

BARBARA GILI FIVELA, SONIA D'APOLITO, ANNA CHIARA PAGLIARO,
VINCENZO SALLUSTIO, MARIA LUISA FIORELLA, MARCO SINISCALCHI

Introduzione

La lingua orale è un oggetto di studio ideale per ricavare informazioni sull'individuo, sul suo sistema sociolinguistico di riferimento, e quindi sulla sua identità linguistica, ma anche sullo stato complessivo dei sistemi fisiologici che gli permettono la produzione e percezione di parlato. L'analisi dell'eloquio offre informazioni preziose e lo sviluppo di applicazioni volte ad automatizzare tale analisi può avere forti ricadute anche in ambito socio-sanitario. L'aumento del numero di soggetti affetti da malattie neurodegenerative, come la malattia di Parkinson e la malattia di Alzheimer, che influiscono in vari modi sulla capacità di comunicare dell'individuo, suggerisce quale possa essere la ricaduta delle indagini interdisciplinari che facciano della voce e del parlato il loro obiettivo d'indagine.

In questo volume è raccolta una selezione di contributi presentati durante il convegno dal titolo "Il parlato in ambito medico: analisi linguistica, applicazioni tecnologiche e strumenti clinici/ Spoken language in the medical field: linguistic analysis, technological applications and clinical tools", svoltosi a Lecce dal 15 al 17 febbraio 2023. Il convegno è stato organizzato prendendo spunto dal progetto PRIN 2017JNKCYZ "Phonetic analysis of dysarthric speech by speakers of different varieties of Italian to develop clinical tools" e ha rappresentato un'occasione di discussione tra studiosi di settori scientifici diversi che affrontano in modo multi- e interdisciplinare lo studio del parlato, con attenzione alle applicazioni tecnologiche in ambito clinico.

I contributi presenti nel volume riguardano vari temi relativi sia al parlato normofasico che al parlato patologico, e mostrano che l'analisi dell'eloquio, automatica o meno, offre informazioni importanti sia sul piano teorico che su quello applicativo.

Il volume, pubblicato grazie ai fondi del PRIN summenzionato, si apre con il contributo di Isabella Laura Simone "Language Impairment in Neurodegenerative diseases", che descrive in modo molto chiaro le basi neurofisiologiche del linguaggio e quali sono gli aspetti del linguaggio e del parlato che possono essere considerati marcatori di sindromi specifiche o di malattie neurodegenerative. Nell'articolo di Serge Pinto "Speech motor control from its impairment: the relevance of studying dysarthria in Parkinson's disease" l'attenzione è rivolta alle disartrie e, in particolare, alla disartria ipocinetica tipica della malattia di Parkinson. Si discute dell'importanza dello studio della dimensione universale e di quella linguo-specifica, perché l'attenzione a questi aspetti può aiutare sia a sviluppare terapie personalizzate, sia a

comprendere ulteriormente il ruolo di alcune strutture neurali nel controllo motorio esercitato durante il parlato.

Nella sezione “Disturbi del parlato e della fluenza”, sono raccolti contributi multi- e interdisciplinari, basati sia su analisi manuali che su elaborazioni computazionali. Diletta Sciancalepore, Valentina Lavermicocca, Tresanna Illuzzi, Laura Lumaca, Simone Dipace e Maria Luisa Fiorella firmano l’articolo dal titolo “Motor speech disorders and dysphagia in neurodegenerative diseases: clinical profiles compared”. Gli autori considerano diverse malattie neurodegenerative e analizzano i principali correlati acustici delle alterazioni nella fonazione e articolazione del parlato nonché le difficoltà nella deglutizione, concludendo che, nel caso disartria e disfagia coesistano, la prima è normalmente peggiore della seconda. Nel contributo “Acoustic biomarkers of Parkinson’s Disease: a pilot study on the *Italian Parkinson’s Voice and Speech* dataset”, Alessandra Ferrera e Gloria Gagliardi presentano i risultati preliminari di un’analisi acustica del parlato di 65 soggetti dell’Italia meridionale, tra i quali 28 pazienti Parkinson, confermando l’alterazione di vari parametri acustici nei Parkinson rispetto ai soggetti di controllo e riportando la tendenza al cambiamento delle misure di durata a seconda del genere del parlante.

Nell’articolo dal titolo “Phonetic analysis and deep neural network examination for dysarthria detection beyond sociolinguistic differences: a comparison”, Barbara Gili Fivela, Raffaello Caserta, Sonia d’Apolito, Anna Chiara Pagliaro, Vincenzo Sallustio e Marco Siniscalchi confrontano l’analisi fonetica e l’uso di *deep learning model* nella classificazione di parlato disartrico e normotipico. L’analisi degli stessi materiali con metodi di fonetica acustica e tramite una rete neurale (una *Convolutional Neural Network* addestrata con dati in lingua spagnola e, tramite *transfer learning*, oggetto di *fine tuning* grazie a dati in italiano) suggerisce che i diversi metodi possano essere integrati con successo. Claudio Ferrante e Vincenzo Scotti, in “Cross-lingual Transferability of Voice Analysis Models: a Parkinson’s Disease Case Study”, propongono nuovamente l’uso dei *deep learning model* per estrarre *feature* acustiche e ottengono buoni risultati nel caso dell’individuazione del Parkinson grazie al transfer cross-linguistico e all’adattamento di modelli dell’inglese al telugu.

Sempre in merito al Parkinson, nell’articolo “Disfluencies in Parkinson’s Disease. A study on Italian early-stage patients”, Marta Maffia, Loredana Schettino, Rosa de Micco e Alessandro Tessitore discutono dati relativi alle disfluenze in parlato monologico, suggerendo che esistano tipi, caratteristiche e usi specifici delle disfluenze già nei primi stadi della malattia. L’articolo successivo si intitola “Amelia Rosselli e l’angoscia del respiro: tra stile e riflessi di parlato patologico” ed è firmato da Valentina Colonna, Federico Lo Iacono e Antonio Romano. Il contributo riguarda l’analisi delle letture della poetessa Amalia Rosselli tra il 1985 e il 1990. Lo stile prosodico della poetessa viene messo in relazione con le informazioni relative al suo stato di salute in quel periodo.

Sulla balbuzie in bambini di età prescolare si concentrano Claudio Zmarich, Simona Bernardini, Sabrina Bonichini, Mirko Bonotto, Federica Chiari, Erika

Ganassin, Caterina Pisciotta, Fabiola Vullo ed Elisa Zacheo nell'articolo "Linguistic proficiency, communicative attitude and severity in a sample of pre-school children who stutter". Discutendo una serie di esperimenti, gli autori riportano che non sussistono differenze significative nella *communication attitude* dei soggetti sperimentali, mentre risultati peggiori si riscontrano nei test di articolazione e ripetizione dei balbuzienti rispetto a quelli del gruppo di controllo.

La sezione "Parlato nelle demenze e nell'autismo" si apre con l'articolo di Elena Martinelli, Vita Garramone, Francesca Mori, Incoronata Nolè, Franca Cameriero, Matilde Martino, Gaetano Di Bello e Gloria Gagliardi, dal titolo "I correlati linguistici del deterioramento cognitivo: raccolta e analisi di un corpus di eloquio patologico di individui anziani lucani affetti da demenza". Lo studio riguarda 20 soggetti affetti da vari tipi di demenza e 20 soggetti di controllo. L'analisi di 9 ore di parlato semispontaneo, effettuata tramite test statistici inferenziali, evidenzia il deterioramento linguistico a livello fonetico, morfosintattico/lessicale, sintattico e semantico, mentre l'analisi qualitativa permette di individuare la carenza di coerenza e coesione, nonché la presenza di ripetizioni e strategie di compensazione. Riguarda le demenze anche l'articolo di Francesco Sigona, Barbara Gili Fivela, Pietro Vigorelli e Andrea Bolioli, dal titolo "A computational linguistic analysis of speech in elderly individuals with cognitive decline: the Anchise-2022". L'analisi computazionale di circa 200 conversazioni tra *caregiver* e anziani con declino cognitivo è condotta con tecniche di Natural Language Processing (NLP) statistiche descrittive e analisi inferenziali, che vengono correlate con dati sulla progressione della demenza misurata con Mini-Mental State Examination (MMSE).

Nell'articolo "Caratteristiche pragmatiche dell'eloquio nel Disturbo dello Spettro Autistico (ASD): uno studio pilota su un campione di bambini italofofoni residenti in Campania", Syria Cira Imparato, Maria Izzo, Olga Liguori, Ernesto Orsino, Donata Sensale, Tina Santarpia, Giulia Perfetto, Giovanna Gison e Gloria Gagliardi esaminano le caratteristiche pragmatiche e conversazionali del parlato semi-spontaneo prodotto da 17 bambini in età scolare con diagnosi di Autismo High-Functioning e 17 con sviluppo tipico. I testi prodotti dai bambini ASD risultano meno articolati e più frammentanti internamente. Un minor numero di connettivi e pause piene confermano il deficit pragmatico, ma suggeriscono anche interessanti interazioni tra pragmatica, morfosintassi e fluency nel parlato. La pragmatica nel parlato di soggetti con ASD è discussa anche nel contributo di Virginia Chiti, Valentina Saccone e Alessandro Panunzi, dal titolo "L'alterazione degli aspetti pragmatici e prosodici nel Disturbo dello Spettro Autistico: analisi del parlato di bambini in età scolare". Gli autori descrivono l'analisi delle caratteristiche acustiche, lessicali e di struttura informativa del parlato semispontaneo di 9 bambini con ASD (8 e i 12 anni), mettendo in evidenza la variazione presente tra parlato monotono ed estremamente enfatico, nonché la frammentazione.

La sezione su "Variazione e apprendimento della seconda lingua" si apre con il contributo di Duccio Piccardi "Reference, sound symbolism, and variation. In search of an elusive link", che riguarda il fonosimbolismo e la proposta dell'esistenza

del *frequency code*, da parte di Ohala, attraverso lo studio della possibile relazione tra la durata dell'aspirazione delle occlusive sorde in fiorentino e la percezione di virilità. Lo studio delle risposte in un test a risposta forzata non conferma la relazione ipotizzata, anche se i risultati sono compatibili con l'esistenza di un transfer di aspetti fonosimbolici tra diversi domini di significato. Nell'articolo "Approssimanti e laterali palatali in italiano: l'effetto della pronuncia sulla conversione fono-grafemica in un gruppo di giovani parlanti torinesi", Anna Anastaseni, Antonio Romano e Sylvain Gerber si occupano della neutralizzazione dell'opposizione /ʎ/-/j/ a favore di foni [j]-like da parte di alcuni parlanti. L'analisi di dati di percezione e acustici mette in evidenza il ruolo della durata nell'identificazione dei fonemi indagati e permette di individuare un gruppo di parlanti che neutralizza la distinzione e un altro che, invece, mantiene la differenza. Roberta Bianca Luzietti propone, invece, un articolo su "La degeminazione delle rotiche in Toscana: indagine preliminare in un archivio orale storico", nel quale descrive i risultati preliminari di un'indagine sociofonetica sulla degeminazione nella varietà di Prato, "fotografata" nell'archivio orale di Angela Spinelli. L'archivio è particolarmente prezioso, perché permette l'analisi di parlato non-standard e più spontaneo di quello di laboratorio, tipicamente usato per indagini acustiche.

Giuditta Avano, Cinzia Avesani, Mario Vayra firmano un articolo dal titolo "Meccanismi di livellamento in contesto universitario: la Gorgia Toscana in soggetti fiorentini fuorisede", nel quale studiano la spirantizzazione delle occlusive in posizione intervocalica da parte di fiorentini che hanno studiato a Bologna. Il confronto con studenti fiorentini che non si sono spostati a Bologna offre l'occasione per considerare vari fattori linguistici e sociolinguistici e osservare un minor numero di casi di spirantizzazione negli studenti fuorisede, spia di un fenomeno di *levelling* che interessa la Gorgia Toscana. Tuttavia, il fenomeno sembra quasi inconscio e fortemente influenzato da aspetti sociolinguistici, differendo da quello precedentemente osservato in lavoratori fiorentini che si sono spostati a Bologna. Kim Groothuis, Giovanni Leo, Giuseppe Magistro e Claudia Crocco presentano un contributo sulle "Prosodic features of pragmatic *quello* in Informal Neapolitan Italian: data from real-world speech", nel quale utilizzano un corpus di parlato raccolto nel canale YouTube con il software Pyrlato. Nell'articolo, gli autori indagano le caratteristiche prosodiche del pronome dimostrativo *quello* che, in Campania, presenta anche usi pragmatici. L'analisi delle 248 istanze rilevanti per l'indagine (tra le 1775 occorrenze estratte da YouTube grazie a Pyrlato) mostra che l'uso pragmatico di *quello* corrisponde alla presenza di un accento intonativo e di un confine prosodico. Gli autori suggeriscono anche un'ipotesi sull'origine dell'uso pragmatico di *quello* e del corrispettivo dialettale.

Sul versante delle tecnologie del parlato, Vincenzo Scotti e Roberto Tedesco discutono di sintesi da testo scritto nell'articolo "On the Role of Dialogue Context in Predicting Speaking Style". Gli autori prendono in considerazione l'efficacia di modelli recenti, che permettono di predire lo stile d'eloquio dal testo della conversazione, pensando alle ricadute per agenti conversazionali e chatbot. In particolare,

descrivono lo sviluppo e addestramento di una rete neurale che, per predire lo stile d'eloquio, funge da collegamento tra la componente testuale e quella di sintesi vocale di un chatbot.

Per quanto riguarda l'apprendimento della seconda lingua, Federica Cavicchio e Mirko Grimaldi affrontano il tema "I training audio e audiovisivi migliorano in misura uguale la percezione segmentale della seconda lingua", verificando l'efficacia dell'*High Variability Phonetic Training*, effettuato usando materiali audio-video, per migliorare la percezione dell'inglese L2 nel caso di italofoeni. L'analisi del post-test dei gruppi sperimentali, che hanno effettuato un addestramento solo audio oppure audio-video, e del gruppo di controllo suggerisce che l'accuratezza nell'identificazione di fonemi migliora con qualsiasi addestramento e che le informazioni audiovisive influenzano l'identificazione delle vocali L2, seppur temporaneamente. Gli autori concludono che, in linea con il modello *Analysis by Synthesis*, le rappresentazioni fonetico-fonologiche si basano principalmente sulle caratteristiche acustiche del segnale. Nicolas Nese, invece, affronta la questione "Arabo lingua straniera. Acquisizione di nuove categorie fonologiche", orientando l'attenzione sull'acquisizione di foni consonantici dell'arabo da parte di 64 studenti italofoeni, in particolare l'acquisizione delle faringalizzate [tʁ], [dʁ], [sʁ], [ðʁ], delle faringali [ħ], [ʕ], e dell'uvulare [q]. I risultati di un compito di lettura, uno di identificazione e uno di discriminazione migliorano in base alla competenza dell'apprendente, mostrando una differenziazione dei foni in termini di difficoltà.

Sempre in questa sezione, Bianca Maria De Paolis, Antonio Romano e Cecilia Andorno presentano un contributo dal titolo "Prosodic and syntactic markers of narrow focus in Italian L1 and L2 (French L1): an experimental study", nel quale analizzano le strategie per la realizzazione del *narrow focus*, correttivo e non, in 41 parlanti italiani e francesi, questi ultimi registrati mentre parlano la L1 e l'italiano L2. Influenze crosslinguistiche emergono sia a livello sintattico che prosodico. Le somiglianze tra i due codici inizialmente aiutano nella segnalazione del *narrow focus* tramite l'uso delle frasi scisse, mentre le differenze prosodiche non favoriscono la segnalazione grazie all'uso della prosodia.

Infine, tre lavori compongono la sezione inerente ad applicazioni tecnologie per il "Riconoscimento automatico del parlato e del parlante", alcune delle quali riguardano l'ambito forense, e non quello medico, ma che possono fornire ulteriori spunti per l'analisi del parlato. Eleonora Rodegher, Alessio Brutti e Daniele Falavigna, con "Automatic speech recognition for low-resource languages: exploring transfer learning from related languages", descrivono alcuni esperimenti relativi all'uso della tecnica del *transfer learning*. Grazie a studi che riguardano italiano, spagnolo, portoghese e tedesco, così come galiziano e vallader, gli autori verificano che il *fine-tuning*, effettuato per adattare ad una specifica lingua o varietà di lingua un modello sviluppato grazie a molti dati relativi ad un'altra lingua, è più funzionale in termini di *error rate* rispetto all'uso di un modello addestrato su pochi dati.

Francesco Sigona e Mirko Grimaldi, con "Calibration and Fusion of Parametric and Non-parametric Statistical Methods for Forensic Semi-Automatic Speaker

Recognition Systems: experimental results” prendono in esame l’accuratezza degli approcci più usati in Italia per il riconoscimento automatico del parlante in ambito forense, emulati nel software IMPAVIDO. L’ipotesi è che la calibrazione della regressione lineare migliori l’accuratezza. La metrica usata per stabilire l’accuratezza è il *likelihood-ratio cost* e i fattori considerati sono molteplici, inclusa la dimensione del database e il modo in cui vengono considerate le caratteristiche disponibili per le vocali. Infine, Francesco Sigona, Giuseppe Vitolo e Mirko Grimaldi si occupano di comparazione automatica della voce in ambito forense, in “Forensic Automatic Speaker Recognition based on Emphasized Channel Attention, Propagation and Aggregation in Time Delay Neural Network”. Dopo aver discusso le differenze tra applicazioni in ambito forense e non, gli autori riportano i confortanti risultati preliminari ottenuti con l’uso dell’*Emphasized Channel Attention, Propagation and Aggregation in Time Delay Neural Network*, che migliorano all’aumentare della durata dei campioni di parlato e alla diminuzione del rumore di fondo.

I curatori ringraziano sentitamente i membri del Comitato Scientifico del convegno svoltosi a Lecce e tutti gli studiosi che hanno agito come revisori degli *abstract* inviati in occasione del convegno e dei contributi pubblicati in questo volume: Giovanni Abete (Università degli Studi di Napoli “Federico II”), Paola Angelelli (Università del Salento), Cinzia Avesani (ISTC-CNR Padova), Valentina Bambini (Scuola Universitaria Superiore IUSS Pavia), Sarah Bigi (Università Cattolica del Sacro Cuore), Silvia Calamai (Università di Siena), Chiara Celata (Università di Urbino “Carlo Bo”), Sonia Cenceschi (Scuola Universitaria Professionale della Svizzera Italiana), Claudia Crocco (Università di Gent, Belgio), Francesco Cutugno (Università degli Studi di Napoli “Federico II”), Valentina De Iacovo (Università di Torino), Anna De Meo (Università degli Studi di Napoli L’Orientale), Francesca Dovetto (Università degli Studi di Napoli “Federico II”), Mariaelena Favilla (Università di Modena e Reggio Emilia), Vincenzo Galatà (ISTC-CNR Padova), Gloria Gagliardi (Università di Bologna), Giovanni Gobber (Università Cattolica del Sacro Cuore), Paola Leone (Università del Salento), Robin Lickley (Queen Margaret University), Paolo Mairano (Université de Lille), Giovanna Marotta (Università di Pisa), Pietro Maturi (Università degli Studi di Napoli “Federico II”), Chiara Meluzzi (Università di Milano), Rosalba Nodari (Università degli Studi di Siena), Antonio Origlia (Università degli Studi di Napoli “Federico II”), Maria Palmerini (Cedat85), Alessandro Panunzi (Università di Firenze), Elisa Pellegrino (Universität Zürich, Svizzera), Tommaso Raso (Universidade Federal de Minas Gerais, Brasile), Antonio Romano (Università di Torino), Michela Russo (University Jean-Moulin Lyon 3, Francia), Stephan Schmid (Universität Zürich, Svizzera), Francesco Sigona (Università del Salento), Patrizia Sorianello (Università di Bari), Lorenzo Spreafico (Libera Università di Bolzano), Mario Vayra (Università di Bologna), Alessandro Vietti (Libera Università di Bolzano), Claudio Zmarich (ISTC-CNR Padova).