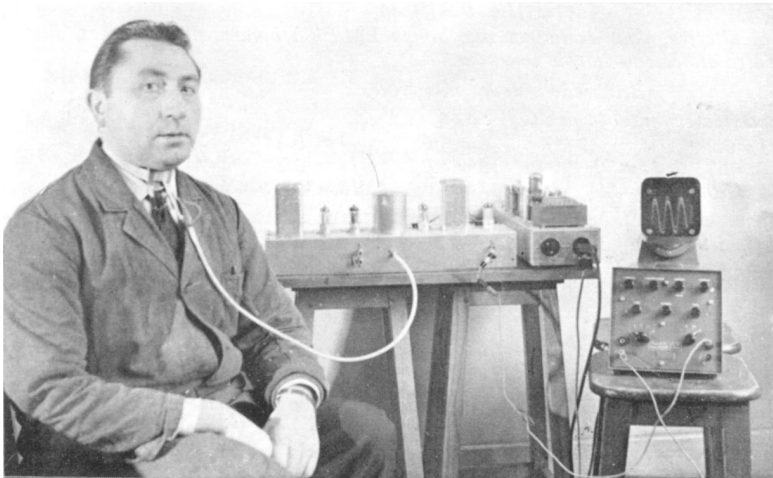


Bollettino del Laboratorio di  
**FONETICA SPERIMENTALE**  
«Arturo Genre»  
dell'Università di Torino



# **Bollettino del Laboratorio di Fonetica Sperimentale**

## **«Arturo Genre»**

### **dell'Università di Torino**

**Pubblicazione semestrale**

---

#### **COMITATO SCIENTIFICO**

- |   |   |
|---|---|
| MARIA GRAZIA BUSÀ – Dip. di Studi linguistici e letterari – Università di Padova          | ANTONIO ROMANO – Dipartimento di Lingue e L.S. e C.M. – Università di Torino            |
| ELISABETTA CARPITELLI – Dép. Parole et Cognition GIPSA-Lab. – Université Grenoble-Alpes   | MATTEO RIVOIRA – Dipartimento di Studi Umanistici – Università di Torino                |
| MARCO GAMBA – Dipart. di Scienze della Vita e Biologia dei Sistemi – Università di Torino | MAURO TOSCO – Dipartimento di Studi Umanistici – Università di Torino                   |
| JOHN HAJEK – Research Unit for Multilingualism University of Melbourne                    | MAURO UBERTI – Comitato scientifico Laboratorio di Fonetica Sperimentale «Arturo Genre» |
| PAOLO MAIRANO – Université de Lille   | FABIÁN SANTIAGO VARGAS – Structures Formelles du Langage – Université de Paris 8        |
| CARLA MARELLO – Dip. di Lingue e Lett. Straniere e Culture Moderne – Università di Torino | STEPHAN SCHMID – Laboratorio di Fonetica – Università di Zurigo                         |
| VICTORIA MARRERO – UNIED Madrid   | MARIE BERTHE VITTOZ – Centro Linguistico di Ateneo – Università di Torino               |
| LORENZO MASSOBRIO – Istituto dell'Atlante Linguistico Italiano – Università di Torino     |   |
- 

Direttore scientifico: ANTONIO ROMANO

#### **COMITATO EDITORIALE**

- |  |   |
|--|---|
| VALENTINA COLONNA – Dipartimento di Lingue e L.S. e C.M. – Università di Torino      | MATTEO RIVOIRA – Dipartimento di Studi Umanistici – Università di Torino                |
| VALENTINA DE IACOVO – Dipartimento di Lingue e L.S. e C.M. – Università di Torino    | ANTONIO ROMANO – Dipartimento di Lingue e L.S. e C.M. – Università di Torino            |
| PAOLO MAIRANO – Univ. de Lille – Laboratorio di Fonetica Sperimentale «Arturo Genre» | MAURO UBERTI – Comitato scientifico Laboratorio di Fonetica Sperimentale «Arturo Genre» |

Direttore responsabile: MAURO UBERTI

---

#### **DIREZIONE E REDAZIONE**

Laboratorio di Fonetica Sperimentale «Arturo Genre» dell'Università di Torino

Via Sant'Ottavio n. 20, 10124 Torino - E-MAIL: [lfsag.unito@gmail.com](mailto:lfsag.unito@gmail.com)

<http://www.lfsag.unito.it/ricerca/phonews/index.html>

Registrazione del Tribunale Ordinario di Torino n° 33 del 22 maggio 2018

Stampato in proprio.

Bollettino del Laboratorio di  
**FONETICA SPERIMENTALE**  
«Arturo Genre»  
dell'Università di Torino

N. 9 – Giugno 2022

## SOMMARIO

ANTONIO ROMANO,	Fonetica costruttiva: costruire ponti, non creare fossati . . .	1
ANNA ANASTASENI,	Dominanza del tratto durata nella classificazione di /λ/ e /j/ . . . . .	5
ANTONIO ROMANO, VALENTINA DE IACOVO, DARIO STRANGIS & SILVESTRO ROATTA,	Primi rilievi EGG sulla voce cantata . . . . .	33
CHIARA PESENTI,	POS tagging of read speech. A tools comparison and an adaptation proposal . . . . .	47
	PHONEWS . . . . .	56
IN MEMORIAM	Vesna Deželjin (Ljiljana Banjanin) . . . . .	57
	Adriana Teresa Damascelli (Antonio Romano) . . . . .	59
	Giovanni Ronco (Valentina De Iacovo e Antonio Romano) . . . . .	61
	Norme editoriali . . . . .	63

# Fonetica costruttiva: costruire ponti, non creare fossati

Presentazione del n° 9 del Bollettino LFSAG

Antonio Romano

In apertura di questo nuovo numero propongo una constatazione sulla complessità della comunicazione scientifica in certi ambiti di ricerca.

Mentre alcuni studiosi sono intenti da decenni a ricercare la famosa interfaccia tra fonetica e fonologia, altri sembrano infatti ritenere – talvolta senza accorgersene – che accentuare la separazione di campi o escludere dalle grandi occasioni congressuali ricercatori con altri interessi e altre finalità possa portare a un reale progresso del loro metodo di studio.

Ad es. si registra da decenni una forma di omologazione ad alcuni modelli di analisi prosodica rigidi e poco espandibili. C'è chi si avvicina con diffidenza e c'è invece chi vi aderisce entusiasticamente, magari cercando di migliorarli dall'interno, ma imbattendosi in pressioni che lo/la riconducono a una certa ortodossia. Questo avviene nell'ambito di una convivenza che include, da un lato coloro che – tanta l'apatia e l'indifferenza – neanche si accorgono delle restrizioni imposte e, dall'altro, quelli/e che sono naturalmente aperti al dialogo e collaborano anche con sostenitori di altri metodi di lavoro, perché – all'esterno – è vero che c'è chi continua per

la sua strada senza accettare compromessi, ma c'è pure chi è disposto ad accogliere i risultati ottenuti con un approccio diverso con altre prospettive di progresso.

Ricordo ad es. il caso di una ricercatrice molto nota come sostenitrice del modello AM/ToBI che aveva a lungo impiegato nelle sue analisi (in particolare quello di un articolo che era circolato in bozze per diversi anni, prima di essere pubblicato in una miscellanea di un certo peso) una notazione<sup>1</sup> che faceva ricorso a un bersaglio tonale medio M (ritenendolo non trascurabile, rispetto agli L e H universalmente accettati)<sup>2</sup>. Tutte le analisi erano state poi riviste per la versione a stampa e il «tono» M era scomparso, forse perché fonte di

- 
1. La comunità internazionale che aderisce a questo modello, a cui appartiene la ricercatrice in questione è quella che ha definito una classificazione intonologica che poggia proprio su una notazione (spesso arbitraria) di eventi locali.
  2. In realtà, ai bersagli tonali si possono opporre movimenti. E succede pure che siano stigmatizzate (o insabbiate) le 'traviature' di sostenitori della prima ora che iniziano a pensare che l'elemento funzionale non sia il bersaglio tonale, ma la transizione o il movimento melodico tra i due (divergenza di vedute non sempre ben esplicitata da chi adotta il modello).

possibili eresie (e confusioni con altri modelli che invece già da prima lo prevedevano). Il messaggio implicito potrebbe essere stato: «se vuoi pubblicare in questo volume, devi adeguarti» (non so; forse penso male: è possibile che l'autrice davvero ci abbia ripensato senza condizionamenti, all'ultimo momento, per non mettersi in evidenza).

Per allargare il campo alla fonetica ricordo però pure un intervento di Nina Grønnum che, in apertura del congresso ICPhS nel 1999, aveva fatto delle proposte interessanti. Infatti, nonostante la relatrice vedesse le cose in modo completamente diverso (avendo definito un modello sovrapposizionale che godeva in quel momento di una certa fortuna), proprio in una sessione plenaria, aveva discusso dell'opportunità di accogliere alcune novità terminologiche e notazionali offerte dal sistema ToBI per sviluppare una sezione di annotazione prosodica IPA diversa da quella tradizionale (effettivamente poco usata) sostituendola con i simboli di quella ormai dominante e intuitivamente più semplice (al punto che molti/e giovani autori/autrici hanno finito per confondere una notazione generica e impressionistica con l'analisi fonologica)<sup>3</sup>.

---

3. Ovviamente la tabella IPA è rimasta invariata a questo riguardo, ma la proposta resta a disposizione di chi voglia accoglierla anche solo per l'utilità procedurale, notazionale e tipologizzante (cioè non necessariamente fonologica).

Si trattava di una proposta che a distanza di anni anche io e il mio gruppo di ricerca, pur adottando un modello analitico distante da quello AM, con cui pure molti aderenti al nostro metodo diffusosi nell'ambito AMPER hanno tentato (e talvolta conseguito il successo di) un dialogo virtuoso, abbiamo deciso di adottare la notazione per tipologizzare una grande quantità di dati di lingue e dialetti diversi (De Iacovo 2021, Romano & De Iacovo, in c. di p., Romano, De Iacovo & Colonna, in c. di p.).

Ecco però che, nel comitato scientifico di un convegno nazionale di un'associazione da sempre orientata all'accoglienza di proposte innovative (eretiche?), appare il revisore (o la revisora) completamente ignaro/a di questa possibilità, e forse anzi ignaro/a dell'esistenza di modelli diversi di lavoro, che si prende l'autorità di dire che non si poteva applicare a quei dati il metodo AM (che, d'altr'onde, nessuno menziona nel testo) e quindi proponendo di non accettare l'articolo.

Mi verrebbe da suggerire, quindi, che un po' di modestia – forse – nel valutare il lavoro altrui non guasterebbe, se la prospettiva fosse quella di costruire ponti in una visione di progresso congiunta.

In ogni caso, ammettendo che persone come queste abbiano tutto il diritto di esistere e spadroneggiare opportunisticamente nell'élite del modello maggioritario, credo che, comunque,

sia ora di prendere le distanze da ricercatori dogmatici che – come si dice degli *bater* «sociali» – «protetti dall'anonimato», ammanniscono conoscenze parziali e immature e censurano la ricerca altrui come se non valesse nulla (che impudenzal!).

L'invito ai giovani (o vecchi) valutatori anonimi è quindi a una maggiore modestia, al sospetto di non essere sempre in grado (per eccesso di parzialità, magari inconsapevole, e quindi per ignoranza) di cogliere tutte le differenze di metodo che permettono di costruire una comunità scientifica in cui si confrontano metodi e risultati per una finalità di crescita collettiva e non per affermare modelli di omologazione a sacrificio di interessi e motivazioni culturali diversi. Occorrerebbe anche informarsi bene sulle finalità e sugli orientamenti più o meno inclusivi dei gruppi storicamente affiatati di ricercatori e ricercatrici a cui si aderisce, prima di giudicare una ricerca (che ha comunque alle spalle un lavoro, magari di mesi) e considerarlo nullo, scarso o inadatto. Ma forse dimentico qui che, dietro una scelta scientifica c'è quasi sempre un'idea politica, un'idea di mondo da valorizzare o costruire che può non avere sempre alle spalle una visione progressista e può, anzi, lasciare emergere visioni oscurantiste.

A coronamento di questo discorso posso ancora aggiungere un recentissimo episodio relativo a due pro-

poste di comunicazione spedite a un convegno internazionale di prosodisti (molto apprezzato dai colleghi italiani) e sicuramente affidate a revisori umorali e distruttivi (con probabili interessi nell'ambito di ricerca in questione). Anziché valutare l'originalità dell'idea e del metodo e la qualità dei risultati, le revisioni insistevano sull'aderenza degli strumenti predisposti per raggiungere un determinato obiettivo ai metodi convenzionali della (loro) scuola di grande visibilità internazionale (anglo-sassone). Adducendo ragioni di poca chiarezza nel metodo (un metodo che presentiamo in diverse sedi da ormai più di vent'anni) e, – soprattutto – chiedendosi quale potesse l'utilità di una simile ricerca, concludevano suggerendo agli organizzatori di scartare le due proposte. Gli stessi materiali (con la stessa formulazione, nella stessa lingua) proposti per un altro evento altrettanto internazionale (con 79% di tasso di accettazione), ma organizzato da un'altra rete di specialisti di analisi prosodica, hanno ottenuto valutazioni entusiastiche (complessivamente 4/5 e 5/5) con apprezzamenti dei/delle valutatori/valutatrici che addirittura si chiedevano come fare per esportare il metodo proposto ad altre lingue e altri contesti. Inutile dire: i sistemi di valutazione *blind- / peer-review* in certi settori non possono ammantarsi di scientificità; soprattutto quando co-

minciano a frapportsi ragioni di scuola e/o elementi di risentimento personale.

Detto questo, anche solo per incoraggiare una migliore disposizione al dialogo dei colleghi, presento i tre articoli di questo numero, dapprima quello di Anna Anastaseni, «Dominanza del tratto durata nella classificazione di /k/ e /j/», che documenta solidamente, mediante osservazioni acustiche e percettive, le condizioni di mantenimento/cedimento del contrasto tra le due palatali in un campione di tre generazioni di parlanti torinesi. Segue un articolo esplorativo a firma congiunta (A. Romano, V. De Iacovo, D. Strangis, S. Roatta) nato per documentare i primi passi di una ricerca in divenire sulle caratteristiche fisiologiche della vocalità: «Primi rilievi EGG sulla voce cantata». Un terzo articolo, più breve «POS tagging of read speech. A tools comparison and an adaptation proposal», di Chiara Pesenti – reduce da un soggiorno di ricerca presso l'Univ. di Nijmegen – testimonia del percorso di avvicinamento di una giovanissima studiosa al problema dell'annotazione di materiali orali, un tema evidentemente già trattato in diverse occasioni anche negli incontri di ricerca nazionali, ma qui definito all'interfaccia

tra i metodi adottati da due laboratori molto distanti per organizzazione, numerosità e finalità, ma verosimilmente destinati ad affrontare le stesse sfide, anche se con mezzi totalmente diversi: l'analisi acustica del parlato disartrico.

Omettendo di dettagliare i contenuti della consueta sezione di notizie *PhoneWS*, aggiungo invece un amaro pensiero in merito alla sezione *In memoriam* che, ricordando la recente perdita di cari/e colleghi/e, dovrebbe far riflettere tutta la comunità scientifica di questo appassionante dominio interdisciplinare (ripenso soprattutto ai/alle franchi/e tiratori/tiratrici delle revisioni anonime) sulla necessità di privilegiare sentimenti autenticamente umanitari, rinunciando a ideali di progresso troppo dogmatici, valorizzando il carico di esperienza e conoscenza che ciascuno porta e seguendo l'esempio di ricercatori e ricercatrici che, con la loro modestia, rappresentano un modello di crescita scientifica e culturale molto meglio di quanto non facciano i/le bempensanti che in queste occasioni credono sia più dignitoso ostentare un loro implicito sprezzante snobismo (spesso impregnato di ottusa saccenteria).



# Dominanza del tratto durata nella classificazione di /ʎ/ e /j/

Anna Anastaseni  
LFSAG

## 1. Introduzione

L'inventario fonologico dell'italiano distingue la laterale palatale /ʎ(ʎ)/ (che rappresenteremo da qui in avanti semplicemente /ʎ/)¹ e l'approssimante palatale /j/, tuttavia tale distinzione è a basso rendimento funzionale. Tra le poche coppie minime si riportano *abbagliare* ~ *abbaiare*; *paglia* ~ *paia*; *maglia* ~ *Maià*².

La palatale /ʎ(ʎ)/ in pronuncia neutra è geminata in posizione postvocalica (Canepari 1992), la sua durata supera in italiano standard e in diverse varietà regionali i 100 ms (Celata 2006); tuttavia non raggiunge la lunghezza della corrispondente geminata lessicale non palatale [l] e, nei dati di Endo & Bertinetto (1999), sembrerebbe non indurre la vocale precedente ad accorciarsi come negli altri casi di sillaba chiusa.

La realizzazione di [ʎ] richiede un complesso meccanismo articolatorio (Bladon & Carbonaro 1978; Canepari 1992; Quilis & Fernández 1969;

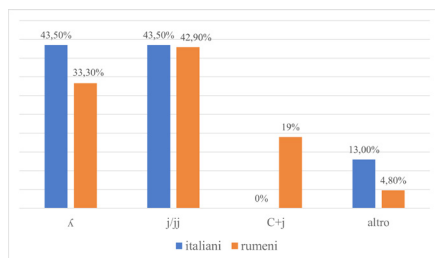


Fig. 1. Realizzazione di /ʎ/ da parte di bambini italiani e rumeni (da Galatà et al. 2012).

Recasens 1984a, 1984b) e il fonema corrispondente si definisce tardi nell'inventario fonologico dei bambini italo-foni (Bortolini 1995; Zanobini et al. 2012)³. Per questo viene spesso semplificata con soluzioni diverse anche nel corso della crescita. I dati comparativi raccolti da Galatà et al. (2012) mostrano infatti che, nella realizzazione di /ʎ/, bambini italiani in età prescolare si comportano in modo simile ai bambini rumeni apprendenti (v. Fig. 1). Il 43,5% produce [j] o [ji] in luogo del fono target. Sembrerebbe che produrre [j] al posto di [ʎ] richieda meno sforzo

3. Bortolini (1995) riporta che meno del 50% dei bambini fino ai 48 mesi da lei analizzati presentano il fonema nel loro inventario. Coerentemente anche Zanobini et al. (2012) hanno rilevato la presenza del fono solo in 3 dei 30 bambini di 36-42 mesi.

1. Riguardo alla scelta tra /ʎ/ o /ʎʎ/ (o /ʎ:/) come fonema sottostante alle realizzazioni scempia (in contesto /#\_/, /C\_/) e geminata (in contesto V\_V) della laterale palatale si veda Romano e Miletto (2017) e Marotta e Vanelli (2021).  
2. V. Mioni (1973). Per una rassegna di fonti successive, v. anche Romano (2009).

al parlante e per questo possa essere favorito (Galatà et al. 2012; Porras 2013; Zampaulo & Haug 2015).

Il fonema /ʎ/ è raro nelle lingue del mondo, è presente solo in 20 delle 451 considerate in UPSID (*UCLA Phonological Segment Inventory Database*)<sup>4</sup>, inoltre realizzazioni non laterali sono molto diffuse. In particolare, la neutralizzazione di /ʎ/ e /j/ è favorita infatti dalla somiglianza acustica<sup>5</sup>.

La [ʎ] condivide le stesse caratteristiche acustiche, ma non articolatorie, con l'approssimante palatale [j] delle parole *aia* ['aja], *caio* ['kajjo], *baia* ['baja]: unica differenza è nella durata che, in [j], è minore (Giannini 2010).

Entrambi i foni possono essere definiti approssimanti, hanno una struttura periodica, ma hanno un'intensità ridotta di circa 10 dB rispetto alle vocali (Pettorino & Giannini 1991, Marotta & Nocchi, 2001). Per quanto riguarda l'italiano, tuttavia, è ancora necessario

descrivere i due foni ([ʎ] e [j]) da un punto di vista acustico, che ne evidenzia con chiarezza differenze e peculiarità.

Tra le lingue che presentano il fonema /ʎ/ nel proprio inventario fonologico c'è il portoghese, in cui il mutamento [ʎ] > [j] caratterizza i dialetti *caipira* del sud-est brasiliano (Zampaulo & Haug 2015) e compare come tratto marcato nei parlanti con basso livello di istruzione (Azevedo 2005).

Parallelamente in spagnolo si assiste al fenomeno dello *yeísmo* con diversi esiti che includono quelli derivanti dalla neutralizzazione dell'opposizione tra /ʎ/ e /j/ (Quilis 1969, p. 123). Questa neutralizzazione dà luogo in regioni diverse a numerosi allofoni classificabili in quattro grandi gruppi: fricativi, affricati, occlusivi e approssimanti (Rost Bagudanch 2013, 2014b). Si può tracciare una *continuum* di variazione nelle realizzazioni fonetiche esito di tale mutamento, tra questi allofoni il più frequente è l'approssimante palatale.

La diffusione dello *yeísmo* è tale che il sistema che non distingue i due foni deve essere considerato maggioritario (Gómez & Molina Martos 2013; RAE 2011). Sebbene le prime attestazioni del mutamento fonetico siano state registrate in Andalusia, le cause del cambiamento non necessariamente devono essere attribuite all'espansione geografica di questa variante, ma possono piuttosto essere ricercate nel sistema fonologico dello spagnolo.

4. L'interfaccia HTML è stata realizzata da Henning Reetz a partire dal database Maddieson (1984). È disponibile online su sito: [http://web.phonetik.uni-frankfurt.de/upsid\\_find.html](http://web.phonetik.uni-frankfurt.de/upsid_find.html).

5. Si è scelto di considerare [j] come un fonema distinto dell'italiano, pur consapevoli del particolare statuto fonologico delle approssimanti [j] e [w]. Per approfondire la questione vedi Marotta e Vanelli (2021: 52-53). Come precisazione ulteriore, Romano e Miletto (2017:66) sottolineano che /j/ è intrinsecamente breve «anche se [...] si presenta soggetto a variazione persino nella pronuncia dei professionisti della dizione».

lo (Gómez & Molina Martos 2013) e in meccanismi di natura strettamente fonetica (Rost Bagudanch 2014b). Un ruolo significativo è da attribuire anche al fatto che percettivamente sembra non essere molto saliente la distinzione tra [ɔ] e [j] (Zampaulo & Haug 2015), distinzione che inoltre a livello fonologico anche in spagnolo è a basso rendimento funzionale.

Parallelamente in italiano la neutralizzazione dell'opposizione tra /ɔ/ e /j/ è molto diffusa: come già detto [j]/[j] sono soluzioni adottate nel parlato infantile. Inoltre, la soluzione articolatoria delateralizzata è diffusa a macchia di leopardo al Centro-sud (con una notevole variazione micro-areale, Colonna & Romano 2018), fatta eccezione per la Toscana e la Sardegna (Canepari 1992; Telmon 1993)<sup>6</sup>.

---

6. Valutando fonti scritte, Granatiero (2021: 15) osserva *j* in romanesco (*pjià* 'piagliare', *fjiò* 'figlio'): «Sebbene da alcuni osteggiata, perché ormai desueta in lingua, la *j* resiste in quasi tutti i dialetti (*vuoje* 'voglio', *maravije* 'meraviglia', *aju* 'ho', [...] *jurnu* 'giorno', *fiju* 'figlio', [...], *sapje* 'sapevo', *proverbij* 'proverbi, [...] *peju* 'peggiore'). Invece, tra le altre grafie usate nel Centro-Sud per /j/ Romano (2016), per *mèju* 'meglio', segnala anche gli esempi *mèju*, *mèju*, diffuse in area salentina. Mentre resta necessaria una rassegna di fonti per le parlate del Nord-Italia (tanto per le aree contigue con territori in cui sono diffusi modelli gallo-romanzi quanto per quelle esposte a modelli slavi), può essere utile ricordare gli interessanti contributi di: Vitali & Pioggia (2014) che, spaziando dai dialetti romagnoli a un'area allargata del Centro-Nord, discutono delle rela-

D'altro canto, nelle stesse regioni, si assiste all'allungamento di /j/ in posizione intervocalica, che invece in pronuncia neutra non allunga mai (Canepari 1992). Camilli (1911) registra che la pronuncia scempia fiorentina di [j] non si riscontra sistematicamente nel resto dell'Italia e persino nelle montagne pistoiesi e lucchesi è possibile una pronuncia di tipo [j̃]<sup>7</sup>.

Canepari (1992) inoltre esplicita la necessità di non confondere /ɔ/ con /lj/ e /llj/. Infatti, nel nord d'Italia [ɔ] non è lunga in posizione intervocalica ([fa'mi:ɔa] in luogo di [fa'mi:ɔa]) e ha alcune volte realizzazione prepalatale [j̃], che si confonde con realizzazione simili di /lj/ e /llj/ con coalescenza (Romano & Miletto 2017)<sup>8</sup>.

---

zioni tra /ɔ/ e /j/ e della resa grafica di quest'ultima; Iannaccaro (2015) che, pur non riferendo esempi in merito a questi contrasti, propone una ricca rassegna di fonti; e, infine, Miola (2015a, 2015b), per i dialetti lombardi e piemontesi. Distinguendolo da quest'ultimo per finalità e motivazione, è inoltre necessario ricordare l'impegno riversato in area piemontese da Arturo Genre (Genre 1979 e materiali inediti).

7. Come recentemente rilevato da Romano (2018), Camilli (1929) segna tre diversi gradi di fortizione di /j/, con esiti fonetici diversi; ad es. a [j̃] de *jje* 'gli' oppone infatti [j] di *ji* 'andare' in contesto di non-RF e [j̃] per la stessa parola in contesto di RF. Si noti che Bertinetto e Loporcario (2005) riportano anche una variante [j̃] nel romanesco non-acrolettale.

8. La difonizzazione di [ɔ] in [lj] e viceversa la monofonizzazione di [lj] in [ɔ] sono riportate da Telmon (1993, pp. 106-107) come caratteristiche di Piemonte, Liguria, Veneto, Friuli

A complicare questo quadro, bisogna ricordare inoltre, il carattere ‘composizionale’ del diasistema italiano. Un’analisi attenta della situazione sociolinguistica contemporanea mostra che spesso, in particolare per le nuove generazioni e nei centri urbani, è difficile caratterizzare l’italiano di un soggetto nei termini di una singola varietà regionale. Al contrario, specie sul piano fonetico, tendono a sommarsi tratti provenienti da varietà differenti (Cerruti, 2013). Per quanto riguarda i fonemi oggetto di interesse, un confronto tra tre *corpora* di italiano parlato a Torino (*ParlaTO*, *KIP* e *CLIPS*) condotto da Fontanot (2019) per la sua tesi magistrale emerge che indipendentemente dal *background* socio-culturale e familiare dei parlanti si alternano [ʌ] e [lj], le variabili attese, a [ʌ:] come nello standard. Inoltre, *ParlaTO* e *CLIPS* mostrano in maniera asistematica la variante intermedia [ʌj] e anche la variante [j:] tipica di alcuni italiani centro-meridionali.

Sembrirebbe di conseguenza che la mancata distinzione tra /ʌ/, /j/, /lj/ e /llj/ possa interessare a vari livelli italo-foni di distinta origine e procedenza. Resta da chiarire se un simile mutamento abbia effetti anche nella capa-

Venezia Giulia, Emilia, Abruzzo e Roma (marcato diastraticamente). D’altro canto sembrerebbe che anche in aree del Centro-sud /llj/ possa essere articolato come [ʌ:] e, forse anche, come [j:] (Canepari, 1992).

cià di classificare i fonemi differenti e quali conseguenze possa avere effetti nella conversione grafo-fonematica.

### 1.1. Interfaccia grafo-fonematica

L’italiano è una lingua con ortografia *trasparente* (o *superficiale*)<sup>9</sup>. Tuttavia, mentre la conversione grafo-fonematica totalmente prevedibile, fatta eccezione unicamente per la posizione dell’accento e per pochissime eccezioni;<sup>10</sup> viceversa nella conversione fonografica ci sono delle zone di opacità. Quindi, per scrivere secondo la norma ortografica attuale parole come *acqua*, *cuore*, *scienza*, è necessa-

9. Il livello di trasparenza di una lingua si valuta principalmente secondo tre misure: la regolarità, la consistenza e l’entropia (*regularity, consistency, entropy*). Per il calcolo della *regolarità*, si parte dal mappare le principali regole di trascrizione grafo-fonematica e, a partire da esse, si misura la percentuale delle parole della lingua alle quali sono applicate. Al contrario, la *consistenza* indica il grado di variabilità nella corrispondenza di unità ortografiche e fonetiche. Infine, l’entropia permette di discriminare non solo tra i casi in cui sono presenti più o meno alternative, ma anche di valutare quanto sia prevedibile la conversione grafo-fonematica e fonografica delle singole (Borleffs et al. 2017). Maggiori sono i livelli di regolarità, consistenza e entropia più una lingua può essere definita trasparente.

10. Tra le poche eccezioni si annoverano alcune forme del verbo *negliere* (*negligente, negligenza*), *siglare* (*sigliamo, sigliate*), e forme neoclassiche come *ipoglicemia* che hanno pronuncia biconsonantica [gʎ] e non laterale palatale [ʎ] (Pettorino 2011).

rio l'accesso al lessico ortografico.

Una pronuncia regionale marcata, o le specificità di pronuncia di un singolo, possono rendere il sistema ancor più sbilanciato e l'ortografia più complessa. L'organizzazione diversa dell'inventario fonologico può portare alla violazione di corrispondenze fono-grafemiche che in italiano standard sono regolari, questo comporta l'esigenza di affidarsi alla via lessicale nella scrittura.

In italiano standard al fonema /k/ e ai nessi /lj/ e /llj/ corrispondono rispettivamente le stringhe grafiche <gl(i)>, <li> e <lli>, si tratta quindi di una conversione trasparente.

Per quanto riguarda /j/ invece la trascrizione è generalmente <ɰ>, la corrispondenza è quasi biunivoca, tuttavia in alcuni casi eccezionali si può trovare <ɰ> o <ɣ>. In particolare l'uso del grafema <ɰ> è stato lungamente discusso nelle diverse fasi di standardizzazione e fissazione dell'ortografia italiana e ha subito diverse oscillazioni nel corso dei secoli. Già a partire dalla fine dell'Ottocento l'uso di <ɰ> ha iniziato un forte regresso<sup>11</sup>. Attualmente è ancora pre-

sente in alcuni nomi propri e toponimi in particolare in posizione iniziale (come *Jolanda* e *Jesolo*) e nei cognomi ha retto maggiormente anche in posizione interna (*Lojacono*, *Rajna*) (Serianni, 1991). Alcune parole presentano ancora sia la forma con <ɰ> sia quella con <ɰ'>.<sup>12</sup>

D'altro canto il grafema <ɣ> con pronuncia [j] è ancora più raro in italiano. È presente anche esso in alcuni cognomi, soprattutto di origine straniera (*Meyer*) e in prestiti non adatti (*yacht* o *papaya*, che presenta anche la forma *papaia*)<sup>13</sup>.

---

12. Un caso particolare è quello di *zabaione* che presenta anche le varianti *zabajone* e *zabaglione*. Quest'ultima è stata così commentata da Coletti (2015): «da pronuncia popolare ha immesso una /l/ tra la *a* e la *i* e quindi ha favorito una pronuncia *zabalione* che poi ha generato *zabaglione*, da ultimo pure accolto nella scrittura». Al contrario il *Dizionario di ortografia e pronunzia* (DOP) la interpreta come ipercorretta rispetto alla forma in [ajɔ] sentita come dialettale. Tutti i dizionari consultati sono concordi nel definirla meno preferibile. La *Grammatica italiana* (2012) commenta invece *zabajone* come ormai «antiquata e letteraria». Sarebbe interessante tuttavia indagare la tendenza a utilizzare la forma *zabajone* nel lessico del marketing, in particolare nella scelta dei nomi da attribuire ai prodotti (*product naming*). In particolare tale grafia potrebbe essere utilizzata come mezzo per evocare una maggiore vicinanza con il dialetto e per favorire la narrazione del prodotto come locale e autoctono. Parallelamente questo avviene anche con *gianduja* e *giandujotto*, grafia preferita da alcune grandi cioccolaterie piemontesi.

13. Per quanto riguarda <ɰ> e <ɣ> si ricordi che i grafemi, in particolare nei forestierismi, non sempre corrispondono alla pronuncia dell'approssimante palatale [j]. In nomi propri come

---

11. Per approfondire l'arco storico di evoluzione del trattamento di tale grafema tra il Cinquecento l'inizio del Novecento si rimanda a Migliorini (2019 [1994]) e a Revelli (2018) per un *excursus* sul modo di affrontare questo tema della trattativa scolastica ottocentesca e novecentesca. Inoltre Malagoli (1905) può essere utile per un quadro sintetico ma efficace che rende l'idea di come venisse concepito il dibattito sull'uso di <ɰ> all'inizio del secolo scorso.

Per quanto riguarda la laterale /ʎ/ può essere realizzata graficamente o con il digramma «gl» o con il trigramma «gli» a seconda del contesto: comporta l'applicazione di una regola di conversione su base sillabica: «gli»+a/e/o/u /«gl»+i. Si tratta di una norma che si applica in modo regolare, la corrispondenza è biunivoca ed è sempre prevedibile.

Tuttavia, lo *status* particolare di geminata intervocalica della laterale palatale potrebbe dare adito a particolari confusioni grafiche (resa grafica in conformità con la regola generale di corrispondenza tra fono lungo e grafia con raddoppiamento [ʎ] → \*«ggl»), complicate ulteriormente dal fatto che non presenta una conversione grafo-fonematica uno a uno.

Inoltre, l'alternanza nella pronuncia [ʎ], [ʎi], [ʎj], [lj] e [ji] e di altre possibili realizzazioni rende potenzialmente ambigue le parole che presentano le stringhe «gl(i)», «» intervocalica «ll» prevocalica e «l» prevocalica, rendendo il sistema di conversione grafo-fonematico significativamente più opaco.

Per questo motivo è necessario interrogarsi sulle capacità di classificazione

di tali suoni: se percettivamente essi non sono distinguibili non è sufficiente la conversione fono-grafemica per recuperarne la grafia, ma è necessario passare per la via diretta. Questo ha importanti conseguenze per la valutazione di eventuali errori di questo genere durante i task di diagnosi per la disortografia, in particolare per quelli che fanno utilizzo del dettato di non parole, e per la classificazione degli errori commessi come «fonologici», «non fonologici» o «superficiali».

## 2. Esperimenti

Visto lo statuto debole di /ʎ/ e /j/ come fonemi distinti, la vicinanza articolatoria e la diffusione della neutralizzazione di /ʎ/ e /j/ in favore di [ji]/[lj], la possibile confusione di [lj] e [li] e le relative problematiche ortografiche che ne potrebbero scaturire, si è deciso di proporre due task percettivi.

Il primo punta a valutare le capacità di classificazione di /ʎ/, /j/, /lj/ e /llj/, il secondo invece è focalizzato solo sull'identificazione di /ʎ/ e /j/ al variare della durata dei foni bersaglio. L'obiettivo principale del presente studio è quindi valutare se a livello percettivo da parte di italofoeni avviene la stessa neutralizzazione tra /ʎ/ e /j/, ma anche la confusione con /lj/ e /llj/, già commentata a livello articolatorio. Inoltre, verranno discussi degli effetti di questo nell'interfaccia grafica.

Nei paragrafi successivi si presenta-

*John* o *Jessica* ad esempio si pronuncia [dʒ] e in prestiti come *abat-jour* [ʒ]. D'altro canto «» con pronuncia pienamente vocalica si trova nell'onomastica italiana in diminutivi di origine inglese (*Jimmy, Mary, Tony*). Inoltre in molti anglicismi lo stesso grafema ha realizzazioni foniche diverse (come in *hockey, by-pass*) (Serianni 1991).

no le specifiche dei task proposti e la metodologia utilizzata.

## 2.1. Metodo

### 2.1.1. Procedura sperimentale



Figura 2: Interfaccia software test di percezione. ESPERIMENTO 1.

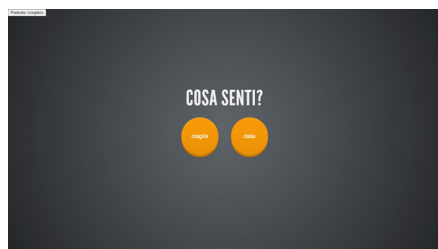


Figura 3: Interfaccia software test di percezione. ESPERIMENTO 2

In entrambi gli esperimenti si è scelto di somministrare un *task* chiuso a un gruppo di uditori italo-foni, che prevedeva l'ascolto di non parole. I partecipanti erano poi forzati ad associare una trascrizione ortografica in una lista chiusa di possibili etichette (grafia con <gb>, <ϕ>, <di> e <li> nel primo esperimento, <gb> e <ϕ> nel secondo; v. Figura 2 e 3).

I task sono stati diffusi online attraverso la piattaforma *Folterpa* (Fernán-

dez Rei et al. 2021), un software elaborato presso l'Istituto da Lingua Galega (*ILGA*) dell'Università di Santiago de Compostela (Galizia, Spagna).

Nel primo task, a ogni partecipante sono state proposte 32 non parole, nel secondo 40, ciascuna delle quali è stata testata due volte (totale stimoli: 64 primo test; 80 secondo). Per evitare effetti di lista, gli stimoli sono stati presentati a ogni riproduzione con ordine aleatorio. Ciascun *item* veniva ripetuto al partecipante una sola volta prima del giudizio.

Prima di cominciare il task veniva presentata la seguente spiegazione:

*Ti verranno fatte ascoltare alcune parole inesistenti, associale alla grafia che ritieni essere più corretta. Se ti è possibile indossa un paio di cuffie e assicurati di essere in un posto silenzioso. Rispondi il più velocemente possibile, senza rifletterci troppo. Non potrai tornare indietro: se ti capita di premere il tasto sbagliato non ti preoccupare e procedi con il test.*

A metà del task era prevista una pausa di durata a scelta del partecipante.

Sono calcolati i tempi di risposta in millisecondi che intercorrevano fra la fine ripetizione di ogni stimolo e la risposta dell'ascoltatore.

### 2.1.2. Partecipanti

Il primo task è stato proposto a 67 studenti universitari tra i 18 e i 29 anni (media: 21,20; DS: 2,52) tutti di madre lingua italiana. Il 62% dei soggetti ha indicato come luogo di provenienza

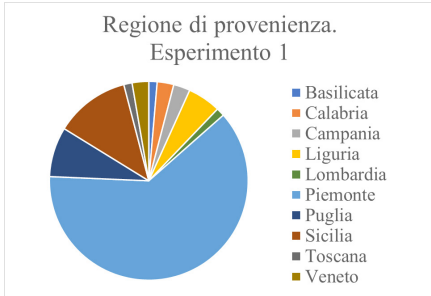


Figura 4: Regione di provenienza dei partecipanti. Esperimento 1.

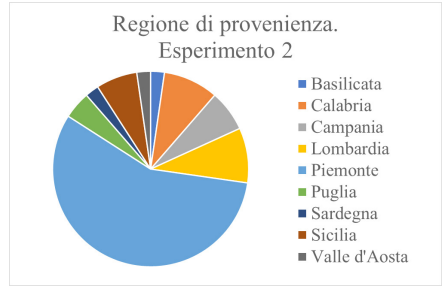


Figura 6: Regione di provenienza dei partecipanti. Esperimento 2.

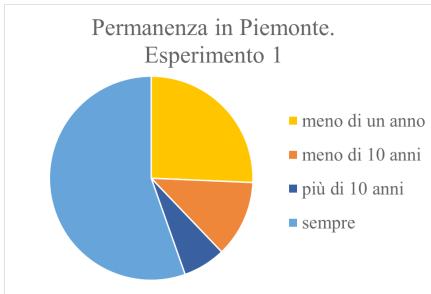


Figura 5: Permanenza in Piemonte dei partecipanti al test. Esperimento 1.

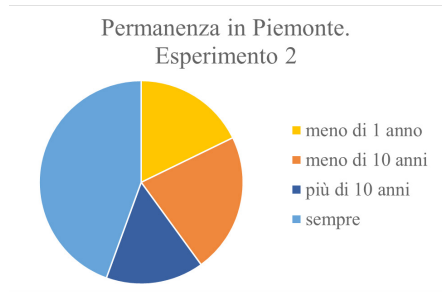


Figura 7: Permanenza in Piemonte dei partecipanti al test. Esperimento 2.

il Piemonte (v. Figura 4, pag. 12) e il 55% ha dichiarato di vivere in questa regione da più di dieci anni (v. Figura 5, pag. 12).

Per quanto riguarda il secondo esperimento invece è stato svolto complessivamente da 46 studenti dell'Università degli studi di Torino di età compresa tra 21 e 31 anni (media 23,93; DS: 2,07), il 57% dei partecipanti ha indicato come regione di provenienza il Piemonte (Figura 6) e il 44% del campione vive in questa regione da più di dieci anni (Figura 7). Come già visto tenere in

considerazione questi dati è necessario perché potrebbero avere un ruolo significativo nell'organizzazione dell'inventario fonologico del singolo e di conseguenza nelle sue capacità di distinzione categoriale tra fonemi differenti. Per lo stesso motivo è stato chiesto ai partecipanti di indicare tutte le lingue di cui hanno competenza attiva e passiva<sup>14</sup>.

14. Alcuni studi comparativi condotti sul catalano e lo spagnolo mostrano che i catalofoni hanno migliori prestazioni nel distinguere i fonemi [ʎ] e [j], distinti nella loro lingua madre rispetto ai castiglianofoni (Rost Bagudanch 2016, 2020).



Tutti i partecipanti al task hanno dichiarato di non aver problemi di udito, né di aver mai sospettato una ipoacusia.

Infine, è stato controllato che nessun soggetto avesse dato più dell'85% delle risposte identiche e sono stati esclusi i risultati di quattro persone che hanno dichiarato di aver avuto problemi tecnici durante lo svolgimento del test.

A ciascuno studente inoltre è stato richiesto di compilare il Questionario Vinegrad Plus (v. n. sotto), secondo la versione modificata di Montesano e Valenti (2017). Si tratta di un adattamento in italiano della *Revised Adult Dyslexia Checklist* (ADCL) di Vinegrad (1994), un questionario *self-report* di 26 domande volto all'autovalutazione delle proprie difficoltà di letto-scrittura e calcolo. Non è uno strumento diagnostico, ma permette di identificare chi ha difficoltà riconducibili a un Disturbo Specifico dell'Apprendimento. Si considerano a rischio, infatti, gli studenti che totalizzano un punteggio superiore o uguale a 9 punti (Cornoldi & Montesano 2020; Montesano & Valenti 2017; Montesano & Valenti 2018)<sup>15</sup>.

---

15. I dati che verranno presentati sono stati raccolti sottoponendo un questionario che sarà presentato più dettagliatamente in altra occasione. Il questionario è composto da 26 affermazioni, per ciascuna delle quali bisogna rispondere se corrisponde al vero oppure no. A ogni risposta positiva viene attribuito 1 punto, 0 per le risposte negative, fatta eccezione per la domanda 3 e la domanda 6 (*Mi piace leggere ad alta voce.* e *Mi piace leggere libri lunghi.*) per le

Nel campione che ha svolto il task di identificazione il 33% dei partecipanti risulta nella fascia di rischio, mentre il 28% ha dichiarato di sospettare di avere un Disturbo Specifico dell'Apprendimento. Questi dati sono particolarmente significativi visto lo stretto legame tra consapevolezza metafonologica e DSA (Banai & Ahissar 2018; Cornoldi 2019; Ramus et al. 2003; Snowling 2019). Molti studi hanno dimostrato infatti la presenza di difficoltà fonologiche anche in persone dislessiche adulte (Cavalli et al. 2017; Svensson & Jacobson 2006). Inoltre, sebbene con l'avanzare degli anni le persone con DSA italiane, che parlano quindi una lingua con ortografia trasparente, tendono a compiere un numero maggiore di errori non fonologici rispetto agli errori fonologici (Bindelli et al. 2009; Cornoldi & Montesano 2020; Fenzi & Cornoldi 2015); tuttavia gli esiti sulle prove volte alla valutazione delle abilità metafonologiche complesse (come il task di inversioni di iniziali «*spoonerism task*») evidenziano il permanere di un deficit significativo a carico delle abilità di processamento fonologico (Danelli et al. 2017).

Inoltre, si deve tenere in considerazione il fatto che il task di identifica-

quali bisogna attribuire i punti all'inverso (0 per le risposte positive, 1 per le risposte negative). Il punteggio finale si calcola sommando tutti i punti parziali, un punteggio superiore a 9 viene considerato a rischio.

zione nelle modalità qui proposte ha una componente visiva importante, aspetto che può rendere più difficile il compito per chi ha un DSA (McCrary et al., 2005). Gli studenti, infatti, hanno dovuto fare uno sforzo di conversione grafo-fonemica per associare gli stimoli ascoltati alle etichette ortografiche proposte.

È interessante quindi approfondire la correlazione tra le capacità di distinzione tra gli /ʌ/ e /j/ e il posizionamento nella fascia di rischio.

Per comprendere meglio le capacità metafonologiche di classificazione di questi fonemi, sia delle persone nella fascia di rischio che del gruppo di controllo, sarà necessario in futuro proporre un task di discriminazione uditiva, che non comporti la fase di etichettatura e che permetta anche di valutare con maggiore accuratezza la capacità di differenziazione dei due fonemi indipendentemente dall'interfaccia grafica.

### 2.1.3. Scelta degli stimoli

Per i task di identificazione sono stati scelti quattro gruppi di non-parole con foni bersaglio [ʌ], [j] [lj] e [lj]. Per il primo task sono stati utilizzati tutti gli stimoli, per il secondo invece solo gli stimoli con [ʌ] e [j].

Tutti gli stimoli presentano la stessa struttura fonotattica: sono tutte non-parole bisillabiche e parossitone (in questi task, mantenere costante queste

caratteristiche è necessario per evitare eventuali effetti di lunghezza dello stimolo e di posizione dell'accento).

Si è scelto di utilizzare parole inesistenti per evitare che variabili lessicali incidessero sulla prova di identificazione. Inoltre, sicuramente, conoscendo la corretta grafia della parola, l'ascolto sarebbe stato orientato.

Il contesto in cui si presenta il fono target è stato limitato a quello intervocalico, con vocale precedente allo stimolo accentata, in modo da non rendere eccessivamente lungo il compito di identificazione. Inoltre, è stata evitata la presenza di vocali palatali poiché uno studio condotto sullo spagnolo proprio riguardo la distinzione tra /ʌ/ e /j/ ha mostrato che esse potrebbero condizionare l'identificazione in favore di [j] (Rost Bagudanch 2014a).

Si è controllato che nessuna delle non-parole scelte fosse il toponimo di una località conosciuta o il nome proprio di una persona famosa.

Gli stimoli scelti sono riportati nella Tabella 1.

Tabella 1: Stimoli selezionati per il task di identificazione (esperimenti 1 e 2).

	Task 1 e 2		Solo task 1	
	ʌ	j	lj	llj
['a]_[a]	ciàglia	ciàia	ciàlia	ciàllia
['a]_[o]	nàgljo	nàio	nàlio	nàllio
['ɔ]_[o]	rògljo	ròio	ròlio	ròllio
['ɔ]_[a]	tòglia	tòia	tòlia	tòllia

### 2.1.4. Trattamento degli stimoli

Una volta selezionate le non-parole per il test, è stato registrato uno speaker professionista, Massimo Giardini, in una camera silente, che ha ripetuto gli stimoli in isolamento.

In seguito, sono state uniformate con l'ausilio del software di editing audio GoldWave le durate delle vocali antecedenti ai foni target [k], [kj], da una parte, e [j], [lj], dall'altra, a 145 ms ( $\pm 5$  ms) e 160 ( $\pm 5$  ms).

Si è poi calcolato il valore medio della durata di [k] nelle quattro non parole (*ciaglia, nàglio, ròglio, tòglia*): 158 ms. Sono quindi stati uniformati a questo valore tutti gli stimoli (con un'approssimazione di 2 ms), e poi mantenendo la vocale precedente identica si è fatta variare per il primo e il secondo task

la durata del fono target riducendola di volta in volta di 40 ms e creando quattro nuovi stimoli con [k] di durata 220 ms, 180 ms, 140 ms e 100 ms (v. Figura 8, a sinistra).

Per il secondo test è stata fatta la stessa operazione con gli stimoli che presentano il fono [j], nell'audio originale di durata media di 180 ms, manipolato in modo da riprodurre lo stesso continuum di [k] (260 ms, 220 ms, 180 ms, 140 ms, 100 ms) (v. Figura 8, a destra).

In questo modo si è potuto sottoporre a ciascun partecipante l'ascolto di stimoli esattamente identici fatta eccezione per la durata del fono target. Questo permette di fare alcune interessanti osservazioni sulla sensibilità dei partecipanti al test rispetto a questo tratto di durata.

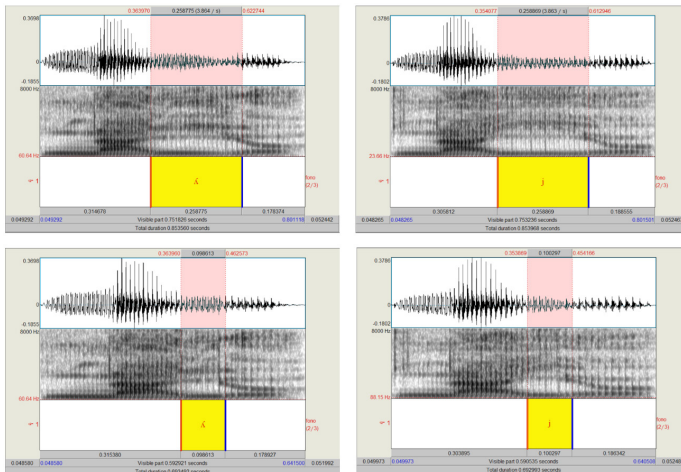


Figura 8: A sinistra, esempi di stimoli di durata variabile per [k] (da 260 ms circa dello stimolo naturale, in alto, a 100 ms circa, in basso) e, a destra, per [j] (da 100 ms circa, in basso, a 260 ms circa, in alto – in questo caso la durata dello stimolo naturale era 180 ms circa).

## 2.2. Analisi dei dati

### 2.2.1. Esperimento 1

Dal primo task è emerso che in linea generale i partecipanti sono stati in grado di distinguere correttamente /lj/ e /llj/ rispetto agli altri due casi in esame (v. Tabella 2, Figura 9). Utiliz-

Tabella 2: Dati riassuntivi risultati dal task di identificazione percettiva. Esperimento 1. Classificazione di /k/, /j/, /lj/, /llj/.

	<gb>	<i>	<li>	<lli>
/k/*	69%	22%	6%	4%
/j/	23%	70%	3%	4%
/lj/	2%	1%	91%	7%
/llj/	3%	1%	5%	92%

\*NB: per quanto riguarda /k/ si è scelto di riportare in questa tabella e nelle successive i risultati ottenuti dalla somministrazione degli stimoli aventi come fono bersaglio [k] con durata di 180 ms, il valore medio tra quelli testati, in modo che i grafici risultino più leggibili. In seguito si discuteranno maggiori dettagli sull'effetto della durata nella classificazione di questo fono. Inoltre inserendo qui i valori di durata di [k] =180 ms si confrontano [k] e [j] con la stessa durata.

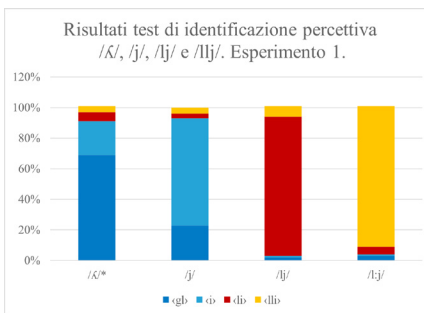


Figura 9: Risultati task di identificazione percettiva preliminare. Classificazione di /k/, /j/, /lj/, /llj/. Esperimento 1.

zando il software SPSS si è calcolato il valore di Chi-quadrato e di V di Cramer da un lato di /k/ e /j/, dall'altro di /lj/ e /llj/, posto  $\alpha=0,5$ . Ne è risultato che il grado di associazione di /lj/ e /llj/ con le rispettive stringhe grafiche <li> e <lli> è decisamente più alto ( $\chi^2=606,696$ ;  $V=0,871$ ) rispetto invece a /k/ e /j/ ( $\chi^2=191,701$ ;  $V=0,490$ ).

Questo dato è tanto più significativo perché la difonizzazione di [k]/[k:] in [lj] e la monofonizzazione del nesso [lj] in [k] sono considerati tra i tratti caratteristici della varietà di italiano regionale piemontese (Ledgeway, 2016; Telmon, 1993) e il 62% dei soggetti ha indicato come luogo di provenienza il Piemonte, dichiarando di vivere in questa regione da più di dieci anni (cfr §2.1.2).

Quindi, poiché nonostante le specifiche tradizionalmente attribuite all'italiano regionale piemontese il confine categoriale sembra essere chiaro, si è potuto stringere il campo di ricerca e concentrarsi su [k] e [j], la cui distinzione già in questo primo esperimento risulta piuttosto incerta in particolare per quanto riguarda le non parole *ciaglia* e *nàglio* (§2.2.1.2).

Nell'ambito di questo primo task oltre a verificare le capacità di etichettare i foni proposti, si è scelto di valutare la sensibilità dei partecipanti rispetto alla durata di [k] valutando se essa incida sulla classificazione. Come si è discusso, infatti uno dei tratti di variazione nella rea-

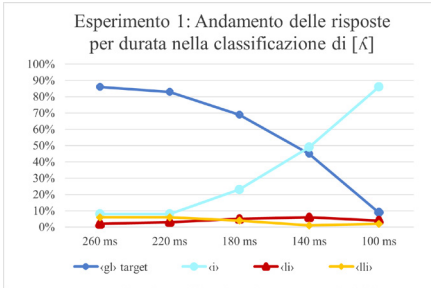


Figura 10: Andamento per durata nella classificazione di [k]. Esperimento 1.

lizzazione di questo fonema è la durata.

Per questa ragione ciascuna non parola è stata proposta con il fono bersaglio [k] di durata rispettivamente: 260 ms, 220 ms, 180 ms, 140 ms e 100 ms.

È risultato che effettivamente il tratto durato avesse un ruolo significativo nel corretto riconoscimento del fono ( $r$  di Pearson=0,947\*\*,  $p=0,01$ )<sup>16</sup>. Al diminuire della durata del fono bersaglio infatti il numero di risposte target (<gl>) è sceso al punto da arrivare appena al 9% per gli stimoli con [k] di durata pari a 100 ms (Figura 10). Dati questi primi risultati, si è scelto di approfondire il discorso proponendo un task in cui si modificasse la lunghezza non solo di [k] ma anche di [j] (v. §2.1.1).

16. Le variabili prese in considerazione per il calcolo della correlazione lineare sono il continuum di variazione della durata del fono target, da 260 ms a 100 ms, e il numero di risposte target (<gl>).

### 2.2.1.1. Tempi di risposta – Esperimento 1

Oltre all’accuratezza nello svolgimento del compito, sono stati registrati i tempi di risposta dei partecipanti al task, per verificare se ci fossero stimoli sui quali essi si sono soffermati di più per via della loro difficoltà.

Poiché non è stato imposto un tempo minimo di risposta, i dati sono molto dispersi. Si è scelto di prendere in considerazione per i test statistici solo i tempi inferiori ai 5000 ms.

Pur imponendo questo limite, il range dei dati è molto vasto (intervallo: 4862,00), quindi, a fianco dei dati di media e deviazione standard, si riporta anche la mediana (Tabella 3).

Tabella 3: Dati riassuntivi statistiche descrittive tempi di risposta. Esperimento 1.

[k]	<gl> target	<i>	<li>	<lli>
260 ms	86%	8%	2%	6%
220 ms	83%	8%	3%	6%
180 ms	69%	23%	5%	4%
140 ms	45%	49%	6%	1%
100 ms	9%	86%	4%	2%

Una volta eliminati gli *outlier*, è stata verificata la normalità della distribuzione dei dati grazie al test Shapiro-Wilk ( $W=0,886$ ,  $p < 0,001$ ). Infine, applicando il test ANOVA univariato è emerso che le medie tra i tempi di risposta ai singoli stimoli sono statisticamente significative ( $F(8,242)=7$ ,  $p < 0,001$ ). Alcuni stimoli, per essere riconosciuti, hanno quindi richiesto significativamente più tempo che altri.

Emerge sorprendentemente che gli stimoli con [lj] e [l:] hanno richie-

sto un tempo maggiore, pur essendo quelli identificati con maggiore chiarezza. D’altro canto sono stati registrati comportamenti simili anche sugli stimoli con [ʌ] di durata 180 ms, 140 ms e 100 ms, che nel task hanno dato risultati progressivamente più devianti rispetto al target fino a non essere riconosciuti nell’81% delle risposte (Figure 11 e 12). Per entrare maggiormente nel dettaglio su quali differenze tra i grup-

pi sono effettivamente significative è stato necessario applicare un test post hoc (il test di Tukey) in modo da operare un confronto multiplo tra le medie. Ne è emerso che i tempi di risposta sugli stimoli con [lj] e [lʲ] si discostano significativamente da tutti gli altri fatta eccezione per quelli con [ʌ] a 180, 140 e 100 ms. Al contrario per questi ultimi il confronto è significativo solo rispetto agli stimoli con [ʌ] a 220 ms.

Questi primi dati sui tempi di risposta sebbene suggestivi, tuttavia risentono di alcuni limiti dovuti al metodo di raccolta. Il task di identificazione è stato somministrato online, questo comporta certamente che alcuni dati possono aver subito le conseguenze di eventuali problemi di connessione. Inoltre, i partecipanti hanno utilizzato i propri dispositivi personali, le risposte sono state quindi date sia utilizzando schermi *touch* sia con l’utilizzo del mouse. Dall’analisi ANOVA univariata è emerso che i tempi di risposta di chi ha utilizzato un dispositivo *touch* sono significativamente inferiori rispetto agli altri ( $F(27,891)=1$ ;  $p < 0,001$ ).

Dato questo limite, potrebbe essere utile ripetere l’esperimento facendo utilizzare ai partecipanti un software che non prevede l’utilizzo di internet e uniformare i dispositivi utilizzati, preferendo o i dispositivi *touch* o un software che permetta di

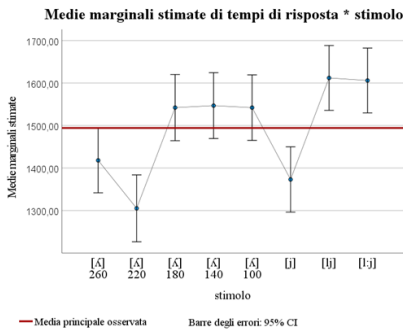


Figura 11: Grafico di confronto tra le medie dei tempi di reazione per stimolo diverso ricavate dall’analisi ANOVA a una via. Esperimento 1.

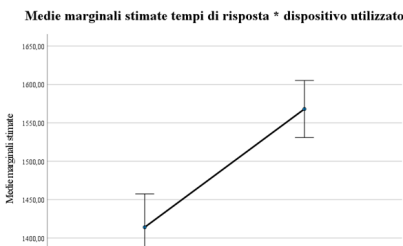


Figura 12: Grafico di confronto tra le medie dei tempi di reazione per dispositivo utilizzato (*touch* o non) ricavate dall’analisi ANOVA a una via. Esperimento 1.

rispondere al quesito con la tastiera del computer, associando a due tasti le possibili risposte.

### 2.2.1.2 Effetto del contesto vocalico – Esperimento 1

Dopo aver presentato i risultati complessivi del primo task di identificazione, è necessario aggiungere qualche considerazione su come variano le risposte dei partecipanti a seconda dello stimolo, al variare quindi del contesto vocalico in cui si trova il fono bersaglio.

Osservando le percentuali di risposta target date dai partecipanti per stimolo, si nota che mentre per quando riguarda gli stimoli con [lj] e [l:j] il comportamento dei partecipanti è stato piuttosto omogeneo, al contrario le risposte per gli stimoli con foni [k] hanno un ampio arco di variazione a seconda del contesto vocalico in cui si trovavano gli stimoli (Tabella 4, Figure 13-16).

Anche per quanto riguarda il tratto durata la soglia di riconoscibilità di /k/ varia a seconda del contesto (v. Figure 17-20).

Tabella 4: Dati riassuntivi del task di identificazione percettiva (Esperimento 1), dati distribuiti per contorno vocalico.

Stimolo	Media	Dev. Std.	Mediana
[k] 260 ms	1418,0	842,0	1186,0
[k] 220 ms	1305,1	712,1	1115,0
[k] 180 ms	1542,3	825,6	1336,0
[k] 140 ms	1546,9	805,7	1330,5
[k] 100 ms	1542,2	773,8	1377,0
[j]	1373,1	790,2	1166,0
[lj]	1612,1	756,0	1406,0
[l:j]	1606,1	676,7	1527,5

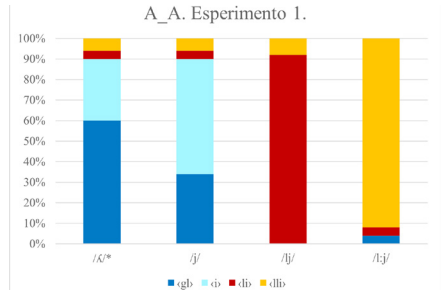


Figura 13: Grafico riassuntivo risposte target \* contesto vocalico A\_A. Esperimento 1.

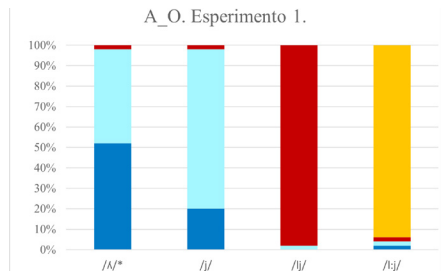


Figura 14: Grafico riassuntivo risposte target \* contesto vocalico A\_O. Esperimento 1.

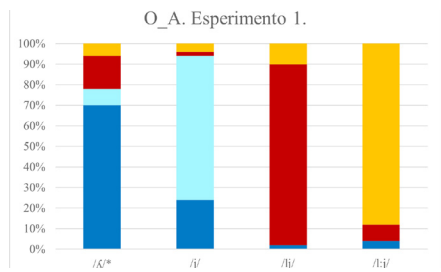


Figura 15: Grafico riassuntivo risposte target \* contesto vocalico O\_A. Esperimento 1.

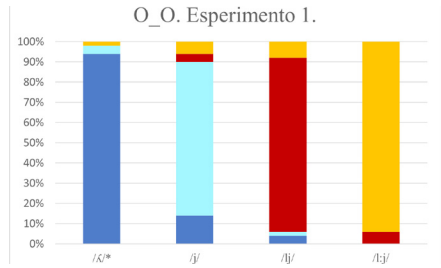


Figura 16: Grafico riassuntivo risposte target \* contesto vocalico O\_O. Esperimento 1.

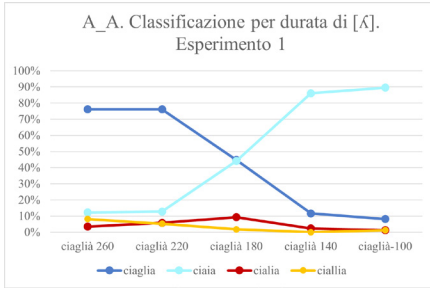


Figura 17: Andamento per durata nella classificazione di [ʎ] \* contorno vocalico A\_A. Esperimento 1.

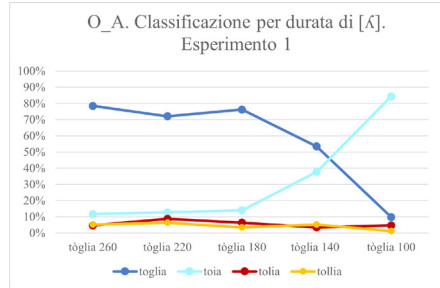


Figura 19: Andamento per durata nella classificazione di [ʎ] \* contorno vocalico O\_A. Esperimento 1.

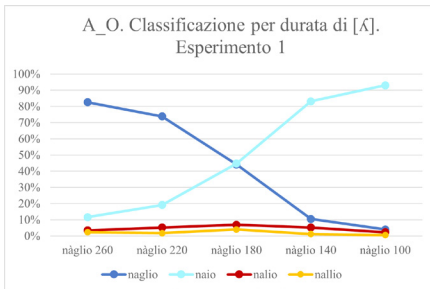


Figura 18: Andamento per durata nella classificazione di [ʎ] \* contorno vocalico O\_A. Esperimento 1.

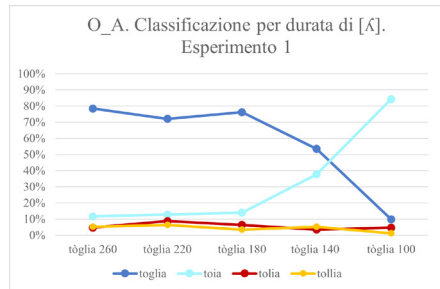


Figura 20: Andamento per durata nella classificazione di [ʎ] \* contorno vocalico O\_O. Esperimento 1.

Si noti come [ʎ] in contesto O\_O viene associato a «g» in più dell'80% delle risposte; al contrario, nei contesti A\_A e A\_O, le risposte target scendono a meno del 60%.

A partire dai dati fin qui presentati sembrerebbe che la vocale che incide maggiormente sui risultati del task sia quella che precede il fono bersaglio. Gli stimoli con fono bersaglio in contesto A\_A e A\_O disegnano profili simili, diversi invece da quelli in contesto O\_A e O\_O. Bisogna tuttavia tenere in considerazione che in tutti gli stimoli tale vocale era anche la tonica, quindi non

è chiaro se abbia un ruolo o meno la posizione dell'accento oppure questo fenomeno dipenda in generale da una maggior incidenza della vocale precedente a [ʎ] e [j].

Poiché in letteratura non è trattata adeguatamente la distinzione acustica tra [ʎ] e [j] è difficile avanzare ipotesi precise sul motivo per cui un timbro posteriore arrotondato della vocale precedente possa favorire il riconoscimento del fono.

Per questo è stata svolta una prima analisi preliminare di confronto acustico tra gli stimoli presenti nel task



pronunciati da due parlanti diversi di sesso maschile e femminile. Ne è emerso che le strategie usate per differenziare i due foni sono differenti tra i due parlanti, ma anche nelle singole riproduzioni dei diversi stimoli. L'unica costante che sembra emergere è che sebbene nel passaggio dalla vocale a [ʌ] e a [j] ci sia un calo di energia di circa 10 dB (come già notato da Pettorino & Giannini, 1991), tuttavia in un caso tale calo di energia sembra essere maggiormente repentino che nell'altro. Questo avviene indistintamente per tutte le non-parole a prescindere dal contesto vocalico. Tuttavia potrebbe giustificare la maggiore attenzione da riservare alla vocale prima rispetto a quella seguente il fono target. Il fatto poi che il contesto O\_ favorisca la corretta identificazione di /ʌ/ e /j/ potrebbe poi dipendere dal maggiore arco che deve compiere la seconda formante nel passaggio tra vocale e approssimante.

Tuttavia per chiarire meglio la questione sarà necessario svolgere uno studio acustico maggiormente approfondito sulle realizzazioni di /ʌ/ e /j/ che coinvolga un numero più rappresentativo di parlanti.

### *2.2.2. Esperimento 2*

Come già visto nel primo esperimento, il fattore durata del fono bersaglio si è rivelato molto determinante nelle scelte dei soggetti esaminati nel task

di identificazione percettiva proposto.

I dati raccolti nel primo task sono stati poi confermati dal secondo esperimento e sono di seguito presentati.

Anche in questo esperimento erano proposti stimoli uditivi manipolati (v. sopra) e si chiedeva se la grafia da associare fosse di parole con ⟨g⟩ o ⟨j⟩.

I risultati dimostrano che quando il fono [ʌ] ha durata superiore a 220 ms viene solidamente associato (in più del 90% delle risposte) alla trascrizione target ⟨g⟩; parallelamente lo stesso avviene per gli stimoli con [j] a 100ms associati a ⟨j⟩.

C'è un rapporto direttamente proporzionale tra la durata di [ʌ] ( $r$  di Pearson = 0,935;  $p$  = 0,02; Figura 21) e l'associazione target: maggiore è la durata del fono migliore è il suo riconoscimento. Mentre, sorprendentemente, un rapporto inversamente proporzionale appare nel riconoscimento di [j] ( $r$  di Pearson = -0,988;  $p$  = 0,002; Figura 22, pag. 22).

Parallelamente a quanto visto al §2.2.1, per quanto riguarda [ʌ] sotto la soglia dei 180 ms non viene più identificato con certezza e arriva a essere scambiato per /j/ dall'81% dei partecipanti quando ha durata di 100 ms. D'altro canto [j] viene quasi univocamente associato a ⟨j⟩ quando ha durata 100 ms e 140 ms.

Significativamente però le curve che descrivono la classificazione di [ʌ] e [j] come ⟨g⟩ e ⟨j⟩ a seconda della durata

hanno un profilo non del tutto simmetrico (v. Tabella 5, Figura 21).

Il dato più significativo riguarda le non-parole aventi [j] di durata maggiore, che vengono associate a grafie alternativamente con <gb> e <ì>.

Sembra che entrino in conflitto parametri diversi: la qualità del fono e la sua durata. Ne consegue che il fono [j] di durata maggiore di 180 ms non è chiaramente classificabile né come /ɰ/ né come /j/. Questo dato suggerisce che la durata del fono venga considerata come indice significativo dagli uditori. In effetti, in italiano la durata caratterizza e distingue i due fonemi.

Ne consegue che al variare del-

la lunghezza degli stimoli target la loro conversione grafo-fonemica non è del tutto trasparente. Questo dato può avere importanti conseguenze sia perché in alcune varianti regionali /ɰ/ non si allunga regolarmente in posizione intervocalica, sia perché nel parlato spontaneo e poco sorvegliato potrebbe non raggiungere la durata standard.

Inoltre, si conferma il fatto che i confini categoriali tra /ɰ/ e /j/ sono piuttosto labili.

### 2.2.2.1. Tempi di risposta – Esperimento 2

Come per l'Esperimento 1, sono stati raccolti i dati sui tem-

Tabella 5: Dati riassuntivi risultati del task di identificazione percettiva per la classificazione di [ɰ] e [j] al variare della durata. Esperimento 2.

A A	<gb>	<ì>	<di>	<lli>	A O	<gb>	<ì>	<di>	<lli>
[ɰ]*	60%	30%	4%	6%	[ɰ]*	52%	46%	2%	0%
[j]	34%	56%	4%	6%	[j]	20%	78%	2%	0%
[lj]	0%	0%	92%	8%	[lj]	0%	2%	98%	0%
[t:j]	4%	0%	4%	92%	[t:j]	2%	2%	2%	94%
O A	<gb>	<ì>	<di>	<lli>	O O	<gb>	<ì>	<di>	<lli>
[ɰ]*	70%	8%	16%	6%	[ɰ]*	94%	4%	0%	2%
[j]	24%	70%	2%	4%	[j]	14%	76%	4%	6%
[lj]	2%	0%	88%	10%	[lj]	4%	2%	86%	8%
[t:j]	4%	0%	8%	88%	[t:j]	0%	0%	6%	94%

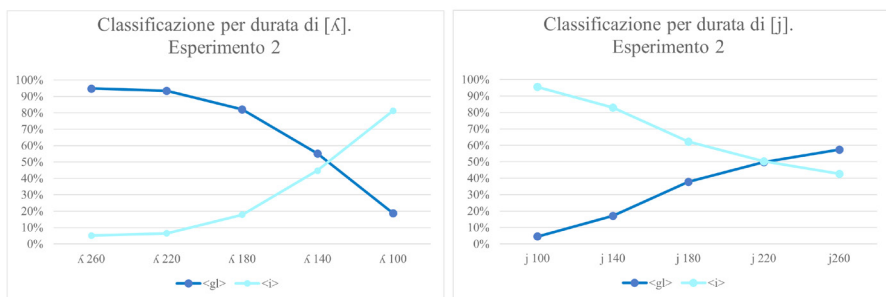


Figura 21: Andamento per durata nella classificazione di [ɰ] (a sinistra), classificazione di [j] (a destra). Esperimento 2.

Tabella 6: Dati riassuntivi statistiche descrittive tempi di risposta. Esperimento 2.

	<gl>	<i>		<gl>	<i>
κ 260ms	95%	5%	j 260ms	57%	43%
κ 220ms	93%	7%	j 220ms	50%	50%
κ 180ms	82%	18%	j 180ms	38%	62%
κ 140ms	55%	45%	j 140ms	17%	83%
κ 100ms	19%	81%	j 100ms	5%	95%

più di risposta dei partecipanti.

Come nel precedente, i dati sono molto disomogenei e il range di variazione è alto. La Tabella 6 riassume le statistiche descrittive per stimolo.

Anche qui sono stati esclusi i valori superiori a 5000 ms ed è stata verificata la normalità della distribuzione dei dati grazie al test Shapiro-Wilk ( $W = 0,803, p < 0,001$ ). Poi è stato applicato il test ANOVA univariato ed è emerso che le medie tra i tempi di risposta ai singoli stimoli sono statisticamente significative ( $F(5,608) = 9, p < 0,001$ ).

Osservando il profilo che disegna le medie marginali stimate, si nota immediatamente che riproduce quello dell'accuratezza delle risposte dei partecipanti. Per quanto riguarda gli stimoli con /k/, più il fono bersaglio è breve, maggiore è il tempo di risposta e parallelamente maggiore è la percentuale di errore. Al contrario per quanto riguarda gli stimoli con /j/ i valori sono più compatti; solo per [j] con 100ms si registrano valori sotto la media complessiva.

Per sapere quali differenze tra i gruppi sono effettivamente significative è stato necessario come prima applicare un test post hoc (il test di Tukey) in modo da operare un confronto multiplo tra le medie.

Come ricordato in §2.2.1.1 i dati relativi ai tempi di risposta per il modo in cui sono stati raccolti non sono del tutto affidabili. Tuttavia, correlano positivamente con l'idea che quando [k] viene presentato con una durata inaspettata rispetto a quella prototipica si presentino problemi nel suo riconoscimento.

### 2.2.2.2. Effetto del contorno vocalico – Esperimento 2

Infine, parallelamente a quanto visto in §2.2.1.2, anche in questo secondo esperimento il contesto vocalico ha un effetto sul comportamento dei partecipanti al task.

La coppia di non-parole per la quale è più valido il quadro descritto fin qui

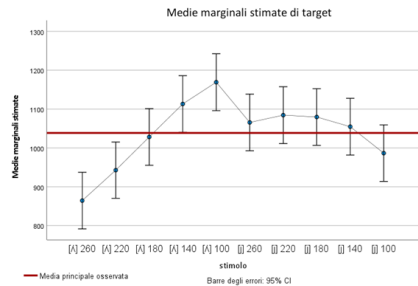


Figura 22: Grafico di confronto tra le medie dei tempi di reazione per stimolo diverso ricavate dall'analisi ANOVA a una via. Esperimento 2.

Tabella 7: Dati riassuntivi del task di identificazione percettiva (Esperimento 2), dati distribuiti per contorno vocale.

	ciaglia	ciaia	nàglio	naio	tòglio	toglia	toia	ròglio	roio		
ciaglia 260 ms	95%	5%	nàglio 260 ms	92%	8%	tòglio 260 ms	95%	5%	ròglio 260 ms	97%	3%
ciaglia 220 ms	93%	7%	nàglio 220 ms	94%	6%	tòglio 220 ms	91%	9%	ròglio 220 ms	95%	5%
ciaglia 180 ms	84%	16%	nàglio 180 ms	65%	35%	tòglio 180 ms	85%	15%	ròglio 180 ms	94%	6%
ciaglia 140 ms	36%	64%	nàglio 140 ms	26%	74%	tòglio 140 ms	75%	25%	ròglio 140 ms	83%	17%
ciaglia 100 ms	3%	97%	nàglio 100 ms	11%	89%	tòglio 100 ms	16%	84%	ròglio 100 ms	44%	56%
ciaia 260 ms	63%	38%	nàio 260 ms	53%	47%	tòia 260 ms	49%	51%	ròio 260 ms	65%	35%
ciaia 220 ms	57%	43%	nàio 220 ms	55%	45%	tòia 220 ms	41%	59%	ròio 220 ms	47%	53%
ciaia 180 ms	55%	45%	nàio 180 ms	43%	57%	tòia 180 ms	30%	70%	ròio 180 ms	24%	76%
ciaia 140 ms	28%	72%	nàio 140 ms	23%	77%	tòia 140 ms	8%	92%	ròio 140 ms	9%	91%
ciaia 100 ms	3%	97%	nàio 100 ms	7%	93%	tòia 100 ms	2%	98%	ròio 100 ms	6%	94%

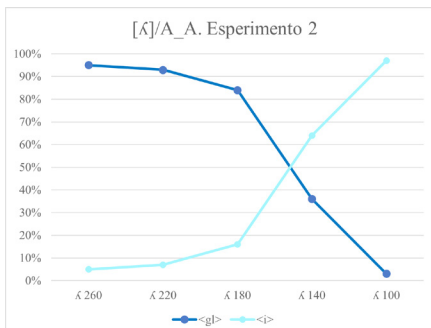


Figura 23: Grafico riassuntivo risposte target ( $[[k] \sim gb)$  \* contorno vocale A\_A. Esperimento 2.

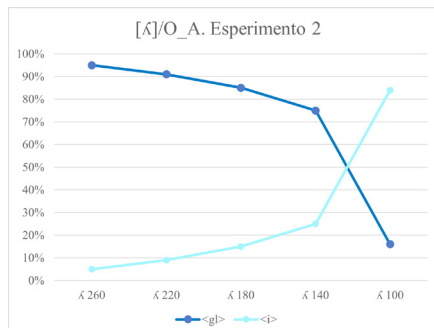


Figura 25: Grafico riassuntivo risposte target ( $[[j] \sim \dot{\iota})$  \* contorno vocale A\_A. Esperimento 2.

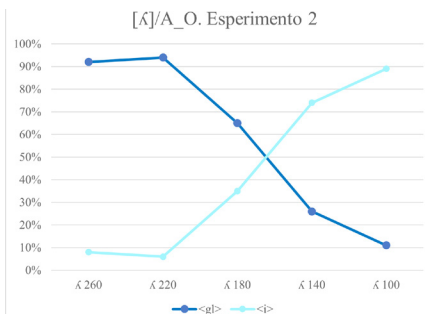


Figura 24: Grafico riassuntivo risposte target ( $[[k] \sim gb)$  \* contorno vocale A\_O. Esperimento 2.

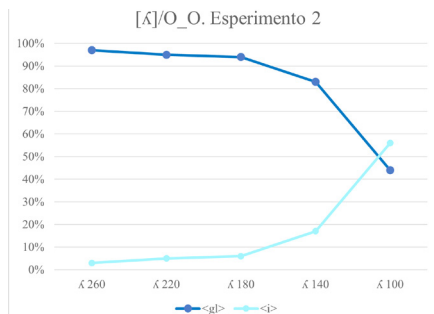


Figura 26: Grafico riassuntivo risposte target ( $[[j] \sim \dot{\iota})$  \* contorno vocale A\_O. Esperimento 2.

2.2.2.2. Effetto del contorno vocalico – Esperimento 2

Infine, parallelamente a quanto visto in §2.2.1.2, anche in questo secondo esperimento il contesto vocalico ha un effetto sul comportamento dei partecipanti al task.

La coppia di non-parole per la quale è più valido il quadro descritto fin qui è *naglio/naio*. Invece, anche in questo secondo esperimento, le risposte sono state più accurate per [ɹ]/O\_A, O\_O (Tabella 7, Figure 23-26 e Figure 27-30).

2.2.2.3. Accuratezza nelle risposte dei partecipanti in fascia di rischio DSA

Come visto nella presentazione dei partecipanti (§2.1.2) è stato somministrato agli uditori dei due esperimenti il *Questionario Vinegrad Plus* per verificare se ci fosse nel campione una percentuale di persone che si collocano nella fascia di rischio di Disturbo Specifico di Apprendimento. Poiché è emerso che il 33% effettivamente si trova in questa fascia e vista la possibile correlazione tra DSA e esito del task, sono stati incrociati i dati relativi

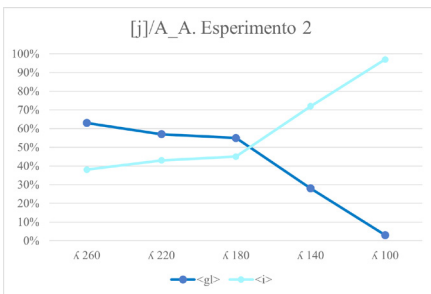


Figura 27: Grafico riassuntivo risposte target ([j] ~ ÷) \* contorno vocalico O\_A. Esperimento 2.

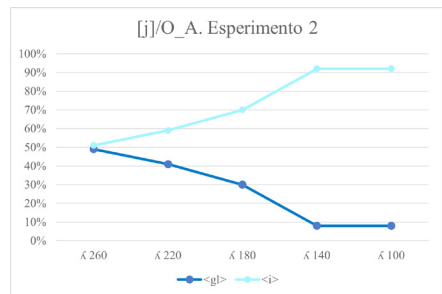


Figura 29: Grafico riassuntivo risposte target ([k] ~ gb) \* contorno vocalico O\_O. Esperimento 2.

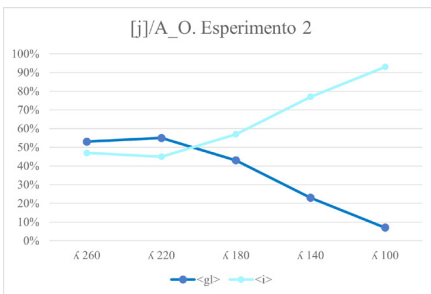


Figura 28: Grafico riassuntivo risposte target ([j] ~ ÷) \* contorno vocalico O\_A. Esperimento 2

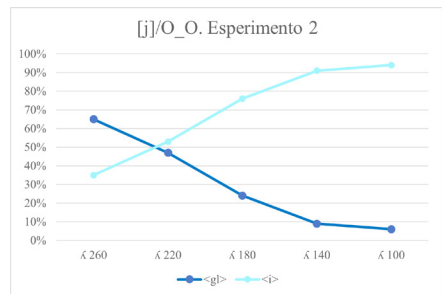


Figura 30: Grafico riassuntivo risposte target ([j] ~ ÷) \* contorno vocalico O\_O. Esperimento 2.

l'accuratezza delle risposte con la presenza o assenza di rischio DSA.

Per verificare questa ipotesi è stato applicato il test del Chi Quadro ed è stato calcolato il valore di  $\chi^2$  e di  $V$  di Cramer. Questa analisi non ha dato risultati significativi ( $\chi^2 = 0,626$ ;  $V = 0,013$ ).

Il fatto suggerisce che le difficoltà nella classificazione di / $\Delta$ / e / $j$ / non siano caratteristiche solo di chi ha difficoltà specifiche, ma più diffuse nella popolazione. Questo dato rinforza l'idea che sia necessario dare un'attenzione specifica a questi fonemi in fase di diagnosi di DSA; in modo da non sovrastimare gli errori compiuti dall'eventuale soggetto al quale si stia somministrando un test di diagnosi di disortografia.

### 3. Conclusioni e prospettive

Sono stati sottoposti a un totale di 113 studenti universitari italo-fonemi due task identificazione percettiva.

Dai risultati di tali task è emerso che mentre [k] e [j] sono stati saldamente identificati dagli uditori, i quali, al contrario, hanno avuto maggiori difficoltà nel riconoscimento di [Δ] e [j].

In particolare, si è notato che la durata ha un ruolo molto importante nel determinare la classificazione di tali fonemi. Inoltre, i dati mostrano che anche il contesto vocalico determina una migliore o peggiore riuscita del task.

Quanto osservato ha importanti conseguenze sull'ortografia. Bisogna tenere in considerazione infatti che se a livello percettivo la distinzione tra i fonemi non è chiara, nella scrittura non può essere utilizzata la via indiretta (di trascrizione fonogrammatica) o comunque rischia di non essere affidabile. Questo comporta la necessità di recuperare l'immagine ortografica della parola per poter scegliere la corretta ortografia. Ne consegue che gli errori che eventualmente emergono devono essere considerati in maniera differente che se la loro conversione fosse del tutto trasparente. Inoltre, bisogna applicare una particolare cautela nella creazione di liste di non parole da dettare durante i test di diagnosi della disortografia, evitando di inserire stimoli che presentino questi fonemi.

Alla luce degli esiti di questi due primi esperimenti, inoltre, emerge la necessità di proseguire la ricerca indagando sulla capacità di discriminazione dei due fonemi a prescindere dalla loro ortografia, attraverso un task che non preveda l'associazione di uno stimolo con una resa ortografica.

Inoltre, sarebbe interessante verificare se il livello di neutralizzazione dell'opposizione tra / $\Delta$ / e / $j$ / abbia anche ragioni anagrafiche, se sia quindi un fenomeno in espansione. Bisognerebbe inoltre verificare la capacità di discriminazione e identificazione dei due fonemi in fasce di età differen-

ti, oltre che estendere geograficamente la ricerca.

Infine, data la scarsità di approfondimento in bibliografia sarebbe interessante condurre uno studio di tipo acustico su un vasto campione di parlanti italofoni.

## Bibliografia

Azevedo M.M. (2005). *Portuguese: A linguistic introduction*. Cambridge: Cambridge University Press.

Banai K. & Ahissar M. (2018). Poor sensitivity to sound statistics impairs the acquisition of speech categories in dyslexia. *Language, Cognition and Neuroscience*, 33(3), 321-332. <https://doi.org/10.1080/23273798.2017.1408851>.

Bertinetto P.M. & Loporcaro M. (2005). The sound pattern of Standard Italian, as compared with the varieties spoken in Florence, Milan and Rome. *Journal of the International Phonetic Association*, 35(2), 131-151. <https://doi.org/10.1017/S0025100305002148>.

Bindelli D., Depretis D., Fasola A., Folisi K., Marzorati D., Profumo E., Serafino R. & Torcellini F. (2009). La comorbidità tra dislessia, disortografia, disgrafia, discalculia nella scuola secondaria di secondo grado. *Dislessia*, 6, 59-76.

Bladon R. & Carbonaro E. (1978). Lateral consonants in Italian. *Journal of Italian Linguistics Amsterdam*, 3(1), 43-54. <https://doi.org/10.1017/S0025100306002398>.

Borleffs E., Maassen B.A., Lyytinen H. & Zwarts F. (2017). Measuring orthographic transparency and morphological-syllabic complexity in alphabetic orthographies: a narrative review. *Reading and writing*, 30(8), 1617-1638. <https://doi.org/10.1007/s11145-017-9741-5>.

Bortolini U. (1995). *Manuale PFLI (Prove per la Valutazione Fonologica del Linguaggio Infantile)*. Padova: Edit Master.

Camilli A. (1911). Dubbi di pronunzia. *Le Maître Phonétique*, 3-4, 50-51.

Camilli A. (1929). Il dialetto di Servigliano (AP), *Archivum Romanicum*, XIII, 220-271 (ripubbl. a cura di F. Paci, Comune di Servigliano, 2002).

Canepari L. (1992). *Manuale di pronuncia italiana: con un pronunciario di oltre 30000 voci e due audiocassette C45*. Bologna: Zanichelli.

Cavalli E., Duncan L.G., Elbro C., El Ahmadi A. & Colé P. (2017). Phonemic-Morphemic dissociation in university students with dyslexia: An index of reading compensation? *Annals of dyslexia*, 67(1), 63-84. <https://doi.org/10.1007/s11881-016-0138-y>.

Celata C. (2006). *Analisi del processo di retroflessione consonantica in area romanza, con dati sperimentali dal corso e dal siciliano*. Tesi di dottorato inedita. Pisa: Scuola Normale Superiore.

Cerruti M. (2013). «Varietà dell'italiano». In G. Iannaccaro (a cura di), *La linguistica italiana all'alba del terzo millennio (1997-2010)*, Roma: Bulzoni, 91-127.

Coletti V. (2015). *Grammatica dell'italiano adulto: l'italiano di oggi per gli italiani di oggi*. Bologna: Il Mulino.

Colonna V. & Romano A. (2018), «La variazione diatopica nel micro-spazio dialettale leccese: il dialetto salentino delle frazioni di Vernole». In G. Caramuscio & A. Romano (a cura di), *Una d'arme, di lingua, d'altare, di memorie, di sangue, di cor - Omaggio a Luciano Graziuso*, Lecce: Grifo, 105-123.

Cornoldi C. (2019). *I disturbi dell'apprendimento*. Bologna: Il Mulino.

Cornoldi C. & Montesano L. (2020). *Nuova batteria per studenti universitari e adulti LSC-SUA*. Trento: Erickson.

Danelli L., Berlingeri M., Bottini G., Borghese N.A., Lucchese M., Sberna M., Price C.J. & Paulesu E. (2017). How many deficits in the same dyslexic brains? A behavioural and fMRI assessment of comorbidity in adult dyslexics. *Cortex*, 97, 125-142. <https://doi.org/10.1016/j.cortex.2017.08.038>.

Endo R. & Bertinetto P.M. (1999). Caratteristiche prosodiche delle così dette 'rafforzate' italiane. In R. Delmonte & A. Bristot (a cura di), *Aspetti computazionali in fonetica, linguistica e didattica delle lingue: modelli e algoritmi. Atti delle IX Giornate di Studio del Gruppo di Fonetica Sperimentale, Venezia 17-19 dicembre 1998*. Venezia: Università Ca' Foscari, 243-255.

Fenzi V. & Cornoldi C. (2015). «Le difficoltà ortografiche di adolescenti con dislessia». *Dislessia*, 12(1), 75-86.

Fernández Rei E., Aguete Cajiao A., Osorio Peláez C. & Cutrín Garabal J.A. (2021). FOLERPA: Ferramenta On-Line para ExpeRimentación PerceptivA. Santiago de Compostela: Instituto da Lingua Galega. <<https://ilg.usc.gal/folepa/>>.

Fontanot A. (2019). *Aspetti di italiano «composito» nella pronuncia di giovani torinesi: un'indagine su corpora* [Tesi di laurea magistrale, Università degli studi di Torino].

Galatà V., Meneguzzi G., Conter L. & Zmarich C. (2012). *Primi dati sull'acquisizione fonetico-fonologica dell'italiano L2 in prescolari rumeni*. In A. Paoloni & M. Falcone (a cura di), *La voce nelle applicazioni*, Roma: Bulzoni, 35-50.

Genre A. (1979). «Appunti sulla grafia del piemontese», *Rivista Italiana di Dialettologia*, 3, 311-342.

Gómez R. & Molina Martos I. (2013). *Variación yeísta en el mundo hispánico*. Madrid: Iberoamericana Editorial Vervuert (Vol. 32).

Granatiero F. (2021). *Scrivere la lingua madre*, Roma: Cofine.

Iannàccaro G. (2015). «Vedere il dialetto. Vocabolari bresciani e ortografie spontanee». In M. Piotti (a cura di), *Dalla scripta all'italiano. Aspetti momenti figure di storia linguistica bresciana*, Brescia: Morcelliana, 225-262.

Ledgeway A. (2016). Italian, Tuscan, and Corsican. In A. Ledgeway & M. Maiden (Eds.), *The Oxford Guide to the Romance Languages*, Oxford University Press, 206-227.



- Marotta G. & Nocchi N. (2001). La liquida laterale nel livornese. *RID. Rivista italiana di dialettologia*, 25, 285-326. DOI: 10.1400/12765.
- Marotta G. & Vanelli L. (2021). *Fonologia e prosodia dell'italiano*. Roma: Carocci.
- Maddieson I. & Ferrari-Disner S. (1984). *Patterns of sounds*. Cambridge University Press.
- Malagoli G. (1905). *Ortoepia e ortografia italiana moderna* (Vol. 143). Milano: Hoepli.
- McCrory E. J., Mechelli A., Frith U. & Price C.J. (2005). «More than words: a common neural basis for reading and naming deficits in developmental dyslexia?» *Brain*, 128(2), 261-267. <https://doi.org/10.1093/brain/awh340>.
- Migliorini B. (2019 [1994]). *Storia della lingua italiana*. Milano: Bompiani.
- Miola E. (2015a). «La tirannia della tastiera: Il caso dell'ortografia piemontese», *Language Problems & Language Planning*, 39(2), 136-153.
- Miola E. (2015b). «Chì pòdom tucc scriv come voeurom. Scrivere in lombardo online». In: S. Dal Negro, F. Guerini & G. Iannàcaro (a cura di), *Elaborazione ortografica delle varietà non standard. Esperienze spontanee in Italia e all'estero*, Bergamo: Bergamo University Press-Sestante ed., 79-96.
- Mioni A.M. (1973). *Fonematica contrastiva*. Bologna: Pàtron.
- Montesano L. & Valenti A. (2017). I Disturbi Specifici dell'Apprendimento nei giovani adulti. Un progetto di screening attraverso l'utilizzo di una versione modificata del questionario Vinegrad. *Difficoltà di apprendimento e didattica inclusiva*, V.2, 201-210.
- Montesano L. & Valenti A. (2018). SLD: promoting continuity between School and University education. Results of a screening program in Calabria. *Form@re-Open Journal per la formazione in rete*, 18(1), 192-206. <http://dx.doi.org/10.13128/formare-22313>.
- Pettorino M. & Giannini A. (1991). «Indagine acustica sulle approssimanti dell'italiano». In *Atti del XIX Convegno Nazionale ALA*, 441-447.
- Porrás J.E. (2013). Spanish Yeísmo: A Cognitive Linguistic Approach to Phonological Change. In R. Gómez & I. Molina Martos (a cura di), *Variación yeísta en el mundo hispánico*, 335-352. <https://doi.org/10.31819/9783954870691-015>.
- Quilis A. & Fernández J.A. (1969). *Curso de fonética y fonología españolas para estudiantes angloamericanos*. 4. Vol. 2. Madrid. Instituto «Miguel Cervantes».
- RAE. Real Academia Española (2011). *Nueva gramática de la lengua española. Fonética y fonología*.
- Ramus F., Rosen S., Dakin S.C., Day B.L., Castellote J.M., White S. & Frith U. (2003). Theories of developmental dyslexia: insights from a multiple case study of dyslexic adults. *Brain*, 126(4), 841-865. <https://doi.org/10.1093/brain/awg076>.

Recasens D. (1984a). *Producció i coarticulació de vocals i consonants del català en el decurs*. Tesi di Dottorato inedita. Universitat de Barcelona.

Recasens D. (1984b). Vowel-to-vowel coarticulation in Catalan VCV sequences. *The Journal of the Acoustical Society of America*, 76(6), 1624-1635. <https://doi.org/10.1121/1.391609>.

Revelli L. (2018). «Questioni d'ab-bicci»: presentazioni e rappresentazioni del sistema alfabetico italiano fra Ottocento e Novecento. *Italiano LinguaDue*, 10(1), 414-434. <https://doi.org/10.13130/2037-3597/10957>.

Romano (2009). *Inventari sonori delle lingue*. Alessandria: dell'Orso (1ª ed. 2008).

Romano A. (2016). «Scrivere il dialetto: lo stile e la tecnica esemplari di N.G. De Donno». In: L. Giannone (a cura di), *La poesia dialettale di Nicola G. De Donno*, Lecce: Milella, 175-203.

Romano A. (2018), «I contributi di Amerindo Camilli all'IPA e alla descrizione della pronuncia dell'italiano e delle sue varietà». *Bollettino dell'Atlante Linguistico Italiano*, 42, Torino (2018), 207-217.

Romano A. & Miletto A.M. (2017). *Argomenti scelti di glottologia e linguistica*. Torino: Omega (edizione aggiornata).

Rost Bagudanch A. (2013). La transcripción fonética en estudios dialectales: propuestas en el caso del yeísmo. *Revista de filología española*, 93(1), 165-192.

Rost Bagudanch A. (2014a). El papel de la percepción en el yeísmo: Estudio preliminar con jueces catalanoha-

blantes, *Loquens*, 1(2). <http://dx.doi.org/10.3989/loquens.2014.010>.

Rost Bagudanch A. (2014b). Una panorámica del yeísmo:¿ un proceso acabado o en construcción? *Una panorámica del yeísmo:¿ un proceso acabado o en construcción?*. XII(1(23)). 141-163.

Rost Bagudanch A. (2016). La percepción de /ɬ/ y /j/ en catalán y en español. Implicaciones en la explicación del yeísmo. *Estudios de Fonética Experimental*, XXV, 40-80.

Rost Bagudanch A. (2020). Bilingualism and sound change: perception in the /ɬ/-/j/merger process in Majorcan Spanish. *Zeitschrift für romanische Philologie*, 136(1), 106-133. <https://doi.org/10.1515/zrp-2020-0005>.

Serianni L. (1991). *Grammatica italiana*. Torino: Utet.

Snowling M.J. (2019). *Dyslexia: A very short introduction*. Oxford University Press.

Svensson I. & Jacobson C. (2006). How persistent are phonological difficulties? A longitudinal study of reading retarded children. *Dyslexia*, 12(1), 3-20. <https://doi.org/10.1002/dys.296>.

Telmon T. (1993). Varietà regionali. In G. Berruto (a cura di), *Introduzione all'italiano contemporaneo. La variazione e gli usi*. II. Roma: Laterza, 93-116.

Vinegrad M. (1994). A revised adult dyslexia check list. In *Educare—London national bureau for handicapped students*. 48, 21-21.

Zampaulo A. & Haug D. (2015). The evolution of the (alveolo)palatal lateral consonant in Spanish and Portuguese. In D. Haug (ed.), *Historical Linguistics 2013: Selected papers from the 21st International Conference on Historical Linguistics* (Oslo, 5-9 August 2013), Amsterdam: John Benjamins, 69-86. <https://doi.org/10.1075/cilt.334>.

Zanobini M., Viterbori P. & Saraceno F. (2012). Phonology and language development in Italian children: An analysis of production and accuracy. *Journal of Speech, Language, and Hearing Research*, 55, 16-31. [https://doi.org/10.1044/1092-4388\(2011/10-0228\)](https://doi.org/10.1044/1092-4388(2011/10-0228)).

Vitali D. & Pioggia D. (2014). *Dialetti romagnoli*. Verucchio: Pazzini.

## Sitografia

DOP: <https://dop.netadcom.com/> (consultato il 03.03.22; il sito è attualmente offline).

UPSID: [http://web.phonetik.uni-frankfurt.de/upsid\\_find.html](http://web.phonetik.uni-frankfurt.de/upsid_find.html).

«Zabaione o zabaglione» su *La grammatica italiana*: <https://www.treccani.it>

Giannini A. (2010). Laterali. In Treccani.it, *Enciclopedia dell'Italiano*: <https://www.treccani.it>

Pettorino M. (2011). Palatali. In: Treccani.it, *Enciclopedia dell'Italiano*: <https://www.treccani.it>

## Software

SPSS for Analytics (by IBM), v. 28.

Praat: doing phonetics by computer 6.2.10 [[www.praat.org](http://www.praat.org)]

Folerpa [<http://ilg.usc.es/FOLERPA>]

Audacity v. 3.1.x

Goldwave v. 4.02



# Primi rilievi EGG sulla voce cantata

Antonio Romano<sup>1</sup>, Valentina De Iacovo<sup>1</sup>, Dario Strangis<sup>2</sup>, Silvestro Roatta<sup>3</sup>

<sup>1</sup>LFSAG - <sup>2</sup>Voce InForma–Queen Margaret Univ. - <sup>3</sup>Ist. Fisiologia, Dip. Neuroscienze, UniTo

## Introduzione

Nell'ambito del progetto «fisiologia emissione vocale parlata e cantata» il LFSAG ha finalmente ottenuto la possibilità di acquistare un sistema di rilevamento elettroglottografico (EGG).

Dopo mesi di attesa, il dispositivo fornito dall'azienda *Laryngograph* Ltd. (Sutton, UK) è stato ricevuto e immediatamente testato<sup>1</sup>. I primi rilevamenti sono stati resi possibili grazie alla pazienza di Xinghui Hu, *general manager* dell'azienda nata dalla collaborazione con alcuni dei pionieri degli studi basati sulla tecnica della glottografia impedenziometrica: Evelyn Abberton & Adrian Fourcin<sup>2</sup>.

1. Apprezziamo per ringraziare il personale di UniTO che ha garantito le condizioni per concludere l'acquisto. Sul piano amministrativo questo è avvenuto con RDA 51865 (<https://www.unito.it/ugov/trasparenza/2406578>).

2. V. Abberton & Fourcin (1997). L'incontro dell'autore AR con le tecniche sviluppate da questi ricercatori era già avvenuto nell'ambito della realizzazione della sua tesi di laurea in telecomunicazioni (*Analyse de la fonction de transfert du conduit vocal dans les intervalles de fermeture de la glotte et extraction de la forme d'onde d'excitation glottale par filtrage inverse*, 1994). In particolare in quella sede i riferimenti erano stati offerti da Fourcin & Abberton (1971) e Fourcin A. (1974). Gli autori dei numerosi studi condotti con questi

La tecnica, generalmente nota come *Impedance Glottography*, recentemente illustrata da Celata & Ricci (2021), trova riscontro in numerosi studi ed è stata implementata in diversi prodotti commerciali. La nostra esperienza era infatti partita dalla valutazione preliminare di altri dispositivi simili (quello della *F-J Electronics*, appartenente al fondo A. Genre del LFSAG, o l'estensione EGG CSL 4400 della Kay Elemetrics, testato presso il Laboratorio di Fonetica del DLLC dell'Univ. di Aveiro, Portogallo, diretto da Lurdes de Castro Moutinho, su cui si basano le considerazioni riportate nel §1).

Dopo un primo collaudo, il dispositivo EGG in dotazione del LFSAG è stato collocato provvisoriamente nella sede dove sono avvenuti i rilievi discussi in questo breve resoconto<sup>3</sup>.

metodi accreditano come precursore quello illustrato da P. Fabre nel 1957 ("Un procédé électrique percutané d'inscription de l'accollement glottique au cours de la phonation: glottographie de haute fréquence. Premiers résultats", *Bulletin de l'Académie Nationale de Médecine*, 141, 66-99); cfr. anche la testimonianza di K. Marasek, nell'ormai sito storico pluridecennale EGG dell'università di Stoccarda, 1997, e, più recentemente, Henrich et al. (2004).

3. Si tratta del laboratorio che si trova sotto la responsabilità dell'autore SR nella strut-

Nel §2 si dettagliano in particolare le caratteristiche del segnale EGG che è stato possibile analizzare sulla base di alcune emissioni vocali controllate, in questa prima fase ancora solo in associazione con una presa di suono microfónica e a una grossolana stima fonometrica<sup>4</sup>.

## 1. I cicli glottidali

### 1.1. Velocità di vibrazione delle pliche, frequenza fondamentale, altezza e funzioni linguistiche

Secondo il modello sorgente-filtro proposto sin da Fant (1960), l'onda di pressione acustica, il cui andamento può essere rappresentato per mezzo dell'ampiezza istantanea rilevata da un microfono (e raffigurata come oscillogramma), è il risultato di effetti risonanziali le cui caratteristiche sono ascrivibili a una sorgente sonora (vibratoria o rumorosa) e alle attività filtranti delle cavità di risonanza e irraggiamento del suono. Assumiamo per tutti i suoni sonori (*voiced*) che la

tura di c.so Raffaello, 30 (*Fisiologia*, Dip. di Neuroscienze dell'Univ. di Torino).

4. I rilievi sono avvenuti a cura dell'autrice VD sulle produzioni controllate dell'autore DS (a sua volta già coinvolto in simili esperimenti presso il *Centro di Foniatria y Logopedia* di Santander, Spagna, diretto dal Medico Foniatra Alfonso Borragán Torre). Il fonometro è un *Sound Level Meter* acquistato sugli stessi fondi dal collega Massimo Spadola Bisetti che ringraziamo in quest'occasione.

sorgente sia localizzata essenzialmente nella laringe (in particolare nella glottide, la superficie tra le pliche vocali) e che gli effetti filtranti, di caratterizzazione del timbro, siano imputabili alle configurazioni assunte dalle cavità epilaringee (condotto vocale, labbra, cavità nasali).

La responsabilità della presenza di un'armonica fondamentale (che fissa il passo di definizione dello spettro acustico del suono) è riconducibile alla ciclicità con cui si alternano configurazioni di apertura e chiusura delle rime, interlegamentosa e interaritenoidica, che definiscono lo spazio glottidale<sup>5</sup>.

La periodicità con cui si presentano questi fenomeni ciclici è dunque alla base della definizione della variabile acustica detta frequenza fondamentale ( $f_0$ ), a sua volta ritenuta all'origine delle sensazioni di altezza dei suoni sonori<sup>6</sup>.

5. In determinati assetti di fonazione, queste possono tuttavia partecipare anche alla definizione del timbro ed essere responsabili di sorgenti secondarie (rumori, sfiati etc.) che contribuiscono a definire la qualità della voce (riferimenti vari in Romano et al. 2011).

6. Ovviamente, sul piano uditivo, tutte le armoniche contribuiscono a quest'effetto e, anche in assenza della prima, se ne potrebbe avere una percezione illusoria (cfr. varie considerazioni e riferimenti in Albera 2019). Ricordiamo che in molta letteratura di stampo anglosassone, tanto la variabile acustica quanto i suoi effetti in termini di percezione sono generalmente confusi nel ricorso al termine iperonimico di *pitch*. Al lettore inesperto ricordiamo anche che  $f_0$  indica l'armonica fondamentale, mentre si usa la lettera

A differenza del canto, in cui l'altezza viene regolata in modo da raggiungere e sostenere nel tempo le note (cioè livelli culturalmente quantizzati) di una melodia prefissata con determinate qualità vibratorie, nel parlato la periodicità (e la stessa qualità) di questi cicli subisce variazioni notevoli nel passaggio da una porzione all'altra degli enunciati ed evolve solitamente nell'arco di singoli segmenti sonori e loro fasi. Questo contribuisce localmente alla categorizzazione di qualità linguistiche (fonologiche) dei segmenti sonori e alla caratterizzazione ritmico-intonativa degli enunciati: le variazioni di  $f_0$  sono infatti suscettibili di codificare relazioni sintattiche e veicolare informazioni semantiche, emotive e caratteriali del parlante.

Contrariamente alla visione generale di questa variabile che può emergere nell'immaginario del non specialista a causa delle scarse informazioni che se ne danno nella manualistica corrente, uno studio rigoroso dei caratteri linguistici dell'intonazione si scontra con le difficoltà della sua misurazione partendo dal segnale acustico, soprattutto in condizioni di basso rapporto segnale/rumore e, appunto, nelle varie condizioni che inducono il parlante a modificare non solo la velocità di

maiuscola ( $F_0$ ) per indicare l'eventuale formante associata a questa nella parte bassa dello spettro (anche se questa, nella maggior parte dei casi, coincide con una sola armonica:  $f_0$  appunto).

vibrazione delle pliche, ma gli assetti fonatori e le modalità di vibrazione all'interno di uno stesso segmento. Questo è causa di irregolarità vibratorie che inducono in errore gli algoritmi di misurazione e mettono in difficoltà lo specialista che deve decidere da un lato se i valori rilevati si riferiscano a bersagli sovrasegmentali prefissati dal parlante e dall'altro se il ruolo funzionale, in termini linguistici, dipenda dall'altezza percepita o da quella indicata dalla misurazione (indecidibile nei casi di diplofonia e in diverse condizioni di cricchiato)<sup>7</sup>.

### 1.2. Definizione del ciclo glottidale

Il monitoraggio delle caratteristiche del ciclo glottidale può avvenire ovviamente con mezzi fibroscopici, ma sono invalsi e diffusi da decenni metodi indiretti non invasivi come l'elettroglottografia (EGG) che contribuisce tradizionalmente anche una misurazione più accurata, direttamente alla sorgente e, quindi, maggiormente esente da errori indotti da effetti di filtraggio.

Di questo metodo daremo qui una dimostrazione, applicandolo a modalità fonatorie tipiche di alcuni stili di

7. In questi casi è inoltre difficile stabilire se l'altezza percepita sia quella di una prima armonica irregolarmente raccolta/rappresentata nel segnale sonoro o dell'insieme di armoniche superiori che attivano determinati *pattern* cocleari (v. n. prec.).

canto diffusi oggi nella musica commerciale occidentale. L'argomento non è nuovo, ma sono sperimentali le condizioni di osservazione e descrizione.

Pare opportuno, però, anticipare le variabili in gioco e i grafici convenzionali su cui si basano le osservazioni tradizionali in riferimento a una rappresentazione ipotetica delle condizioni fisiologiche in cui avviene un passaggio d'aria attraverso la glottide (Fig. 1, in cui le dimensioni delle frecce indicano arbitrariamente flussi d'aria, in verde, e forze di pressione, in blu).

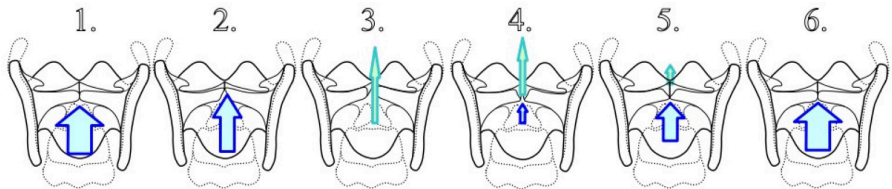


Figura 1. Fasi di vibrazione delle pliche vocali (tratto da Romano 2008 IS).

Il ciclo glottidale è ben noto sin dagli anni '50 (Fant 1960, Hirano 1989) ed è riprodotto in numerose opere scientifiche anche a carattere didattico o divulgativo.

Alle distinte fasi ricostruite schematicamente si possono associare i valori istantanei di diverse variabili fisiche rilevabili nelle forme d'onda raccolte ed elaborate con varie procedure analitiche.

A ciascuna fase del ciclo di apertura-chiusura di Fig. 1 corrisponde una variazione che si ripercuote su diverse variabili tra le quali le più studiate sono

il volume d'aria istantaneo che attraversa la glottide, forma d'onda glottidale o *fdog* (misurabile solo attraverso sonde flussometriche, ricostruibile dal segnale microfonico per *inverse filtering* v. Fig. 2a; v. Romano 1994; cfr. Ferrero et al. 1994). In corrispondenza di questo grafico proponiamo quello di Fig. 2b relativo a una tipica curva di variazione istantanea (velocità) del segnale EGG, associato a un segnale reale EGG (Fig. 2c), ottenuto in un rilievo pilota eseguito tramite EGG CSL 4400 della Kay Elemetrics, testato presso il Laboratorio di Fonetica

del DLLC dell'Univ. di Aveiro, Portogallo (diretto da Lurdes de Castro Moutinho).

Il grafico di «velocità della glottide» (Fig. 2b), basato sulla funzione  $u_g(t)$  per la quale si rimanda ora a Henrich et al. (2004) e Roubeau et al. (2009), mostra un insieme di valori relativamente stabili all'inizio del ciclo evidenziato di durata  $T_0$ ; la fase di glottide chiusa (1-2), il cui inizio è qui indicato con  $t_p$ , corrisponde infatti a velocità nulla. Si assume positiva invece nella fase di apertura, che avviene piuttosto gradualmente fino a un massimo raggiun-



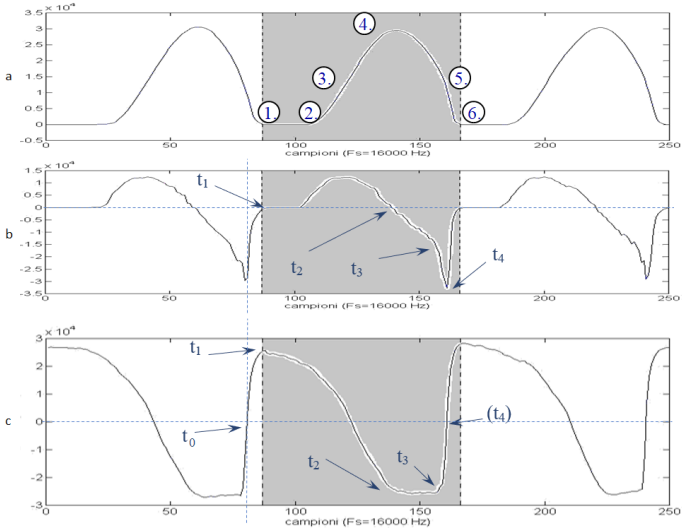


Figura 2: a) forma d'onda glottidale ( $fdog$ ) di tre cicli ricostruita per filtraggio inverso (stima del glottal flow, volume d'aria istantaneo che attraversa la glottide). L'area evidenziata stabilisce le corrispondenze con il ciclo di cui si dettagliano varie fasi in Fig. 1; b) diagramma di velocità della glottide di tre cicli (funzione  $u_g(t)$ , generalmente derivata dal segnale EGG; c) segnale elettroglottografico (EGG,  $g(t)$ ) rilevato per tre cicli simili.

to in un punto di flesso della  $fdog$  (3) e tornare a zero nel momento in cui si ha una glottide massimamente aperta ( $4 \approx t_2$ )<sup>8</sup>. Dopo un certo tempo di latenza inizia la fase di riavvicinamento con velocità moderata all'incirca costante nel verso opposto (durante una fase in cui si può considerare la glottide ancora relativamente aperta) fino a un momento in cui subisce un'accelerazione è più decisa (dall'istante  $t_3$ ) e raggiunge il suo massimo nell'ultima fase, impul-

8. I momenti in cui le pliche vocali sono ferme sono quindi due: quello, relativamente più lungo, in cui restano accollate e quello, di durata più breve e variabile, in cui raggiungono il massimo distanziamento nella porzione mediana (massima apertura della glottide).

siva, di chiusura ( $5 \approx t_4$ ) che la riporta a zero nel punto in cui (6) si torna al  $t_1$ .

Nonostante sia più promettente l'analisi della curva di velocità della glottide (funzione  $u_g(t)$ , Henrich et al. 2004) in questa prima fase di studio ci siamo concentrati sul segnale elettroglottografico (EGG),  $g(t)$ , di cui sono riprodotti tre cicli in Fig. 2c. Per un trattamento quantitativo (automatizzato) sono stati infatti proposti vari metodi di parametrizzazione proprio di questa forma d'onda (Rothenberg 1992, Marasek 1995, Vieira et al. 1996) e ci riferiamo quindi qui in particolare a questi ultimi, evidenziando nei nostri rilievi i tre punti (tre tempi) che definiscono le durate delle principali fasi osservabili

sul segnale EGG di una voce modale ( $t_1$  corrisponde al punto 3 di Marasek,  $t_2$  al 6 e  $t_3$  al punto 1; il  $t_0$  di riferimento è l'intersezione del segnale  $g(t)$  con l'asse delle ascisse)<sup>9</sup>.

## 2. Dispositivi e materiali analizzati

Come anticipato nell'introduzione, per i nostri rilievi abbiamo usato un dispositivo di rilevamento elettroglottografico (EGG), acquistato presso l'azienda *Laryngograph*. L'elettroglottografo della *Laryngograph* comprende tre collari con sensori laringometrici (elettrodi di taglia diversa) e un microfono a cravatta per il rilevamento dell'onda di pressione acustica che si diffonde nell'ambiente. Si collega direttamente tramite porta USB a un PC su cui sia installata una *Voice Suite* che prevede vari software tra cui *Speech Studio* che consente la visualizzazione in tempo reale (e differito) della forma d'onda istantanea catturata dal microfono ( $S_p$ ) e di quella EGG ( $Lx$ )<sup>10</sup>. Gli elettrodi, opportunamente posizionati sul col-

lo, in corrispondenza dei due lati della cartilagine tiroidea, registrano le rapide, minime variazioni di conduttanza tra di essi che dipendono dai cambiamenti in natura e quantità del contatto che si verifica istantaneamente tra le pliche vocali durante la fonazione.

### 2.1. Materiali

Sono state eseguite 18 prese di suono con registrazione del segnale glottografico e valutazione fonometrica. Il soggetto all'origine delle produzioni vocali è un professionista con una buona estensione vocale e con particolari abilità nell'esecuzione canora in vari assetti fonatori<sup>11</sup>. Con l'aiuto di un diapason<sup>12</sup>, nel corso di una prima registrazione, ha prodotto in succes-

11. Il soggetto riferisce tuttavia la presenza di solcature nella porzione anteriore delle sue pliche vocali, descritta per mezzo di rilievi laringostroboscopici.

12. Assumiamo qui come riferimento il LA<sub>3</sub> a 440 Hz (Anche in Italia la legge 170 del 3/5/1989 stabilisce che "Il suono di riferimento per l'intonazione di base degli strumenti musicali è la nota la<sub>3</sub>, la cui altezza deve corrispondere alla frequenza di 440 Hertz (Hz), misurata alla temperatura ambiente di 20 gradi centigradi" - *Gazzetta Ufficiale* [www.gazzettaufficiale.it/eli/id/1989/05/12/089G0200/sg]). In base a questa convenzione: RE<sub>1</sub> = 73 Hz, RE<sub>2</sub> = 147 Hz, RE<sub>3</sub> = 294 Hz, RE<sub>4</sub> = 587 Hz e RE<sub>5</sub> = 1175 Hz. I valori di frequenza associati alle singole note sono quindi quelli di Romano et al. (2012) in relazione a una classificazione novecentesca delle voci della tradizione europea del *bel canto*.

9. Alle valutazioni del quoziente di chiusura (*closing quotient*), una cui stima è ricavabile dal rapporto tra le differenze  $t_4 - t_2$  e  $t_2 - t_1$ , contiamo di aggiungere misure del quoziente di *skewing* (in riferimento "all'inclinazione" dell'onda) basate sul rapporto tra il tempo di incremento (stimato da  $t_2 - t_1$ ) e tempo di diminuzione del flusso (stimato da  $T_0 - t_2$ ).

10. Un'altra utile interfaccia operante su dati precedentemente salvati è QA che consente l'esecuzione di un'analisi quantitativa complessiva (QA = *Quantitative Analysis*).

sione i seguenti 11 suoni con caratteristiche timbriche di [a] (per una durata media di 5 s ciascuno): un RE<sub>1</sub> tenuto, un RE<sub>1</sub> messa di voce, un RE<sub>2</sub> tenuto, un RE<sub>2</sub> messa di voce, un RE<sub>3</sub> tenuto, un RE<sub>3</sub> messa di voce, un RE<sub>4</sub> tenuto, un RE<sub>4</sub> messa di voce, un segmento di voce fischiata (FA#<sub>4</sub>), un glissando RE<sub>1</sub>-RE<sub>4</sub>-RE<sub>2</sub> e un glissando RE<sub>1</sub>-RE<sub>4</sub>-RE<sub>1</sub><sup>13</sup>. Tutte le registrazioni sono avvenute in presenza di un rumore di fondo stimato fonometricamente a circa 45 dB.

## 2.2. Osservazioni sperimentali

Sono state rilevate le forme d'onda riprodotte nelle figure 3-21 con le seguenti caratteristiche.

Il primo RE<sub>1</sub> tenuto ha presentato un'altezza media di 72 Hz e 62 dB<sup>14</sup>, con una forma d'onda elementare EGG caratterizzata come in Fig. 3 (con  $T_0$  pari a poco più di 13 ms). La fase di chiusura è risultata relativamen-

13. Nel corso di una seconda registrazione, il soggetto ha invece prodotto 7 modalità fonatorie: un *(vocal) fry* espiratorio, un *(vocal) fry* inspiratorio, un [a] modale (MI<sub>3</sub>), un [a] soffiato (MI<sub>3</sub>), un [a] nello stile *grunt*, un [a] nello stile *rattle* e un [a] nello stile *growl*. In questa sede dettagliamo solo gli 11 campioni modali, rinviando la discussione delle 7 produzioni non-modali a un lavoro a venire che sarà svolto in riferimento alla tassonomia proposta da Gerratt & Kreiman (2001) e alla luce delle riflessioni di Aaen et al. (2020).

14. Nel corso di tre osservazioni gli intervalli di intensità sonora sono stati di 50÷61,6; 50,3÷61,4; 49,4÷63,4 dB.

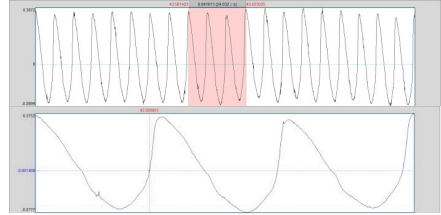


Figura 3. Segmento di forma d'onda EGG rilevata per un RE<sub>1</sub> tenuto (circa 250 ms) con zoom su tre cicli (in basso).

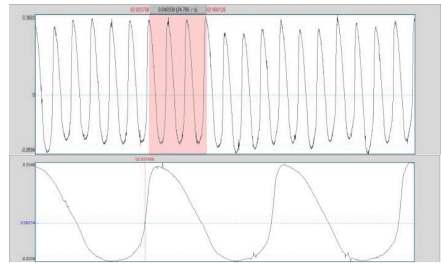


Figura 4. Segmento di forma d'onda EGG rilevata per una messa di voce su RE<sub>1</sub> (ca. 250 ms) con zoom su tre cicli (in basso).

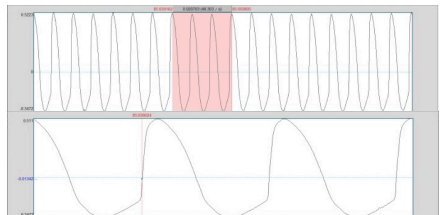


Figura 5. Segmento di forma d'onda EGG rilevata per un RE<sub>2</sub> tenuto (circa 134 ms) con zoom su tre cicli (in basso).

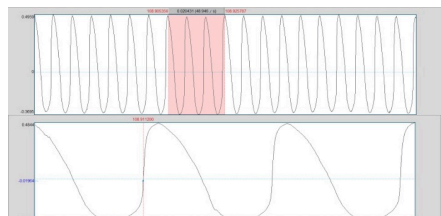


Figura 6. Segmento di forma d'onda EGG rilevata per una messa di voce su RE<sub>2</sub> (ca. 134 ms) con zoom su tre cicli (in basso).

te veloce ( $t_1 = 1,1$  ms) con una tenuta della chiusura istantanea e una fase di apertura più lunga ( $t_2 = 9,3$  ms).

La messa di voce sul RE<sub>1</sub> ha presentato un'altezza con maggiori oscillazioni (tra 66 e 82 Hz con *jitter* = 1,48%) e una media di 74 Hz, ma con energia crescente (da 56 a 66 dB nelle fasi di maggiore stabilità, *shimmer* = 4,46%)<sup>15</sup>. In una porzione centrale, la forma d'onda elementare EGG è caratterizzata, come si può vedere in Fig. 4 da un  $T_0$  pari a poco più di 13 ms (cfr. Fig. 3). Anche in questo caso la fase di chiusura è risultata relativamente veloce ( $t_1 = 0,9$  ms) con una tenuta della chiusura istantanea e una fase di apertura che si esaurisce a  $t_2 = 8,4$  ms.

Il primo RE<sub>2</sub> tenuto ha presentato un'altezza media di 144 Hz (126÷158) e intensità 42÷72,2 dB, con una forma d'onda elementare EGG caratterizzata come in Fig. 5 (con  $T_0$  pari a 7 ms). La fase di chiusura presenta in questo

caso due diverse velocità ( $t_1 = 0,85$  ms, ma in presenza di un  $t_3 = 1,8$  ms e fase di massimizzazione della velocità di durata inferiore a 1 ms) con una tenuta della chiusura istantanea e una fase di apertura più lunga ( $t_2 = 4,8$  ms).

La messa di voce del RE<sub>2</sub> ha presentato un'altezza media di 146 Hz (nell'intervallo 139÷158 Hz) e intensità variabile (nell'intervallo 49÷84,3 dB), con una forma d'onda elementare EGG caratterizzata come in Fig. 6 (con  $T_0$  pari a circa 7 ms e  $t_3 \approx t_4$ ). La fase di glottide chiusa si presenta come nel caso di Fig. 4 con fasi di chiusura e apertura relativamente veloci (con tempi dimezzati:  $t_1 = 0,85$  ms,  $t_2 = 5$  ms e fase di massimizzazione della velocità da  $t_3 = 6,8$  ms)<sup>16</sup>.

Il RE<sub>3</sub> tenuto ha presentato un'altezza media di 295 Hz (min 261 max 303, per via di un transitorio iniziale più instabile, con intensità che si stabilizza su circa 55 dB). La forma d'onda elementare EGG si caratterizza come in Fig. 7 (con  $T_0$  pari a 3,5 ms). Le fasi di chiusura e apertura tendono a presentare durate simili (con  $t_1 = 0,6$  ms e  $t_2 = 2,5$  ms). Seppur visibile, la

15. Le misure di *jitter* (microperturbazioni della frequenza fondamentale) e di *shimmer* (microperturbazioni dell'ampiezza dell'onda) sono state rilevate grazie al *Voice Report* di PRAAT (lo stesso software usato per ottenere i grafici qui riprodotti). I valori rilevati sono generalmente al di sotto delle soglie tipiche rilevate per voci eufoniche di altri spazi linguistici (Brockmann et al. 2009). Un confronto sarebbe auspicabile con i valori riportati da fonti italiane. Quelle da noi reperite si basano però su altre variabili e offrono un riferimento meno agevole che rinviamo a studi più approfonditi.

16. Il confronto tra tenuta e messa di voce in questo caso non mostra le stesse differenze osservabili per RE<sub>1</sub> e RE<sub>3</sub> (v. dopo): la riduzione della fase di apertura è minima e non consente di osservare l'aumento della velocità della chiusura che sarebbe stato in accordo con l'aumento di pressione (e flusso) che si presume avvenga durante la messa di voce.

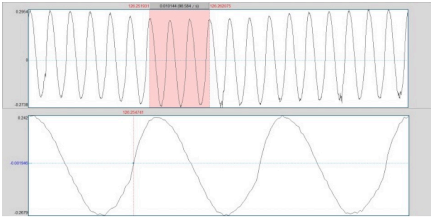


Figura 7. Segmento di forma d'onda EGG rilevata per un RE3 tenuto (circa 68 ms) con zoom su tre cicli (in basso).

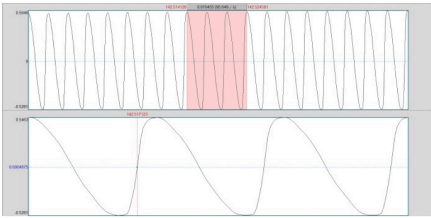


Figura 8. Segmento di forma d'onda EGG rilevata per una messa di voce su RE3 (ca. 66 ms) con zoom su tre cicli (in basso).

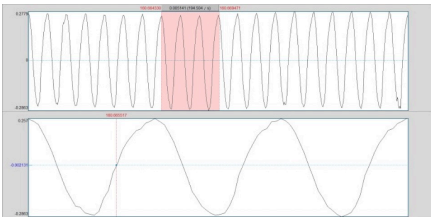


Figura 9. Segmento di forma d'onda EGG rilevata per un RE4 tenuto (ca. 33 ms) con zoom su tre cicli (in basso).

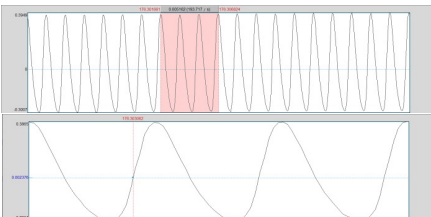


Figura 10. Segmento di forma d'onda EGG rilevata per una messa di voce su RE4 (ca. 34 ms) con zoom su tre cicli (in basso).

presenza di un punto di aumento di velocità nella fase di chiusura ( $t_3$ ) risulta poco rilevante rispetto ai casi visti sopra. Tuttavia il secondo rilevamento di questo RE<sub>3</sub>, in una messa di voce (con intensità variabile 49,1÷88,2 dB), ha prodotto un risultato lievemente diverso (v. Fig. 8;  $T_0 \sim 3,5$  ms): anche a queste frequenze, infatti, a un tempo di chiusura ridotto ( $t_1 \sim 0,6$  ms) si associano tempi di apertura più lunghi ( $t_2$  dell'ordine dei 2,9 ms).

Un'osservazione variabile si conferma anche tra RE<sub>4</sub> tenuto e messa di voce. Se nel caso del primo si ha un transitorio iniziale particolarmente mosso (altezza media 588 Hz, min 499 max 596, e intensità 56,1÷83,5 dB), la forma d'onda si caratterizza come in Fig. 9 e differisce da quella del secondo (Fig. 10) che ha una media 586 Hz e minimi e massimi (rispettivamente 563 e 656) distanti anche per via di particolari instabilità cicliche (associate a uno *shimmer* più elevato, 1,33%<sup>17</sup>). In entrambi i casi si ha  $T_0$  pari a 1,6 ms ma nel primo caso la fase di chiusura si conclude a  $t_1 \sim 0,5$  ms mentre quella di apertura culmina a  $t_2 \sim 1,3$  ms; nel secondo – compatibilmente con quanto osservato nel caso di RE<sub>1</sub> e RE<sub>3</sub> – la proporzione cambia, con un tempo di chiusura più corto ( $t_1 = 0,6$  ms e  $t_2 = 1,4$  ms circa).

17. L'intensità rilevata dal fonometro è infatti nell'intervallo 75÷91,2 dB.

Una voce fischiata è infine stata osservata a circa 734 Hz (FA#<sub>4</sub>,  $T_0 = 1,4$  ms circa, v. Fig. 11) con intensità variabile ma generalmente piuttosto elevata (70,1÷78 dB). A un tempo di chiusura piuttosto breve ( $t_1 = 0,5$  ms) si associa un tempo di apertura più lungo ( $t_2 = 1,1$  ms) con proporzioni simili a quelle già descritte per l'esempio di Fig. 10 (v. §1.2).

In merito alle osservazioni svolte sui glissando, ci riferiamo qui (Figg. 12 e

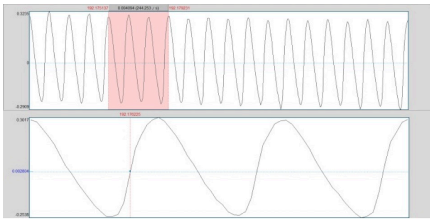


Figura 11. Segmento di EGG rilevato per una voce fischiata su FA#4 (ca. 26 ms) con zoom su tre cicli (in basso).

13) solo alla fase crescente del primo e a quella decrescente del secondo. Il primo glissando presenta un progressivo aumento di intensità sonora da 48,3 a 83,7 dB e una successiva diminuzione fino a 65,2 dB, con rilevanti perturbazioni nella cosiddetta fase di *passaggio*<sup>18</sup>.

Queste instabilità sono state studiate recentemente proprio su voci teno-

18. Si noti che, malgrado l'energia sonora rilevata dal fonometro presenti questi valori, la misura stimata da PRAAT sul segnale microfonico varia da circa 60 a circa 82 dB, mentre quella rilevata sull'EGG resta all'incirca co-

rili da Echternach et al. (2017) e Selamtzis (2018) che le hanno osservate e descritte a partire dal FA<sub>4</sub> (circa 350 Hz) fino a una frequenza individuale che stabilisce un altro intervallo stabile (almeno in termini di intensità) nel quale si definisce un tipo vocale detto *SVaP* (*Stage Voice above the Passaggio*)<sup>19</sup>. Si riportano in Fig. 12 due campioni di cicli raccolti rispettivamente nella fase grave iniziale ( $T_0 = 13$  ms circa, con  $t_1 = 1$  ms,  $t_2 = 8,6$  e  $t_3 = 12$  ms) e di tenuta centrale ( $T_0 = 1,7$  ms circa, con  $t_1 = 0,6$  ms e  $t_2 = 1,1$  ms).

La parte decrescente del secondo glissando ha presentato una progressiva diminuzione di intensità sonora da

stante, sugli 82 dB, con un'unica flessione (min. 75 dB) in corrispondenza del *passaggio*. Discrepanze simili si osservano tra i valori rilevati per il secondo glissando.

19. Le qualità performative degli interpreti a questo riguardo sono oggetto delle preoccupazioni di C. Ambrosini che, in riferimento alla tradizione colta europea, nel suo contributo dal titolo “Voce cantata e condizionamenti linguistici” (in Ambrosini et al. 2013) si pone il problema di comporre un'aria tenendo conto delle sue caratteristiche promeliche o antimeliche, in funzione delle possibilità dell'interpretazione (definendo condizioni in cui questa possa definirsi comoda e congeniale alla vocalità del cantante, proprio in termini di tessitura e aree di passaggio tra i registri). Nella stessa sede si veda anche l'intervento di G. Tisato che discute della non-linearità dei meccanismi fonatori e analizza le possibilità classificatorie delle modalità vocali documentate in alcune tradizioni ed esplorate in applicazioni sperimentali (soprattutto sin dai primi contributi di D. Stratos).



circa 82 dB a 48 dB. Si riportano in Fig. 13 due campioni di cicli raccolti rispettivamente in corrispondenza della nota acuta iniziale ( $T_0 = 1,6$  ms circa, con  $t_1 = 0,5$  ms,  $t_2 = 1,1$  ms) e di quella finale ( $T_0 = 13,6$  ms circa, con  $t_1 = 1,7$  ms,  $t_2 = 8,5$  ms e  $t_3 = 11,4$  ms). In questo caso, nel *passaggio*, nel corso della lunga fase d'instabilità di ampiezza/intensità sonora, l'algoritmo di stima di  $f_0$  usato in PRAAT ha prodotto anche alcuni errori di rilevamento nell'ultima fase di transizione verso RE<sub>1</sub>.

Alla somiglianza tra le forme d'onda elementari dell'EGG nel caso dei due RE<sub>1</sub> si contrappone la marcata differenziazione di quelle dei due RE<sub>4</sub>.

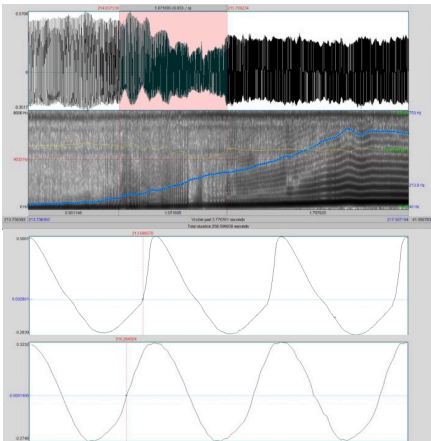


Figura 12. Oscillogramma e spettrogramma dell'onda EGG della fase crescente del primo glissando (in alto). Si noti la lunga instabilità di ampiezza/energia in corrispondenza del passaggio (evidenziato). Al centro, zoom su tre cicli rilevati nel corso del RE1 iniziale; in basso, zoom su tre cicli del RE4 finale.

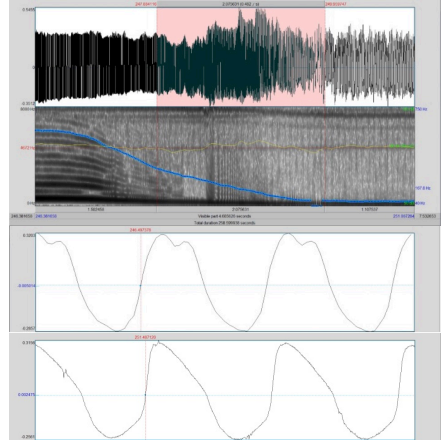


Figura 13. Oscillogramma e spettrogramma della fase decrescente del secondo glissando (in alto). Si noti la lunga instabilità di ampiezza/intensità sonora in corrispondenza del passaggio (evidenziato). Al centro, uno zoom su tre cicli di EGG in corrispondenza della nota più alta (RE4); in basso, zoom su tre cicli in corrispondenza di quella più bassa (RE1).

### 3. Interpretazione e conclusioni parziali

Sulle emissioni vocali controllate di voce modale qui presentate è stata eseguita una prima accurata osservazione dei tempi con cui si organizza la forma d'onda EGG alle varie frequenze a cui abbiamo associato una prima valutazione impressionistica delle possibilità di ritrovare una stessa forma d'onda elementare EGG per una stessa nota prodotta da uno stesso cantante in condizioni diverse.

Nonostante la disponibilità delle misure di una selezione dei tempi parametrizzati in letteratura, non è stata ancora condotta la discussione sui rapporti tra fasi di durata e di chiusura

(sulla base di soglie che restano da stabilire). Data però la presenza di profili con un diverso distanziamento degli eventi rilevabili nelle forme d'onda elementari analizzate (cfr. ad es. Fig. 5 e Fig. 12 destra) sembrerebbe evidente che si tratti di una variabilità tale da rendere necessario il supporto di altre variabili, come i volumi e le strategie respiratorie (Folino et al. 2017) o l'ispezione di curve derivate più facilmente caratterizzabili (Henrich et al. 2004). Una rappresentazione multidimensionale potrà includere anche osservazioni svolte con modalità di presentazione dei dati che prevedano il ricorso al fonetogramma (Ginepro 2022) in grado di tracciare le dinamiche che consentono il raggiungimento e la tenuta dell'altezza osservata.

Anche tenendo conto di alcune anomalie rilevate (come quella dell'oscillazione nei valori della fase di glottide chiusa di Fig. 12 sinistra, ancora da quantificare adeguatamente in base a un'opportuna definizione delle soglie), resta comunque ancora impressionistica l'osservazione di un principio di minor tempo di chiusura (cioè valori inferiori del *closing quotient*) associato alle tonalità più alte.

Un altro fatto notevole riguarda le proporzioni generali che regolano i rapporti tra le misurazioni della qualità EGG di una nota e quelle della nota corrispondente tre ottave più in alto. Passando ad es. da una forma d'onda

elementare con un  $T_0 = 13$  ms a una con un  $T_0 = 1,6-1,7$  ms si arriva a condizioni in cui l'errore di misura non è più trascurabile.

Un tentativo di interpretazione più organica delle variazioni osservate è quindi rimandato alle successive fasi di lavoro in cui, a un'osservazione più accurata dei *passaggi* (da un meccanismo all'altro), sarà associata l'analisi dei campioni di voce cantata prodotta in condizioni non-modal di fonazione tipiche delle voci estreme (sul modello di Aaen et. al. 2020).

## Bibliografia

Aaen M., McGlashan J. & Sadolin C. (2020). «Laryngostroboscopic Exploration of Rough Vocal Effects in Singing and their Statistical Recognizability: An Anatomical and Physiological Description and Visual Recognizability Study of Distortion, Growl, Rattle, and Grunt using laryngostroboscopic Imaging and Panel Assessment». *Journal of Voice*, 34(1), 162.e5-14.

Abberton E. & Fourcin A. (1997), «Electrolaryngography». In Martin J. Ball & Chris Code (eds.), *Instrumental Clinical Phonetics*, London: Whurr, 119-148.

Ambrosini C., Bravi P., Proto T., Tisato G. & Romano A. (2013). «Speaking voice, singing voice, and performance». In: V. Galatà (a cura di), *Multimodalità e multilinguallità: la sfida più avanzata della comunicazione orale* (Atti del IX Convegno Nazionale dell'Asso-



ciazione Italiana di Scienze della Voce, Venezia 21-23/01/2013), Roma: Bulzoni, 3-11 (abstract p. 3).

Brockmann M., Drinnan M.J., Carding P.N. (2009). «Comparison of voice intensity effects on electroglottographic versus acoustic jitter and shimmer». In: *Proc. of the 157<sup>th</sup> meeting of the Acoustical Society of America* (Portland, Oregon, USA, April 2009), 2532.

Celata C. & Ricci I. (2021). «Electrolaryngography/Electroglottography». In Martin J. Ball (ed.), *Manual of Clinical Phonetics*, London: Routledge, 306-321.

Echternach, M., Burk, F., Köberlein, M., Herbst, C. T., Döllinger, M., Burdumy, M., and Richter, B. (2017). Oscillatory Characteristics of the Vocal Folds Across the Tenor Passaggio. *Journal of Voice*, 31(3), 381.e5-e14.

Fant G. (1960). *Acoustic Theory of Speech Production*. The Hague: Mouton (new ed. Dordrecht: Kluwer, 2004).

Ferrero F., Pierucci P. & Mesirca A. (1994). «Determinazione automatica degli istanti di chiusura glottale», *Rivista Italiana di Acustica*, 18, 19-26.

Folino A., Benzo M., Pasquero P., Laguzzi A., Mesin L., Messere A., Porta M., Roatta S. (2017). Vena cava responsiveness to controlled isovolumetric respiratory efforts, *J. Ultras Med.*, 36, 2113-2123 (Q2 Scopus Radiology, Nuclear medicine).

Fourcin A. (1974). «Laryngographic Examination of Vocal Fold Vibration, Ventilatory and Phonatory Control

Mechanisms». In B. Wyke (ed.), *Ventilatory and Phonatory Control Systems* (International Symposium, Papers 19), London: OUP, 316-333.

Fourcin A. (1986). «Electrolaryngographic Assessment of Vocal Fold Function», *Journal of Phonetics*, 14, 435-442.

Fourcin A. (2005). «Clinical Voice Measurement Using Egg/Lx Signals». In C. Manfredi (ed.), *Models and Analysis of Vocal Emissions for Biomedical Applications. Proc. of the 4<sup>th</sup> International Workshop* (Firenze, 29-31 Oct. 2005), 111-114.

Fourcin A. & Abberton E. (1971). «First application of a new laryngograph». *Med. Biol. Ill.*, 21, 172-182,

Gerratt B. & Kreiman J. (2001). Toward a taxonomy of nonmodal phonation, *Journal of Phonetics*, 29, 365-381.

Ginepro R. (2022). «L'impiego dei fonetogrammi nell'analisi di voci cantate in molteplici lingue». Univ. di Torino: Tesi di Laurea in Scienze linguistiche (rel. Antonio Romano).

Henrich Bernardoni N. (2021). La voce umana, dal respiro al canto. *Bollettino del LFSAG*, 7, 43-57 [www.lfsag.unito.it/ricerca/phonews/07/7\_3.pdf]

Henrich N., d'Alessandro Chr., Doval B. & Castellengo M. (2004). On the use of the derivative of electroglottographic signals for characterization of nonpathological phonation. *The Journal of the Acoustical Society of America*, 115, 1321-1332.

Marasek K. (1995). «A semi-automated Lx-based method for the measurement of voice onset time», *Proc. of XIIIth ICPHS* (Stockholm, 14-19 August 1995), vol. 3, 310-313.

Marasek K. (1997) *EGG & Voice quality*: <https://www2.ims.uni-stuttgart.de/EGG/>

Romano A. (1994). Analyse de la fonction de transfert du conduit vocal dans les intervalles de fermeture de la glotte et extraction de la forme d'onde d'excitation glottale par filtrage inverse / Analisi della funzione di trasferimento del condotto vocale durante gli intervalli di chiusura della glottide ed estrazione della forma d'onda glottidale per filtraggio inverso, Grenoble: Institut de la Communication Parlée/Politecnico di Torino (resp. G. Feng & E. Biglieri).

Romano A., Cesari U., Mignano M., Schindler O. & Vernerio I. (2012). «Voice Quality» / «La qualità della voce». In: A. Paoloni & M. Falcone (a cura di), *La voce nelle applicazioni* (Atti

dell'VIII Convegno dell'Associazione Italiana Scienze della Voce, Roma, 25-27 gennaio 2012), Roma: Bulzoni, 75 (art. int. CD 35 pp.)

Rothenberg M. (1992). «A multi-channel electroglottograph», *Journal of Voice*, 6 (1), 36-43.

Roubeau B., Henrich N. & Castellingo M. (2009). Laryngeal Vibratory Mechanisms: The Notion of Vocal Register Revisited. *Journal of Voice*, Vol. 23, No. 4, 425-438.

Selamtzis A. (2018). Analyses of voice and glottographic signals in singing and speech. Doctoral Thesis in Speech and Music Communication, Stockholm: KTH Royal Institute of Technology School of Electrical Engineering and Computer Science

Vieira M.N., McInnes F.R. & Jack M.A. (1996). «Analysis of the effects of electroglottographic baseline fluctuation on the F0 estimation in pathological voices», *The Journal of the Acoustical Society of America*, 99, 3171-3178.

# POS tagging of read speech

## A tools comparison and an adaptation proposal

Chiara Pesenti<sup>1,2</sup>

<sup>1</sup>Department of Humanities, University of Turin

<sup>2</sup>CLST (Center for Language and Speech Technology), Radboud University

### Abstract

This study aims to investigate the effectiveness of the POS tagging process with regard to read speech, focusing on two main aspects. On one hand, it wants to compare the functioning and the output's differences of two different parsing tools for Dutch language. On the other hand, it focuses on how the divergences between standard written language and read speech could possibly affect a correct POS tagging of read speech. The POS tags of read speech and standard written language obtained from the two parsers are observed and compared, and adaptation regarding the punctuation relevance on read speech is proposed and discussed. More specifically, the POS tagging process improved with the automatic insertion of punctuation marks in the orthographic transcriptions of read speech. The present research showed both expected and unexpected results: the POS tagging of read speech reported less accuracy than the POS tagging of standard written language, and the two parsing tools showed different outputs not only within the read speech but also

within the standard written language. Moreover, a parsing tool proved to perform more accurately than the other regarding the correctness of the POS tags, and at the same time seemed more sensitive to the insertion of punctuation in read speech. Therefore, the results showed that the specific mechanism underlying the two Dutch parsing tools sometimes can lead to misleading results. The findings also showed how written language is more easily interpreted by parsing tools than read speech, and confirmed that punctuation marks play a rather important role in the POS tagging process.

### 1 Introduction

#### 1.1 Background

Part-Of-Speech tagging (POS tagging) is the process of assigning a special label to each token in a text corpus, by specifying the part of speech of the token according to the word's definition and syntactic context [1, 2, 3]. As pointed out in [4], a parsing tool might perform differently according to the specific type of language analysed (i.e., written or spoken language). Read speech shares some cha-

racteristics with both written language and spoken language. The syntactic structure of a read speech is rather consistent and not fragmented, similarly to written language. Yet, some typical elements of spoken language are also present in read speech, such as no punctuation, stuttering, repetitions, bribes, filled pauses, and lengthened vowels.

In this study, two parallel comparisons are made: the comparison between the POS tagging outputs of two different parsing tools for Dutch language, and the comparison between the POS tags obtained from written language and the POS tags obtained from orthographic transcriptions of read speech. More specifically, a text containing the sentence prompts (SPs) is first parsed by the tools, as standard written language, and then a text containing the orthographic transcriptions (OTs) from recordings of speakers reading the SPs is parsed as read speech, with the same parsing tools.

Afterwards, to obtain a better performance of the POS tagging tools, a computational adaptation involving the automatic addition of punctuation marks is applied on the OTs, and its results are discussed.

## 1.2 Research questions and hypothesis

The research questions (RQs) of this study are the following:

\* do the two parsers for Dutch language perform POS tagging similarly?

\* is the POS tagging of SPs more accurate than the POS tagging of OTs?

\* how and to what extent the presence of punctuation affects the POS tagging process?

According to the initial assumptions, the two parsing tools are expected to return a similar output. Moreover, since the parsing tools used in this study are designed with a reference point more akin to written language rather than read speech [4, 2, 3], like most parsing tools, greater accuracy is assumed to be performed for the POS tags of the SPs rather than for the POS tags of OTs. Also, it is expected that the POS tagging of the OTs would return a more accurate output after the insertion of punctuation marks in the OTs. A similar behavior is expected from the two parsing tools.

## 2 Materials and method

### 2.1 Materials and method: parsing tools, read speech, scripts

#### 2.1.1 Parsing tools

The software used in this research are named Frog [5] and Alpino [6], and they are both available at CLST, Center for Language and Speech Technology of Radboud University (<https://webservices.cls.ru.nl/>). Alpino is a hybrid dependency parser for Dutch, which uses rule-based constraints combined with corpus-based

statistics, while Frog is an NLP suite based on memory-based learning and trained on large quantities of manually annotated data. A basic knowledge of the SSH (Security SHell) and the access to LaMachine, a unified Natural Language Processing (NLP) open-source software distribution, were required to proceed with the parsing. Once the POS tags were extracted with Alpino and Frog, they were moved to an Excel file. For Alpino, this last step required the use of the Python script FoLiA [7] [8] written by Van Gompel and Bloem [9].

### *2.1.2 Read speech*

The read speech used in this study comes from a larger research project which involves atypical speech analysis [10]. A group of 8 Dutch native speakers suffering from dysarthria was asked to read a set of sentences before and after a therapy treatment. From the recordings, orthographic transcriptions (OTs) were later annotated manually. The OTs, together with the text file of the SPs, were POS tagged with Frog and Alpino. Note that the OTs do not report phenomena like filled pauses and lengthened vowels and do not have any punctuation marks, but contain repeated, stuttered, or fragmented words. In this work, only the OTs of the pre-processed recordings are taken into account, and the adaptation proposed in the second stage of

this study case concerns only the punctuation, while all the other typical elements of the read speech reported by the OTs (repeated words, stuttering, fragmented words) remain unchanged.

The sentence prompts are seven, and are taken from the story «Papa en Marloes» and from apple pie recipes also used by Ganzeboom et al. [11] [12]. The prompts include 32 sentences for a total of 250 words. Therefore, besides the prompts text, all the words annotated in the OTs, uttered by 8 speakers who were asked to read the prompts, were POS tagged separately for each speaker.

### *2.1.3 Scripts*

Two Python scripts were created specifically for this work. One script was used to compare the different POS tags, grouped by speakers, tools, and kind of text (SPs and OTs), and one script was used to automatically insert the punctuation in the OTs.

The first code was made with the purpose to return the number of mismatches between two given groups of POS tags. More specifically, what the code did was to observe the words and POS tags columns of two different Excel files, e.g., a file reporting the SPs POS tags obtained with Alpino and a file reporting the SPs POS tags obtained with Frog. The code was meant to check row by row if, for instance, the POS tag «ALPINO-tag1» corre-

sponding to the word» «*SPsWord1*» in the first Excel file equaled or differed from the POS tag «*FROG-tag1*» corresponding to the word «*SPsWord1*» of the second Excel file. Similarly, for the comparison between an Excel file showing POS tags obtained from SPs and a file showing POS tags obtained from OTs (with the POS tags being extracted by the same parsing tool), the code would check row by row whether the POS tag «*ALPINOTag1*», corresponding to the word «*SPsWord1*» in the first Excel file equaled or differed from the POS tag «*ALPINOTag1*» corresponding to the word «*OTsWord1*» of the second Excel file. To do this, the script also considered the order in which the words were listed, thus avoiding an overlap between identical words. Henceforth, those POS tags that according to the script differed between each other (e.g. «*ALPINOTag1*» from the first file differs from «*FROGtag1*» or «*ALPINOTag1*» from the second file) will be referred as *mismatches*.

The second script was made to automatically insert the punctuation in the OTs. The code considered the text of the SPs, which contained punctuation, and the text of the OTs, with no punctuation. The code first located and identified the punctuation marks showed in the SPs and then reinserted them into the OTs. It did so by establishing the insertion spot on the ba-

sis of the first three words preceding and the first three words following the punctuation marks. Considering the sequence of the three words preceding and following the full stops and the commas was crucial for a proper functioning of the coding. Otherwise, by looking only at the first word preceding and following the punctuation, there would have been the concrete risk to automatically add extra punctuation whenever a word was repeated more than once in the sentences. The same script was used for full stops and commas.

## 2.2 Pipeline

According to the adopted pipeline, the first step was to extract all POS tags at SPs and OTs level through Frog and Alpino.

Once the POS tags were collected for each word in Excels files, the comparison code was used for two main comparisons: a comparison between SPs POS tags obtained with Alpino and Frog, to observe the functioning of the two tools, and a comparison between the SPs and the OTs POS tags before and after the insertion of the punctuation, to evaluate the punctuation impact. The accuracy of the SPs POS tags was further checked by Dutch native speakers studying at Radboud University.

Afterwards, the punctuation was added to the OTs with the second script,

and the coding effectiveness was evaluated by counting the inserted punctuation. The comparison code was used again for the SPs POS tags and the OTs obtained both with Alpino and Frog, before and after the insertion of the punctuation.

Finally, all the comparison results were collected, and the behavior of the parsing tools was observed together with the impact of the punctuation in the OTs.

### 3 Observations and results

#### 3.1 Alpino vs. Frog

The SPs POS tags were used as control sample to evaluate the functioning of the two parsing tools. The comparison code revealed that amongst the 250 words included in the prompts, 27 were tagged differently by Alpino and Frog, hence 27 mismatches were found.

Table 1: Prompts POS tags mismatches. Comparison between parsing software

Detected mismatches	Alpino correct mismatches	Frog correct mismatches
27 (out of 250)	23	4

To understand the possible causes of the mismatches and which tool performed better, Dutch grammar skills proved necessary, therefore some native Dutch students from Radboud University were consulted. As shown in Table 1, with their help it turned out that amongst the 27 mismatches

revealed, 24 of them were due to a Frog misinterpretation, while only 3 of them were attributed to Alpino.

Table 2: Comparison between prompts and one (random) speaker

Parsing tool	Punctuation in OTs	POS tag mismatches
Frog	no punctuation	20 out of 250
Frog	punctuation	17 out of 250
Alpino	no punctuation	24 out of 250
Alpino	punctuation	7 out of 250

Thus, Alpino seemed to report a higher number of correct POS tags with respect to Frog. A further confirm of these results was given by the accuracy scores showed in the outputs of Frog. The scoring indeed was always low in correspondence of the mismatches.

#### 3.2 SPs vs. OTs POS tagging, before and after the insertion of punctuation

Also the comparison between SPs and OTs POS tags revealed interesting results. Table 2 shows the number of mismatches revealed from 4 different comparisons: first the number of mismatches revealed with the comparison between the SPs POS tags and the OTs POS tags of a random speaker before and after the insertion of punctuation marks. This first comparison was made twice, once with Alpino and once with Frog. Table 2 reports the number of POS tag mismatches of only one random sampled speaker, since it can be considered significant and



indicative for the entire set of speakers.

The number of mismatches confirmed again a different behavior of the two parsing tools. Before the automatic insertion of the punctuation marks, indeed, Alpino reported 4 mismatches more than Frog. However, once the punctuation was added in the OTs, the results showed a decrease of the number of mismatches for both Alpino and Frog, but the change was much more evident for Alpino. Indeed, if with Frog only 3 mismatched POS tags changed to correct POS tags after the insertion of punctuation, thus decreasing from 20 to 17 mismatches, with Alpino 17 mismatches turned correct, thus decreasing from 24 to 7 after the addition of punctuation.

#### 4 Discussion

The findings of this study case partly met the given assumptions and partly diverged from them.

The most unexpected result is undoubtedly such a noticeable difference between the two software's behaviour. While it was assumed that the results of Alpino and Frog would have been roughly similar, the findings proved otherwise, with regard to the comparison between the POS tagging of standard written language (SPs) and of read speech (OTs), as well as the sensitivity to punctuation marks in read speech. Indeed, it seemed that Alpino performed a more accurate POS tag-

ging than Frog on written language. Considering only the read speech instead, Alpino and Frog's POS tags did not show great difference, but they both reported a 9% of mismatches. After the insertion of punctuation, the difference between the two tools was again accentuated: although both parsers proved to be sensitive to punctuation, Alpino's POS tagging showed a stronger and more significant improvement than the Frog's minimal one. This clear difference between the two tools probably lies in the different operating mechanisms underlying the parsing functioning, which unfortunately are not investigated in detail in this work.

Nevertheless, a closer observation of the mismatches can lead to some interesting considerations. Out of the 27 mismatches revealed amongst SPs, 12 tags concern cases where a verb, correctly interpreted by Alpino, was recognised as a noun or an adjective by Frog. Since this seemed to be a recurrent error in Frog, its origins can be traced back to the syntactic structure of the sentences. Indeed, the sentence prompts from Apple Pie Recipe, also used in [11], are part of culinary recipes, therefore very often start with a verb in the infinitive form in the left-most position of the sentence. Thus, considering the order of the syntactic elements of a sentence in Dutch language (SVO/ SOV)[12], the elements



in those sentences appear in a non-standard order, since the subject is omitted. That could be a cause of the software misinterpretation of the verb as a noun.

On the other hand, the read speech proved to be more problematic for the POS tagging process, thus requiring some adaptations for a more correct parsing.

In short, the answers to the RQs of this study are as follows: the two parsing tools showed different behaviours, and proved to be more appropriate for written language than for read speech. Alpino turned to be the parser that most has met the initial assumptions, also showing that adding the punctuation to the OTs of a read speech could be an adaptation capable of clearly increasing the accuracy of POS tagging.

Table 3: Punctuation code accuracy in Ots

Speakers	Full stops	Commas
sp01	26	3
sp02	26	3
sp03	25	3
sp04	27	4
sp05	27	3
sp06	27	2
sp07	28	3
sp08	26	4
<b>Prompts</b>	<b>30</b>	<b>4</b>

The following observations on the limits of this research should be made:

- the accuracy of the punctuation

script is limited to the size of the small corpus used for the experiment. By way of example, Table 3 reports the number of commas and full stops that were successfully added with the script. Therefore, this script can be extended only to corpora of similar size (32 sentences), since the previous and subsequent words of the text are the parameters used to state where punctuation should be added. The more words the corpus have, the greater the risk of script malfunctioning.

- a statistical analysis would be the most appropriate way to analyse the data of this study and to obtain a scientifically valid result, but given the small sample size even a manual observation has allowed some interesting considerations.

- this work investigates the read speech of dysarthric speakers, but for a complete analysis and more defined conclusions, also healthy speakers should be examined.

This study provides an insight on only one of the possible methods for a more accurate POS tagging of read speech, and points out some flaws of the POS tagging process, that in the case of Dutch language did not seem to have a uniform mechanism based on standardized criteria. Future research amongst different kinds of language might extend the explanations of the punctuation relevance and of the aspects that distinguish a kind of

language from the other (written language, spoken language, read speech etc.). Future studies on POS tagging could also investigate whether it is possible to have a unified parsing mechanism working properly for every kind of language or whether it would be more appropriate to have different mechanisms for each kind of language.

## 5 References

- [1] Kumawat D. Jain V. (2015) POS tagging approaches: A comparison. *International Journal of Computer Applications*, 118(6).
- [2] Brill E. (1994) Some advances in transformation-based part of speech tagging. arXiv preprint [cmp-lg/9406010](https://arxiv.org/abs/1904.06010).
- [3] Manning C.D. (2011) Part-of-speech tagging from 97% to 100%: is it time for some linguistics? Proc. of the *International conference on intelligent text processing and computational linguistics*, 171-189.
- [4] Van den Bosch A., Schuurman I., Vandeghinste V. (2006) Transferring PoS-tagging and lemmatization tools from spoken to written Dutch corpus development. *Proceedings of the Fifth International Conference on Language Resources and Evaluation (LREC'06)*. Available from: [http://www.lrec-conf.org/proceedings/lrec2006/pdf/167\\_pdf.pdf](http://www.lrec-conf.org/proceedings/lrec2006/pdf/167_pdf.pdf).
- [5] Bosch A. van den, Busser B., Canisius S., Daelemans W. (2007) An efficient memory-based morphosyntactic tagger and parser for Dutch. *LOT Occasional Series*, 7, 191-206.
- [6] Bouma G., Van Noord G., Malouf R. (2000) Alpino: Wide-coverage Computational Analysis of Dutch, 37, 45-59.
- [7] Van Gompel M., Reynaert M. (2013) FoLiA: A practical XML format for linguistic annotation – a descriptive and comparative study. *Computational Linguistics in the Netherlands Journal*, 3, 63-81.
- [8] Van Gompel M., Sloot K., Reynaert M., Van den Bosch A. (2017) FoLiA in Practice. The Infrastructure of a Linguistic Annotation Format.
- [9] Van Gompel M., Bloem J. (2021) folia2columns. Script available at <https://github.com/proycon/foliatools/blob/master/foliatools/folia2columns.py>.
- [10] Pesenti C., van Bommel L., van Hout R., Strik H. (2022). The effect of eHealth training on dysarthric speech. Paper accepted for the LREC 2022 workshop RAPID.
- [11] Ganzeboom M., Bakker M., Beijer L., Rietveld T., Strik H. (2018) Speech training for neurological patients using a serious game. *British Journal of Educational Technology*, 49(4), 761-774.
- [12] Ganzeboom M., Bakker M., Beijer L., Strik H. & Rietveld T. (2022) A serious game for speech training in dysarthric speakers with Parkinson's disease: Exploring therapeutic efficacy and patient sati-

sfaction. *International Journal of Language & Communication Disorders*, forthcoming.

[13] Koster J. et al. (1975) Dutch as an SOV language. *Linguistic analysis*, 1(2), 111-136.

# PHONEWS



## GIORNATA MONDIALE DELLA VOCE

GIOVEDÌ 28 APRILE 2022

Aula Mole - CLA (Palazzo Nuovo, piano seminterrato)

h 14.30 - 15.00

**DARIO STRANGIS**

*Balbutie e cluttering:  
orgoglio (stereotipi) e pregiudizio*

h 15.00 - 15.30

**ALICE PIANO**

*Fatti sentire (anche dal fondo della platea)*

h 15.30 - 16.00

**FIAMMETTA FANARI**

*La voce della donna transgender*

h 16.00 - 16.30

INTERMEZZO MUSICALE

**GIULIA IMPACHE E BIANCA DE PAOLIS**

*Tanti generi, tante voci*

h 16.30 - 17.00

**MASSIMO SPADOLA BISETTI**

*La tutela della voce nel contesto lavorativo*

h 17.00 - 17.30

**MAX GIARDINI**

*L'artigiano attore e la lettura interpretata*

h 17.30 - 18.00

**DIEGO COSSU**

*La voce del bambino:  
tra possibilità di crescita e rischio performativo*

h 18.00 - 18.30

**VALENTINA COLONNA**

*Parlami ancora come fosse mia la voce*



*Lift Your Voice*  
world voice day 16th april



e in diretta YouTube!

[https://www.youtube.com/watch?v=e9\\_9IIvqMb0](https://www.youtube.com/watch?v=e9_9IIvqMb0)

## In memoriam



**VESNA DEŽELJIN**

(Zagabria, 1958-2021)

Vesna Deželjin era nata Zagabria (Croazia) dove si era formata all'Università, presso la Facoltà di Filosofia, laureandosi in lingue e letterature straniere con la specializzazione in inglese e in italiano. Già durante gli studi aveva soggiornato presso le università italiane (Udine, Milano, Perugia, Trieste, Bari) dove aveva perfezionato le sue conoscenze, diventando un'ottima linguista.

Presso la Facoltà di Filosofia di Zagabria aveva concluso anche gli studi post-universitari, con la tesi intitolata *Talijanska književnost u hrvatskom tisku u razdoblju od 1918.do 1930.* [La letteratura italiana nella stampa croata nel periodo dal 1918 fino al 1930], mentre a partire dal 2003 aveva frequentato la Scuola di dottorato della Facoltà di Filosofia di Lubiana (Slovenia) dove si

era addottorata nel 2006, con una tesi sui contatti culturali e linguistici.

Vesna aveva iniziato la sua lunga e proficua esperienza lavorativa come insegnante di lingua inglese e italiana in diverse scuole private a Zagabria, fino al 1993, anno in cui era diventata lettrice di Lingua italiana presso il Dipartimento di Lingua e letteratura italiana della Facoltà di Filosofia di Zagabria, dove aveva prestato il servizio anche successivamente come collaboratore scientifico (ricercatore e ricercatore confermato) e poi professore associato presso l'Università di Fiume.

Nel corso della sua carriera, Vesna ha partecipato a diversi convegni sia in Croazia sia all'estero (Italia, Danimarca, Germania, Macedonia, Polonia, Romania, Spagna, Stati Uniti, Ungheria). È stata inoltre invitata come *visiting professor* nell'ambito di scambi universitari e di Erasmus+ presso numerose università in Italia e all'estero (Firenze, Torino, Pescara, Pecz, Varsavia). È stata apprezzata come consulente didattica della scuola estiva di Gemona organizzata da: Università Cattolica di Milano, Università degli Studi di Udine e Università degli Studi di Trieste. Si è distinta anche partecipando a numerosi progetti di ricerca: «Hrvatsko-talijanski književni odnosi» (1994-2006), «Hrvatsko-talijanski kulturni odnosi» (2006-2013), «Labo-

ratorio internazionale della comunicazione» che era il risultato di collaborazione con l'Università Cattolica di Milano e con l'Università Statale di Milano. Per diversi anni (2014-2018) è stata vicedattrice della prestigiosa rivista croata «Strani jezici» e membro attivo di associazioni di linguisti e filologi: SILFI, A.I.P.I., AISLLI, Hrvatsko društvo za primjenjenu lingvistiku (Società croata di linguistica applicata), Hrvatsko udruženje profesora engleskog jezika (Società croata di professori di lingua inglese).

Le svariate e numerose attività di Vesna Deželjin comprendono anche i suoi impegni presso il Ministero croato di educazione e sport, dove era coinvolta nella valutazione dei manuali

di lingua italiana e dei materiali didattici legati alle nuove tecnologie.

Ha pubblicato un centinaio di contributi in italiano, in croato e in inglese che trattano argomenti filologici, in particolare di linguistica italiana, un volume tratto dalla tesi di dottorato *Elementi alloglotti nella prosa dialogata degli scrittori triestini Carpinteri e Fataguna come riflesso di contatti culturali e linguistici* (2012), le recensioni e le traduzioni.

La prematura e inaspettata scomparsa di Vesna Deželjin ha colpito profondamente i suoi colleghi italiani, italianisti e slavisti, interrompendo la collaborazione segnata dall'amicizia e da tanta energia che Vesna trasmetteva a tutti coloro che l'hanno conosciuta.

LJILJANA BANJANIN



### **ADRIANA TERESA DAMASCELLI**

(Torino, 11 maggio 1966 – 2 gennaio 2022)

Adriana Damascelli era responsabile operativa del Centro Linguistico di Ateneo (CLA-UniTO) quando – dopo diversi anni di sofferenza e di strenua lotta contro la malattia che le aveva progressivamente impedito di deambulare – ci ha lasciati il 2 gennaio 2022. Era Dottoressa di ricerca in Anglistica (indirizzo Linguistica inglese) e aveva anche conseguito l'Abilitazione scientifica nazionale (ASN 2012-2013) a Professore Associato per il *ssd* L-Lin/12, Lingua e traduzione inglese. Era stata docente a contratto di «Valutazione e autovalutazione nella metodologia CLIL» (presso il CLA-UniTO) e aveva svolto seminari e tenuto moduli didattici di «Tecnologie applicate alla traduzione» (presso i corsi di vari docenti). Aveva insegnato «Inglese propedeutico» a specialisti di altre discipline (presso le Facoltà di Medicina e Chirurgia, Scienze della Forma-

zione e Psicologia, Agraria e Medicina Veterinaria dell'Università degli Studi di Torino).

Adriana aveva iniziato la carriera nel 1999-2000 come Consulente Lessicografa c/o DIMA LOGIC (Torino), occupandosi della revisione dei lemmi del dizionario italiano/inglese Oxford-Paravia, ma aveva progressivamente acquisito conoscenze rilevanti nel campo della linguistica computazionale e della traduzione automatica.

Dal 2001-2002 aveva creato una prima versione del sito *internet* del CLA (ex-CLIFU: [www.clifu.unito.it](http://www.clifu.unito.it)) di cui è rimasta Redattrice *web* fino all'ultima revisione che ha portato al sito attuale (<https://www.cla.unito.it>).

Tra i tanti punti di contatto col *LFSAG* (spesso coinvolto anche nell'organizzazione degli eventi della Giornata Europea delle Lingue che si svolge annualmente presso il CLA), ricordo in particolare il lavoro svolto, insieme alla Prof.ssa Vittoz e al suo staff e con Valentina De Iacovo, in occasione di alcune attività delle Scuole Estive e del XX seminario nazionale dell'AICLU (Torino, CLA-UniTO, 29-31 maggio 2014). Avevamo anche trascorso una giornata insieme a Grenoble, partecipando alla conferenza biennale *Précopil – Innovalangues* (organizzata da Monica Masperi il 21/01/2015) che aveva consolidato l'idea di un percorso di ricerca congiunto sulla fonodidattica

della prosodia (poi sfociato, tardivamente, in *CALL-UniTO*, dopo che, dal 2017, le delusioni accademiche e la malattia l'avevano allontanata da vari progetti).

Tra le sue diverse pubblicazioni, ricordo in particolare le seguenti:

A. Martelli & A. Damascelli (2002). *Corpus Linguistics and Computational Linguistics: an Overview with Special Reference to English*, Torino: CELID.

V. Pulcini & A. Damascelli (2005). A corpus-based study of the discourse marker «okay», in A. Bertacca (a cura di), *Plus - Università di Pisa Historical Linguistic Studies of Spoken English* (Papers read at the 11<sup>th</sup> Italian Conference on the History of the English Language, Pisa 5-7 June 2003), Pisa:

Edizioni PLUS, 231-243.

A. Damascelli (a cura di) (2017). *Digital Resources, Creativity, Innovative Methodologies and Plurilingualism: New Approaches to Language Teaching and Learning*, Newcastle upon Tyne: Cambridge Scholar Publishing.

A lei, il CLA ha dedicato una pagina di ricordi con diverse toccanti testimonianze di colleghi e amici ([www.cladrianadamascelli.unito.it](http://www.cladrianadamascelli.unito.it)), ma un importante evento è programmato per il 26 settembre 2022 in occasione di una Giornata di studio in suo ricordo quando, al termine della mattinata, sarà posata una targa nell'aula SMART con dedica ad «Adriana Teresa Damascelli, già Responsabile Operativa del Centro».

ANTONIO ROMANO





### GIOVANNI RONCO

Torino 27 febbraio 1953 – 27 maggio 2022

A distanza di un mese, ci sembra ancora assurdo che Giovanni Ronco se ne sia andato. Giovanni ha rappresentato l'anima dell'Istituto dell'*Atlante Linguistico Italiano (ALI)*, in cui aveva cominciato a collaborare sin dal conseguimento della sua Laurea in Lettere (Dialectologia Italiana) presso l'Università di Torino nell'a.a. 1975-76, partecipando ai progetti di ricerca guidati in quegli anni da C. Grassi, G.L. Beccaria, A. Genre e L. Massobrio. Il suo primo impegno scientifico era stato quello di redattore del *Grande Dizionario della Lingua Italiana UTET (GDLI)*, di S. Battaglia), sotto la direzione di G. Barberi Squarotti, dal vol. XI fino alla redazione degli indici degli autori, nel 2004. In qualità di lessicografo, aveva partecipato anche ad altre imprese come le collaborazioni col *Dizionario della Lin-*

*gua Italiana Moderna (Garzanti)*, il *Lessico Etimologico Italiano (LEI)* e il *Repertorio Etimologico piemontese (REP)* e, negli stessi anni, era stato occasionalmente coinvolto anche in attività legate alla fonetica sperimentale (di cui era stato cultore) e alla (socio-)linguistica giudiziaria.

Nel suo ruolo di redattore, caporedattore e, dal 2011, condirettore dell'*Atlante Linguistico Italiano* aveva dedicato all'Istituto la sua costante, immancabile quotidiana presenza. Il suo apporto resta essenziale non solo per le sue ricerche in ambito dialettologico, ma anche per aver contribuito alla pubblicazione dei volumi I-IX. Aveva anche partecipato puntualmente, dal 1979 al 2021, a tutte le relazioni annuali sul lavoro d'équipe dell'*ALI* pubblicate sul *Bollettino dell'ALI*, rivista annuale di cui era stato segretario di redazione (dal 1987 al 1993) e a cui aveva affidato diversi contributi scientifici individuali.

Sempre nell'ambito dialettologico, operava sul territorio al fianco di varie associazioni e centri di tutela del patrimonio linguistico locale, ma non era mancato il suo investimento in grandi cantieri regionali, come l'*Atlante Toponomastico del Piemonte Montano (ATPM)* o l'*Atlante Linguistico-Etnografico del Piemonte Occidentale (ALEPO)*. Per quanto spesso temporanee e/o indirette, ricordiamo anche le sue collaborazioni con progetti internazionali come l'*Atlas Linguistique*

Roman (*ALiR*) e l'*Atlas Linguarum Europae* (*ALE*). Oltre che membro di diverse società scientifiche (*SLI*, *ASLI*), era stato referente d'area di varie riviste locali o nazionali (come la *RID*, *Rivista Italiana di Dialettologia*) e aveva partecipato a numerose iniziative di terza missione, spesso relative all'insegnamento di elementi di fonetica, grafia e grammatica dialettale.

Giovanni era stato assunto inizialmente come Funzionario Tecnico (1991) presso il *Centro Linguistico d'Ateneo* (all'epoca *CLAU*) con distacco presso l'Istituto *ALI*, ma aveva poi vinto un concorso da Ricercatore (2004) e uno da Professore Associato (2007) nel settore L-FIL-LETT/12, Linguistica Italiana, di cui assicurava vari insegnamenti presso il Dip. di Lingue e Lett. Straniere e C.M. (all'interno del quale era coordinatore della sezione di Linguistica). Dal 2006 teneva anche un insegnamento di Lingue e Culture del Piemonte nel Master di I livello in Promozione Turistico-culturale della Regione Piemonte.

A una ricca, pluridecennale bibliografia di studi (con decine di recensioni e segnalazioni) su questi temi aveva aggiunto un interessante contributo, in collaborazione con altre figure prominenti dell'*ALI* (F. Cugno & M. Rivoira): il primo volume del *PALP*, *Piccolo Atlante Linguistico del Piemonte* (2018).

Il rapporto con il *LFSAG* è stato fin

dall'inizio fondamentale ed è rimasto saldo nel tempo: Giovanni era infatti consulente linguistico per le aree piemontese e toscana (anche per via di una sua familiarità con la Val d'Arno) e aveva assicurato la prima o la seconda relazione di diverse tesi di laurea insieme ad A. Romano. Non solo: nel 2015 l'equipe del laboratorio aveva preso parte insieme a lui, E. Carpitelli e J. Bucci al Progetto «Galileo UIF G14-126» per allestire una base di dati dialettali per lo studio del dialetto di Corato (Puglia) e i suoi consigli nell'allestimento dell'archivio (ora disponibile online) erano stati di grande aiuto. Qualche anno dopo, c'era stata una nuova collaborazione tra il *LFSAG* e l'*ALI* per presentare le basi di dati dialettali e gli archivi vocali di lingue locali e minoritarie nell'opera «*Documenti sonori. Voce, suono, musica in archivi e raccolte*» (a cura di Brunetti, D. Robotti e E. Salvalaggio, 2021) e anche in quel caso, oltre a contribuire documentando i progetti in cui era coinvolto, aveva lavorato all'armonizzazione generale delle diverse sezioni.

Insomma, Giovanni si è sempre speso per testimoniare quello stretto rapporto che intercorre tra le parole e la quotidianità di una comunità linguistica, con quel suo atteggiamento garbato, ospitale e mai banale, che traspariva subito, dalle parole che sceglieva e dal quale abbiamo potuto soltanto imparare.

VALENTINA DE IACOVO  
e ANTONIO ROMANO

# Norme editoriali

La formattazione dell'articolo inviato (25000 battute massimo, comprese note e bibliografia, spazi inclusi) deve avvenire in pagine di formato A5 con l'impiego del carattere Garamond Unicode<sup>1</sup>, corpo 11; interlinea semplice, senza sillabazione. Dopo il titolo, indicare nome e cognome dell'autore e affiliazione corrente. Testo indentato: rientro 0,5 cm.

La numerazione (e il formato) dei titoli dei paragrafi (senza punti finali, non indentati) deve uniformarsi ai seguenti esempi: **1. Xxxxx**, 1.1. Xxxxxy, ..., 1.2. Xxxyy, 1.2.1. Xxyyy, 1.2.2 Xyyy, **2. Yyyyy...**

I grassetti nel testo vanno usati con molta moderazione; i corsivi solo per simboli isolati e forestierismi.

Il testo dev'essere articolato in paragrafi. In generale: un'introduzione al tema, trattato con rimando ai principali contributi di fonti autorevoli sui diversi argomenti attinenti con la ricerca presentata (stato dell'arte), un secondo paragrafo di presentazione degli strumenti e dei dati usati nell'ambito della tesi per arricchire le conoscenze in quel dato settore (protocollo d'indagine) e un terzo di presentazione, commento e interpre-

tazione dei risultati ottenuti, anche questo in riferimento a risultati simili di altre fonti (o in disaccordo con queste).

Il riferimento alle fonti avviene nel testo con l'indicazione del nome dell'autore (data) (es: «nella descrizione riservata a questo fenomeno da Mereu (2004), Vayra *et alii* (2007)...»).

Le citazioni testuali vanno virgolettate (con l'indicazione della fonte: tra parentesi il nome dell'Autore Data: Pagina).

Es.:

«In queste condizioni, ci si può chiedere quale spazio sussista per effettuare indagini fonetiche circa la realizzazione e la ricezione degli elementi prosodici» (Bertinetto 1981: 37).

Figure e tabelle devono essere corredate da una didascalia (nella quale, se l'immagine o i dati non sono frutto di un lavoro originale, si rinvia alla fonte). Le immagini, rigorosamente in bianco e nero, oltre a essere inserite nel testo, devono essere allegate separatamente verificando che siano mantenuti i contrasti tra le diverse tonalità prescelte.

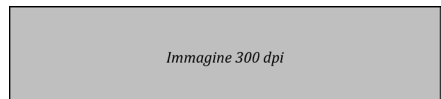


Figura 1. Immagine raffigurante uno schema delle funzioni dell'accento (tratta da Bertinetto 1981: 43).

<sup>1</sup> Scaricabile gratis, e.g., dal sito [www.wfonts.com/font/garamond](http://www.wfonts.com/font/garamond).

Traduzioni, commenti e rimandi bibliografici ritenuti secondari vanno in nota a piè di pagina (di cui è conigliato l'impiego con moderazione).

Accorgimenti tipografici: evitare spazi doppi e tabulazioni; l'apostrofo (') è diverso dall'apice (^) e dalla virgoletta semplice aperta ("); le virgolette sono di diverso tipo, ma devono essere usate coerentemente (aperta-chiusa: " " o « » etc.); l'aferesi e l'elisione si indicano con ('). I ganci semplici (< >) possono essere usati per evidenziare le forme grafiche, mentre le rappresentazioni fonologiche sono precedute e seguite da una barra obliqua (/) e le forme fonetiche racchiuse tra parentesi quadre ([ ]). Evitare pseudo-eufonismi come <ad>, <ed>, <od> (riservandoli solo al necessario; es. «ed eventuali», «ad altri», «od opportuni»). Si noti ancora che si ha <perché> e non \*<perchè>, <ciò> e non \*<ciòé>, <po> e non \*<po> etc.; il maiuscolo di è non corrisponde a \*<E>, ma a <È>...

## Riferimenti bibliografici

(in fondo al testo).

Esempi:

Bertinetto P.M. (1981). *Strutture prosodiche dell'italiano*. Firenze, Accademia della Crusca.

Bertinetto P.M. & Magno Caldognetto E. (1993). Ritmo e intonazione, In A.A. Sobrero (a cura di) (1993a), 141-192.

Cho T. & Ladefoged P. (1999). Variations and universals in VOT: evidence from 18 languages, *Journal of Phonetics*, 27, 207-229.

Levinson S.C. (1983). *Pragmatics*, Cambridge, Cambridge University Press (trad. it. *La Pragmatica*, Bologna, Il Mulino, 1985).

Mereu L. (2004). *La sintassi delle lingue del mondo*, Roma-Bari, Laterza.

Sobrero A.A. (a cura di) (1993a). *Introduzione all'italiano contemporaneo: le strutture*, Roma-Bari, Laterza.

Sobrero A.A. (a cura di) (1993b). *Introduzione all'italiano contemporaneo: la variazione e gli usi*, Roma-Bari, Laterza.

Vayra M., Avesani C. & Fowler C. (1984). Patterns of temporal compression in spoken Italian. *Proceedings of the 10<sup>th</sup> ICPbS* (Utrecht, The Netherlands, 1983), 2, 541-546.

## Sitografia

AMPER-ITA - *Atlas Multimédia Prosodique de l'Espace Roman*. [www.lf-sag.unito.it/amper-ita](http://www.lf-sag.unito.it/amper-ita) (ultimo accesso 20/03/2018).

**SUJET EN EXPÉRIENCE**  
**AVEC LE GLOTTOGRAPHE ÉLECTRIQUE DE PH. FABRE**  
Fig. 10 da *Physiologie de la phonation*  
di RAOUL HUSSON, Paris, Masson et. C<sup>e</sup>, 1962.

L'écran de l'oscillographe, à droite, montre trois périodes du glottogramme donné par le sujet.

