

ALESSANDRO VIETTI, LORENZO SPREAFICO, DANIELA MEREU,  
VINCENZO GALATÀ

## Prefazione

Questo volume prende le mosse dall'attività congressuale del XIV Convegno Nazionale AISV organizzato presso la Libera Università di Bolzano tra il 25 e il 27 gennaio 2018. Le relazioni tenute in quell'occasione vertevano, da diverse prospettive, attorno ai temi della rappresentazione e dello studio del parlato nel contesto naturale.

Per la comunità dei fonetisti, studiare la lingua nel contesto naturale in cui viene usata significa, in primo luogo, allontanarsi dal cosiddetto parlato di laboratorio, ovvero da un dato deliberatamente poco “rumoroso”, raccolto da campioni di norma ridotti e, preferibilmente, composti da soggetti socialmente omogenei. Le mutate condizioni tecniche di acquisizione e trattamento di dati linguistici e una nuova consapevolezza teorica e metodologica permettono al fonetista di andare alla ricerca della lingua parlata prodotta in contesti comunicativi reali. Tali contesti sono caratterizzati, per esempio, dalla spontaneità della produzione, dalla natura dialogica dell'interazione, dall'esecuzione in ambienti rumorosi, dalla simultaneità delle fonti di informazione nonché dalla natura dinamica di tali segnali. La “scoperta” dei contesti naturali non si limita evidentemente a questi aspetti, ma comprende anche la possibilità di analizzare in modo strumentale varietà e lingue poco descritte o rappresentate nel panorama della ricerca fonetica e fonologica.

Il volume raccoglie pertanto le diverse declinazioni nelle quali questo ampio tema poteva essere trattato, proponendo una selezione rappresentativa dei contributi presentati durante il convegno. La diversità degli studi e delle riflessioni metodologiche proposte ha reso difficile, o forse superflua, una strutturazione del volume in sezioni tematiche. Si è preferito pertanto ordinare i diversi contributi lungo un *gradatum* all'interno del quale ogni elemento della sequenza condivide localmente alcune proprietà con gli elementi adiacenti. Il volume contiene inoltre anche alcune relazioni a tema libero, ovvero che, secondo lo spirito dei convegni AISV, non trattavano le tematiche suggerite dal convegno.

Aprè il volume il contributo di Khalil Iskarous che ha proposto, a partire dalla sua relazione su invito<sup>1</sup>, una riflessione sul tema “The encoding of vowel features in Mel-Frequency Cepstral Coefficients”. L'articolo di Iskarous affronta il tema del convegno da una prospettiva metodologica. All'interno della ricerca fonetica la parametrizzazione delle caratteristiche acustiche del parlato è avvenuta principalmente attraverso la selezione di indici all'interno del dominio della frequenza. Tuttavia,

---

<sup>1</sup> La seconda relazione su invito dal titolo “Biosignal-based Spoken Communication”, presentata da Tanja Schulz, non è contenuta in questo volume.

questi indici, pur essendo molto informativi, sono allo stesso tempo estremamente sensibili alle caratteristiche fisiche del parlante e alle perturbazioni indotte dall'ambiente nel quale il segnale si propaga. Al contrario, l'analisi del segnale vocale adottata nell'ambito delle tecnologie del parlato si basa da molti anni su LPC (*Linear Predictive Coding*) e MFCC (*Mel-Frequency Cepstral Coefficients*) che si sono mostrate fonti di informazione più robuste e meno condizionate dalle caratteristiche individuali dei parlanti. L'obiettivo dell'articolo di Iskarous è quello di esplorare il significato linguistico degli MFCC attraverso uno studio di caso sulle vocali, ovvero osservare quali relazioni sussistano tra i coefficienti e la geometria del tratto vocale.

La raccolta di contributi prosegue con due studi che mirano a valutare l'apporto differenziale degli articolatori nella produzione di parlato: da un lato, la batteria di esperimenti condotti da Carignan mostra quali siano gli effetti indipendenti della nasalizzazione sulla struttura spettrale delle vocali; dall'altro, l'indagine presentata da Dipino e Celata fornisce una prima descrizione del ruolo della configurazione linguale nella realizzazione di contrasti di durata e sonorità nelle occlusive dell'italiano toscano.

A questi primi due studi seguono tre contributi distinti per obiettivi di ricerca, ma senz'altro accomunabili per il generale interesse per la variazione e il cambiamento linguistico, lo studio di varietà e dialetti romanzi e la prospettiva metodologica orientata all'analisi di parlato spontaneo. Mereu, nel suo studio sulla palatalizzazione di /k, g/ nella varietà di sardo di Cagliari, sottolinea l'importanza di osservare dati ecologicamente validi per poter elicitar variabili sociofonetiche marcate diastraticamente. Tordini, Galatà, Avesani e Vayra mettono in evidenza il ruolo del contatto linguistico nella produzione di fricative coronali nella varietà di veneto bellunese parlata da emigrati veneti in Australia. Infine, Bernardasci e Negrinelli utilizzano la distinzione tra parlato spontaneo e controllato come variabile esplicativa in due processi di cambiamento in atto che colpiscono, da un lato, il contrasto tra affricata alveo-palatale e occlusiva palatale e, dall'altro, la distribuzione delle vocali medie anteriori in due dialetti lombardo-alpini.

Il contributo di Schirru presenta invece un'indagine strumentale del consonantismo nel dialetto armeno di Gavar. Lo studio mostra, attraverso un'analisi acustica, la reale natura fonologica dei contrasti all'interno del sistema delle occlusive. L'analisi rivela infatti che l'opposizione tra la serie delle sorde aspirate e quella delle sonore deve essere piuttosto reinterpretata come una distinzione di tensione laringea.

I tre studi che seguono sono contraddistinti dallo studio dell'intonazione come dominio di indagine. In particolare, la ricerca condotta da Gili Fivela e Nicora ha tra i propri obiettivi quello di determinare la natura dei *pattern* intonativi in zone di transizione tra aree dialettali nelle quali sistemi intonativi distinti entrano in contatto e, non di rado, in conflitto. Il secondo contributo, a opera di Pinelli, Avesani e Poletto, si concentra sull'interfaccia tra sintassi e prosodia nelle frasi scisse dell'italiano. Le autrici esaminano la presenza e natura di indizi prosodici in grado di testare l'ipotesi che tali costruzioni condividano in maggior grado i tratti strutturali

delle frasi mono-clausola. Il terzo elemento in questa sotto-sequenza prosodica è lo studio di Wehrle, Cangemi, Krüger e Grice che propone un metodo per misurare la nozione, altrimenti impressionistica, di intonazione “monotonica” (*robotic*) e “cannerina” (*singsongy*). Il metodo illustrato si basa sul cambiamento dinamico di F0 nel tempo e arricchisce in questo modo il repertorio dei metodi di misurazione della distribuzione di F0 su domini molto estesi, ovvero eccedenti i costituenti prosodici più ampi.

Il contributo di Origlia, Rodà, Zmarich, Così, Nigris, Colavolpe, Brai e Leorin discute dello sviluppo di un sistema ludicizzato per la somministrazione di un test di discriminazione fonologica ad apprendenti di italiano L1. Il sistema è capace di tenere traccia delle prestazioni dei parlanti e, quindi, di regolarsi dinamicamente generando nuove serie di stimoli alla luce sia delle risposte sino a quel punto ricevute, sia dell'utilità e del valore informativo che gli stimoli possono avere per valutare le abilità di discriminazione fonologica del bambino.

Il lavoro di Di Nardi, Turrisi, Inuggi, Riva, Mauri e Badino è invece dedicato alla presentazione di un sistema di riconoscimento automatico dell'italiano parlato da soggetti affetti da sclerosi laterale amiotrofica. Il contributo mostra chiaramente come il tipo di architettura e il paradigma di addestramento della rete neurale possano incidere sulla precisione del sistema di riconoscimento, in particolare come ciò interessi più il parlato atipico (disartrico) che non quello tipico.

La ricerca di Di Maro, Falcone e Cutugno riguarda l'analisi delle richieste di utenti umani che interagiscono con agenti virtuali. Gli autori confrontano alcune caratteristiche fonetiche di narrazioni semi-spontanee con quelle rintracciabili nelle interazioni rivolte dagli stessi parlanti a due distinti agenti conversazionali. I risultati mostrano che mentre alcuni parametri fonetici sono modificati in egual misura e indipendentemente dall'agente con cui si interagisca (per esempio la velocità di eloquio), altri (per esempio la riconfigurazione dello spazio vocalico) subiscono l'influenza della familiarità dell'utente con il tipo di agente virtuale e sono quindi variabilmente modificati.

L'articolo di Schettino, Origlia e Cutugno tratta di prominenza prosodica e suggerisce un miglioramento della scala di valutazione nota come metodo PromDrum. In particolare, gli autori mostrano come ricorrendo all'algoritmo Dynamic Time Warping sia possibile non solo ricomprendere nelle analisi sulla prominenza file (solitamente scartati, seppur altamente informativi) in cui il numero di battiti annotati non equivalga alle sillabe attese, ma anche recuperare le informazioni sulle metriche di distanze tra battiti e sillabe per valutare la qualità del lavoro del valutatore nonché delle singole valutazioni di un valutatore.

Il contributo di Gretter, Omologo, Cristoforetti e Svaizer descrive lo sviluppo di una piattaforma robotica a basso costo con cui gli utenti possono interagire a distanza. Il sistema, sempre in ascolto, funziona in tempo reale e senza necessità di connessione a servizi remoti. La gestione dell'interazione dialogica è basata su una tecnologia di riconoscimento vocale che sfrutta un modello di Markov nascosto (HMM). La piattaforma è dotata di una serie di microfoni per la localizzazione del

parlante e una migliore acquisizione del segnale parlato anche in ambiente rumoroso, tant'è che le verifiche sperimentali dimostrano che le prestazioni del sistema sono soddisfacenti sia in termini di velocità di riconoscimento che di comprensione, anche quando l'utente si trovi a diversi metri di distanza dal robot.

Gli ultimi due contributi del volume presentano soluzioni che semplificano il processo di annotazione e analisi di dati fonetici, rendendoli accessibili anche a utenti non esperti. In particolare, Piccardi e Becattini propongono VOTEUS, un'interfaccia grafica concepita per assistere il ricercatore durante il processo di allineamento, annotazione ed estrazione delle durate del tempo di attacco della sonorità. Bravi invece presenta Prosit, un plug-in per Praat progettato per permettere ai ricercatori di gestire raccolte di file audio contenenti annotazioni TextGrid, effettuare delle ricerche e, infine, ascoltare, visualizzare e analizzare i segmenti restituiti dalla ricerca.

Questa breve introduzione mostra la varietà e la ricchezza di temi, approcci e finalità degli autori che hanno deciso di trattare del parlato in contesto naturale come da invito degli organizzatori del convegno e curatori del volume. Questi ultimi, da parte loro, non avrebbero mai potuto svolgere il lavoro di supervisione tanto del convegno quanto del volume che ne è scaturito se non fossero stati supportati nelle varie fasi del processo dal fattivo contributo dei membri del comitato scientifico, ovvero Cinzia Avesani, Leonardo Badino, Chiara Bertini, Maria Grazia Busà, Silvia Calamai, Francesco Cangemi, Chiara Celata, Piero Cosi, Conceição Cunha, Francesco Cutugno, Maria Paola D'Imperio, Silvia Dal Negro, Anna De Meo, Mauro Falcone, Barbara Gili Fivela, Mirko Grimaldi, Phil Hoole, Khalil Iskarous, Constantijn Kaland, Paolo Mairano, Pietro Maturi, Daniela Müller, Maurizio Omologo, Antonio Origlia, Elisa Pellegrino, Marianne Pouplier, Irene Ricci, Antonio Romano, Luciano Romito, Stephan Schmid, Jim Scobbie, Antonio Stella, Jane Stuart-Smith, Mario Vayra, Daniela Veronesi, Claudio Zmarich, Enrico Zovato. A ciascuno di loro va il più sentito ringraziamento dei curatori.