brazil, May 6-9, 2008, 427-430.

Bertinetto, P.M. & Bertini, C. (forthcoming), Towards a unified predictive model of Speech Rhythm, Manuscript.

Bertini, C. & Bertinetto, P.M. (2009), Prospezioni sulla struttura ritmica dell'italiano basate sul corpus semispontaneo AVIP/API, in *La fonetica sperimentale. Metodo e applicazioni* (L. Romito, V. Galatà & R. Lio, editors), Atti del 4° Convegno Nazionale dell'Associazione Italiana di Scienze della Voce, Arcavacata di Rende (CS), 3-5 Dicembre 2007, Torriana: EDK Editore, 3-21.

Mairano, P. & Romano, A. (2007), Inter-Subject Agreement in Rhythm Evaluation for Four Languages (English, French, German, Italian), in *Proceedings of the 16th International Congress of Phonetic Sciences*, Saarbrücken, Germany, August 6-10, 2007, 1149-1152.

Mairano, P. & Romano, A. (2010), Un confronto tra diverse metriche ritmiche usando Correlatore, in *La dimensione temporale del parlato* (S. Schmid, M. Schwarzenbach & D. Studer, editors), Atti del 5° Convegno Nazionale dell'Associazione Italiana di Scienze della Voce, Zurigo, Svizzera, 4-6 febbraio 2009 (this volume).

**LINGUISTICA,
FONETICA
E FONOLOGIA**

UN CONFRONTO TRA DIVERSE METRICHE RITMICHE USANDO CORRELATORE

Paolo Mairano, Antonio Romano
Laboratorio di Fonetica Sperimentale 'Arturo Genre', Università degli Studi di Torino
paolo.mairano@unito.it, antonio.romano@unito.it

SOMMARIO

Basandosi sulle teorie di percezione del parlato da parte dei bambini (v. Mehler *et al.*, 1996), Ramus *et al.* (1999) hanno proposto tre correlati ritmici (ΔC , ΔV and %V) che permetterebbero di distinguere i due gruppi di lingue. Questo nuovo approccio alla tipologia ritmica ha dato un impulso alla ricerca in questo campo cosicché, come è noto, sono stati proposti nuovi correlati: i PVI di Grabe & Low (2002), i Varco di Dellwo & Wagner (2003 e seguenti) e, recentemente, i CCI di Bertinetto & Bertini (2008).

Un primo obiettivo che ci siamo posti in questo lavoro è stato quello di testare il mutamento dei risultati al variare di alcuni fattori: sono stati calcolati %V, ΔV , ΔC , nPVI(V), rPVI(C), Varco(V), Varco(C), CCI(V) e CCI(C) per 36 campioni de *Il vento di tramontana e il sole* in varie lingue. Questa scelta rispecchia la nostra convinzione che sia opportuno studiare campioni di parlato controllato prima di affrontare il parlato spontaneo e che non sia necessario utilizzare campioni estremamente lunghi poiché alcuni test di discriminazione hanno dimostrato che il cervello umano distingue lingue isosillabiche e isoaccentuali anche con campioni di pochi secondi (v. Ramus *et al.*, 1999). Proponiamo anche di evitare il ricorso alla segmentazione automatica in quanto il guadagno in termini di tempo è controbilanciato da una perdita di precisione; al contrario, abbiamo pensato di automatizzare il processo di calcolo dei correlati: a questo scopo l'autore PM ha realizzato *Correlatore*, disponibile sul sito del Laboratorio di Fonetica Sperimentale 'Arturo Genre' di Torino; si tratta di un programma in Tcl/Tk che calcola i valori di %V, ΔV , ΔC , nPVI(V), rPVI(C), Varco(V), Varco(C), CCI(V) e CCI(C) e costruisce i grafici a partire dai file di annotazione di Praat (TextGrid) etichettati semplicemente come CV (consonante-vocale) o in SAMPA.

I dati ottenuti ci permettono di analizzare il variare dei risultati a seconda: a) dei diversi correlati utilizzati; b) dei diversi parlanti di una stessa lingua; c) di chi segmenta. Non sono stati ancora debitamente valutati gli eventuali effetti della velocità d'eloquio (argomento discusso in numerosi studi, v. per es. Dellwo & Wagner, 2003) perché, comunque, i nostri campioni sono piuttosto omogenei a questo riguardo (5-6,5 syll/s). In tutti i casi, si nota un certo grado di sovrapposizione tra gruppi di lingue che graviterebbero tradizionalmente attorno ai due poli sillabico e accentuale, presumibilmente dovuto al fatto che i diversi correlati sono sensibili a fenomeni differenti. Tuttavia, è difficile stabilire quali correlati rispecchino meglio la tradizionale dicotomia di lingue isosillabiche/isoaccentuali in quanto non sembra possibile avere un riscontro oggettivo del punto esatto in cui un determinato campione debba situarsi all'interno del continuum.

BIBLIOGRAFIA

- Bertinetto, P.M. (1989), Reflections on the dichotomy 'stress' vs. 'syllable-timing', *Revue de Phonétique Appliquée*, 91-92-93, 99-130.
- Bertinetto, P.M. & Bertini, C. (2008), On modeling the rhythm of natural languages, in *Proceedings of the 4th International Conference on Speech Prosody*, Campinas, Brazil, May 6-9, 2008, 427-430.
- Dauer, R.M. (1983), Stress-timing and syllable-timing reanalyzed, *Journal of Phonetics*, 11, 51-62.
- Dellwo, V. & Wagner, P. (2003), Relations between language rhythm and speech rate, in *Proceedings of the 15th International Congress of Phonetics Sciences*, Barcelona, Spain, August 3-9, 2003, 471-474.
- Grabe, E. & Low, E.L. (2002), Durational Variability in Speech and the Rhythm Class Hypothesis, in *Papers in Laboratory Phonology 7* (C. Gussenhoven & N. Warner, editors), Berlin: Mouton de Gruyter, 515-546.
- Ramus, F., Nespors, M. & Mehler, J. (1999), Correlates of linguistic rhythm in the speech signal, *Cognition*, 73, 265-292.

VARIABILITÀ RITMICA DI VARIETÀ DIALETTALI DEL PIEMONTE

Antonio Romano ^{a,b,c}, Paolo Mairano ^{a,c}, Barbara Pollifrone ^{b,c}

^aLFSAG – Laboratorio di Fonetica Sperimentale ‘Arturo Genre’

^bFacoltà di Lingue e Letterature Straniere

^cUniversità degli Studi di Torino

antonio.romano@unito.it, paolomairano@gmail.com, polbarbara@yahoo.it

ABSTRACT

In un suo recente contributo, Schmid (2004) ha mostrato come il torinese presenti caratteristiche più isoaccentuali (IA) rispetto ad altre varietà italo-romanze più isosillabiche (IS).

In questo studio, riferendoci allo stesso quadro metodologico e adottando una tecnica sperimentale già presentata in studi precedenti su altre lingue (v. Mairano & Romano, 2007), ci proponiamo di contribuire alla collocazione ritmica di alcuni dialetti del Piemonte, approfondendo le caratteristiche ritmiche di parlate piuttosto distanti tra loro, appositamente scelte alla periferia di questa regione linguistica.

La nostra ricerca si concentra sulle proprietà ritmiche presenti nelle versioni de *La tramontana e il sole* tradotte, adattate e interpretate da sei parlanti delle varietà di Roccaforte Ligure (AL), Briga Alta (CN), Exilles (TO) e Capanne di Marcarolo (AL), raccolte nell’ambito di lavori dell’*Atlante Toponomastico del Piemonte Montano (ATPM, 2005-2008)* e di Campertogno (VC) e Bagnolo Piemonte (CN), raccolte nell’ambito della recente monografia di Molino & Romano (2008) e della tesi di laurea inedita di Piccato (2007).

Alla classificazione suggerita dagli indicatori di Ramus *et al.* (1999) abbiamo affiancato quelle basate sugli indici nPVI(V) e rPVI(C) (*Pairwise Variability Index* di Grabe & Low, 2002), VarcoV e VarcoC (Dellwo & Wagner, 2003), CCI(V) e CCI(C) (*Control and Compensation Index* di Bertinetto & Bertini, 2008).

I brani sono stati segmentati in intervalli vocalici e consonantici grazie al programma *Praat* per mezzo di un’accurata procedura che valuta, fonologicamente, la categoria di appartenenza del suono atteso e, foneticamente, quella naturale del suono realizzato. I file *TextGrid* prodotti sono analizzati automaticamente mediante un applicativo in *Tcl/Tk* (v. Mairano & Romano, in questo volume) che produce l’analisi dei diversi correlati e ne permette una rappresentazione grafica.

Limitandoci alla lettura dei grafici basati sulle metriche di Ramus *et al.* (1999), i risultati collocano agli estremi opposti le due varietà di Roccaforte Ligure (fortemente caratterizzata per via dei suoi dittonghi discendenti), che si situa infatti tra quelle più IA (alti ΔV e ΔC), e quella di Capanne di Marcarolo (che conserva meglio il vocalismo atono finale ed evita gli allungamenti in sillaba chiusa), la quale si situa tra quelle IS (medio ΔV e basso ΔC). Campertogno, gravitante in area lombarda, si mostra quella col più alto ΔC mentre Exilles e Briga Alta (rispettivamente di area occitana e ligure) si caratterizzano per un ΔC medio ma un alto ΔV . Anche la varietà di Bagnolo Piemonte si avvicina a una posizione IA, restando tra quelle con valori più alti di rPVI(C) e VarcoC e CCI(C).

La disposizione dei valori ottenuti per le produzioni analizzate secondo i vari correlati differisce talvolta in modo significativo, mantenendo però sempre costanti alcune caratteristiche della disposizione generale (ad esempio, i valori relativi a Capanne di M. sono sempre i più bassi contribuendo a mantenere separati i nostri dati di questa varietà). Nel grafico CCI si osserva una diversa distribuzione di molte delle varietà considerate, che pur restano tra quelle a compensazione, tranne i valori riscontrati per il parlante di Campertogno che si avvicinano di più alla bisettrice del grafico.

Le distinzioni non mancano di seguire il gradiente di altri fenomeni, come appunto quelli legati alle riduzioni postaccentuali: se infatti a Exilles *forte* perde totalmente la sua ultima sillaba (come accade, con dati incostanti, per Bagnolo), a Campertogno si ha ancora la perdita della sola vocale finale e a Capanne la conservazione (con Briga e Roccaforte ancora propense a perderla o a desonorizzarla più spesso).

RIFERIMENTI BIBLIOGRAFICI

ATPM - Atlante Toponomastico del Piemonte Montano (2005-2008: 27-Roccaforte Ligure; 28-Briga Alta, 30-Exilles e 33-Capanne di Marcarolo).

Bertinetto, P.M. & Bertini, C. (2008), On modeling the rhythm of natural languages, in *Proceedings of the 4th International Conference on Speech Prosody*, Campinas, Brazil, May 6-9, 2008, 427-430.

Dellwo, V. & Wagner, P. (2003), Relations between language rhythm and speech rate, in *Proceedings of the 15th International Congress of Phonetics Sciences*, Barcelona, Spain, August 3-9, 2003, 471-474.

Grabe, E. & Low, E.L. (2002), Durational Variability in Speech and the Rhythm Class Hypothesis, in *Papers in Laboratory Phonology 7* (C. Gussenhoven & N. Warner, editors), Berlin: Mouton de Gruyter, 515-546.

Mairano, P. & Romano, A. (2007), Inter-Subject Agreement in Rhythm Evaluation for Four Languages (English, French, German, Italian), in *Proceedings of the 16th International Congress of Phonetic Sciences*, Saarbrücken, Germany, August 6-10, 2007, 1149-1152.

Mairano, P. & Romano, A. (2010), Un confronto tra diverse metriche ritmiche usando Correlatore 1.0, in *La dimensione temporale del parlato*, Atti del 5° Convegno Nazionale dell'Associazione Italiana di Scienze della Voce, Zurigo, Svizzera, 4-6 febbraio 2009 (S. Schmid, M. Schwarzenbach & D. Studer, editors).

Molino, P. & Romano, A. (2008), *Il dialetto valsesiano nella media Valgrande*, Alessandria: Dell'Orso.

Piccatò, E. (2007), *La parlata di Bagnolo Piemonte*, Tesi di Laurea (rel. A. Romano), Facoltà di Lingue e Letterature Straniere dell'Università degli Studi di Torino, inedita.

Ramus, F., Nespò, M. & Mehler, J. (1999), Correlates of linguistic rhythm in the speech signal, *Cognition*, 73, 265-292.

Schmid, S. (2004), Une approche phonétique de l'isochronie dans quelques dialectes italo-romans, in *Nouveaux départs en phonologie* (T. Meisenburg & M. Selig, editors), Tübingen: Narr, 109-124.

TEMPI E MODI DI CONSERVAZIONE DELLE *R* ITALIANE NEI *FRIGORIFERI CLIPS*

Alessandro Vietti ^a, Lorenzo Spreafico ^a, Antonio Romano ^b

^aCentro di Ricerca Lingue, Libera Università di Bolzano

^bLaboratorio di Fonetica Sperimentale 'Arturo Genre', Università degli Studi di Torino
alessandro.vietti@unibz.it, lorenzo.spreafico@unibz.it, antonio.romano@unito.it

SOMMARIO

Come noto, le realizzazioni di /r/ in italiano risentono di una certa variabilità socio-linguistica che si associa in parte a fattori diatopici e diastratici e in parte a imprevedibili tendenze individuali e familiari. Scopo della ricerca è offrire una caratterizzazione acustica di alcune realizzazioni di /r/ nell'italiano contemporaneo ricavabili dai dati CLIPS e da un campione raccolto a Bolzano con metodologia analoga.

Nel panorama degli studi condotti sulle varietà d'italiano, se si escludono le osservazioni di Canepari (1986 e 1999), le ricerche sulle modalità di realizzazione di /r/ sono al momento relativamente poco avanzate: studi acustici preliminari hanno soltanto sottolineato alcune caratteristiche salienti di rese piuttosto standard (cfr. tra gli altri Vaggies *et al.*, 1978) oppure osservato dati dialettali specifici nell'ambito di studi con finalità più ampie (Soriano, 2003; Felloni, 2006).

Mentre per altri domini linguistici l'argomento, già esplorato preliminarmente, incomincia ad essere affrontato più estensivamente (cfr. ad es. Meyer-Eppler, 1959; Delattre, 1966 e 1971; Schiller, 1988; Recasens, 1991; Espy-Wilson *et al.*, 1997; Solé, 1999; Wiese, 2001; Docherty & Foulkes, 2001; Blecua Falgueras, 2001; Kouznetsov & Pamies Bertrán, 2008) per quello italiano disponiamo del solo quadro di riferimento offerto da Romano (2002; in prep.) che ha osservato alcune delle numerose realizzazioni possibili col metodo dei *loci* acustici e nel quadro della teoria della perturbazione. Questo riferimento può tuttavia risultare inadeguato quando si tratti di rendere conto di articolazioni multiple e di strategie di realizzazione che, bisognose di verifiche articolatorie, sfuggano a rappresentazioni certe in quest'ottica.

Ancora maggiori sono le incertezze che sorgono quando l'osservazione si basa su materiali in cui diversi fenomeni concorrenti si sovrappongono a quelli indotti da variabili ambientali o sociolinguistiche: alla varietà di realizzazioni può corrispondere infatti anche una notevole variabilità acustica.

L'analisi che qui proponiamo si basa sui dati di italiano letto ricavati dalle liste di parole contenute nei materiali CLIPS e da quelle usate in un'indagine sull'italiano a Bolzano (cfr. Vietti & Spreafico, 2008). In particolare ci siamo concentrati sulle sequenze /VrV/ della parola *frigorifero*.

In termini generali, nei materiali CLIPS, nei quali è rappresentata da 120 realizzazioni (da parte di 8 parlanti per le 15 città del campione), abbiamo rilevato ad es. per la realizzazione della seconda /r/ di *frigorifero* (in attacco di sillaba accentata) una realizzazione dominante monovibrante (vibrata apico-alveolare, presente nel 40,8%; ben esemplificata nei dati di Palermo o, in presenza di una sospetta componente secondaria di uvularità, di Genova) che, rispetto a quella riportata nei dati di varietà iberiche (Recasens, 1991; Solé, 1999), si caratterizza per una maggiore durata e per una certa rigidità energetica prima e

dopo che la fanno percepire (seppur non polivibrante) come più forte di una normale monovibrante (cfr. Kouznetsov & Pamies Bertrán, 2008).

Quanto alle numerose varianti, i dati si presentano con una diversa distribuzione, con pochi casi di presunta cancellazione (3,3%) o, comunque, con sole forme di rotacismo vocalico (1,7%; con esempi isolati, da Napoli a Bergamo). Rese approssimanti pure sono state riconosciute nel 6,7% dei casi (in luoghi d'articolazione diversi), mentre di gran lunga più frequenti si sono rivelate le realizzazioni approssimanti interrotte da localizzati cali di energia (18,3%). Un altro 5,8% è rappresentato da realizzazioni monovibranti di durata particolarmente significativa (da 33 a 50 ms) durante la quale le caratteristiche acustiche di energia denunciano una tenuta simile a quella di un'occlusiva sonora e sono confermate da un discreto scoppio (che anche all'ascolto lascia pensare a brevi /d/ alveolari o post-alveolari). Realizzazioni velari, uvulari e faringali (uvularizzate o faringalizzate), approssimanti o costrittive, compaiono a Genova (1), Parma (5), Torino (4), Milano (1) e Cagliari (1) per un complessivo 10% (l'unica chiara costrittiva uvulare è di Parma, mentre sono risultate più comuni le monovibranti o approssimanti alveolari uvularizzate; cfr. Canepari, 1999). Una vibratile (*flap*), talvolta lateralizzata, è infine dominante nei dati di Venezia (per un residuo 5% complessivo sui dati nazionali), rendendo i *frigoriferi* di questa località gli unici del corpus la cui provenienza geografica sia facilmente riconoscibile.

Per queste rese, come per quelle dei materiali bolzanini, vista la rilevanza dei contributi delle altre formanti (F_3 e F_4) piuttosto che descrivere il suono cercando di determinare il *locus*, abbiamo osservato le caratteristiche temporali della transizione, in particolare la maggiore o minore rapidità del movimento acustico. Questo è definito attraverso la misurazione dei valori delle formanti delle vocali precedente e seguente nei punti stazionari e di transizione. Le variazioni di frequenza formantiche così ricavate, rapportate ai tempi del loro sviluppo e integrate con le informazioni relative alla concavità o convessità delle curve, permettono infatti di avanzare una descrizione più completa delle transizioni formantiche. Accanto alla descrizione delle caratteristiche spettro-temporali si osservano la presenza e la consistenza numerica di eventuali *burst*, rumori o frizioni legati alle diverse strategie di articolazione di /r/.

RIFERIMENTI BIBLIOGRAFICI

Si rimanda alla bibliografia riportata in calce all'articolo contenuto nel CD-ROM.

NOTE SULLE OPPOSIZIONI DI QUANTITÀ VOCALICA

Arianna Uguzzoni
Alma Mater Studiorum – Università di Bologna
arianna.uguzzoni@unibo.it

SOMMARIO

Queste note affrontano in modo succinto e certamente non esaustivo il tema della caratterizzazione e classificazione di alcune lingue europee che fanno un uso distintivo delle differenze temporali.

Nella prima nota si fa una separazione tra lingue con sistema quantitativo doppio, che riguarda sia vocali sia consonanti, e lingue con sistema quantitativo unico. Per queste ultime vengono delineati, con l'appoggio di esempi, tre scenari tipologicamente diversi tra di loro, a seconda che le opposizioni di quantità abbiano il loro ambito nella sequenza /vocale+consonante/, nella /consonante/ o nella /vocale/.

Argomento della seconda nota è la considerazione dei principali fattori che determinano l'organizzazione delle lingue in merito alla utilizzazione libera o vincolata delle distinzioni di quantità. Sono illustrate le condizioni fonotattiche e prosodiche entro le quali operano le opposizioni di quantità vocalica e/o consonantica: l'accento lessicale, la forma della sillaba, la posizione nella parola.

Riguardo alla indipendenza vs. dipendenza dall'accento lessicale si osserva che la distinzione fonologica tra vocali brevi e vocali lunghe nella maggior parte delle lingue europee odierne è limitata alle sillabe accentate, sia in area germanica, sia in area romanza. Per quel che concerne la distribuzione della quantità vocalica nella sillaba e nella parola sono esaminate tre situazioni: sillaba aperta all'interno di parola, sillaba chiusa, sillaba aperta in fine di parola. Una chiara differenziazione tipologica emerge dal diverso comportamento delle lingue nella prima e nella terza condizione.

La quarta nota presenta e discute aspetti della teoria che interpreta le opposizioni di quantità vocalica di alcune lingue germaniche come opposizioni prosodiche di taglio sillabico (*Silbenschnittgegensätze*). Secondo alcuni il taglio sillabico porta a una netta dicotomia fra due tipi di lingue: le *Quantitätensprachen* e le *Silbenschnittsprachen*. Secondo altri esso fornisce un metodo per descrivere adeguatamente modalità di differenziazione temporale che caratterizzano un tipo 'speciale' di sistema quantitativo.

Vengono di solito considerate lingue con taglio sillabico il tedesco (precisamente il tedesco settentrionale), l'olandese, l'inglese, in cui si oppongono due modalità di interazione tra la vocale accentata e la consonante successiva: *scharfer* vs. *sanfter Schnitt*. In alcune versioni della teoria brevità e lunghezza delle vocali accentate assumono il ruolo di concomitanti fonetici dei due modi di taglio sillabico, rispettivamente, del taglio brusco e del taglio piano.

FENOMENI D'ARMONIA VOCALICA IN AREA FRIULANA E IBERICA E LE SORTI DI -A FINALE LATINA

Renzo Miotti

Dipartimento di Romanistica – Università di Verona, Italia

renzo.miotti@univr.it

SOMMARIO

L'armonia vocalica è un fenomeno assimilatorio che consiste nell'estensione di tutti o d'alcuni tratti d'una vocale ad altri segmenti vocalici, normalmente adiacenti. L'armonia può manifestarsi in una duplice direzione: da una posizione forte, vale a dire prominente dal punto di vista percettivo, verso posizioni più deboli; viceversa, da posizioni poco prominenti verso gli altri segmenti vocalici. In letteratura, le cause dell'armonia vengono riportate al conseguimento di benefici d'ordine strutturale: semplificazione articolatoria (Pulleyblank, 2002), benefici percettivi (soprattutto se da posizioni deboli; cfr. Walker, 2005, 2006), semplificazione articolatoria e benefici percettivi (Cole & Kisseberth, 1994). Per una sintesi della questione, cfr. Jiménez & Lloret (c.d.s.).

Sulla scorta di queste premesse, il presente lavoro intende presentare e discutere i risultati (ancora provvisori, in vista d'essere corroborati da una base più ampia di dati) d'un'indagine condotta sul friulano centro-orientale, che provverebbero l'esistenza di processi riconducibili al primo dei due modelli d'assimilazione (da posizioni forti a posizioni deboli): il tratto [\pm ATR] (ma non il punto d'articolazione) delle vocali medie semichiusate accentate verrebbe esteso alla vocale media non-accentata finale *-e/* (normalmente [ɛ]) che, diacronicamente, rappresenta il normale esito di *-A* latina, avvenuto per innalzamento e anteriorizzazione, nella maggior parte dei dialetti centro-orientali, più innovativi (laddove le varietà più marginali conservano *-a/*). Va detto che il fenomeno cui si fa riferimento non ha sinora trovato riscontro in letteratura, la quale si limita a rilevare la generale tendenza all'abbassamento ([ɛ̄, ɔ̄]) delle medie non-accentate in posizione finale (Canepari, 2006³; Miotti, 2002), senza però prestar attenzione alle significative differenze riscontrabili, con regolarità, in dipendenza dai condizionamenti visti.

Avremo dunque [ɛ̄] dopo /'e, 'o/ ma [ɛ] nei contesti non-armonici (cioè dopo qualsiasi altra vocale): [ku'ɛ̄tɛ, 'mɔ̄tɛ] 'cotta, mossa' ma ['sɛ̄ɛ̄ɛ̄, 'kɔ̄ɔ̄ɛ̄] 'seta, coda'. Si tenga presente che pure /'e, 'o/ arrivano a [ɛ̄, ɔ̄] (seppure in minor grado rispetto alle non-accentate finali). S'ipotizza che gli stessi meccanismi che operano in sincronia siano stati responsabili, in diacronia, del mutamento lat. *-A* > friul. centro-orient. moderno *-e/*, attraverso i suddetti processi d'innalzamento e anteriorizzazione, a partire da una situazione d'instabilità ed estrema variabilità fonetica che dovette interessare /a/ finale nell'udinese medievale. Il quadro, per queste prime fasi evolutive, sarebbe del tutto simile a quello che è possibile delineare per dialetti valenziani meridionali attuali (vedi oltre).

In friulano centro-orientale, il fenomeno armonico sembrerebbe condizionato da fattori contestuali e da fattori legati alla posizione dei segmenti all'interno dell'enunciato: solo le vocali finali in tonia (cioè alla fine d'enunciato intonativo, in posizione prepausale) sono interessate dal processo descritto, laddove, all'interno dell'enunciato stesso, l'effetto tende a ridursi fino a scomparire (con una notevole tendenza alla riduzione/centralizzazione dei timbri vocalici, fenomeno che verrà discusso in uno specifico paragrafo).

I fenomeni friulani vengono confrontati con quelli descritti per altre varietà romanze, in particolare iberiche (varietà valenziane meridionali: Jiménez, 1998, 2001; Jiménez & Lloret, c.d.s.). In valenziano (dove -A latina > -/a/), il processo è limitato a -/a/ preceduta da /'ɛ, 'ɔ/, che agiscono sulla vocale finale innalzandola e propagando il tratto 'punto d'articolazione' (coronale e labiale, rispettivamente): [is'tɛrjɛ] "isteria", [is'tɔrjɔ] "storia".

Il confronto friulano-valenziano mette in evidenza alcuni punti problematici d'ordine interpretativo, per quanto riguarda l'interpretazione articolatoria delle cause dei processi esaminati: la semplificazione articolatoria si manifesta all'interno di domini omogenei, con estensione dei tratti a vocali contigue, solo in valenziano, mentre ciò non si verifica necessariamente per il friulano, che conosce minori restrizioni in tal senso. Mentre in valenziano l'armonia interessa solo vocali contigue (nei proparossitoni la vocale postaccentata interna blocca il processo: ['tɛtrika], non *-[kɛ] "tetra"), in friulano, al contrario, non sembrano esserci restrizioni in tal senso. Se in valenziano il dominio dell'armonia è la parola prosodica e le vocali dei pronomi enclitici non sono perciò interessate dal processo armonico, in friulano, al contrario, queste ultime partecipano al processo a livello postlessicale: valenziano ['pɛlɛla] (e non *-[lɛ]) "pelala" vs. friulano ['mɛtɪlɛ] (e non *-[lɛ]) "mettila".

RIFERIMENTI BIBLIOGRAFICI

- Canepari, L. (2006³), *Manuale di fonetica*, München: Lincom.
- Cole, J. & Kisseberth, C. (1994), An Optimal Domains Theory of Vowel Harmony, *Studies in the Linguistic Sciences*, 34, 101-114.
- Jiménez, J. (1998), Valencian Vowel Harmony, *Rivista di Linguistica*, 10, 137-161.
- Jiménez, J. (2001), L'armonia vocàlica en valencià, in *Actes del Novè Colloqui d'Estudis Catalans a Nord-Amèrica*, Barcelona, Spagna, 1998, Barcelona: Publicacions de l'Abadia de Montserrat, 217-244.
- Jiménez, J. & Lloret, M.-R. (c.d.s.), Entre la articulación y la percepción: armonías vocálicas en la península Ibérica, in *Actes du XXV Congrès International de Linguistique et de Philologie Romanes*, Innsbruck, 3-8 septembre 2007 (M. Iliescu, H. Siller & P. Danler, editors), Tübingen: Niemeyer.
- Miotti, R. (2002), Friulian, *Journal of the International Phonetic Association*, 32, 237-247.
- Pulleyblank, D. (2002), Harmony drivers: no disagreement allowed, in *Proceedings of the Twenty-eighth Annual Meeting of the Berkeley Linguistics Society* (J. Larson & M. Paster, editors), Berkeley, California: Berkeley Linguistics Society, 249-267.
- Walker, R. (2005), Weak Triggers in Vowel Harmony, *Natural Language and Linguistic Theory*, 23, 917-989, <http://roa.rutgers.edu/>.
- Walker, R. (2006), *Long-distance Metaphony: A Generalized Licensing Proposal*, Lavoro presentato al *PhonologyFest Workshop*, Indiana University, Bloomington. <http://www-rcf.usc.edu/~rwalker/pubs.html>

COARTICOLAZIONE E MUTAMENTO. UNA RICERCA SUL VOCALISMO ATONO FINALE NELL'ENTROTERRA MACERATESE

Tania Paciaroni
Università di Zurigo
paciaron@rom.uzh.ch

SOMMARIO

Come è noto, l'aspetto tipizzante nel vocalismo atono finale dell'area mediana è la distinzione di /-u/ ed /-o/ finali, che non sono confluite in /-o/ come in italiano standard. All'interno dell'area maceratese-fermana-camerte, fin dai primi decenni del secolo scorso, tuttavia, Erich Mengel (1936) segnalava, nelle varietà del litorale costiero, la sostituzione della /-u/ con /-o/, per effetto dell'influenza dello standard, di contro alla saldezza della distinzione nell'interno. L'attualità di questo quadro è a tutt'oggi confermata dalla letteratura linguistica (Balducci, 2000: 28-29). Ne è *exemplum*, tra gli altri, la parlata di Matelica, in provincia di Macerata, ove i parlanti riconoscono nel fatto di "avere più -u" il tratto caratteristico del proprio dialetto. Tuttavia, a dispetto di questa autovalutazione e delle descrizioni disponibili (cfr. Bricchi, 1984; Traballoni, 2002-2003), a chi ascolti il matelicese non sfuggerà la variazione timbrica delle /-u/ finali.

Per verificare se tale variabilità sia un tratto evolutivo recente, dettato da interferenza con l'italiano, o se costituisca invece una tappa intermedia verso una configurazione diversa, eventualmente di tipo già noto e presente nell'area (cfr. Paciaroni & Loporcaro, in stampa), si è proceduto alla raccolta di materiale in inchieste sul campo con tre informatori dialettologi in tre stili enunciativi diversi: parole in isolamento, entro frasi e in parlato (semi)spontaneo. Dal *corpus* sono state estratte le sequenze contenenti: i) /o/ tonica; ii) /u/ tonica; iii) /-o/ finale < -ō, -ō; iv) /-u/ finale dell'articolo, del clitico oggetto e del dimostrativo; v) /-o/ finale dell'articolo, del clitico oggetto e del dimostrativo; v) /-u/ delle altre categorie lessicali con base etimologica -Ū(M). Per ogni vocale, in tutti e tre gli stili di parlato, sono state misurate la frequenza delle prime due formanti (F1, F2), la durata del segmento, l'intensità massima. Complessivamente sono stati misurati 612 stimoli.

Dalle analisi è risultato che l'abbassamento di /-u/ che oggi si presenta a Matelica è un fenomeno di variazione soggetto a condizioni coarticolatorie, senza ripercussioni sul sistema (e perciò non accessibile alla coscienza metalinguistica del parlante), dipendente dalla tipologia di parlato. L'indagine ha anche mostrato che la variazione non è generalizzata, ma ristretta alle sole parole lessicali, mentre non riguarda le parole funzionali, ove l'opposizione /-u/ ≠ /-o/ svolge la funzione morfologica di marcamento di genere maschile ≠ neutro.

Il dialetto di Matelica sembra quindi rispecchiare una tappa intermedia del percorso diacronico che, a partire da una configurazione di tipo reatino, con realizzazione di /-u/ univoca ([-u]), ha portato all'insorgere di sistemi con armonia vocalica di tipo cervaròlo (Merlo, 1922), ampiamente documentati nell'area mediana, presenti nel maceratese, tra gli altri, a San Severino M., con /-u/ che si abbassa a [-o] dopo vocale tonica media, indipendentemente dal fattore diafasico.

RIFERIMENTI BIBLIOGRAFICI

Balducci, S. (2000), *Marche*, Pisa: Pacini.

Bricchi, A. (1984), *Matelica. I suoi abitanti. Il suo dialetto*, Matelica: Associazione Pro Matelica.

Mengel, E. (1936), *Umlaut und Diphthongierung in den Dialekten des Picenums*, PhDThesis, Univ. of Köln, Germany.

Merlo, Cl. (1922), *Fonologia del dialetto della Cervara in provincia di Roma*, Roma: Società Filologica Romana.

Paciaroni, T. & Loporcaro, M. (in stampa), Funzioni morfologiche della distinzione fra *-u* e *-o* nei dialetti del Maceratese, in *Actes du XXV Congrès International de Linguistique et de Philologie Romanes*, Innsbruck, 3-8 septembre 2007 (M. Iliescu, H. Siller & P. Danler, editors), Tübingen: Niemeyer.

Traballoni, M. (2002-2003), *Glossario del dialetto di Matelica*, MThesis, Univ. of Macerata, Italy.

DURATA E STRUTTURE FORMANTICHE NEL PARLATO TOSCANO: INDAGINI PRELIMINARI SU UN CAMPIONE DI DIALOGHI SEMISPONTANEI

Nadia Nocchi ^a, Silvia Calamai ^b

^a Università di Zurigo, ^b Università degli Studi di Siena (sede di Arezzo)
nocchi_nadia@yahoo.com, calamai@unisi.it

SOMMARIO

A partire da un *corpus* articolato e particolarmente complesso anche in considerazione della sua diversificazione geografica ‘fine’ (da dialettologi), il contributo analizza il vocalismo del parlato di area toscana osservando un parametro prosodico (la durata) e alcuni parametri segmentali (le prime due formanti). L’architettura del disegno sperimentale mira a fornire un contributo in merito sia a interrogativi di fonetica generale, sia a interrogativi di carattere più squisitamente geolinguistico, essendo tre le città toscane di cui si presentano i dati acustici. Il materiale sonoro analizzato è di tipo semispontaneo (dialoghi *map task*): sebbene sia stato raccolto con un medesimo protocollo sperimentale nelle tre località, pone inevitabilmente questioni cruciali sia di rappresentatività statistica, sia di difficoltà oggettive nel rilevamento dei valori formantici. Anche per questa ragione, ad osservazioni quantitative affianchiamo, nella seconda parte, considerazioni squisitamente qualitative.

In primo luogo il *corpus* è indagato, da un punto di vista quantitativo, nella sua interezza: viene esplorata l’influenza di alcuni fattori linguistici – ‘accento’ (accento di parola *vs.* accento di frase), ‘posizione’ (iniziale, interna, finale di enunciato) e, più parzialmente, ‘tipo di parola’ – nella distribuzione dei valori formantici e temporali per il vocalismo tonico. I fattori ‘posizione’ e ‘tipo di parola’ vengono poi osservati nel vocalismo atono. Infine, il vocalismo tonico viene confrontato con quello atono. In secondo luogo sono osservati i valori formantici e temporali suddivisi per località, al fine di indagare la possibile influenza del fattore ‘spazio’ sia nel vocalismo tonico che in quello atono. Per il sistema tonico scorporato per località si fornisce anche un quadro relativo all’influenza del fattore ‘accento’. Un confronto tra vocalismo tonico e vocalismo atono nei tre punti d’indagine chiude l’analisi quantitativa. In conclusione, viene proposto un tentativo di ispezione fine dell’evoluzione temporale dei movimenti formantici all’interno di un numero ristretto di dati – sempre suddivisi per località – per capire se gli aspetti dinamici dei segmenti vocalici permettono, anch’essi, di rendere conto di alcune differenze geolinguistiche.

Dal momento che il quadro del vocalismo toscano – anche nei suoi aspetti più tipizzanti – ci è noto, dalla letteratura, negli stili più controllati (parlato letto di varie tipologie), il presente contributo permette anche di valutare in quale misura viene preservata la tipicità del vocalismo pisano e soprattutto livornese (abbassamento delle vocali medio-basse e posteriorizzazione di /a/) in stili di eloquio più liberi (parlato semispontaneo). I dati riferiti al vocalismo atono – per la prima volta presentati su un campione di parlato semispontaneo di base toscana – intendono offrire un contributo in merito al controverso tema che va sotto il nome di ‘riduzione vocalica’.

**DIAGNOSTICA FONOLOGICA E DIAGNOSI FONETICA.
OSSITONI LUNGI IN SILLABA LIBERA
NEI DIALETTI DI SAMBUCA PISTOIESE (PT)**

Lorenzo Filipponio, Nadia Nocchi
Università di Zurigo
filippon@rom.uzh.ch; nocchi_nadia@yahoo.com

SOMMARIO

Le isoglosse che compongono la linea La Spezia-Rimini (o, meglio, Carrara-Fano) risultano nell'area appenninica tra Pistoia e Bologna pressoché sovrapposte tra loro: ciò non impedisce però di rilevare anche in questo territorio una scansione dei fenomeni la cui distribuzione può rendere conto della loro cronologia (§1). Un luogo di indagine privilegiato è il territorio di Sambuca Pistoiese (PT), politicamente toscano ma idrograficamente adriatico, costituito da piccole frazioni oramai disabitate, dislocate lungo valli parallele e isolate tra loro. Questa situazione orografica ha di fatto permesso, pur nello spazio ridotto di 77 chilometri quadrati, il mantenimento di caratteristiche linguistiche peculiari e differenziate, in un contesto generale di forte conservatività. Qui sopravvivono (o agonizzano) colonie garfagnine (Treppio), varietà pienamente toscane (Torri), varietà analoghe a quelle dell'alto Appennino bolognese (Pàvana) e altre più marcatamente di transizione (Lagacci, Stabiazioni, Cavanna, Castello di Sambuca), in cui a un fondo gallo-italico si sono sovrapposti tratti toscani, come è dimostrato, per esempio, dalla sovraestensione della gorgia ai contesti di degeminazione protonica e di mancato raddoppiamento fonosintattico. Abbiamo dunque concentrato la nostra attenzione su queste ultime varietà, per verificare se anche in esse fosse insorta la quantità vocalica distintiva, già operante nelle limitrofe località dell'area bolognese così come a Pàvana. Vista la resistenza della geminazione postonica, la verifica è stata condotta misurando la durata della vocale tonica degli ossitoni in sillaba libera, priva, come insegna André Martinet (1975), dei condizionamenti dovuti al nesso con la consonante successiva. Molti dialetti gallo-italici hanno infatti sviluppato degli ossitoni secondari con vocale tonica lunga, che formano coppie minime con ossitoni proto-romanzi con vocale tonica parametricamente breve. Questi ossitoni secondari si ritrovano, con significative differenze di distribuzione, anche nelle varietà da noi analizzate. La nostra indagine (§3) ha evidenziato per quelle geograficamente più vicine al toscano, Lagacci e Stabiazioni, un reintegro degli ossitoni nei parametri dell'italiano standard, che escludono la presenza di vocale tonica lunga finale di parola. Diversa invece è la situazione prospettata dalle parlate di Cavanna e di Castello: in questo caso, infatti, abbiamo registrato la presenza di ossitoni con vocale tonica lunga, costante in contesto di isolamento e finale di frase, saltuaria in contesto interno di frase. Ciò lascia supporre che la lunghezza della vocale tonica degli ossitoni secondari sia foneticamente ammessa, ma non sia integrata strutturalmente, vigendo ancora la quantità consonantica distintiva. Nei casi di presenza in contesto interno di frase di ossitoni con vocale tonica lunga si osserva addirittura la messa in atto di strategie di riparazione, come l'allungamento della consonante iniziale della parola successiva, che oscura la lunghezza vocalica producendo una sorta di raddoppiamento fonosintattico del tutto inaspettato. I dati confermano dunque (§4) un quadro tipicamente di transizione, che aggiunge un nuovo dettaglio alla casistica di Martinet, mostrando un indebolimento in ossitonia del parametro di quantità, che prelude ai successivi sviluppi gallo-italici visibili appena più a nord.

ELISIONE OBBLIGATORIA, VARIABILE E POCO FREQUENTE NEL FIORENTINO: UN CASO DI ALLOMORFIA FRASALE PRECOMPILATA CON FORME PREFERENZIALI

Luigia Garrapa
Università del Salento & Universität Würzburg
luigiagarrapa@yahoo.it

SOMMARIO

Questo lavoro intende esplorare e chiarire l'applicazione dell'elisione nella varietà di italiano parlata a Firenze, individuando i fattori che ne determinano l'applicazione. I contesti presi in esame sono costituiti dai determinanti monosillabi (*l(o)*, *la* e *le*) e bisillabi (*un(o)/una*, *quell(o)/quella*, *questo/questa*, *quelle* e *queste/questi*) seguiti da sostantivi che iniziano per vocale e dai proclitici monosillabi (*lo/la*, *li/le_{acc}*, *mi/ti*, *ci/vi* e *le_{dat}*) seguiti da verbi lessicali che iniziano per vocale.

I dati analizzati provengono da due fonti: il corpus *C-Oral-Rom* (Cresti & Moneglia, 2005) e le inchieste sul campo condotte a Firenze in dicembre 2007. Sebbene sia i dati del corpus che quelli raccolti sul campo siano rappresentativi per il fiorentino, i primi sono da intendersi come parlato spontaneo, mentre i secondi sono caratterizzabili come parlato elicitato. In questa sede si discutono i dati congiunti del corpus e delle inchieste e si avanza una proposta (finora assente in letteratura) per la rappresentazione dell'elisione nei determinanti e nei proclitici.

L'elisione nel fiorentino risulta condizionata dalla tipologia delle parole funzionali (determinanti vs. proclitici), dal tratto morfologico di numero e dallo stile discorsivo. In primo luogo, le vocali atone finali dei determinanti subiscono l'elisione con maggiore frequenza rispetto alle vocali finali dei proclitici. In secondo luogo, le vocali finali dei determinanti singolari, dei proclitici accusativi singolari e dei clitici di persona vengono elise più frequentemente rispetto alle vocali finali dei determinanti plurali e dei clitici accusativi plurali. In terzo luogo, l'elisione viene applicata con maggiore probabilità nel parlato informale rispetto a quello formale.

Nel complesso, tre processi di elisione possono essere individuati: l'elisione obbligatoria nei determinanti maschili singolari *un(o)*, *l(o)* e *quell(o)*, l'elisione variabile nei determinanti singolari e plurali e l'elisione poco frequente con i proclitici. Per quanto riguarda l'elisione variabile nei determinanti, /o/ ed /a/ dei determinanti singolari vengono elise nel 96% dei casi, mentre /i/ ed /e/ dei determinanti plurali subiscono l'elisione nel 13% dei casi. Le vocali finali dei proclitici vengono elise poco frequentemente. Le vocali /o/ ed /a/ dei clitici accusativi singolari subiscono l'elisione nel 43% dei casi, /i/ dei clitici di persona viene elisa nel 32% dei casi, e /i/ ed /e/ dei clitici accusativi plurali vengono cancellate nel 5% dei casi. Il clitico dativo *le*, invece, rifiuta categoricamente l'elisione.

La proposta avanzata in questo lavoro è che l'elisione nelle parole funzionali analizzate non sia il risultato di un 'vero' processo di cancellazione vocalica, ma che sia le forme terminanti per vocale che le forme elise delle parole funzionali analizzate siano elencate individualmente nel lessico mentale assieme al contesto in cui possono apparire ed alle eventuali preferenze di selezione (fra le forme elise e quelle terminanti per vocale) in contesto prevocalico. Sebbene una tale proposta contribuisca a rendere piuttosto ridondante

il lessico mentale, essa risulta plausibile in prospettiva psicolinguistica. In effetti, determinanti e proclitici possiedono un'elevata frequenza di occorrenza nel parlato spontaneo e le 'parole' che vengono processate più frequentemente nel lessico mentale tenderebbero a sviluppare una propria rappresentazione lessicale indipendentemente dalla loro eventuale complessità morfologica (Schreuder & Baayen, 1995).

Sulla scia di Hayes (1990), Mascaró (2007) e Bonet *et al.* (2007), l'elisione viene interpretata come un processo di allomorfia frasale precompilata, in cui la selezione dell'allomorfo appropriato non è condizionata internamente nel lessico, ma è determinata fonologicamente, ossia dalla forma fonologica della parola adiacente e, più precisamente, dal fatto che la parola adiacente inizi per vocale o per consonante. L'allomorfia frasale è *categorica* in una serie ristretta di contesti e *graduale* nella maggior parte dei contesti. Per quanto riguarda le preferenze di selezione fra gli allomorfi, si suggerisce che, per i determinanti ed i proclitici analizzati, sia le forme terminanti per vocale che quelle elise siano elencate nel lessico mentale in qualità di allomorfi e che le preferenze selettive fra gli allomorfi in contesto prevocalico siano indicate nelle entrate lessicali delle parole funzionali in questione.

Per quanto riguarda l'elisione obbligatoria su *un(o)*, *l(o)* e *quell(o)*, essa viene rappresentata come allomorfia frasale precompilata *categorica*, in cui le forme elise *un*, *l'* e *quell'* sono elencate nel lessico mentale assieme alle forme *uno*, *il/lo* e *quell/quello*. Dunque, l'elisione obbligatoria con *un(o)*, *l(o)* e *quell(o)* sarebbe il risultato della selezione categorica di *un*, *l'* e *quell'* davanti a sostantivi che iniziano per vocale.

L'elisione variabile nei determinanti viene interpretata come un caso di allomorfia frasale precompilata *graduale*, in cui sia le forme terminanti per vocale che quelle elise dei determinanti singolari e plurali sono elencate nel lessico mentale assieme ai rispettivi contesti di occorrenza. Tuttavia, anche le preferenze di selezione in contesto prevocalico sono codificate nell'entrata lessicale dei determinanti in questione. Per quanto concerne i determinanti singolari, gli allomorfi elisi costituiscono le forme preferenziali da inserire in contesto prevocalico, ad esempio *quest' > questa oliva*, il che risulta nell'apparente frequente applicazione dell'elisione. Per quanto riguarda i determinanti plurali, invece, gli allomorfi terminanti per vocale rappresentano le forme preferenziali da inserire in contesto prevocalico, ad esempio *queste > quest'olive*, dando luogo all'applicazione poco frequente dell'elisione.

L'elisione poco frequente nei proclitici, invece, viene rappresentata come un caso di allomorfia frasale precompilata *graduale* in cui le forme terminanti per vocale costituiscono le forme preferenziali in contesto prevocalico sia per i clitici accusativi singolari, che per quelli plurali e per i clitici di persona, ad esempio *lo > l'usava* e *mi > m'amava*.

RIFERIMENTI BIBLIOGRAFICI

Si rimanda alla bibliografia riportata in calce all'articolo contenuto nel CD-ROM.

CONTINUUM DIAFASICO E DINAMICHE DIAGENERAZIONALI NEL BASSO E ALTO CASERTANO ORIENTALE

Edoardo Mastantuoni
Università degli Studi di Torino
emastantuoni@hotmail.com

SOMMARIO

Il presente contributo illustra alcuni esiti di un'inchiesta dialettologica in corso in sei comuni orientali della provincia di Caserta (basso casertano: San Nicola la Strada, Castel Morrone, Ruviano; alto casertano: Baia e Latina, Piedimonte Matese, Letino), della quale si premettono qui di seguito le principali linee-guida e metodologiche.

1. I sei punti sono disposti lungo un asse sud-nord, che dal circondario del capoluogo di provincia sale fino ai monti del Matese (ca. 65 km). Il materiale linguistico raccolto include l'intero continuum diafasico che unisce le varietà locali d'italiano (il substandard) ai dialetti e consente inoltre un'analisi diagenazionale. Il lavoro finale sarà inoltre incentrato anche sugli aspetti attinenti alla diatopia, oltre che su possibili osservazioni contrastive rispetto alle *koinai* napoletane.

2. Pur concorrendo a colmare in parte la mancanza di documentazione sui dialetti di Terra di Lavoro, l'inchiesta mira a descrivere, con una 'fotografia' delle dinamiche linguistiche in atto, gli usi linguistici di due gruppi sociali: gli anziani e i giovani.

3. Una peculiarità del campione consiste nel fatto che esso è composto quasi esclusivamente da donne, a differenza di quanto è accaduto per grandi opere geolinguistiche come l' AIS e l' ALI, nelle quali era in stragrande maggioranza maschile.

4. Per poter essere facilmente accettato nelle comunità, e nel tentativo di stabilire un rapporto con i soggetti da intervistare quanto più possibile cordiale, rilassato e amichevole, mi sono avvalso della collaborazione di alcuni *insider*. Grazie a questi ultimi, individuandoli tra i loro parenti, amici e conoscenti, ho scelto cinque informatrici e un informatore per ciascuna delle sei località (per un totale di trentasei informatori: trenta donne e sei uomini), come illustrato nella tabella 1:

	≤ 40	≥ 60
F	2	3
M	1	-

Tabella 1: Costituzione numerica dei due gruppi sociali di giovani (≤ 40) e anziani (≥ 60)

5. Le interviste sono state libere, sono state richieste adducendo un pretesto sociologico, hanno avuto una durata media di circa venti minuti e sono state registrate con dispositivo digitale visibile agli informatori.

6. Il piano descrittivo prescelto privilegia il livello fonetico: vocalismo tonico, atono, consonantismo. L'analisi del livello fonetico viene svolta con metodologia uditiva e descrittiva, ma non si escludono, in un'altra sede, possibili affondi di analisi anche strumentale su singoli aspetti di particolare complessità. Il lavoro finale considererà anche fonosintassi e prosodia, oltre agli aspetti morfologici e sintattici più rilevanti.

CONFINI PROSODICI E VARIAZIONE SEGMENTALE. ANALISI ACUSTICA DELL'ALTERNANZA MONOTTONGO/DITTONGO IN ALCUNI DIALETTI DELL'ITALIA MERIDIONALE

Giovanni Abete & Adrian Simpson
Friedrich-Schiller-Universität Jena
giovanni.abete@libero.it, adrian.simpson@uni-jena.de

SOMMARIO

Il presente contributo espone i primi risultati di una ricerca quadriennale che ha riguardato la variabilità del vocalismo tonico in quattro dialetti dell'Italia meridionale (Pozzuoli e Torre Annunziata in Campania, Belvedere Marittimo in Calabria, Trani in Puglia). In particolare, ci si è concentrati su un fenomeno di alternanza sincronica tra esiti monottongali e esiti dittongali, che è tra le caratteristiche più interessanti dei dialetti su menzionati. Come l'analisi evidenzierà, questa alternanza è particolarmente sensibile alla presenza di determinati confini prosodici, con le varianti dittongali che emergono in corrispondenza dei confini di ordine gerarchicamente superiore. Alla questione della variazione tra esiti dittongali e esiti monottongali si aggiunge quella delle variazioni di durata, anch'esse dipendenti in larga misura dalla posizione della variabile nella struttura prosodica. Lo studio delle relazioni tra posizione prosodica, variazioni di durata e alternanza monottongo/dittongo costituisce dunque l'obiettivo primario di questo lavoro. La trattazione sarà limitata in questa sede ai dati relativi a Pozzuoli e Torre Annunziata.

Tale studio si inserisce nel filone delle ricerche che negli ultimi anni hanno indagato gli effetti della struttura prosodica non solo a livello soprasegmentale ma anche a livello segmentale (cfr. ad es. Fougeron & Keating, 1997; Keating *et al.*, 2003; Cho, 2004; Cho *et al.*, 2007). Queste ricerche hanno messo in evidenza la mole di variazione fonetica sistematica, di livello anche molto fine, che si accompagna a diverse posizioni nella struttura prosodica, attribuendo in genere a tale variazione una funzione importante nel processo di segmentazione della catena parlata in unità prosodiche di livelli diversi. Rispetto ai lavori citati, generalmente condotti su parlato di laboratorio, la presente ricerca si contraddistingue per l'uso di parlato spontaneo e di varietà substandard quali i dialetti italiani.

Questo studio si avvale di un corpus di parlato spontaneo raccolto sul campo da uno degli autori, mediante una versione modificata dell'intervista libera (cfr. Como, 2006). Il corpus è stato segmentato ed etichettato in maniera parziale, limitando questo lavoro a una lista di parole precedentemente selezionate. L'etichettatura ha previsto diversi livelli: segmenti, parole, sintagmi intonativi, enunciati. Altre informazioni prosodiche sono state inserite nella trascrizione al livello segmentale. Questa impalcatura ha consentito di distinguere tra *tokens* vocalici in diverse posizioni prosodiche: interna al sintagma fonologico, finale di sintagma fonologico ma interna al sintagma intonativo, finale di sintagma intonativo.

Sulla base dei *tokens* vocalici etichettati sono state effettuate analisi acustiche della durata e della struttura formantica. Le durate sono state ottenute direttamente dai *files* di etichettatura, mentre per l'analisi della traiettoria dittongale è stato realizzato uno *script* per la stima automatica dei valori formantici, valori che sono poi stati controllati manualmente e, quando necessario, corretti. Un coefficiente di dittongazione per ciascun *token* è stato

calcolato sulla base dei valori in Bark delle prime due formanti. Il metodo di descrizione della traiettoria dittongale utilizzato nella presente ricerca costituisce una evoluzione del metodo di Holbrook & Fairbanks (1962), e segue essenzialmente Simpson (1998), con qualche differenza rispetto all'algoritmo per la stima delle traiettorie formantiche.

L'analisi delle durate ha mostrato la presenza di un forte allungamento dei segmenti vocalici tonici nella posizione finale di sintagma intonativo rispetto alle durate in posizione interna (l'incremento è del 92% per Pozzuoli e del 116% per Torre Annunziata). Un allungamento statisticamente significativo, anche se più modesto, è stato rilevato nella posizione finale di sintagma fonologico rispetto alle durate in posizione interna a tale costituente (l'incremento è del 35% per Pozzuoli e del 26% per Torre Annunziata). L'allungamento è dunque di tipo incrementale: le durate si fanno progressivamente più lunghe a mano a mano che si sale nella gerarchia dei costituenti prosodici.

L'analisi del coefficiente di dittongazione ha evidenziato una polarizzazione piuttosto netta tra le realizzazioni della variabili vocaliche in posizione interna e quelle in posizione finale di sintagma intonativo. Mentre in posizione interna il coefficiente si mantiene su livelli piuttosto bassi, in posizione finale esso supera tendenzialmente (e spesso in maniera consistente) il valore di 1.8, che su base impressionistica è stato individuato come soglia critica per la percezione di dittongazione. L'alternanza monottongo/dittongo in questi dialetti rientra quindi tra i fenomeni di variazione fonetica che dipendono dalla posizione nella struttura prosodica e che forniscono al parlante indici acustici per la segmentazione della catena palata in costituenti prosodici.

BIBLIOGRAFIA

Cho, T. (2004), Prosodically conditioned strengthening and vowel-to-vowel coarticulation in English, *Journal of Phonetics*, 32, 141-176.

Cho, T., McQueen, J.M. & Cox, E.A. (2007), Prosodically driven phonetic detail in speech processing: the case of domain-initial strengthening in English, *Journal of Phonetics*, 35, 210-243.

Como, P. (2006), Elicitation techniques for spoken discourse, in *Encyclopedia of language and linguistics* (K. Brown, editor), second edition, vol. 4., Amsterdam: Elsevier, 105-109.

Fougeron, C. & Keating, P.A. (1997), Articulatory strengthening at edges of prosodic domains, *Journal of the Acoustical Society of America*, 101, 3728-3740.

Holbrook, A. & Fairbanks, G. (1962), Diphthong formants and their movements, *Journal of Speech and Hearing Research*, 5, 38-58.

Keating, P., Cho, T., Fougeron, C. & Hsu, C. (2003), Domain-initial articulatory strengthening in four languages, in *Phonetic interpretation* (J. Local, R. Ogden, R. Temple, editors), Papers in Laboratory Phonology 6, Cambridge: Cambridge University Press, 143-161.

Simpson, A. (1998), Characterizing the formant movements of German diphthongs in spontaneous speech, in *Computer Linguistik und Phonetik zwischen Sprache und Sprechen* (B. Schröder, W. Lenders, W. Hess & T. Portele, editors), Tagungsband der 4. Konferenz zur Verarbeitung natürlicher Sprache – KONVENS – 98, Frankfurt: Lang, 192-200.

PHONETIC DETAIL IN INTONATION CONTOUR DYNAMICS

Francesco Cangemi
Laboratoire Parole et Langage – Université de Provence
francesco.cangemi@lpl-aix.fr

ABSTRACT

The Autosegmental-Metrical theory of intonation investigates the relationship between f0 contours and post-lexical meaning. Phonetic data are represented in the phonology as a sequence of discrete, local events. The properties of the *transitions* between an event and the next are considered to be phonologically irrelevant.

We present data on Neapolitan Italian which show a significant correlation between the shape of these transitions and the pragmatic context in which a sentence is uttered.¹ This correlation is stronger than the one displayed by traditional autosegmental-metrical indices.

In the conclusions, we discuss the usefulness of our findings as a step towards the fine-tuning of the autosegmental-metrical theory.

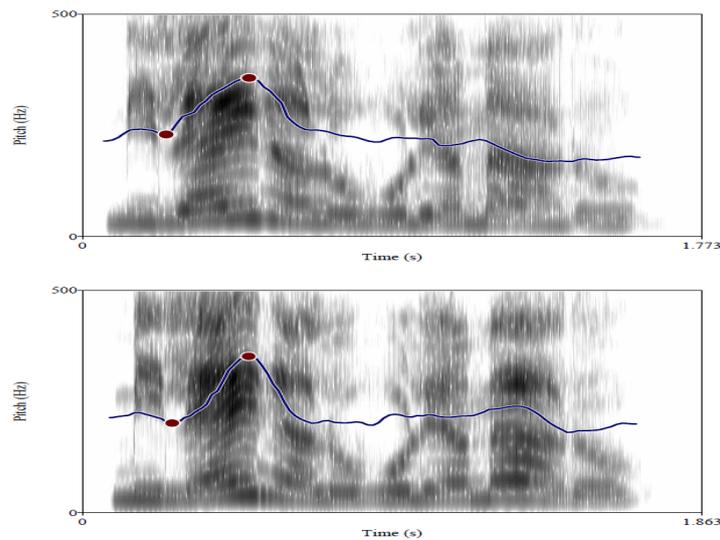


Figure 0: Spectrogram and f0 contour (in blue) for the sentence *Milena lo vuole amaro* uttered as a partial topic statement (top) and as a narrow focus question (bottom)

¹ For instance, figure 0 shows that the transition between the two tonal events (red dots) is concave in the first case and convex in the second.

INTERROGATIVE E ASSERTIVE IN UN CORPUS DIALETTALE RECUPERATO (BOMARZO)

Amedeo De Dominicis
Università degli Studi della Tuscia (Viterbo)
dedomini@unitus.it

SOMMARIO

Il contributo si propone di analizzare un piccolo *corpus* di frasi interrogative e assertive con lo scopo di individuare un possibile descrittore fonologico e linguistico in grado di rendere conto della differenziazione tra i due tipi grammaticali.

Il *corpus* è estratto da un archivio sonoro relativo alla parlata di Bomarzo – un piccolo comune a nord di Viterbo – che fu registrato su cassette magnetiche nel 1979 e poi depositato presso l'audioteca della Provincia di Viterbo, che ne è il legittimo proprietario. Nel 2005 il materiale venne digitalizzato dal Laboratorio di Fonetica dell'Università della Tuscia,¹ sulla base di un accordo con l'ente detentore. L'archivio comprende circa 32 ore di registrazione. Si tratta di materiale eterogeneo (parti parlate si alternano a parti cantate, dialoghi a monologhi) e registrato in condizioni spesso acusticamente infelici (si trovano parti di segnale saturato, rumoroso, affetto da disturbi ambientali e da sovrapposizioni di voci). I dettagli storico-documentari sull'archivio sono illustrati in un precedente lavoro di De Dominicis e Mattana (2009).

L'interesse offerto da questo materiale è duplice. Innanzitutto, è un primo esempio di recupero di un *corpus* vocale secondo gli auspici di progetti come *Calliope* e *Covaid* (cfr. De Dominicis, 2002). Affronta e discute quindi problemi che potrebbero incorrere nell'analisi di qualsiasi altro *corpus* sottoposto a analogo recupero. Un secondo elemento di interesse è costituito dalla relativa profondità cronologica delle registrazioni: dati relativi a una varietà linguistica non ancora descritta e risalenti a trent'anni fa possono rappresentare un piccolo giacimento da cui attingere, sia per l'analisi sincronica che per la eventuale comparazione diacronica.

I risultati dell'indagine mostrano che la tipizzazione ToBI dell'andamento intonativo non consente di differenziare interrogative e assertive in modo consistente. Al contrario, la durata normalizzata di alcuni costituenti della gerarchia metrica appare un buon descrittore linguistico per classificare le due modalità di frase del *corpus*.

¹ Il lavoro di documentazione e riversamento digitale venne portato a termine dalla dottoressa Pamela Mattana.

BIBLIOGRAFIA

De Dominicis, A. (2002), Co.Va.I.D. (**C**Onservazione e **V**Alorizzazione degli archivi vocali dell'Italiano e dei suoi **D**ialetti), *La comunicazione* (G. Rinaldo & R. Piraino, editors), Atti della Conferenza TIPI (Tecnologie Informatiche nella Promozione della lingua Italiana), LI, 97-98.

De Dominicis, A. & Mattana, P. (2009), Il Progetto Bomarzo, in *La Fonetica Sperimentale. Metodo e Applicazioni* (L. Romito, V. Galatà & R. Lio, editors), Atti del 4° Convegno dell'Associazione Italiana di Scienze della Voce, Arcavacata di Rende (CS), 3-5 dicembre 2007, Torriana: EDK Editore, 405-411.

BALBUZIE E COARTICOLAZIONE

Caterina Pisciotta^a, Massimiliano Marchiori^b, Claudio Zmarich^a

^a Istituto di Scienze e Tecnologie della Cognizione (ISTC), C.N.R., Sede di Padova;

^b Psicologo, libero professionista

caterina.pisciotta@pd.istc.cnr.it, marchiori.massimiliano@yahoo.it, claudio.zmarich@pd.istc.cnr.it

SOMMARIO

La coarticolazione, intesa genericamente come l'influenza di un fono su un altro, nei soggetti balbuzienti è stata oggetto di numerosi studi: nel nostro, siamo partiti dalla considerazione che alcune ricerche sulla balbuzie suggeriscono che la coarticolazione nel parlato percettivamente fluente dei balbuzienti può essere diversa da quella riportata per i non balbuzienti – minor coarticolazione negli adulti (cfr. Robb & Blomgren, 1997), ma maggior coarticolazione dei bambini di età prescolare (cfr. Subramanian *et al.*, 2003).

Esiste poi un altro filone di studi fonetici sulla balbuzie che si sofferma sulle influenze prosodiche nel linguaggio dei balbuzienti, che mostrano come essi siano effettivamente in grado di realizzare le differenze prosodiche tra parole focalizzate e non focalizzate, ma aumentano la frequenza delle disfluenze sulle parole in focus (vedi per es. Bergmann, 1986; Marchiori, Zmarich, Avesani & Bernardini, 2005).

Questo lavoro prende le mosse dagli studi di Zmarich & Marchiori (2005) e Marchiori *et al.* (2005). In Zmarich & Marchiori (2005), per quanto riguarda la coarticolazione, le analisi su tutte le sillabe toniche (indipendentemente dalla posizione della sillaba nella parola) mostravano che i balbuzienti, rispetto ai non balbuzienti, hanno un coefficiente di *slope* della regressione lineare maggiore per i valori di $F2_C$ e $F2_V$ per le sillabe atone sotto condizione di 'focus ampio' e 'non in focus' (maggior coarticolazione), ma minore (minor coarticolazione) per le sillabe toniche sotto condizione di 'focus ristretto' ossia coarticolarono di più le sillabe sotto condizione di 'focus ampio' e 'non in focus', ma di meno quelle in condizione di 'focus ristretto'.

Con il presente studio, al contrario, ci siamo proposti di determinare il grado di coarticolazione anticipatoria di V su C solo sulla prima sillaba della parola, pronunciata in modo percettivamente fluente, in due contesti prosodici che favoriscono opposti livelli di coarticolazione, minimo per le sillabe non accentate e non in focus, e massimo per le sillabe accentate in focus ristretto (Zmarich, Avesani & Marchiori, 2006). I risultati evidenziano come, in condizioni massimamente critiche per il balbuziente (richiesta per un grado minimo di coarticolazione, e grado massimo di suscettibilità alle disfluenze; cfr. Marchiori *et al.*, 2005), i parlanti balbuzienti esibiscono un livello di coarticolazione significativamente maggiore rispetto a quello dei parlanti non balbuzienti. Questi risultati confermano lo studio di Subramanian, Yairi & Amir (2003), che esaminarono le transizioni di F2 nel parlato percettivamente fluente di bambini prescolari registrati subito dopo l'inizio della balbuzie.

Questo risultato è però difficilmente interpretabile dato che non si capisce se questo livello eccessivo di coarticolazione è un sintomo diretto della balbuzie, oppure una reazione reattiva, di tipo secondario, ad essa. Questo studio si propone di rispondere a questo interrogativo. Abbiamo pertanto analizzato, attraverso l'analisi acustica, le produzioni disfluenti di un balbuziente: se anche le ripetizioni e i prolungamenti di suono e le ripetizioni

di sillaba fossero risultate caratterizzate da un livello eccessivo di coarticolazione, sarebbe stato facile concludere che essa è allora un sintomo diretto della balbuzie.

Abbiamo analizzato le ripetizioni di sillaba nelle proposizioni disfluenti del soggetto (B) che produceva più disfluenze rispetto agli altri (54 parole disfluenti in totale). Abbiamo quindi misurato con la tecnica del *Locus of Equation* i valori di F2 al rilascio di C e al centro di V nella sillaba ripetuta immediatamente prima della sua produzione fluente (sillaba target). L'equazione della retta di regressione che interpola le coppie di valori F2 C /d/ e F2 V /a, i/ per le sillabe ripetute e per le sillabe target non evidenzia una slope significativamente diversa tra la sillaba disfluente e quella fluente. Ciò significa che il grado di coarticolazione di V su C non è significativamente diverso (e si attesta in entrambi i casi su valori piuttosto bassi). Tuttavia, confrontando tra loro i valori assoluti, abbiamo riscontrato una differenza statisticamente significativa al t-test tra i valori di F2 C delle sillabe ripetute e i valori di F2 C delle sillabe target seguite da /i/. Possiamo cautamente interpretare questi dati dicendo che le sillabe iniziali ripetute in contesto /i/ sono meno coarticolate rispetto alle sillabe target, e che dunque probabilmente la lingua compie un maggior spostamento a maggior velocità passando da C a V.

Inoltre, ipotizzando, come sostenevano già Wingate (1976) e Howell & Au-Yeung (2002), che la sillaba su cui i balbuzienti incontrano difficoltà non sia la prima, ma la seconda, sono stati calcolati, selezionando solo le prime due sillabe atone di quelle parole disfluenti che erano accentate sulla terza sillaba, i valori dei coefficienti ('k' e 'c') della retta di regressione per i 3 tipi di sillaba (ripetizione iniziale, prima sillaba fluente, seconda sillaba fluente) prodotti dal soggetto balbuziente: essi appaiono piuttosto bassi, segnalando una scarsa coarticolazione.

La stessa analisi è stata applicata alle prime due sillabe, atone, delle produzioni fluenti di 'dididi' e 'dadada' dei 4 soggetti normali e 4 soggetti balbuzienti. Il confronto tra i valori di *slope* delle sillabe delle parole disfluenti prodotte dal soggetto balbuziente B e i valori di *slope* delle sillabe delle parole fluenti prodotte dagli altri soggetti, normali e balbuzienti, evidenzia come le sillabe prodotte dal soggetto balbuziente B siano caratterizzate da un grado di coarticolazione anormalmente basso. In conclusione, pensando che queste sillabe appartengono a parole disfluenti, si conclude che il grado di coarticolazione significativamente maggiore di quello dei non balbuzienti evidenziato dai quattro soggetti balbuzienti nella produzione delle sillabe iniziali accentate in focus contrastivo possa attribuirsi a una strategia di reazione e compensazione che il soggetto balbuziente attua per tenere sotto controllo la propria balbuzie e riuscire fluente.

BIBLIOGRAFIA

Si rimanda alla bibliografia in calce all'articolo contenuto nel CD-ROM

CANTO

OSSERVAZIONI PRELIMINARI SUGLI ASSETTI INTERVALLARI NEL CANTO A *MUTETUS* DELLA SARDEGNA MERIDIONALE

Paolo Bravi
Conservatorio di Musica 'G. P. da Palestrina', Cagliari
pa.bravi@tiscali.it

SOMMARIO

La ricerca sui caratteri delle scale musicali, oltre ad essere stata uno degli aspetti più rilevanti dell'analisi in ambito etnomusicologico fin dalle origini della disciplina, è un elemento apparentemente ineludibile nella definizione dei caratteri di un 'sistema musicale'. È stato frequentemente osservato, infatti, che le strutture intervallari relative a stili di canto non assimilabili ai generi della musica 'classica' (o di sua derivazione) diffusi in ambito extraeuropeo e folklorico hanno caratteristiche diverse rispetto al sistema euroculto, in particolare per quanto riguarda l'ampiezza degli intervalli, raramente assimilabile a quelli della scala temperata che da circa tre secoli costituisce il sistema di riferimento in ambito occidentale.

Il presente lavoro prende in esame uno stile di canto utilizzato dai poeti improvvisatori dell'area meridionale della Sardegna per l'esecuzione di componimenti nella forma metrica del *mutetu*. Il canto *a mutetus* viene praticato in maniera prevalente da poeti improvvisatori (*cantadoris*) semi-professionisti, principalmente in occasione di gare poetiche organizzate in concomitanza con feste patronali. Le esecuzioni si svolgono in buona parte in forma solistica, mentre il passaggio dall'una all'altra sezione della forma metrico-musicale è segnato dagli interventi di un coro bivocale composto dalla voce gutturale del *basciu* (con emissione simile a quella del *bassu* dei cori a tenore diffusi nell'area centrale della Sardegna) e dalla voce – anch'essa talvolta in parte gutturale – della *contra*, cui in genere si sovrappone la voce del poeta improvvisatore.

Nel caso del canto *a mutetus*, pur essendoci una terminologia specifica relativamente articolata, non esiste una 'etnoteoria' sistematica che permetta di inquadrare i dati rilevabili all'ascolto relativi agli assetti intervallari di questo stile di canto nel quadro di una teoria emica. Per questo motivo la ricerca si è svolta unicamente in forma sperimentale, ed è stata basata sulla metodologia di analisi inaugurata, agli inizi degli anni Settanta, da Tjernlund, Sundberg e Fransson. Il presupposto di base di tale metodologia è che le frequenze prevalenti in una esecuzioni siano il correlato acustico della scala in uso, e che perciò l'analisi di diagrammi di frequenza relativi alla distribuzione del pitch possa fornire indicazioni essenziali sugli assetti intervallari del canto. Il metodo da un lato presenta il vantaggio di offrire una valutazione di sintesi oggettiva sui comportamenti intonativi dei cantori, dall'altro presenta limiti legati al fatto che la distribuzione delle frequenze non permette di interpretare i fenomeni di dispersione e di distinguere aspetti particolari ma assai comuni nell'intonazione cantata quali, ad es. il vibrato o il portamento.

Osservati attraverso questa metodologia, gli assetti intervallari del canto *a mutetus* evidenziano caratteristiche significativamente diverse rispetto agli intervalli definiti dal sistema temperato. Non si tratta, peraltro, di connotazioni stabili e assolute, ma di linee di tendenza che, seppur non costanti, sono condivise da diversi cantori dei nostri giorni e del recente passato.

PERCEZIONE E APPRENDIMENTO

FUNCTIONS OF THE LEFT AND RIGHT POSTERIOR TEMPORAL LOBE DURING SEGMENTAL AND SUPRASEGMENTAL SPEECH PERCEPTION

Cyrrill Ott, Martin Meyer

Institute of Neuropsychology, University of Zurich
c.ott@psychologie.uzh.ch, m.meyer@psychologie.uzh.ch

ABSTRACT

This manuscript reviews evidence from neuroimaging studies on elementary processes of speech perception and their implications for our understanding of the brain-speech relationship. Essentially, differential preferences of the left and right auditory-related cortex for rapidly and slowly changing acoustic cues that constitute (sub)segmental and supra-segmental parameters, e.g. phonemes, prosody, and rhythm. The adopted parameter-based research approach takes the early stages of speech perception as being of fundamental relevance for simple as well as complex language functions. The current state of knowledge necessitates an extensive revision of the classical neurologically oriented model of language processing that was aimed at identifying the neural correlates of linguistic components (e.g. phonology, syntax and semantics) more than at substantiating the importance of (supra)segmental information during speech perception.

**PHONETIC CONTRASTS
IN FOREIGN LANGUAGE PERCEPTION:
A NEUROPSYCHOLOGICAL STUDY ON SERBIAN AFFRICATES**

Nuria Kaufmann ^a, Martin Meyer ^a, Stephan Schmid ^b

^a Institute of Psychology (Division of Neuropsychology), University of Zurich

^b Phonetics Laboratory, University Zurich

nuria_kaufmann@access.uzh.ch, m.meyer@psychologie.uzh.ch, schmidst@pholab.uzh.ch

ABSTRACT

This study addresses the question to which extent phonetic contrasts of a foreign language are perceived more easily by speakers of a native language that shares similar phonetic categories. The focus lies on two postalveolar and two alveolo-palatal affricates of Serbian: [tʃ] (postalveolar, voiceless), [tɕ] (alveolo-palatal, voiceless), [dʒ] (postalveolar, voiced) and [dʒ̥] (alveolo-palatal, voiced). Swiss-German dialects have the postalveolar voiceless affricate [tʃ] only, while the Rhaeto-Romance variety of Sursilvan has three different affricates, i.e. [tʃ], [tɕ], and [dʒ̥].

In a EEG experiment using a Multi-Deviant Mismatch Negativity (MMN) paradigm, 15 Swiss-German speaking adults and 15 Rhaeto-Romance speaking adults between the ages of 20 to 30 years were instructed to focus on a random reading while not paying attention to the auditory stimuli. The hypothesis is a significant difference in processing between the two groups: Swiss-German speakers will not be able to reliably distinguish the four Serbian affricates. Rhaeto-Romance speakers on the other hand are expected to be able to distinguish all four affricates as they share three of the four phonetic categories.

A significant group-effect was found to corroborate that Rhaeto-Romance speakers process the Serbian affricates differently from the Swiss-German speakers.

DOES A TALKER'S OWN RATE OF SPEECH AFFECT HIS/HER PERCEPTION OF OTHERS' SPEECH RATE?

Sandra Schwab
Université de Genève
sandra.schwab@unige.ch

ABSTRACT

Many studies have investigated the factors that may affect the perception of speech rate (e.g. Grosjean & Lane, 1976; Feldstein & Bond, 1981; Kohler, 1986; Greene, 1987; Crown & Feldstein, 1991), but very few studies have examined the role that the talker's own rate might play in his/her perception of others' rate. Among them, Lass & Cain (1972) investigated the hypothesis that a speaker's preferred rate depended on his actual rate. They indeed showed that speakers who produced slow rates preferred listening to slow rates, whereas fast speakers tended to prefer fast rates. This conclusion raises the question whether speech rate production affects not only speech rate preference, but also speech rate perception. To our knowledge, very few studies have tried to answer this question. Gósy (1991) formulated the hypothesis that "the speaker's own speech tempo determines his judgements concerning that of other people: the faster his own speech the less fast he perceives that of others" (p. 101). Gósy showed that speakers with different speech rates (very slow, slow, moderate, fast, very fast) did not perceive speech rate in a similar way. In the same direction, Koreman (2006) hypothesized that listeners' own speaking habits may affect their perception of speech rate. Nevertheless, his results failed to show an effect of the listener's rate on his/her perception of rate.

Considering the lack of totally conclusive results on the role that the talker's rate might play in rate perception, the objective of this research is to explore more deeply the hypothesis that speakers with different speech rates do not perceive speech rate in a similar way. To this end, we conducted a perception experiment.

In this experiment, participants were asked to listen to and estimate various samples at different speech rates (normal, fast and slow), using a magnitude-estimation task (Stevens, 1957). Results firstly showed a negative correlation between rate estimation and own rate at normal and slow rates (respectively, $r = -0.45$, $r = -0.39$, $p < 0.05$), but no correlation at fast rate ($r = -0.11$, ns): speakers with fast speech rate tended to under-estimate the sample speech rates (i.e. to give a lower numeric estimation) in comparison with slow speakers (at normal and slow rates). Secondly, and more interestingly, a regression analysis revealed that the own rate has a moderator effect on rate estimation, at all rates (normal: $t(781) = -5.67$, $p < 0.001$; fast: $t(781) = -2.06$, $p < 0.05$; slow: $t(781) = -6.46$, $p < 0.001$): the faster a listener speaks, the less his/her rate estimations raise as a function of heard rates, especially at normal and slow rates.

In sum, the present research has shown that the talker's rate plays a role in rate perception: fast speakers not only tend to under-estimate speech rate in comparison with slow speakers, but they are also less sensitive to rate changes.

REFERENCES

- Crown, C.L. & Feldstein, S. (1991), The perception of speech rate from the sound-silence patterns of monologues, *Journal of Psycholinguistic Research*, 20, 47-63.
- Feldstein, S. & Bond, R. (1981), Perception of speech rate as a function of vocal intensity and frequency, *Language and Speech*, 24, 387-395.
- Gósy, M. (1991), The perception of tempo, in *Temporal Factors in Speech. A collection of Papers* (M. Gósy, editor), Budapest: Research Institute for Linguistics, HAS, 63-107.
- Green, K.P. (1987), The perception of speaking rate using visual information from a talker's face, *Perception & Psychophysics*, 42, 587-593.
- Grosjean, F. & Lane, H. (1976), How the listener integrates the components of speaking rate, *Journal of Experimental Psychology: Human Perception and Performance*, 2, 538-543.
- Kohler, K.J. (1986), Parameters of speech rate perception in German words and sentences: Duration, F0 movement, and F0 level, *Language and Speech*, 49, 115-139.
- Koreman, J. (2006), Perceived speech rate: The effect of articulation rate and speaking style in spontaneous speech, *Journal of the Acoustical Society of America*, 119, 582-596.
- Lass, N.J. & Cain, C.J. (1972), A correlational study of listening rate preferences and listeners' oral reading rates, *Journal of Auditory Research*, 12, 308-312.
- Stevens, S.S. (1957), On the psychophysical law, *Psychological Review*, 64, 153-181.

CROSS-LANGUAGE SPEECH PERCEPTION: LEXICAL STRESS IN SPANISH WITH ITALIAN AND FRANCOPHONE SUBJECTS¹

Iolanda Alfano ^a, Sandra Schwab ^b, Renata Savy ^c, Joaquim Llisterri ^a
^a Universitat Autònoma de Barcelona; ^b Université de Genève; ^c Università di Salerno
Iolanda.Alfano@campus.uab.cat, sandra.schwab@unige.ch, rsavy@unisa.it,
Joaquim.Llisterri@uab.cat

ABSTRACT

The present work analyses the perception of lexical stress in Spanish by Italian and French native speakers, trying to take into account the differences between Italian, French and Spanish stress systems both in perception and in production.

We have designed two perception experiments using a similar procedure of other works with native subjects (Llisterri *et al.*, 2005; Alfano, 2006) and non-native subjects (Alfano *et al.*, 2007; Alfano *et al.*, 2009). The corpus consisted of couples and triplets of meaningful three syllable words and meaningless three syllable words (pseudo-words), with three different possible stress patterns: proparoxytone (PP), paroxytone (P) and oxytone (O). It was analyzed using the Praat software (Boersma & Weenink, 2003); for each of the three vowels of the words, we measured fundamental frequency (f_0) and vowel duration (D). F_0 and D values of the original words were systematically manipulated and a resynthesis was performed to create the stimuli used in the test, so that the role of these cues in perception could be studied. The test stimuli were created in the following way: in proparoxytone words, f_0 and duration values for each vowel were replaced by the corresponding f_0 and duration values found in the equivalent paroxytone words (PP>P); in the same way, in P words, f_0 and duration values for each vowel were replaced by the corresponding f_0 and duration values found in the equivalent oxytone words (P>O).

Three groups of Italian subjects and two groups of Francophone subjects, divided according to their competence in Spanish, have participated in the experiments performing an identification test.

Italian subjects perceive correctly the stressed syllable in almost 100% of the cases for the original PP and P, but make a mistake in more than 15% of the cases for the original O. The isolated manipulation of D or f_0 does not trigger a change in stress pattern perception. When f_0 and duration values are simultaneously modified, subjects perceive a change in stress location in a high percentage of cases for PP>P (between 56,7% and 90%) but in less than 40% for P>O.

Francophone subjects recognize the stressed syllable in original stimuli in more than 70% of the cases, obtaining the best results in PP perception and the worst performance for O. Such as the previous case with Italian speakers, the isolated manipulation of one of the two acoustic cues does not produce a clear change in stress pattern perception, while the combined manipulation produces a change in stress perception up to 77% of the cases for PP>P and in percentages ranging from 32 to 44% of the cases for P>O.

¹ The work is a result of the collaboration among the authors; nevertheless, we owe to Iolanda Alfano the paragraphs 1 and 2, to Sandra Schwab the paragraph 4, to Renata Savy the paragraph 3 and to Joaquim Llisterri the paragraphs 5 and 6.

Our results indicate that native language influence is not sufficient to explain the perception process in a foreign language, suggesting that non-native subjects use not only purely 'linguistic' perception strategies, since their choices seem to be also determined by psycholinguistic factors and by acoustic properties of the signal.

BIBLIOGRAFIA

Alfano, I. (2006), La percezione dell'accento lessicale: un test sull'italiano a confronto con lo spagnolo, in *Analisi prosodica. Teorie, modelli e sistemi di annotazione*, Atti del 2° Convegno Nazionale dell'Associazione Italiana di Scienze della Voce, Salerno, 30 novembre-2 dicembre 2005 (R. Savy & C. Crocco, editors), Torriana: EDK Editore, 632-656.

Alfano, I., Llisterri, J. & Savy, R. (2007), The perception of Italian and Spanish lexical stress: A first cross-linguistic study, in *Proceedings of the 16th International Congress of Phonetic Sciences*, Saarbrücken, Germany, August 6-10, 2007 (J. Trouvain & W.J. Barry, editors), 1793-1796.

Alfano, I., Savy, R. & Llisterri, J. (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* (L. Romito, V. Galatà & R. Lio, editors), Atti del 4° Convegno Nazionale dell'Associazione Italiana di Scienze della Voce, Arcavacata di Rende (CS), 3-5 Dicembre 2007, Torriana: EDK Editore, 22-39.

Boersma, P. & Weenink, D. (2003), *Praat: doing phonetics by computer* (V. 4.0.4), <http://www.praat.org/>

Llisterri, J., Machuca, M., de la Mota, C., Riera, M. & Ríos, A. (2005), La percepción del acento léxico en español, *Filología y lingüística* 1 (Madrid: CSIC-UNED-U. de Valladolid), 271-297.

PERSISTENZA DELL'ACCENTO STRANIERO. UNO STUDIO PERCETTIVO SULL'ITALIANO L2

Giovanna Marotta ^a, Philippe Boula de Mareuil ^b

^a Dipartimento di Linguistica, Università di Pisa, ^b LIMSI-CNRS, Orsay
gmarotta@ling.unipi.it, mareuil@limsi.fr

SOMMARIO

La ricerca fonologica sull'acquisizione di L2 si è finora concentrata sul versante della produzione, trascurando quello della percezione, nonostante sia da tempo nota la rilevanza dei processi percettivi anche nella resa fonetica dei segmenti; in particolare, risulta ancora poco indagata la tematica relativa alla percezione del cosiddetto 'accento straniero'.

All'interno di questa area di ricerca, un problema specifico concerne il peso dei tratti fonetici sulla percezione del *foreign accent*. Presentiamo qui i risultati di un test percettivo in cui alcuni frammenti di parlato italiano prodotto da parlanti con diversa L1 (francese, spagnolo, tedesco, inglese) sono stati valutati da parlanti nativi italiani.

I soggetti sono stati chiamati ad ascoltare gli stimoli acustici naturali, uno per volta, e a giudicare se il parlante era italiano o straniero; se valutato come straniero, i soggetti dovevano indicare la lingua madre del parlante tra le quattro lingue sopra elencate, valutando anche il grado di accento straniero su una scala a tre gradini, che va da 0 (poco accento) a 2 (accento forte). Ogni ascoltatore italiano è stato preliminarmente invitato ad autovalutare sia la sua competenza nelle quattro lingue straniere, che il suo grado di familiarità con l'accento delle stesse lingue.

I risultati mostrano che nella maggioranza dei casi gli ascoltatori sono in grado di percepire la differenza tra parlanti italiani nativi e parlanti non nativi, anche in caso di ottima competenza dell'italiano. Più complesso si è invece rilevato il compito relativo all'identificazione della lingua materna dei parlanti. Soltanto gli stimoli prodotti da parlanti inglesi sono stati identificati con una percentuale di riconoscimento soddisfacente, mentre quelli relativi a parlanti spagnoli presentano i valori di corretto riconoscimento più bassi. Inoltre, gli stimoli prodotti da parlanti tedeschi sono stati spesso confusi con quelli relativi ai parlanti inglesi.

Il grado di successo nel riconoscimento della L1 appare dunque inversamente proporzionale alla vicinanza strutturale e fonologica tra L1 e L2: italiano e spagnolo sono discriminati con difficoltà, mentre il parlato dei tedeschi tende ad essere confuso con quello degli inglesi più che con quello degli spagnoli.

Tuttavia, dai nostri dati non risulta una buona corrispondenza tra l'autovalutazione dell'ascoltatore e la sua *performance* nel test percettivo. In maniera abbastanza prevedibile, soltanto nel caso dell'inglese si osservano valori comparabili tra autovalutazione e percezione, mentre per le altre lingue straniere si rileva una discrasia più o meno marcata tra il supposto livello di familiarità con un accento straniero e la corretta identificazione della lingua straniera nel test sperimentale. In altri termini, la percezione dell'accento straniero può esser indipendente dal corretto riconoscimento della lingua materna parlata da colui che è stato identificato come straniero.

PERCEZIONE E PRODUZIONE DEI FONEMI DELL'INGLESE AMERICANO IN PARLANTI CON UN SISTEMA PENTAVOCALICO

Bianca Sisinni, Mirko Grimaldi

Centro di Ricerca Interdisciplinare sul Linguaggio (CRIL) – Università del Salento
bianca.sisinni@unisalento.it, mirko.grimaldi@unisalento.it

SOMMARIO

Con questo lavoro ci proponiamo di studiare i processi di produzione e percezione durante l'acquisizione dei fonemi vocalici dell'Inglese Americano (AE) in un gruppo di studenti universitari della Facoltà di Lingue dell'Università del Salento (parlanti nativi dell'Italiano Salentino (IS) con un sistema a 5 vocali e tre gradi di apertura.

In letteratura è ancora dibattuta la questione se lo sviluppo percettivo della L2 preceda lo sviluppo della produzione, oppure se una idonea percezione non sia condizione necessaria per una corretta produzione (revisioni in Listerrri, 1995; Leather, 1999; Escudero, 2005; Hansen Edwards-Zampini, 2008). L'idea prevalente ritiene che per spiegare appieno l'acquisizione fonologica di una L2 bisogna prima spiegare il modo in cui i parlanti della L2 riescono a sviluppare una percezione appropriata e quindi una rappresentazione cognitiva dei segmenti della L2: la produzione corretta sarebbe una diretta conseguenza del corretta rappresentazione astratta.

In quest'ottica abbiamo incrociato gli strumenti di analisi dello *Speech Learning Model* (SLM) di Flege (1995) e del *Perceptual Assimilation Model* (PAM) di Best (1995), anche per capire se gli studenti di Lingue risultano assimilabili alla categoria dei *naïve L2 listeners* oppure a quella dei *fluent L2 listeners* (Best & Tyler, 2006).

Sono state prese in esame 18 studentesse universitarie (SU) – età media 20,4 anni, con alle spalle una media di 10 anni di studio scolastico dell'inglese anche con insegnanti madrelingua – al secondo mese del corso di inglese tenuto da un lettore madrelingua USA.

Delle SU, tramite lettura di frasi-cornice (Dico 'CVCCV adesso, sei ripetizioni per fonema), sono state rilevate le aree di esistenza (proiettate su assi cartesiani) della L1. Allo stesso modo sono state rilevate le aree di esistenza di tre parlanti nativi (NS) di sesso femminile dell'Inglese Americano (AE; frase cornice: I say 'CVC now), come rappresentato nella Fig. 1. Si è quindi proceduto a rilevare 11 fonemi dell'AE (/i:/, /ɪ/, /e/, /æ/, /ɜ:/, /ʌ/, /ɑ:/, /ɔ:/, /ʊ/, /u:/: Ladefoged, 2005) prodotti dalle SU, utilizzando il metodo della *delayed repetition* (Flege *et al.*, 2003; Piske *et al.*, 2001), in cui un NS produce una frase cornice che le SU devono ripetere dopo l'ascolto di elementi distrattori e *interventing speech material* (garantendo così una ripetizione del fonema bersaglio distanziata dall'ascolto del fonema nativo), come in (1):

- (1) NS: I say 'CVC now (FRASE-CORNICE).
 Could you repeat, please? (DISTRATTORE)
 (PAUSA – BIP – PAUSA – BIP) (DISTRATTORI)
 SU: Of course I could (INTERV. SPEECH MAT)
 I say 'CVC now (FRASE-CORNICE)

F0, struttura delle frasi-cornice di L1 ed L2, contesti consonantici, qualità delle sillabe sono stati mantenuti il più possibile costanti, per un totale di 54 stimoli per il corpus di IS e 66 stimoli per il corpus di AE. Le registrazioni sono state effettuate con il sistema CSL 4500 (frequenza di campionamento di 22.05 KHz). La segmentazione è stata eseguita con *Praat* 4.6.29, calcolando per ogni fonema la durata totale (in secondi) e le F0, F1, F2 e F3 nel punto medio (0,02 s).

La capacità di percezione è stata invece testata attraverso la somministrazione di due classici test percettivi: l'*identification test* (IT) e l'*oddity discrimination test* (ODT) (Flege & MacKay, 2004; Tsukada *et al.*, 2005). Gli stimoli per questi test sono stati ricavati dalle produzioni delle NS per un totale di 36 stimoli uditivi. Le parole sono state normalizzate in intensità e segmentate con *Praat* 4.6.29. Le SU sono state testate singolarmente in una stanza insonorizzata e gli stimoli sono stati somministrati tramite PC. L'ODT è stato fatto eseguire anche ad un gruppo di 10 parlanti native dell'AE per ottenere un gruppo di controllo.

Il quadro che emerge è abbastanza complesso, ma le analisi statistiche condotte sui dati acustici ci autorizzano ad ipotizzare che a livello produttivo i soggetti analizzati si trovano in una fase di interlingua in cui l'apprendimento fonologico della L2 risulta ancora mediato dalle caratteristiche implicite nella L1. Ciò tuttavia non ha impedito che fonemi come /u/ ed /ʌ/ siano stati realizzati con valori formantici vicini a quelli dei NS. A livello percettivo, i risultati ottenuti nell'ODT evidenziano come non tutti i contrasti di L2 hanno generato la stessa difficoltà nell'essere discriminati da parte delle SU, le quali hanno raggiunto differenti livelli di accuratezza nella loro discriminazione.

I risultati dei due test percettivi sono stati estremamente funzionali per se il PAM si possa applicare anche alla categoria 'studenti universitari'. Il PAM è un modello percettivo il cui scopo è quello di studiare la discriminazione di contrasti non nativi da parte di una tipologia ben definita di soggetti, i *naïve listeners*, ovvero parlanti monolingue totalmente estranei alla lingua target. In base al modo di assimilazione dei fonemi non nativi ai fonemi nativi, il PAM prevede sei diversi tipi di assimilazione di contrasti non nativi ai fonemi nativi, per ciascuno dei quali è attesa, da parte dei *naïve listeners*, una determinata capacità di discriminazione. In questo lavoro, le modalità di assimilazione dei fonemi dell'AE ai fonemi dell'IS sono state ricavate dai risultati dell'IT mentre la capacità di discriminazione dei contrasti individuati è stata testata nell'ODT. I risultati così ottenuti hanno dimostrato, nel complesso, che le SU hanno discriminato i vari contrasti di L2 in conformità con quanto predetto dal PAM. Conseguentemente, sembra possibile affermare che il framework del PAM può essere applicato alle SU che, percettivamente, si comportano come *naïve listeners* nonostante un lungo background di L2, sebbene esclusivamente scolastico.

Infine l'incrocio dei dati in produzione con quelli in percezione, in particolare con i dati ottenuti nell'IT, ci ha portato ad elaborare una ipotesi che vede la comparazione acustico-percettiva sistematica fra gli spazi acustici della L1 con quelli della L2 come un momento importante dell'individuazione di salienze percettive fra contrasti fonologici. Da qui l'apprendente partirebbe per costruire l'abilità a discriminare i suoni della L2 e ad estrarre dal segnale acustico tratti articolatori invariati (i tratti distintivi) per generare una rappresentazione mentale astratta di un set di categorie fonetiche prima e fonologiche poi. La fase finale di questo processo porterà gli apprendenti a proiettare adeguatamente tali rappresentazioni motorie astratte sull'apparato fonatorio nella produzione dei suoni della L2.

LA DIMENSIONE TEMPORALE IN TRE TIPI DI PARLATO: UN CONFRONTO TRA ARABO E ITALIANO

Dalia Gamal
Università di Ain Shams (Il Cairo)
daliagamal60@hotmail.com

SOMMARIO

Il presente lavoro ha lo scopo di condurre una prima analisi di alcuni fenomeni temporali in arabo e italiano in tre stili diversi: parlato letto di giornali e telegiornali, parlato semispontaneo di dialoghi *Map Task* e parlato meno controllato di interviste televisive.

In un lavoro precedente (Gamal, 2007) sono stati analizzati alcuni fenomeni ritmici che possano determinare il tipo di isocronia dell'italiano L2 per indagare l'eventuale influsso della lingua madre degli informatori (l'arabo) sulla produzione nella lingua straniera. Dal punto di vista degli studi sull'isocronia l'arabo si considera una lingua ad isocronia accentuale, mentre l'italiano viene classificato come una lingua ad isocronia sillabica. Il materiale analizzato constava di turni dialogici elicitati tramite il metodo *Map Task*. Non è stata rilevata una chiara influenza della lingua prima sulle produzioni in italiano degli apprendenti, che hanno dimostrato di rispettare la struttura sillabica italiana; inoltre, è stata riscontrata nelle due lingue una notevole differenza nelle strategie di riduzione in funzione della variazione della velocità d'eloquio. Per esempio, in merito alla riduzione delle sillabe le pretoniche hanno manifestato una maggiore riduzione in arabo, mentre sono state le postoniche e le toniche a ridursi di più in italiano; e, contrariamente alle aspettative, le sillabe collocate tra due toniche si sono mostrate le meno ridotte. All'interno della sillaba le consonanti sono state manipolate diversamente rispetto alle vocali: la conservazione dell'estensione temporale delle sillabe nelle velocità di eloquio maggiori si realizzava in arabo grazie alle vocali, che resistevano maggiormente agli accorciamenti, e in italiano, invece, per via delle consonanti che a volte aumentavano di durata a scapito delle vocali, generalmente più ridotte tranne nelle toniche e le postoniche.

In questa sede, dunque, per un ulteriore approfondimento di questo inatteso distacco dalla L1 si ricorre all'esame della distribuzione del tempo tra sequenze articolate e pause in diversi stili di parlato. Qui la durata è sempre il parametro cardine nell'indagine. Come indici dell'organizzazione ritmica vengono considerate le durate delle sillabe, delle catene articolate, delle pause piene e vuote, così come vengono osservate le strategie di raggruppamento delle parole in gruppi di respiro, in un materiale non controllato a livello di struttura sintattica.

La presente indagine ha inoltre lo scopo di abbinare nel corpus materiale controllato e brani di parlato progettato *on-line* allo scopo di rilevare eventuali differenze dovute alla natura della situazione di produzione. Alla ricerca di ottenere una maggiore omogeneità tra il materiale in italiano L1 e L2 sono state registrate varie letture di uno stesso paragrafo di un editoriale italiano con parlanti nativi e apprendenti egiziani; fanno parte del materiale letto italiano altri due brani del telegiornale prodotti dagli apprendenti. Inoltre, il parlato semispontaneo è stato elicitato con l'ausilio delle stesse mappe italiane, precedentemente progettate e utilizzate nella raccolta del *corpus* AVIP-API. I brani tratti dal TG e dal giornale arabo sono stati registrati con gli stessi informatori del materiale in italiano L2.

Nel *corpus* arabo si aggiungono al parlato letto alcune stringhe di recitazioni del Corano, perché il ritmo del Corano è il modello base della ritmicità, che, anche se non si riscontra in pieno nella lettura moderna, fa sempre parte del sottofondo prosodico degli abitanti dei paesi arabi e islamici. Nella recitazione coranica le durate vocaliche seguono un'organizzazione particolare, in quanto sono stabilite in modo sistematico a seconda del contesto segmentale successivo alla vocale. Oltre alla presenza di vocali brevi e lunghe, queste ultime vengono raddoppiate di durata prima della oclusiva glottidale, che in arabo rappresenta un fonema, e triplicate prima delle consonanti doppie. Tali regole di 'allungamento' ci offrono dunque tre classi temporali da aggiungere alla quarta delle vocali brevi. In arabo, va notato, è stato sempre riconosciuto il valore delle vocali e delle loro durate nella creazione della regolarità ritmica del parlato e nella distinzione tra le parlate delle varie tribù, nonché nella differenziazione tra varie situazioni comunicative, come si osserva, tra gli altri, nello studio di Ibn-Jinni che risale all'undicesimo secolo d.C.

Il parlato dei mass media arabi, inoltre, si presenta all'ascolto più pacato rispetto a quello italiano. Infatti, dalle analisi dei brani del telegiornale nelle due lingue emerge che la durata e la frequenza delle pause costituiscono gli indici determinanti, ancora più della velocità d'eloquio.

Il passaggio nell'analisi dall'arabo recitato – del Corano e quello moderno del giornale e del tg – alla varietà parlata al Cairo ci può fornire un'immagine più chiara dello 'sfondo ritmico' dei parlanti.

Il materiale di parlato arabo spontaneo, malgrado sia ridotto, offre dati interessanti sull'uso 'mirato' delle pause da parte del parlante che riesce tramite la collocazione delle pause e la manipolazione della f_0 a mantenere la parola per spiegare le sue idee.

Riguardo agli apprendenti di italiano i risultati ottenuti nel parlato letto rivelano che, rispetto ai nativi e a parità di testo, il tempo totale richiesto in L2, a un livello avanzato di conoscenza linguistica, resta più esteso, ma si rivela anche sfruttato e organizzato. L'occorrenza dei silenzi è generalmente simile in nativi e apprendenti e la collocazione delle pause è molto funzionale alla trasmissione del senso e al superamento delle difficoltà lessicali e coincide in molti casi con le posizioni scelte dai nativi. Perciò si può affermare che gli apprendenti leggono i brani scritti e rispondono al compito della mappa con grande sensibilità alla natura della situazione in cui si trovano e di conseguenza la loro prosodia risulta varia e ricca.

BIBLIOGRAFIA

Gamal, D. (2007), Sul ritmo in italiano L2: l'eventuale isocronia, in *Scienze vocali e del linguaggio. Metodologie di valutazione e risorse linguistiche* (V. Giordani, V. Bruseghini & P. Cosi, editors), Atti del 3° Convegno Nazionale dell'Associazione Italiana di Scienze della Voce, 29 novembre – 1 dicembre 2006, Povo (Trento), Torriana (RN): EDK editore, 101-118.

Ibn-Jinni, A.-F. (1999⁴), *صناعات* [Le caratteristiche], Cairo: Organizzazione Egiziana del Libro, 3° volume.

TECNOLOGIE DEL PARLATO

ALCUNE CONSIDERAZIONI SULL'IMPORTANZA DEGLI ASPETTI DINAMICI NELLA PERCEZIONE, PRODUZIONE ED ELABORAZIONE DEL PARLATO

Piero Cosi

Istituto di Scienze e Tecnologie della Cognizione
Sede di Padova 'Fonetica e Dialettologia'
Consiglio Nazionale delle Ricerche
e-mail: piero.cosi@pd.istc.cnr.it

SOMMARIO

La dimensione temporale è un elemento costitutivo non solo dei meccanismi di produzione del parlato, intervenendo, a livello segmentale, nella determinazione delle durate e nella pianificazione e nel controllo di tutti i gesti articolatori e, a livello sopra-segmentale, nell'allineamento dei contorni intonativi con le parti dell'enunciato, ma anche, nella percezione del segnale verbale e, più in generale, nell'interpretazione di un qualsiasi atto comunicativo.

Inoltre, sia la configurazione delle caratteristiche facciali che la sincronizzazione delle azioni facciali sono importanti nell'espressione e nel riconoscimento delle emozioni (Cohn, 2007). La configurazione delle azioni facciali (espressioni relative sia a specifiche emozioni sia ad unità di azione individuali) rispetto alle emozioni ed all'intenzione comunicativa è stata un importante tema di ricerca. Meno invece si conosce circa la sincronizzazione delle azioni facciali, anche perché la misurazione manuale della sincronizzazione è assai complicata e laboriosa. Tuttavia, sappiamo che (Cohn, 2007) siamo altamente sensibili alla sincronizzazione delle azioni facciali nelle interazioni sociali (Edwards, 1998). Le azioni facciali più lente, ad esempio, sembrano essere più genuine e naturali (Krumhuber & Kappas, 2005), come pure lo sembrano essere quelle più sincrone nei loro movimenti (Frank & Ekman, 1997). In particolare, le espressioni facciali più sottili diventano visibili soltanto quando le informazioni di movimento sono a disposizione di chi le percepisce (Ambadar, Schooler & Cohn, 2005).

La dinamica è cioè particolarmente importante per inferire l'intenzione comunicativa. Alcuni studi condotti dal gruppo di ricerca CMU/Pitt utilizzando tecniche automatiche di analisi di immagini facciali per misurare la sincronizzazione delle azioni facciali, hanno provato che le caratteristiche dinamiche riescono a discriminare fra i sorrisi intenzionali e quelli spontanei con un livello di precisione dell'89% (Cohn & Schmidt, 2004). Usando caratteristiche simili, il divertimento, l'imbarazzo ed il sorriso 'gentile' sono stati discriminati con una precisione dell'83% (Kanade, Hu & Cohn, 2005), che è paragonabile a quella umana. Lavori più recenti suggeriscono inoltre che la coordinazione multimodale dell'espressione facciale, del movimento della testa e dei gesti sono caratteristiche specifiche dell'imbarazzo (Keltner, 1995).

In questo lavoro vengono illustrati alcuni dei più significativi apporti tecnologici che nel corso degli ultimi anni sono stati influenzati dalla dimensione temporale del parlato nel campo dell'analisi del segnale vocale, della sintesi della voce da testo scritto, del riconoscimento automatico del segnale verbale (Furui, 2005). Per quanto riguarda la realizzazione di facce parlanti animate, saranno discussi poi alcuni esempi dell'influenza degli aspetti

dinamici nella percezione e nell'interpretazione delle espressioni facciali e più in generale degli intenti comunicativi, nella trasmissione di emozioni, stati d'animo e atteggiamenti, nell'interazione faccia a faccia (Cohn, 2007).

RIFERIMENTI BIBLIOGRAFICI

Ambadar, Z., Schooler, J. & Cohn, J.F. (2005), Deciphering the enigmatic face: The importance of facial dynamics to interpreting subtle facial expressions, *Psychological Science*, 16, 403-410.

Cohn, J.F. & Schmidt, K.L. (2004), The timing of facial motion in posed and spontaneous smiles, *International Journal of Wavelets, Multiresolution and Information Processing*, 2, 1-12.

Cohn, J.F. (2007), Foundations of human-centered computing: Facial expression and emotion, *Proceedings of the International Joint Conference on Artificial Intelligence 2007*, Hyderabad, India, 5-12.

Edwards, K. (1998), The face of time: Temporal cues in facial expressions of emotion, *Psychological Science*, 9, 270-276.

Frank, M.G. & Ekman, P. (1997), The ability to detect deceit generalizes across different types of high-stakes lies, *Journal of Personality and Social Psychology*, 72, 1429-1439.

Furui S. (2005), 50 years of progress in speech and speaker recognition, in *Proceedings of 10th International Conference on Speech and Computer 2005*, Patras, Greece, October 17-19, 1-9.

Kanade, T., Hu, C. & Cohn, J.F. (2005), Facial expression analysis, Paper presented at the *IEEE International Workshop on Modeling and Analysis of Faces and Gestures*, Beijing, China, October 16, 2005.

Keltner, D. (1995), Signs of appeasement: Evidence for the distinct displays of embarrassment, amusement and shame, *Journal of Personality and Social Psychology*, 68, 441-454.

Krumhuber, E. & Kappas, A. (2005), Moving smiles: The role of dynamic components for the perception of the genuineness of smiles, *Journal of Nonverbal Behavior*, 29, 3-24.

RECENTI SVILUPPI DI 'SONIC' PER L'ITALIANO: RICONOSCIMENTO AUTOMATICO DEL PARLATO INFANTILE

Piero Cosi

Istituto di Scienze e Tecnologie della Cognizione, Sede di Padova 'Fonetica e Dialettologia'

Consiglio Nazionale delle Ricerche

piero.cosi@pd.istc.cnr.it

SOMMARIO

In questo lavoro vengono descritti i risultati dei più recenti esperimenti di riconoscimento automatico di parlato infantile effettuati, mediante l'utilizzazione del sistema denominato SONIC, su un corpus di parlato letto da bambini di età compresa fra i 7 e i 13 anni. Il *corpus* utilizzato è stato raccolto presso alcune scuole del Trentino da parte dell'ITC-IRST (Giuliani & Gerosa, 2003) ora FBK (Fondazione Bruno Kessler), nell'ambito di un progetto europeo denominato *PF-STAR*.

Il tasso di errore di riconoscimento iniziale di 15,1% per un insieme di 33 unità fonetiche (21,8% considerando un insieme di 40 unità fonetiche) è stato successivamente ridotto al 12,2% (18,6% considerando 40 unità) utilizzando una combinazione delle più aggiornate tecniche di adattamento comprendenti la normalizzazione di lunghezza del tratto vocale (*Vocal Tract Length Normalization*, VTLN), la normalizzazione della varianza dei coefficienti Cepstrali (*Cepstral coefficients Variance Normalization*, CVN) e l'utilizzazione di modelli fonetici addestrati in modalità indipendente dal parlante utilizzando le più recenti strategie iterative denominate *Structural MAP Linear Regression* (SMAPLR) e *Speaker Adaptive Training* (SAT).

Mentre il tasso di errore del sistema allenato su voci di bambini è paragonabile o addirittura migliore di quello ottenuto da sistemi simili sullo stesso *corpus* (ad esempio paragonabile al 22,7% ottenuto da un sistema analogo con 28 unità fonetiche come quello utilizzato in Giuliani & Gerosa, 2003), esiste ancora un significativo margine di miglioramento per un sistema che utilizzi modelli acustici allenati su parlato adulto e utilizzati per decodificare parlato infantile. Infatti quando sono state applicate entrambe le tecniche VTLN e SMAPLR in una condizione di disallineamento adulti/bambini il sistema finale ha ottenuto un tasso di errore fonetico del 19,3% dimostrando di ridurre l'errore fonetico iniziale del 28%. Ciò nonostante, persiste ancora un notevole 30% di differenza relativa fra l'utilizzazione di modelli acustici allenati su parlato adulto e modelli acustici allenati su parlato infantile per la decodifica di quest'ultimo.

Lo sviluppo di sistemi di riconoscimento di parlato infantile spesso si presenta come un compito di ardua soluzione a causa della spesso totale mancanza di risorse acustiche utilizzabili per l'allenamento dei modelli acustici. In questo lavoro, il sistema di riconoscimento denominato SONIC e sviluppato per l'inglese è stato adattato all'italiano ed in particolare è stato considerato il caso del parlato infantile di bambini compresi nella fascia di età compresa fra i 7 e i 13 anni.

Questi nuovi modelli acustici per il parlato infantile italiano sono stati incorporati nel *Colorado Literacy Tutor* (CLT)¹ (Cole *et al.*, 2003) sviluppato al *Centre for Speech and Language Research* (CSLR) della *University of Colorado* di *Boulder* per la lingua inglese,

¹ <http://www.colit.org/>

quale primo passo verso lo sviluppo della sua corrispondente versione italiana l'*Italian Literacy Tutor*, un sistema interattivo e personalizzato per l'apprendimento della lingua italiana (Cosi *et al.*, 2004; Cosi & Pellom, 2005).

BIBLIOGRAFIA

Cole R., van Vuuren S. & Pellom B. (2003), Perceptive Animated Interfaces: First Steps Toward a New Paradigm for Human Computer Interaction, in *Proceedings of the IEEE*, 91, September 2003, 1391-1405.

Cosi P., Delmonte R., Biscetti S., Cole R., Pellom B. & van Vuuren, S. (2004), Italian Literacy Tutor: tools and technologies for individuals with cognitive disabilities, in *Proceedings InSTIL/ICALL Symposium*, Venice, Italy.

Cosi P. & Pellom B. (2005), Italian Children's Speech Recognition For Advanced Interactive Literacy Tutors, in CD-Rom *Proceedings INTERSPEECH 2005*, Lisbon, Portugal, 2201-2204.

Giuliani D. & Gerosa M. (2003), Investigating Recognition of Children's Speech, in *Proceedings of the ICASSP*, Hong Kong, 2003.

TEST FONETICO DELLA PRIMA INFANZIA PER BAMBINI DAI 18 AI 36 MESI: ANALISI CON ‘PHON’ DEI PRIMI DATI RACCOLTI

Claudio Zmarich ^a, Serena Bonifacio ^b, Maria Pia Bardozzetti ^a, Caterina Pisciotta ^a,

^a Istituto di Scienze e Tecnologie della Cognizione (ISTC), C.N.R., Sede di Padova;

^b IRCCS ‘Burlo Garofolo’, Trieste

claudio.zmarich@pd.istc.cnr.it, logopedia@burlo.trieste.it, mariapia.bardozzetti@libero.it,

caterina.pisciotta@pd.istc.cnr.it

SOMMARIO

Questo articolo si propone il duplice obiettivo di descrivere un nuovo test per la valutazione delle capacità fonetico-fonologiche di bambini dai 18 ai 36 mesi, e un nuovo programma, *Phon*, per la codifica e l’analisi semiautomatica degli aspetti segmentali del parlato. Il nesso tra i due oggetti dell’articolo sarà dato dall’applicazione di *Phon* ai primi dati raccolti tramite il *Test Fonetico per la Prima Infanzia*.

Il test, progettato da S. Bonifacio e C. Zmarich, è ancora in fase di sviluppo, e la versione che qui viene descritta risale alla fine del 2007 ed è già stata parzialmente modificata nei modi che illustreremo in seguito. Attualmente in Italia non si dispone di uno strumento che permetta di valutare le capacità fonetiche in età così precoci. Sebbene il test PFLI (Bortolini, 1995) si proponga di farlo dai 24 mesi in poi, tuttavia, in ambito clinico, il suo uso incontra molti problemi. Il test di Bonifacio e Zmarich si propone la stesura dell’inventario fonetico, basato sulla produzione verbale del bambino (stimolata ma non ripetuta), per accertare se un fono o un gruppo consonantico risultano acquisiti in sede iniziale e non iniziale di parola (cfr. Zmarich & Bonifacio, 2005). Il test è composto da una lista di parole suddivise per fascia d’età, che rappresentano gli item lessicali più prodotti dai bambini per quella fascia, in base alle ricerche sul Primo Vocabolario del Bambino (PVB: Caselli & Casadio, 1995; Caselli, Pasqualetti & Stefanini, 2007). Le parole sono state scelte in base a un criterio fonetico, e cioè che tutti i foni della lingua italiana siano attestati in almeno 3 parole per ciascuna posizione lessicale (vedi anche Zmarich, Dispaldro, Rinaldi & Caselli, accettato per la pubblicazione). Il test considera le denominazioni e le produzioni spontanee ottenute su presentazione di oggetti giocattolo riguardanti gli item lessicali.

Sono state analizzate, in questo studio pilota, le produzioni verbali audioregistrate di 12 bambini con sviluppo linguistico tipico e cioè con dimensione del vocabolario espressivo > al 5° percentile, rilevato con il PVB (Caselli *et al.*, 2007). Il campione è composto da tre gruppi d’età: 18-23; 24-29; 30-36 mesi, e ciascun gruppo contiene 4 bambini. Tutti i soggetti sono stati registrati in due asili nido. La produzione verbale relativa agli item del test è stata successivamente trascritta in simboli fonetici IPA e codificata al PC con programma freeware *Phon* (© 2006-2008 *The Phon Project*), versione 1.3R500, che è stato sviluppato all’interno della comunità scientifica CHILDES per informatizzare la ricerca sull’acquisizione fonologica.

La codifica in *Phon* del materiale raccolto è consistita nella creazione di un *record* per ogni parola realizzata dal bambino, codificata in caratteri alfabetici come glossa, direttamente in simboli IPA (Charis SIL) secondo la pronuncia del bambino e secondo la pronuncia adulta. In questo modo le forme delle due pronunce ricevono un allineamento

automatico dei segmenti nei tipi sillabici costituenti la parola, che si basa sulle regole di sillabificazione dell'italiano. Ogni record è associabile facilmente e in modo permanente alla selezione del segnale acustico che lo riguarda. Oltre a prevedere numerose funzioni di interrogazione, *Phon* consente diversi tipi di analisi di uso più comune nello studio dell'acquisizione fonologica. Nel presente studio verranno esemplificati solo quelli relativi alle statistiche di frequenza ripartite per posizione del fono rispetto alla parola e alla sillaba.

BIBLIOGRAFIA

Bonifacio, S. & Zmarich, C. (2007), *Creazione di un test per la valutazione dello sviluppo fonetico in età precoce*, Progetto di ricerca corrente dell'IRCCS Burlo Garofolo di Trieste, responsabile del progetto: dott.ssa Elisabetta Zocconi.

Bortolini, U. (1995), *PFLI Prove per la valutazione fonologica del linguaggio infantile*, Padova: Edit Master Srl.

Caselli, M.C. & Casadio, P. (1995), *Il primo vocabolario del bambino*, Milano: Franco Angeli.

Caselli, M.C., Pasqualetti, P. & Stefanini, S. (2007), *Parole e frasi nel 'Primo vocabolario del bambino'*, Milano: Franco Angeli.

Phon: <http://phon.ling.mun.ca/phontrac>

Zmarich, C. & Bonifacio, S. (2005), Phonetic inventories in Italian children aged 18-27 months: a longitudinal study, in *Proceedings of INTERSPEECH 2005-EUROSPEECH*, Lisboa, September 4-8, 757-760.

Zmarich, C., Dispaldro, M., Rinaldi, P. & Caselli, M.C. (accettato per la pubblicazione), Caratteristiche fonetiche del 'Primo Vocabolario del Bambino', *Psicologia Clinica dello Sviluppo*.

ENFASI E CONFINI PROSODICI IN DUE STILI DI ELOQUIO EMOZIONALE

Pier Luigi Salza, Enrico Zovato, Morena Danieli
Loquendo S.p.A. – Torino

pierluigi.salza@loquendo.com, enrico.zovato@loquendo.com, morena.danieli@loquendo.com

SOMMARIO

La nuova frontiera delle tecnologie di sintesi del parlato è la capacità di generare segnali vocali con caratteristiche espressive in grado di variare in modo analogo a quanto avviene nella voce umana. Molti studi hanno evidenziato correlazioni significative tra lo stile di eloquio e le variazioni di alcuni parametri prosodici e spettrali; inoltre, alcuni sistemi di sintesi vocale sono in grado di replicare in parte queste variazioni. Nel presente lavoro si illustra uno studio pilota condotto su alcuni fenomeni di natura ritmica e intonativa, relativamente a due stili espressivi, a partire da registrazioni effettuate in laboratorio da parlanti madrelingua inglesi. Si sono annotati, in ciascun enunciato, i confini prosodici e i fenomeni di enfasi eventualmente realizzati, cioè quei fenomeni di prominenza (a livello intonativo, di durata e di intensità) tesi ad evidenziare alcune parole nel contesto della frase. L'obiettivo di questo lavoro è cercare di individuare le correlazioni tra i fenomeni prosodici presi in considerazione e lo stile di eloquio, in modo da riprodurre, nel sistema di sintesi vocale, analoghi meccanismi per una più accurata caratterizzazione espressiva dei segnali generati. Gli stili emozionali studiati in questo progetto sono quelli corrispondenti allo stile triste e allo stile allegro. La scelta di ridurre lo studio a due soli stili è stata dettata da motivazioni pratiche legate anche ai domini di applicazione dei prototipi che sono oggetto di sviluppo nell'ambito del progetto citato. Queste applicazioni mirano alla realizzazione di un assistente virtuale capace di conversare con l'utente in modo affettivo, riproducendo, tramite la voce e altre modalità, comportamenti emozionali legati al contesto del dialogo.

Al fine di ottenere del materiale vocale con due stili espressivi ben caratterizzati, un esperto madrelingua inglese ha selezionato, da romanzi e racconti, testi il cui contenuto fosse in grado di indurre nel lettore l'atteggiamento emotivo relativo ai due stili considerati. Si sono quindi effettuate registrazioni in laboratorio da un parlante maschile madrelingua britannico e due parlanti madrelingua americani, uno maschile e uno femminile, cui si è chiesto di leggere i brani in modo appropriato e naturale, conformemente ai contenuti. Si è poi chiesto ai parlanti di leggere gli stessi brani, in sessioni separate, adottando uno stile di lettura neutro indipendentemente dal contenuto dei testi, ottenendo così un *corpus* di riferimento con stile neutro. Per ciascun parlante, il *corpus* originale consiste in 42 frasi espressive (14 in stile allegro e 28 in stile triste) e in 42 corrispondenti versioni neutre. In totale, dunque, il *corpus* ammonta a $(42+42) \times 3 = 252$ frasi. Il *corpus* è stato sottoposto ad una valutazione percettiva utilizzando una scala MOS a 7 punti da parte di 3 soggetti madrelingua inglesi, tesa a verificare l'effettiva resa emozionale degli enunciati. Si è così selezionato un sottoinsieme complessivo di 90 enunciati espressivi e 86 neutri. I testi degli enunciati selezionati sono stati annotati manualmente e separatamente da due operatori. Essi hanno etichettato ciascun confine di parola percepito come confine prosodico e ciascuna parola percepita come avente la presenza di enfasi. Riguardo all'enfasi, gli operatori sono stati addestrati a segnalare ciascuna parola sulla quale fosse percepibile una

discontinuità rispetto al contesto precedente e/o seguente (rallentamento o accelerazione della velocità di eloquio e/o variazione dell'intonazione e/o particolare inflessione intonativa e/o variazione dell'intensità) tale da far ritenere la parola come pronunciata in maniera prominente, marcata o enfatica. La congruenza tra le annotazioni dei due operatori è stata valutata per ciascun enunciato e per ciascun parametro separatamente (confine prosodico e enfasi) tramite il 'K test', usato per misurare il livello di concordanza tra le classificazioni di due valutatori rispetto alla concordanza *per caso*. Si è quindi ulteriormente selezionato il materiale su cui svolgere le successive analisi sulla base del livello di concordanza, prendendo in considerazione soltanto quegli enunciati il cui valore di 'K' (per uno o per entrambi i parametri considerati) è risultato superiore a 0,6 (concordanza sostanziale).

Sul *corpus* definitivo, in realtà composto da due corpora parzialmente sovrapposti, uno per i confini di parola (76 enunciati espressivi e 66 neutri), l'altro per le enfasi (73 enunciati espressivi e 48 neutri), si sono quindi calcolati due indici della frequenza di occorrenza dei fenomeni in esame, C_P per i confini prosodici e C_{EM} per l'enfasi, dati dal rapporto, rispettivamente, tra il numero dei confini prosodici annotati e il numero di confini di parola e tra il numero di enfasi annotate e il numero di parole. Per quanto concerne i confini prosodici, si nota che solo uno dei parlanti tende ad aumentare il numero di confini nello stile triste. Riguardo all'enfasi risulta una maggiore coerenza tra i tre parlanti: il fenomeno appare più frequente nello stile allegro e meno in quello triste, in cui comunque è significativamente più presente rispetto allo stile neutro. Considerando l'elevato grado di concordanza tra le due annotazioni incluse nel *corpus* definitivo, si è svolta un'analisi statistica ANOVA su serie di dati relativi ai coefficienti C_P e C_{EM} , classificate per stile di eloquio, basandosi sulle annotazioni di uno solo dei due valutatori. Il risultato dell'analisi conferma una variazione significativa della frequenza delle enfasi negli stili allegro e triste rispetto allo stile neutro, mentre per quanto riguarda i confini prosodici i dati confermano la maggiore incertezza già evidenziata dalla precedente analisi, anche se lo stile triste sembra differenziarsi in modo più significativo rispetto allo stile neutro e a quello allegro, con valori più elevati del coefficiente C_P e quindi con una frequenza maggiore di pause. Sono state infine analizzate alcune realizzazioni acustiche dell'enfasi, in particolare per quanto riguarda i valori di F0, esaminando i nuclei vocalici delle sillabe toniche. In quasi tutti i casi si verifica un picco intonativo, più pronunciato nello stile allegro con valori massimi di F0 più accentuati. Nello stile triste i massimi di F0 sono di ampiezza minore, coerentemente con i valori mediamente più bassi di F0. Una misura di queste variazioni è stata condotta sul segnale segmentato ed elaborato con strumenti automatici che forniscono anche il calcolo dei parametri acustici. Sia nello stile allegro che in quello triste si registrano variazioni positive di F0 nelle vocali toniche oggetto di enfasi rispetto a vocali toniche senza enfasi, anche se, come anticipato, di entità decisamente superiore nello stile allegro. Queste variazioni, seppur di entità diversa, sono comuni a tutti e tre i parlanti.

Alla luce di questi primi e parziali risultati, si renderebbero necessarie ulteriori indagini su un campione di dati acustici più esteso. Per quanto riguarda i confini prosodici, potrebbe apparire utile distinguere tra confini forti e confini dovuti a movimenti intonativi che non implicano l'interruzione del parlato. Riguardo all'enfasi, potrà essere sicuramente interessante un'analisi di tutti i parametri acustici (oltre ad F0, anche durata e intensità) delle sillabe toniche, confrontando quelle delle parole su cui è stata percepita l'enfasi con quelle delle altre parole. Infine, sarà presa in considerazione anche l'eventuale correlazione tra presenza di enfasi e/o confine prosodico e la categoria grammaticale delle parole.

UN CORPUS SPERIMENTALE PER LO STUDIO CROSS-LINGUISTICO EUROPEO DELLE EMOZIONI VOCALI

Vincenzo Galatà, Luciano Romito
Laboratorio di Fonetica, Università della Calabria
vgalata@libero.it, luciano.romito@unical.it

SOMMARIO

La presente proposta costituisce il primo stadio di una ricerca attualmente in corso, volta allo studio cross-linguistico europeo delle emozioni vocali in quattro lingue: italiano, francese, inglese e tedesco.

Se da un lato si rilevano innumerevoli studi sul parlato emotivo nelle singole lingue, dall'altro gli studi di tipo cross-linguistico e cross-culturale risultano essere assai sparuti (per gli studi cross-linguistici volti all'*encoding* si vedano, ad esempio, Anolli *et al.*, 2008a, 2008b; Braun & Oba, 2007; Kori & Magno Caldognetto, 2003; Piôt, 1999; per quelli volti al *decoding* si vedano, invece, gli studi citati e riportati in Pavlenko, 2005, a cui vanno aggiunti Pell *et al.* 2009; Sawamura *et al.*, 2007; Shochi *et al.*, 2007; Droomey *et al.*, 2005; Tickle, 1999, 2000; Magno Caldognetto & Kori, 1983). Ancora meno sono quelli che hanno affrontato entrambi gli aspetti di *encoding* e *decoding* (ad es. Chung, 1999, 2000; Abelin & Allwood, 2000, 2002; Breitenstein *et al.*, 2001; Thompson & Balkwill, 2006). I motivi sono prevalentemente imputabili alla difficoltà che lo studio delle emozioni vocali impone, difficoltà ulteriormente esacerbate nello studio di tipo cross-linguistico-culturale.

Esaminando i risultati degli studi appena menzionati, si evince come le emozioni vocali siano, alla pari di quelle facciali, riconosciute cross-linguisticamente e cross-culturalmente, con risultati nettamente al di sopra della semplice casualità. Da una meta-analisi condotta da Laukka (2004) volta alla verifica del riconoscimento cross-culturale delle emozioni e alla verifica dell'esistenza di specifici *patterns* acustici della voce per categorie discrete di emozioni, è emerso che: a) l'accuratezza di riconoscimento è superiore a quella della casualità per categorie di emozioni più ampie; b) il *decoding* cross-culturale è inferiore al *decoding* intra-culturale del 7%; c) esistono specifici *patterns* acustici nella voce delle emozioni che vengono utilizzati per comunicare emozioni discrete.

Una ricognizione su ben 64 *database* di parlato emotivo (cfr. Ververidis & Kotropoulos, 2006) ha rilevato l'assenza di *corpora* di parlato emotivo mistilingue utili per uno studio cross-linguistico-culturale delle emozioni. Ciò ha reso necessario la raccolta di produzioni verbali emotive nelle quattro lingue (italiano, francese, inglese e tedesco) con riferimento alle emozioni definite da Ekman (1992) come *basic* (*happiness, anger, fear, sadness, disgust, surprise*).

Gli obiettivi di questa ricerca sono essenzialmente:

- a. motivare e illustrare le caratteristiche del *corpus* raccolto, con particolare riferimento al protocollo di elicitazione adottato;
- b. fornire i risultati della procedura di validazione. Questa operazione è assolutamente necessaria per l'attuazione delle successive fasi della ricerca, come l'analisi acustica dei campioni raccolti per la caratterizzazione delle emozioni nelle lingue in esame; la somministrazione di esperimenti percettivi nella direzione proposta da Scherer *et al.* (2001: 88) per appurare la capacità di soggetti di lingua diversa a decodificare

espressioni emotive in una lingua differente dalla propria; la stima di quanto la conoscenza della lingua possa influire sulla corretta identificazione delle emozioni presentate ecc..

Con riferimento al primo obiettivo, illustreremo nei dettagli il protocollo di elicitazione costituito da tre fasi distinte e tra loro conseguenti ispirate al 'paradigma degli scenari' e al 'contenuto standard' di Anolli *et al.* (2008a, 2008b), Anolli & Ciceri (1992) e Scherer *et al.* (1991) con la raccolta di produzioni da parte sia di soggetti *naif* che di *attori*.

Con riferimento al secondo obiettivo, invece, presenteremo i risultati della fase di validazione del corpus effettuata con due test, rispettivamente di identificazione, attraverso la verifica di eventuali differenze di riconoscimento da parte dei soggetti ascoltatori coinvolti nelle registrazioni con riferimento a ciascuna delle tre fasi del protocollo di elicitazione; e di rappresentatività delle produzioni emotive, per valutare il contributo, in termini di materiale utile e di riconoscibilità delle produzioni, da parte di soggetti *naif* da un lato e di *attori* dall'altro.

BIBLIOGRAFIA

Si rimanda alla bibliografia in calce all'articolo contenuto nel CD-ROM.

**STABILITÀ DEI PARAMETRI NELLO *SPEAKER RECOGNITION*.
LA VARIABILITÀ INTRA E INTER PARLATORE:
F0, DURATA E *ARTICULATION RATE***

Luciano Romito ^a, Rosita Lio ^a, Pier Francesco Perri ^b, Sabrina Giordano ^b

^a Laboratorio di Fonetica, ^b Dipartimento di Economia e Statistica
Università della Calabria

*luciano.romito@unical.it, lio.rosita@libero.it,
pierfrancesco.perri@unical.it, sabrina.giordano@unical.it*

SOMMARIO

La tendenza della ricerca attuale in ambito di *Speaker Recognition* (SR) è volta a individuare informazioni quanto più oggettive possibili presenti nella voce umana analizzando la produzione di un parlatore senza occuparsi della sfera semantica, della produzione linguistica o della struttura sintattica e morfologica. In aggiunta i metodi noti come semiautomatici e parametrici si occupano di dati considerati statici. Tale scelta in primo luogo è giustificata dalla relativa facilità della misura e dal trattamento di un ristretto numero di parametri (cfr. Barlow & Wagner, 1998) e in secondo luogo perché la misura di dati statici è la naturale evoluzione di una tradizionale analisi linguistica (cfr. McDougall 2006).

Sono i segmenti statici quelli utilizzati per lo studio delle lingue, si pensi agli inventari fonologici, alle aree di esistenza delle vocali costruite su porzioni stazionarie (*mid point* o *steady state*), alle rotazioni consonantiche o alle regole fonologiche. Tale analisi prende lo spunto dalla necessità di differenziare due lingue, due dialetti o una lingua da un dialetto. Così, grande spazio nelle riviste, occupano concetti quali isoglosse o isofone utilizzati per identificare confini ideali tra due lingue o tra due dialetti.

Quanto detto risulta funzionale per differenziare ma non per riconoscere, o addirittura identificare. Di fatto anche il concetto di isoglossa oggi viene sostituito dall'idea più 'analoga' di corridoio di transizione, una larga fascia dove coesistono variabili differenti che caratterizzano entrambe le lingue o i dialetti contigui.¹

Un parlante nel produrre un messaggio o un atto comunicativo attraverso un meccanismo astratto (linguistico), organizza *target* e *goal* che, in seguito, verranno tradotti in azioni che si realizzeranno in un 'progetto fonetico'. Il meccanismo linguistico è essenzialmente l'insieme delle regole e della grammatica del parlante; è la lingua costituita dal lessico, dalla morfologia, dalle opposizioni fonologiche, dalla sintassi, ecc. Tale meccanismo è fortemente influenzato dall'età, dal sesso, dal controllo fonologico, da fattori sociali quali l'origine geografica, lo stato economico, il contesto, la scolarizzazione, ecc. Nolan, a tal proposito, nel 1997 (p.749) scrive: "In implementing the resources of their linguistic mechanism, speakers have to map them onto their individual anatomy. Whilst the requirements of communication may determine many of the details of speech articulation, we may hypothesize that there may be aspects of speech production where each individual is free to find his or her own articulatory solution. The speaker's behavior here is not

¹ In un paese della preSila catanzarese (Soveria Mannelli) coesistono due variabili per il passato e l'imperfetto: la variabile catanzarese ['ji:vi] e quella cosentina [ˌsɪju 'ju:tu] "sono andato".

'learned' as part of the shared knowledge of the linguistic community; rather it is acquired, probably by trial and error".

Due differenti parlanti possono eseguire progetti fonetici differenti per lo stesso scopo linguistico e le conseguenze acustiche di tali progetti possono aiutare molto nel differenziare, anche se, a nostro avviso le modifiche non riguarderanno la parte statica del segnale.

Lo scopo di questo progetto di ricerca, i cui primi risultati sono stati presentati ai Convegni AISV 2006 e 2007, è quello di studiare la variabilità interna di alcuni parametri acustici, di verificare la correttezza di un confronto o di una comparazione basata sul progetto fonetico e quindi su parametri dinamici e di comparare i risultati ottenuti con quelli basati su parametri considerati statici. Verrà analizzato soprattutto l'effetto prodotto da differenti canali di registrazione, da differenti stili di parlato e da differenti software di analisi. In questa ricerca l'attenzione non è focalizzata sul numero degli intervistati bensì sulla varietà dei canali di registrazione investigati e degli stili di parlato considerati.

BIBLIOGRAFIA

Barlow, M. & Wagner, M. (1998), Measuring the dynamic encoding of speaker identity and dialect in prosodic parameters, in *Proceedings of the 5th International Conference on Spoken Language Processing*, Sydney, Nov 30 – Dec 4, 1998 (R.H. Mannell & J. Robert-Ribes, editors), Sydney: Australian Speech Science and Technology Association, 81-84.

McDougall, K. (2006), Dynamic features of speech and the characterization of speaker: Toward a new approach using formant frequencies, *The International Journal of Speech, Language and the Law*, 13, 89-126.

Nolan, F. (1997), Speaker recognition and forensic phonetics, in *The Handbook of Phonetic Sciences* (W.J. Hardcastle & J. Laver, editors), Cambridge: Cambridge University Press, 744-67.

LOUDNESS E ‘LIVELLO DEL DIALOGO’ NELLE TRASMISSIONI RADIOTELEVISIVE

Mauro Falcone ^a, Antonino Barone ^b, Alessandro Bonomi ^b,
Alessandro Balestri ^b, Anna Grazia Santoro ^b, Maria Dell’Osso ^b

^a Fondazione Ugo Bordoni

^b Istituto Superiore delle Comunicazioni e delle Tecnologie dell’Informazione
malfalcone@fub.it, antonino.barone@sviluppoeconomico.gov.it

SOMMARIO

Il segnale audio che riceviamo attraverso i media (radio, tv, internet, ecc.) può essere, e di fatto lo è, pesantemente, affetto da diversi tipi di elaborazioni e alterazioni. È un fatto ben noto che questi segnali, e quindi anche la voce, sono codificati, e quindi compressi perdendo parte della loro originaria informazione, secondo diversi standard (mpeg2, mpeg4, ecc.). Meno noto è il fatto che questi segnali possono essere elaborati in modo da modificarne, con diversi fini, il loro contenuto energetico. In particolare il segnale è sicuramente manipolato (rispetto ai suoi naturali livelli) in fase di *mixing*, ma ulteriori e più o meno arbitrarie modifiche sono possibili successivamente. Solo negli ultimi anni si è iniziato a studiare il problema del loudness nelle trasmissioni radiotelevisive cercando di risolvere il problema sia del dislivello *channel to channel*, sia del dislivello *program to program* e infine quello del *program to advertising*. Quest’ultimo caso in particolare è stato oggetto di diverse indagini sia in quanto soggetto a normative giuridiche sia perché percepito come particolarmente fastidioso dagli ascoltatori. Con la raccomandazione internazionale ITU-R 1770 e le sue successive modifiche si è dato un primo fondamentale contributo alla soluzione del problema, risolvendo che tipo di misura deve essere effettuata per misurare il livello del segnale audio. L’unità di misura secondo la predetta raccomandazione è il *Loudness Unit* (LU), che è pur sempre una misura in decibel. Per semplicità di lettura nell’articolo riporteremo tutte le misure con l’abbreviazione ‘dB’ anche le misure di loudness secondo la ITU-R 1770. Definito pertanto ‘come’ misurare il livello, rimane (almeno) ancora un secondo e fondamentale punto da risolvere, ovvero ‘quando’ effettuare la misurazione. Non è infatti corretto misurare il livello indiscriminatamente su tutto il segnale audio trasmesso per quantificare correttamente il livello del loudness, ovvero del volume percepito dall’ascoltatore, ma devono essere selezionati solamente quelle parti percettivamente rilevanti, trascurando tutto il rimanente. A tal fine sono oggi utilizzati, nella maggior parte dei casi, due diversi approcci: il *dialogue intelligence* ed il *gating*. Nel primo caso il sistema opera una caratterizzazione del segnale in ‘parlato’ e ‘non parlato’ per poi eseguire la misurazione solo sul primo tipo (ipotizzando appunto che in ogni caso la veicolazione maggioritaria delle informazioni avviene attraverso la voce), nel secondo caso invece si definisce una soglia di riferimento tale che tutto ciò che ha valore superiore viene considerato di interesse per la misura, mentre tutto ciò che è inferiore viene tralasciato. Ognuno di questi metodi ha specifici vantaggi, ed ovviamente svantaggi. Se il *dialogue intelligence* può essere operato online senza conoscere il livello medio dell’intensità del segnale, a suo svantaggio c’è il fatto che non può essere utilizzato su segnali musicali, e che risulta computazionalmente complesso. La tecnica del *gating*, al contrario, è di facile realizzazione, può essere applicata su qualsiasi tipo di segnale audio. Non ha infatti senso

un gating assoluto senza conoscere quale sia il livello medio di intensità sonora dell'audio che stiamo considerando.

In questo lavoro si mostrano i risultati di un'ampia campagna di misura effettuata attraverso strumentazione professionale e l'utilizzo dello strumento LM100 della Dolby, che è oggi il riferimento internazionale per la misura del loudness attraverso la tecnica di *dialogue intelligence*. Per quanto riguarda il *gating*, invece, è stato sviluppato un apposito software dagli autori. Sfortunatamente nell'utilizzo del *dialogue intelligence*, essendo vincolati al funzionamento dello strumento utilizzato, non è possibile variare alcuna funzionalità in quanto il software utilizzato per il *voice detection* è in ogni caso protetto da brevetto. Al contrario il software sviluppato permette una facile configurazione dei parametri e realizza sia la misura di intensità in RMS, sia la normativa ITU-T 1770. La misura di loudness proposta in questa normativa è subito stata accettata con entusiasmo sia dalla comunità scientifica, sia dalle industrie legate alla produzione e diffusione dell'audio nel broadcast, ed è stata implementata nel nostro sistema di misura. Questa misura tiene conto sia dell'effetto di interferenza dell'ascoltatore, o meglio della sua testa che viene approssimata come una sfera di circa 21 cm, sia di una curva di adattamento relativa alla sensibilità dell'apparato uditivo umano per i suoni della stessa classe di quelli comunemente trasmessi dalle televisioni. Queste curve di 'adattamento' (curve denominate con le lettere: A per la telefonia, B per i segnali di media qualità, e via così per C, D ecc.) sono ben note in psicoacustica, e nella raccomandazione in questione, in particolare, si propone una revisione della curva B.

Le due tecniche, *dialogue intelligence* e *gating*, vengono messe a confronto su un'ampia ed eterogenea quantità di programmi, relativi a tutte le principali emittenti nazionali. Da questa prima analisi si evince come sia possibile trovare una equivalenza tra le due metodologie solo parzialmente, e sotto vincoli molto stringenti sia della tipologia dei contenuti e sia della correttezza, o perlomeno omogeneità, realizzazione del materiale nella fase di missaggio. Vengono riportati i risultati di campagne di misura atte a studiare e quantificare tutte e tre le situazioni di confronto del loudness (C2C, P2P, P2A). Il materiale utilizzato è stato acquisito nell'anno corrente principalmente dalle emittenti RAI e Mediaset, e nell'orario di prima serata e comunque sempre nell'arco di maggior ascolto. Tutto il materiale audio è manualmente etichettato a due livelli: un primo livello individua inizio e fine programma secondo i palinsesti relativi, e un secondo livello dove il segnale viene diviso in due classi ovvero A (il programma vero e proprio) ed R (che contiene tutto il resto come spot, promozioni, jingle, prossimamente, ecc.).

I risultati sperimentali riportati costituiscono un punto di riferimento importante, in quanto non ci risulta siano disponibili pubblicamente studi simili a questo. Inoltre lo studio tra le diverse tecniche di selezione è, anch'esso, una novità nel panorama degli studi sul loudness, e vuole essere un primo contributo alla soluzione di questo problema che già vede schierati i diversi produttori di strumentazione di misura su due fronti opposti. Infine le investigazioni pilota sulle differenze dello stesso contenuto audio attraverso diversi media, e lo studio delle possibili alterazioni dei livelli di parlato rispetto ad una acquisizione lineare di laboratorio costituiscono i punti di partenza per lo sviluppo di nuove attività sperimentali di ricerca.

SONORITY BASED SYLLABLE SEGMENTATION

Bogdan Ludusan, Serena Soldo

Department of Physical Sciences, 'Federico II' University, Naples

ludusan@na.infn.it, soldo@na.infn.it

ABSTRACT

Automatic speech segmentation is a topic of great interest in nowadays speech related literature due to its multiple use. One of its most important application areas is Automatic Speech Recognition (ASR), in which speech segmentation techniques are applied for obtaining the units used for recognition. In the recent ASR literature, the syllable is a frequent choice for such a unit because it offers a good representation of the variability present in the speech signal while retaining a also good trainability. This is the reason behind our proposal for an algorithm for automatic syllable segmentation.

The most widely used syllable definition in phonology is based on the sonority scale. A recent definition of the sonority links it to the loudness of a sound, which is related to its acoustic energy relative to other sounds having the same length, stress, and pitch (Ladefoged, 1993). The Sonority Sequencing Principle (SSP) (Clements, 1990) is used as the principle for syllabification and it states that the sounds inside a syllable increase in sonority from the onset to the nucleus, with a maximum value corresponding to the nucleus and decrease in sonority from the nucleus to the coda.

The sonority was previously used as a feature for segmentation in speech processing, but it was either used to detect only syllable nuclei (Kawai & van Santen, 2002) or to detect syllable boundaries, but combined with other features and in conjunction with statistics from previous segmentations (Mayora-Ibarra & Curatelli, 2002). Recent work (Galves *et al.*, 2002) has proved the usefulness of the sonority in other areas, like rhythmic class discrimination. In this paper, the authors propose a formulation for the sonority function, defined on the interval $[0,1]$. The proposed function has values close to 1 for sounds displaying regular patterns, characteristic of sonorant portions of the signal and close to 0 for regions characterized by obstruency.

We propose an algorithm for the detection of syllable boundaries using a similar method, but with a distance function more suitable for our task. The algorithm uses exclusively speech processing techniques (both frequency and time domain), having no knowledge about the phonetic content of the signal, in order to obtain the syllable boundaries from the continuous speech signal.

In a first step, the sonority is computed as a distance function between the columns of the spectrogram of the signal. Then, the syllable boundaries will be placed in accordance to the SSP. Thus, the maxima in the sonority function would correspond to the syllable nuclei, while the minima between two consecutive maxima would correspond to the syllable boundaries. A post processing step is applied to the first segmentation, in order to correct some of the errors that might appear. This includes errors due to violation of the SSP and due to the presence of several sonorant segments close to each other.

The results obtained are encouraging, having obtained higher segmentation accuracies than other sonority based system and comparable accuracies with a segmentation system

based on the energy of the signal. Further work would include the combination of this system with one based on the energy (Petrillo & Cutugno, 2003).

REFERENCES

Clements, G. (1990), The role of the sonority cycle in core syllabification, in *Papers in laboratory phonology 1: between the grammar and physics of speech* (J. Kingston & M. Beckman, editors), Cambridge: Cambridge University Press, 283-333.

Galves, A., Garcia, J., Duarte, D. & Galves, C. (2002), Sonority as a basis for rhythmic class discrimination, in *Proceedings of the Speech Prosody 2002*, Aix en Provence, France, 323-326.

Kawai, G. & van Santen, J. (2002), Automatic detection of syllabic nuclei using acoustic measures, in *2002 IEEE Workshop on Speech Synthesis*, Santa Monica, California, 39-42.

Ladefoged, P. (1993), *A Course in Phonetics*, 3rd edition (International Edition). Orlando: Harcourt Brace & Company.

Mayora-Ibarra, O. & Curatelli, F. (2002), Time-Domain Segmentation and Labelling of Speech with Fuzzy-Logic Post-Correction Rules, in *Proceedings of the Second Mexican International Conference on Artificial intelligence: Advances in Artificial intelligence*, 1-14.

Petrillo, M. & Cutugno, F. (2003), A syllable segmentation algorithm for English and Italian, in *EUROSPEECH-2003*, Geneva, Switzerland, 2913-2916.

STATICO VS. DINAMICO. UN POSSIBILE RUOLO DELLA SILLABA NEL RICONOSCIMENTO AUTOMATICO DEL PARLATO

Serena Soldo, Bogdan Ludusan
Università degli studi di Napoli "Federico II"
soldo@na.infn.it, ludusan@na.infn.it

SOMMARIO

Il continuum fonico su cui un sistema automatico di riconoscimento deve lavorare viene normalmente segmentato in piccole porzioni sulle quali algoritmi basati su tecniche statistiche operano sia per l'identificazione dell'informazione linguistica in essi contenuta, sia per ricostruire a posteriori il contenuto complessivo dell'enunciato contenuto nel segnale acustico. Mentre tradizionalmente fino a pochi anni fa le porzioni minime di analisi avevano dimensioni che linguisticamente potremmo definire subfoniche, sempre più spesso, ormai, i sistemi di riconoscimento del parlato fanno uso di analisi di segmenti di parlato superiori ai 150-200 ms.

Sebbene la definizione 'classica' di sillaba (ma i linguisti sanno bene quanto trovare una definizione condivisa da tutti sia difficile) tende a mettere in evidenza le caratteristiche dinamiche del segnale vocale, come ad esempio la coarticolazione, in questo lavoro si è cercato di prendere in considerazione una ipotesi alternativa. L'idea è quella di vedere la sillaba come una rappresentazione statica di un pezzo di parlato, una sorta di istantanea che contenga in sé unitariamente informazione che solitamente si ritiene di tipo temporale, che si estende su un determinato intervallo di tempo. Alla luce di questo tipo di rappresentazione, la variabile indipendente rispetto alla quale i fenomeni che osserviamo evolvono e sulla quale basare un sistema di riconoscimento del parlato non risulterà più essere il tempo ma la sequenza di unità sillabiche. Lo scopo di questo lavoro è quello di capire se questo genere di rappresentazione è in grado di fornire buoni risultati e, eventualmente, di evidenziarne i punti deboli.

Le tecniche utilizzate per realizzare questo tipo di rappresentazione sono due: il numero di frames considerati per ciascun segmento può variare al variare della dimensione del segmento stesso; oppure il numero di frames può essere fissato a priori scegliendo una maggiore o minore sovrapposizione delle finestre di avanzamento durante l'estrazione delle *features* in modo da adattarsi a ciascun segmento. Si è scelto di utilizzare per la rappresentazione i 13 coefficienti MFCC (*Mel Frequency Cepstral Coefficients*). Per valutare l'efficacia di queste rappresentazioni, si è deciso di provare a classificare le sillabe così rappresentate tramite un classificatore lineare SVM (*Support Vector Machines*) opportunamente addestrato. La prima tecnica non ha fornito risultati interessanti; invece i risultati ottenuti dalla classificazione delle sillabe rappresentate con la seconda tecnica sono stati decisamente migliori. In particolare, abbiamo ottenuto le migliori prestazioni utilizzando 15 frames per sillaba con un'accuracy pari a 88,25%.

In conclusione questo lavoro mira ad indagare nuove tecniche di rappresentazione delle sillabe. Abbiamo mostrato come un approccio teso a estrarre informazioni sulle caratteristiche statiche del segnale, piuttosto che quelle dinamiche, può fornire buoni risultati.

