

JACOPO SATURNO

# Continuo e discreto nell’intercomprensione tra lingue romanze: distanze e corrispondenze interlinguistiche

Il presente contributo simula il processo di comprensione del lessico catalano da parte di italoфoni privi di competenza nella lingua bersaglio. In particolare, lo studio si propone di mettere sistematicamente a confronto diverse variabili sia continue, sia discrete, così da stabilirne l’eventuale impatto sul riconoscimento di parole imparentate. Ai 231 informanti coinvolti si chiedeva la traduzione italiana di 24 pseudo-parole, ottenute manipolando alcuni nomi italiani relativamente ai seguenti parametri discreti: a) modo di articolazione o luogo di articolazione o grado di sonorità di una consonante; b) sillaba comprendente la consonante modificata; c) lunghezza in termini di sillabe. Le misure continue in esame comprendono invece la distanza di Levenshtein normalizzata rispetto al modello italiano, nonché alcune statistiche estratte dal corpus *PhonItalia*. Al netto della presenza di un contesto, l’analisi statistica rivela un’interazione complessa tra parametri discreti e misure continue.

*Parole chiave:* intercomprensione, distanza interlinguistica, parametri articolatori, sillabe, frequenza.

## 1. Introduzione

Il presente lavoro si colloca nell’ambito degli studi sull’intercomprensione (IC), qui definita come la possibilità di comprendere almeno parzialmente una lingua straniera che non si è mai studiata intenzionalmente, ma che è prossima a una lingua già presente nel proprio repertorio (tra cui in primo luogo la lingua madre). Tale facoltà è resa possibile dalla conoscenza da parte degli interlocutori di una “lingua ponte” prossima alla lingua bersaglio in termini di lessico e grammatica (Marx & Mehlhorn 2010). È dunque evidente come il concetto di somiglianza, o piuttosto di distanza, ricopra un ruolo cruciale nello studio delle dinamiche dell’intercomprensione.

L'indagine si propone di mettere a confronto due diverse concettualizzazioni di distanza, analizzandone l'effetto sul riconoscimento di parole etimologicamente imparentate, ma appartenenti a due varietà distinte. Da un lato si prenderanno in considerazione differenze discrete, cioè specifici tipi di corrispondenze interlinguistiche, per cui a un segmento o tratto della parola di una lingua corrisponde un diverso tratto o segmento nell'altra lingua, es. it. /kj/ - fr. /kl/ in *chiave*, *clé*. Dall'altro lato si esprimerà la distanza interlinguistica mediante un unico valore continuo capace di sintetizzare la proporzione di elementi grafici o fonologici divergenti tra le due parole.

### 1.1 *Quadro teorico*

La distanza tra varietà linguistiche è comunemente misurata mediante la “distanza di Levenshtein” (1966), determinata dal numero minimo di operazioni (inserzione, cancellazione o sostituzione di segmenti) necessarie per trasformare una stringa (di elementi grafici o fonologici) in un'altra (Heeringa 2004; Nerbonne & Heeringa 2009). Dividendo il numero di operazioni per la lunghezza totale dell'allineamento si ottiene un valore “normalizzato”, cioè compreso tra 0 e 1 e perciò confrontabile con altri. La Tabella 1 mostra come la distanza di Levenshtein possa differire a seconda del fatto che venga calcolata sulla base della rappresentazione ortografica o fonologica degli elementi considerati. Al tempo stesso, il medesimo numero di operazioni porta a valori di distanza più elevati nel caso di allineamenti più brevi (cioè, tendenzialmente, di parole più brevi).

Tabella 1 - *Distanza di Levenshtein tra it. lancia e cat. llança*

	ortografia							fonologia1				
it. lancia	-	l	a	n	c	i	a	l	a	n	ʃ	a
cat. llança	l	l	a	n	ç	-	a	ʎ	a	n	s	ə
operazioni <sup>2</sup>	I	=	=	=	S	C	=	S	=	=	S	S
dist.					3						3	
dist. norm.					3/7 = 0,43						3/5 = 0,60	

<sup>1</sup> L'accento lessicale non è qui rappresentato per ragioni di semplicità, ma può essere facilmente implementato.

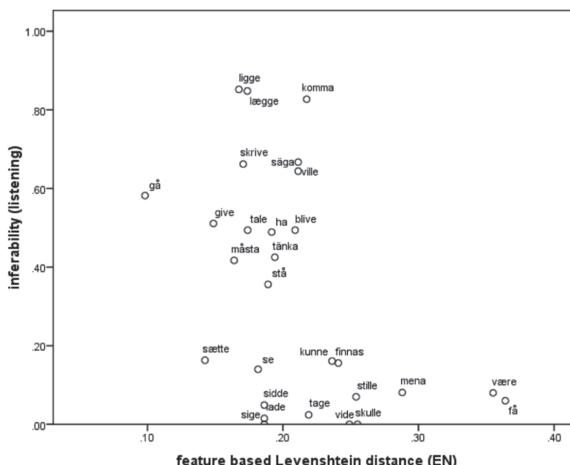
<sup>2</sup> I = inserzione, S = sostituzione, C = cancellazione.

Diversi studi mostrano che una volta controllata l'esposizione alla lingua bersaglio, l'intelligibilità di una lingua sconosciuta è fortemente correlata alla sua distanza da una lingua nota (Gooskens & van Heuven 2019). Ciononostante, ogni misura globale neutralizza le differenze esistenti tra diversi tipi di corrispondenze interlinguistiche, che potrebbero verosimilmente esercitare un peso diverso nel processo di intercomprensione: "since only aggregate phonetic distances were included in the investigation by Gooskens (2007), no conclusions could be drawn about the nature of the phonetic differences that contribute most to intelligibility" (Gooskens *et al.* 2008: 64).

In uno studio sull'intellegibilità di 17 varietà scandinave da parte di soggetti danesi, gli autori citati mostrano che solo le sostituzioni relative alle consonanti ( $r = -.57$ ,  $p < .05$ ) e l'abbreviamento delle vocali ( $r = -.49$ ,  $p < .05$ ) risultano correlati in modo significativo con i punteggi di intellegibilità.

In un esperimento sul riconoscimento di parole danesi da parte di soggetti con conoscenza del tedesco L1 e dell'inglese L2, Berthele (2011: 205) individua una soglia di distanza fra le parole bersaglio e la controparte inglese prossima al 22%, al di sopra della quale i punteggi di riconoscibilità diminuiscono drasticamente. Anche per valori di distanza inferiori a tale soglia e apparentemente sovrappponibili, tuttavia, si rileva un'estrema variabilità (Figura 1).

Figura 1 - *Relazione tra distanza di Levenshtein e intellegibilità nello studio di Berthele (2011: 205)*



L'autore suggerisce un possibile ruolo di variabili discrete quali specifiche corrispondenze fonologiche tra le lingue in contatto: “not mere distance/difference, but rather particular types of differences [may be] the key to the empirical differences across the items in this interlingual guessing puzzle” (Berthele 2011: 202).

L'analisi statistica che segue queste considerazioni evidenzia in effetti varie differenze interlinguistiche significativamente correlate all'intellegibilità, per quanto il quadro generale rimanga dominato da interazioni complesse con altri fattori relativi alle vocali, alla distanza interlinguistica globale e alle caratteristiche individuali dei rispondenti.

Nello studio di Möller & Zeevaert (2015), a 75 partecipanti tedesconi si chiedeva di tradurre una serie di 38 nomi tratti da varie lingue germaniche e divergenti rispetto a uno o due grafemi da varie parole tedesche, tra cui la controparte etimologica, es. rispettivamente olandese *lof* e danese *slugt* vs. tedesco *Lob* ‘lode’ e *Schlucht* ‘gola’. I risultati mostrano che in due diversi tipi di test la concentrazione delle risposte dei partecipanti intorno a specifiche corrispondenze interlinguistiche è significativamente maggiore di quanto ci si potrebbe attendere sulla base della sola distanza interlinguistica globale. Per esemplificare, anche se la parola bersaglio olandese *hommel* ‘bombo’ è ugualmente distante dalle possibili alternative tedesche *Himmel* ‘cielo’ e *Hummel* ‘bombo’, solo il 9% dei partecipanti opta per la prima.

Tra le generalizzazioni avanzate dagli autori, alcune sono direttamente rilevanti per il presente studio. In particolare, i rispondenti tendono a privilegiare le possibili alternative che rispetto alla controparte tedesca differiscono relativamente ai seguenti parametri:

- un segmento diverso dal primo;
- vocali (anziché consonanti);
- modo di articolazione (rispetto al luogo);
- sostituzioni (rispetto a inserzioni o cancellazioni).

Non si rilevano invece tendenze univoche relative alla posizione della corrispondenza interlinguistica all'interno della parola.

Nessuno studio di cui sia a conoscenza lo scrivente, tuttavia, ha finora tentato di identificare in maniera rigorosa l'eventuale ruolo di specifiche corrispondenze interlinguistiche: al contrario, le conclusioni appena riassunte sono state tipicamente ottenute sulla base della codifica a posteriori delle risposte dei partecipanti a stimoli non elab-

borati a questo scopo. Möller e Zeevaert evocano esplicitamente tale limite, insieme all’embrione della soluzione adottata in questo lavoro, cioè l’uso di pseudo-parole, che pure non sviluppano in maniera sistematica: “if the test items are to be pairs of cognates which really exist [...], often it is not possible to avoid inextricable combinations of factors” (Möller & Zeevaert 2015: 324).

Gli studi sulle variabili più rilevanti per il riconoscimento di parole tendono a basarsi su parole presentate in isolamento, nonostante l’inevitabile artificiosità dell’esperimento risultante, in quanto quest’ultima variabile ha spesso un effetto dirimente e difficilmente controllabile (Berthele 2011: 216; Möller & Zeevaert 2015: 315; Stenger, Jagrova & Avgustinova 2020: 46). Diversi lavori mostrano come anche scelte lessicali corrette – riguardanti perfino internazionalismi molto simili nelle lingue in contatto – possano essere messe in discussione perché ritenute non adeguate al contesto (Heinz 2009; Jágrová *et al.* 2019). Il numero di possibili traduzioni di un termine ignoto diminuisce infatti all’augmentare dell’informazione contestuale, in quanto le potenziali opzioni devono adattarsi a una cornice sintattica e semantica sempre più specifica (Muikku-Werner 2014). Möller & Zeevaert (2015: 345) concludono che “even the phonetic intuitions that turned out to be available in the recognition of isolated words seem to take a back seat as soon as semantic context is available”.

Mediante un opportuno approccio sperimentale, in questo lavoro si cercherà di gettare luce anche sull’eventuale interazione tra specifiche corrispondenze interlinguistiche e la presenza o assenza di un contesto.

## 1.2 *Domande di ricerca*

Sullo sfondo del quadro teorico appena delineato, il presente contributo intende individuare le variabili più determinanti per il riconoscimento di parole a) appartenenti a una lingua non nota, ma genealogicamente prossima a una lingua nota al rispondente, e b) etimologicamente imparentate alla controparte nella lingua nota. In continuità con molti degli studi appena citati, si prenderà dapprima in esame una variabile continua, cioè la distanza di Levenshtein tra la parola bersaglio e la controparte italiana. Lo studio valuterà tuttavia anche l’eventuale impatto delle seguenti variabili discrete, le quali potrebbero contribuire a spiegare parte di quella varianza che i modelli comprendenti la sola distanza interlinguistica non riescono a chiarire:

- corrispondenze interlinguistiche relative a parametri articolatori come luogo, modo e grado di articolazione;
- lunghezza in termini di sillabe rispetto alla controparte italiana;
- posizione della sillaba contenente la manipolazione;
- presenza di un contesto.

## 2. Metodologia

Per affrontare le domande di ricerca sopra delineate, si è chiesto ai partecipanti di compilare un questionario online composto di due parti. Nella prima si chiedeva la traduzione di 20 stimoli – presentati ai partecipanti come reali parole catalane – in assenza di contesto (1a: condizione “PAROLA”). Nella seconda parte, i medesimi stimoli comparivano in una frase catalana (1b: condizione “FRASE”).

(1) a. CONDADIN.

b. *El meu avi era un CONDADIN i cada dia es despertava d' hora.*  
 (‘mio nonno era un contadino e ogni giorno si alzava presto’)

In aggiunta alle due condizioni appena descritte, a un piccolo gruppo di partecipanti (n=15) è stato chiesto di indicare la parola a loro parere più adatta per colmare una lacuna presente nelle frasi stimolo (indicata da \*\*\*\*\* in 2). Gli stimoli utilizzati in quest’ultima condizione – presentata per prima nell’ambito dell’esperimento – erano i medesimi della condizione “FRASE”, a cui era stato rimosso l’elemento bersaglio. Si è deciso di introdurre questo ulteriore passaggio al fine di valutare la prevedibilità dell’elemento bersaglio alla luce della frase che lo contiene. I valori risultanti (consultabili nella Tabella 6, colonna “pred”) sono poi stati inclusi nel modello statistico come variabile di controllo.

(2) *El meu avi era un \*\*\*\*\* i cada dia es despertava d' hora.*

Pare ora opportuno chiarire per quale motivo una domanda di ricerca pertinente alla fonologia sia stata affrontata attraverso il mezzo scritto.

In primo luogo, il fatto che in un contesto di intercomprensione gli stimoli scritti siano immediatamente verbalizzati è esplicitamente riportato da Möller & Zeevat’s (2015: 339-340):

nearly all subjects started the different types of tasks by reading aloud the respective unknown words, and [...] many of them repeated this

during the free search. [...] All subjects pronounced the words according to the L1 rules

In secondo luogo, la ricerca psicolinguistica ha dimostrato che la rappresentazione grafica di una parola ne attiva la rappresentazione fonologica (Rastle & Brysbaert 2006; Comesáñ *et al.* 2016; Schmandt *et al.* 2022): le differenze grafematiche tra due elementi dovrebbero dunque rimandare alle differenze fonologiche da essi rappresentate.

Infine, lo studio intendeva simulare un compito di intercomprensione scritta, largamente predominante nella maggior parte degli approssimi didattici dedicati a questa pratica e probabilmente nell'esperienza concreta di intercomprensione.

### 2.1 *Elementi bersaglio*

Gli elementi bersaglio dello studio (si veda la Tabella 6 nell'Appendice) sono costituiti da 24 pseudo-parole ottenute a partire da parole italiane prive di una controparte etimologica in catalano, tutte coerenti con la fonotassi e la morfologia del catalano.

In particolare, gli stimoli sono stati ottenuti manipolando le parole italiane di partenza rispetto ai tre seguenti parametri discreti (Tabella 2):

- un solo tratto articolatorio (tra modo, luogo e grado di articolazione) di una sola consonante (riga “parametro”);
- sillaba comprendente la consonante modificata (riga “sillaba modificata”);
- lunghezza in termini di sillabe rispetto alla parola italiana (riga “numero sillabe”).

Tabella 2 - *Elaborazione di pseudo-parole dalle parole italiane contadino e dogana*

<i>italiano</i>	contadino	dogana
<i>pseudo-parola</i>	<i>CONSADIN</i>	<i>TOGANÀ</i>
<i>parametro</i>	<i>modo (/t/ &gt; /s/)</i>	<i>grado (/d/ &gt; /t/)</i>
<i>sillaba modificata</i>	<i>2</i>	<i>1</i>
<i>numero sillabe</i>	<i>-1</i>	<i>=</i>
<i>distanza</i>	<i>0,22</i>	<i>0,17</i>

Le sezioni che seguono descrivono nel dettaglio le manipolazioni effettuate per ciascuno dei tre parametri.

Per quanto riguarda la manipolazione delle consonanti in termini articolatori, si è prestata particolare attenzione a far sì che gli stimoli potessero essere letti in maniera univoca sulla base dell'ortografia italiana. Si sono in particolare esclusi quei grafemi che nelle lingue europee corrispondono a diversi suoni, es. <ch> (it. /k/, sp. /tʃ/, ted. /χ/ ecc.). Alla luce del requisito appena citato e del repertorio fonematico dell'italiano, si è scelto di manipolare i suoni presenti nella Tabella 3.

Tabella 3 - *Suoni italiani considerati e loro manipolazioni*

<i>suono</i>	<i>luogo</i>	<i>modo</i>	<i>grado</i>
/b/	/d/	/m/	/p/
/d/	/b/, /g/	/n/, /r/	/t/
/t/	/k/, /p/	/s/	/d/
/z/ <sup>3</sup>	/v/	/n/	/s/

Le manipolazioni appena descritte sono state applicate alle consonanti comprese nella prima *oppure* nella seconda sillaba della parola italiana modello. Si sono invece lasciate inalterate le consonanti comprese nella sillaba finale.

Relativamente al numero di sillabe di cui sono composte rispetto all'originale italiano, le pseudo-parole si possono dividere in tre categorie:

- 0: identico numero di sillabe, es. ASPA, cfr. it. *asta*;
- -1: la non parola è priva della vocale finale, es. GAD, cfr. it. *dado*. Si tratta di pseudo-parole coerenti con la morfologia delle parole catalane di genere maschile;
- +1: eccettuata la vocale finale, che è assente come nella precedente categoria, la pseudo-parola presenta una vocale in più rispetto alla controparte italiana, es. VAGUARD, cfr. it. *sguardo*. Per quanto il numero di sillabe sia il medesimo nella pseudo-parola e nell'originale italiano, cambia radicalmente la struttura sillabica. Anche in

<sup>3</sup> Il suono /z/ è stato ottenuto selezionando i contesti ortografici in cui la lettera <s> ha obbligatoriamente questa pronuncia in tutte le varietà di italiano, cioè in posizione iniziale di parola e seguita da una consonante sonora, es. *sguardo*.

questo caso si tratta di pseudo-parole coerenti con la morfologia delle parole catalane di genere maschile.

Le modifiche discrete appena descritte sortiscono anche un effetto sulla variabile continua della distanza interlinguistica, qui rappresentata dalla distanza di Levenshtein (normalizzata) fra la rappresentazione fonologica delle pseudo-parole e la controparte italiana.

Si è ritenuto opportuno controllare statisticamente alcune variabili note per il loro impatto sul riconoscimento di parole in condizioni disturbate. Sulla base del corpus *PhonItalia* (Goslin *et al.* 2014), ciascuna parola italiana modello è stata descritta nei termini delle seguenti statistiche lessicali:

- fqTotL: logaritmo naturale della frequenza totale della forma di parola;
- Phon\_N: “vicinato fonologico” (ingl. *phonological neighbourhood*), cioè il numero di parole divergenti dalla forma in questione per un solo segmento<sup>4</sup>;
- Phon\_N\_MFreq: logaritmo naturale della frequenza totale delle parole comprese nel gruppo precedente.

## 2.2 *Partecipanti*

Lo studio prende in considerazione le risposte di 231 partecipanti, in massima parte iscritti al Dipartimento di Lingue, Letterature e Culture Straniere dell’Università di Bergamo, oppure a vario titolo coinvolti nel settore dell’istruzione e interessati all’italiano L2. Con l’eccezione di cinque partecipanti, tutti avevano l’italiano come lingua madre. Per quanto riguarda la conoscenza del catalano, solo un’esigua minoranza dichiara una competenza di livello 2 (n=10) o 3 (n=2) su una scala da 1 (nessuna conoscenza) a 5 (madrelingua). Si è deciso tuttavia di non escludere dai dati questi partecipanti, principalmente perché nella condizione “PAROLA”, sulla quale si concentrerà l’analisi quantitativa, anche una conoscenza approfondita del catalano non sarebbe stata evidentemente di alcun aiuto per la decodifica delle pseudo-parole.

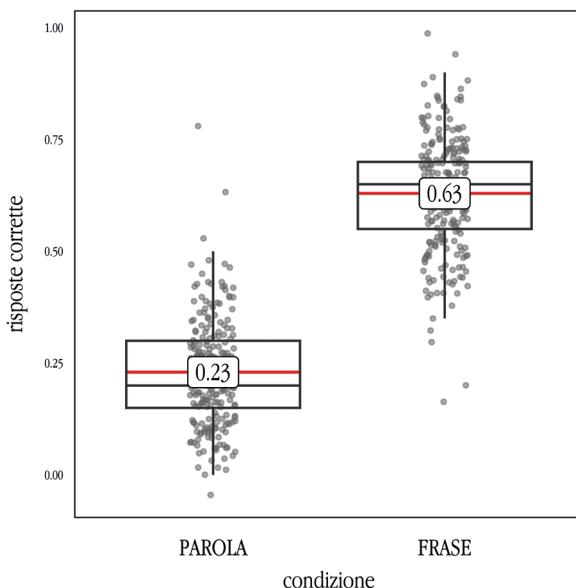
---

<sup>4</sup> In una lingua come l’italiano, questa definizione di vicinato fonologico comprende anche le forme di parola del medesimo lessema, es. *contadino*, *contadini*, nonché lessemi separati, ma semanticamente prossimi, es. *contadino*, *contadina*.

### 2.3 Risultati

Si rileva in primo luogo un ruolo macroscopico per la modalità di presentazione degli elementi bersaglio (Figura 2<sup>5</sup>).

Figura 2 - *Punteggi medi per condizione*

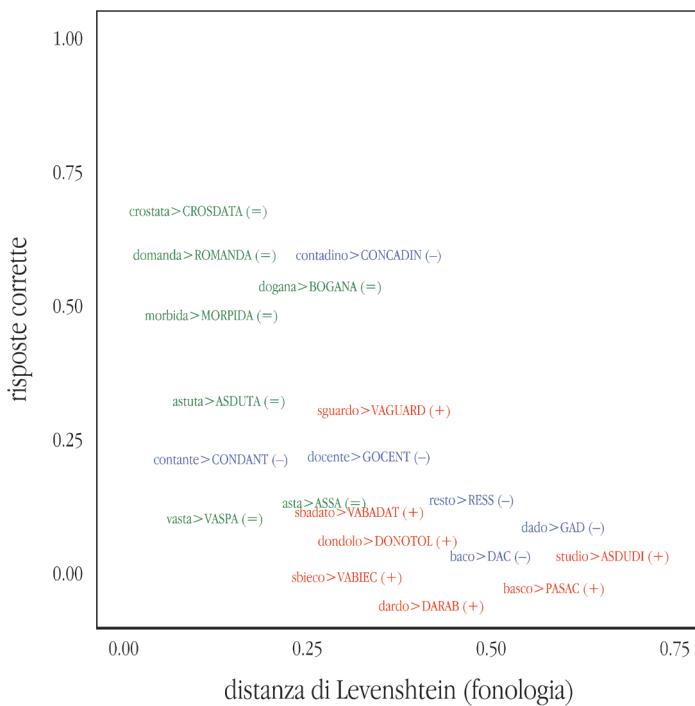


Come si argomenterà nel §3.2, nella condizione “FRASE” non è possibile stabilire con certezza in quale misura il riconoscimento della pseudo-parola catalana sia da attribuire alla distanza dall’originale italiano o piuttosto alla comprensibilità della frase in cui è inserita. L’analisi quantitativa che segue si fonderà quindi sui soli risultati ottenuti per le parole presentate in isolamento.

La Figura 3 illustra la relazione tra numero di sillabe della pseudo-parola rispetto all’originale italiano, la sillaba contenente la modifica e la distanza interlinguistica.

<sup>5</sup> In questo grafico e nei seguenti, i singoli punti rappresentano i punteggi medi di ciascun partecipante per gli elementi bersaglio descritti dalla combinazione di variabili di volta in volta specificata.

Figura 3 - Relazione tra riconoscibilità e distanza di Levenshtein (fonologia), condizione "PAROLA"<sup>6</sup>



Particolarmente cruciale sembrerebbe il numero di sillabe della pseudo-parola rispetto al modello italiano. Ai punteggi di intellegibilità più elevati corrispondono infatti gli elementi bersaglio composti del medesimo numero di sillabe del modello (gruppo “0”); seguono quelli del gruppo “-1”, composti di una sillaba in meno per effetto del troncamento della vocale finale; infine, con punteggi quasi invariabilmente prossimi a zero, si trovano le pseudo-parole del gruppo “+1”, nelle quali al troncamento della vocale finale si aggiunge l'inserimento di una vocale epentetica, e perciò di una nuova sillaba.

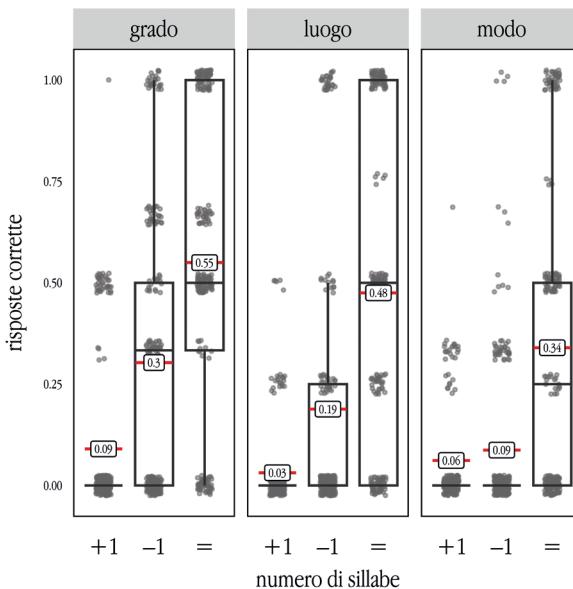
Meno chiaro appare il ruolo della sillaba interessata dalla manipolazione sperimentale.

<sup>6</sup> I colori corrispondono al numero di sillabe della pseudo-parola rispetto alla controparte italiana (rosso ~ +1, blu ~ -1, verde ~ =), indicato anche tra parentesi. Per ciascuna parola italiana è fornito un esempio di pseudo-parola.

Quanto alla distanza interlinguistica, è abbastanza evidente la relazione di proporzionalità inversa rispetto all'intellegibilità. Si può identificare una soglia (25% circa), al di sopra della quale il riconoscimento delle pseudo-parole diviene estremamente problematico. Anche al di sotto di tale soglia, tuttavia, è frequente il caso per cui a identici valori di questa misura continua corrispondano punteggi radicalmente diversi.

Passando ai parametri articolatori manipolati, la Figura 4 illustra i punteggi medi in funzione di tale variabile e del numero di sillabe della pseudo-parola rispetto all'originale italiano.

Figura 4 - *Punteggi medi per parametro articolatorio modificato e numero di sillabe, condizione "PAROLA"*



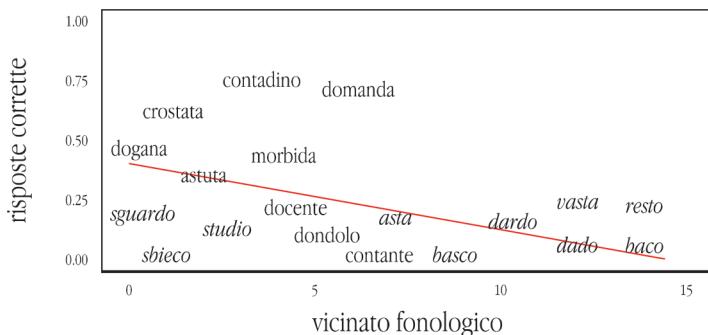
Proprio il numero di sillabe si conferma come la variabile più determinante: le pseudo-parole contenenti una vocale aggiuntiva (gruppo "+1") sono infatti correttamente riconosciute con frequenza al più occasionale; anche quelle prive della vocale finale (gruppo "-1") presentano un tasso di riconoscimento tendenzialmente basso e comun-

que inferiore a quello riscontrabile per le pseudo-parole caratterizzate da identica struttura sillabica rispetto all’italiano (gruppo “=”).

Per quanto riguarda invece i tre parametri articolatori considerati, i punteggi più elevati sono associati a manipolazioni del grado di articolazione, seguiti a breve distanza da modifiche al luogo di articolazione. La manipolazione del modo di articolazione, infine, produce punteggi decisamente più bassi.

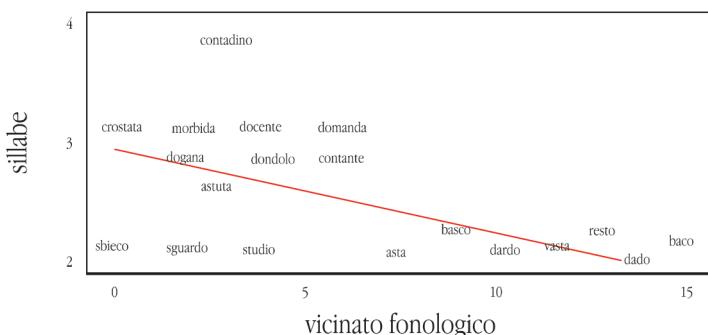
Si rilevano infine due significative correlazioni. La Figura 5 mostra come all’aumentare del numero di vicini fonologici diminuisca il tasso medio di risposte corrette.

Figura 5 - *Relazione tra intelligenza e ampiezza del vicinato fonologico*



La Figura 6 mostra invece che il numero di vicini fonologici è inversamente proporzionale alla lunghezza della parola, qui operazionalizzata in termini di numero di sillabe.

Figura 6 - *Relazione tra numero di sillabe e ampiezza del vicinato fonologico*



### 2.3.1 *Analisi statistica*

L’analisi statistica si basa su un modello lineare generalizzato misto (GLMM) postulante una distribuzione binomiale<sup>7</sup>. Il modello include le seguenti variabili:

- interazione tra la modalità di presentazione degli stimoli (parole isolate o frasi), il numero di sillabe della pseudo-parola rispetto al modello italiano (identico, -1, +1), la sillaba contenente la consonante manipolata (1, 2) e il parametro articolatorio modificato (luogo, modo, grado), es. “PAROLA\_meno1\_1\_modo”;
- interazione tra modalità di presentazione e distanza di Levenshtein normalizzata, es. “dist\_parola”;
- interazione tra modalità di presentazione e logaritmo naturale della frequenza totale della forma di parola italiana, es. “fqTotL\_parola”;
- interazione tra modalità di presentazione e numero di vicini fonologici, es. “Phon\_N\_parola”;
- interazione tra modalità di presentazione e logaritmo naturale della frequenza totale dei vicini fonologici, es. “Phon\_N\_MFreq\_frase”;
- prevedibilità degli elementi bersaglio alla luce della frase che li contiene (solo per la condizione “FRASE”): “pred”.

La ragione per la diffusa presenza delle interazioni risiede nel fatto che l’eventuale effetto di ciascuna delle variabili considerate potrebbe differire a seconda del valore assunto dalle altre.

A questa struttura di effetti fissi si aggiungono gli effetti casuali sotto forma di intercette per ciascun partecipante ed elemento bersaglio.

La Tabella 4 (Appendice) presenta i coefficienti stimati dal modello. In primo luogo, si nota come gran parte dei coefficienti relativi alla condizione “FRASE” abbia segno positivo, laddove la condizione “PAROLA” induce normalmente il segno negativo. Come si argomenterà diffusamente nella sezione 3.1, un’eventuale risposta corretta in presenza di un contesto può essere attribuita tanto al minore impatto delle manipolazioni considerate, quanto alla prevedibilità dell’elemento bersaglio nell’ambito della frase che lo contiene: in effetti, il parametro “pred” è solidamente positivo. Alla luce di queste considerazioni, nell’analisi che segue ci si concentrerà esclusivamente sulla condizione “PAROLA”.

---

<sup>7</sup> Il modello è stato impostato in *Stan* (Stan Development Team 2022) utilizzando *R* (R Core team 2021) e l’interfaccia *cmdstanr* (Gabry & Češnovar 2022)

Tra i coefficienti dal valore assoluto più elevato spicca in primo luogo la distanza interlinguistica complessiva (riga 1). Seguono quelli relativi alle pseudo-parole modificate rispetto alla struttura sillabica (4-6; 10; 13; 16): successivamente, tra le pseudo-parole dalla stessa struttura sillabica della controparte italiana, si trovano i coefficienti relativi alla manipolazione del modo (8-9), luogo (11-12) e grado (18) di articolazione. Meno chiaro appare il ruolo della sillaba contenente la consonante modificata. Infine, raggiungono la significatività statistica anche due variabili di controllo, cioè la frequenza della parola modello italiana e il numero di vicini fonologici ad essa riconducibili (19-20).

I contrasti post-hoc, infine, tralasciando quelli che oppongono le due condizioni di presentazione degli stimoli (17; 20-26; 28; 33-34; 36), rivelano valori significativi rispetto sia al parametro articolatorio considerato (1; 3-5; 8; 11-16), sia alla sillaba contenente la manipolazione (2; 7), sia alla struttura sillabica della pseudo-parola rispetto alla controparte italiana (6; 9-10).

### *3. Discussione*

I risultati dell'esperimento si possono sintetizzare come segue. In primo luogo, come era facile prevedere, il riconoscimento degli elementi bersaglio è decisamente meno efficace quando questi sono presentati come parole isolate che non nell'ambito di una frase di contesto, per quanto non necessariamente del tutto comprensibile. Si è inoltre evidenziato un ruolo importante sia per alcune delle misure di controllo (frequenza totale, ampiezza del vicinato fonologico), sia per una delle variabili di interesse dello studio, cioè la distanza di Levenshtein, intrinsecamente continua, la quale anzi fra tutte risulta quella di maggior impatto. Al tempo stesso si registra un effetto statisticamente significativo per alcune combinazioni delle variabili discrete considerate (struttura sillabica, parametro articolatorio, sillaba contenente la consonante manipolata), le quali come si è visto interagiscono tra loro in maniera estremamente complessa. Determinare il peso specifico di ciascuna nell'ambito dell'interazione postulata dal modello statistico richiederebbe un'analisi approfondita che non è possibile proporre in questa sede. Ai fini della presente discussione basterà la constatazione che l'interazione appare ben giustificata, in quanto l'impatto di una variabile sul risultato dipende in maniera significativa dal valore assunto dalle altre.

### 3.1 *Discreto e continuo*

Come riassunto graficamente dalla Figura 3, esiste una relazione complessa tra variabili discrete e continue, che rende sfuggente la definizione di un modello esplicativo, prima ancora che statistico. Da un lato, la lunghezza di una parola (qui misurata in termini di numero di sillabe) è fortemente correlata con l'ampiezza del vicinato fonologico, il quale nel contesto del presente studio si sovrappone (pur non esaurendolo) all'insieme delle possibili controparti italiane di un elemento bersaglio. A sua volta, questo valore risulta inversamente proporzionale al tasso di riconoscimento delle pseudo-parole, in quanto un numero elevato di vicini fonologici, specie se frequenti, è sinonimo di un elevato numero di potenziali concorrenti, alcuni dei quali egualmente legittimi in termini di distanza interlinguistica.

Per quanto riguarda la distanza interlinguistica stessa, tutte le manipolazioni sperimentali delle consonanti sono state concepite in modo tale da modificare in maniera uniforme questo parametro, con un impatto pari a 1 in termini assoluti: in tutti i casi, infatti, si tratta della sostituzione di una singola consonante con un'altra. Le inserzioni o elisioni utilizzate per manipolare la struttura sillabica provocano però evidentemente un ulteriore incremento. Si crea così una forte correlazione tra una variabile discreta e una continua, ulteriormente complicata dal fatto che la distanza interlinguistica *normalizzata* è sensibile alla lunghezza dell'allineamento, il quale rappresenta il denominatore del rapporto usato per ricondurre la distanza assoluta a un valore compreso tra 0 e 1. Tutto ciò rende difficile attribuire a una specifica variabile le tendenze osservabili nei dati: i punteggi più bassi riscontrati per le pseudo-parole modificate nella struttura sillabica sono indotti dal diverso numero di sillabe rispetto alla controparte italiana, o piuttosto dalla maggiore distanza interlinguistica? Sulla base dei dati a disposizione non è possibile rispondere.

Si evidenzia così la natura asimmetrica del rapporto tra variabili discrete e continue nell'ambito del problema qui considerato: a identiche variazioni della distanza interlinguistica può corrispondere un gran numero di combinazioni di modifiche discrete (es. inserzione di una vocale e modifica del luogo di articolazione di una consonante; elisione e modifica del modo di articolazione, ecc.). Nel caso della distanza normalizzata è vero anche il contrario, giacché identiche manipolazioni di variabili discrete possono portare a valori di distanza interlinguistica diversi a seconda della lunghezza della coppia di parole considerate. Tale conclusione in verità era già intuibile dall'analisi della Figura 1: Berthele (2011) procede infatti

a un'analisi esplorativa delle diverse combinazioni di corrispondenze fonologiche o ortografiche capaci di illuminare la varianza lasciata inspiegata dalla sola distanza interlinguistica. Il presente studio si proponeva di affrontare il medesimo problema controllando sperimentalmente la natura e la combinazione delle variabili in esame, laddove invece Berthele (2011) utilizzava parole realmente esistenti, e quindi più difficilmente controllabili. I risultati mostrano da un lato che tale approccio porta a risultati apprezzabili, dall'altro che la complessità delle relazioni tra le variabili coinvolte nel processo in esame è comunque tale che un'analisi di tutte le loro possibili combinazioni richiederebbe una quantità irrealistica di dati.

È bene ricordare, infine, che i parametri articolatori delle consonanti, qui considerati valori discreti, corrispondono in realtà alle varie configurazioni che può assumere il cavo orale, il quale è evidentemente uno spazio continuo: un problema che riporta alla definizione stessa di fonema<sup>8</sup>. A rigore, inoltre, anche quelle variabili che indicano una quantità (es. numero di sillabe, o di coppie minime) sarebbero discrete, in quanto composte evidentemente di numeri interi: la categorizzazione che si fa qui delle variabili in continue e discrete si confonde quindi in parte con quella che in statistica oppone variabili *categoriche*, da un lato, e *quantitative discrete*, dall'altro.

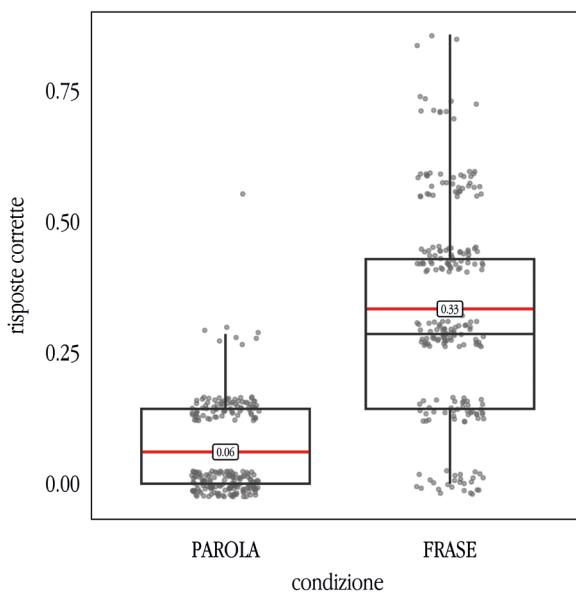
### 3.2 *Il ruolo del contesto*

Meritevole di una certa attenzione è anche il ruolo del contesto, intuitivamente cruciale, ma estremamente difficile da modellizzare, tanto che la maggior parte degli studi dedicati al problema qui in esame scelgono di ricorrere a stimoli presentati in isolamento proprio per sopperire a questa difficoltà. Bisogna però considerare che quest'ultimo approccio non è forse il più consono alla simulazione di un reale contesto di intercomprendere, in cui raramente l'oggetto dell'inferenza è una parola isolata: ben più comune è il caso in cui sia richiesta l'interpretazione di un testo composto di più parole variamente trasparenti. Per questo motivo nel presente studio gli stimoli sono stati anche presentati nel contesto di una frase catalana, una modalità che si è dimostrata di grande aiuto per l'intellegibilità degli elementi bersaglio (Figura 2). Al tempo stesso, sulla base dei dati a disposizione non è possibile stabilire che cosa precisamente abbia indotto tale miglioramento. L'esercizio chiedeva in ogni caso la

<sup>8</sup> Si ringrazia Mario Vayra per il fruttuoso scambio che ha portato a questa riflessione.

traduzione del solo elemento bersaglio, mentre non ci sono informazioni su quanto la frase di contesto sia stata effettivamente compresa. Ci si dovrà quindi limitare a riconoscere che la presenza di un contesto aumenta notevolmente la riconoscibilità della parola, mentre resteranno oscure le cause: in particolare, non è possibile escludere che la parola non sia stata effettivamente riconosciuta, ma piuttosto *intuita* sulla base del senso generale ricavabile dalla frase (una probabilità parzialmente quantificata dalla variabile “pred” nella Tabella 6). Tale scenario appare particolarmente verosimile nel caso delle pseudo-parole del gruppo “+1”, la cui riconoscibilità tende a zero nella condizione “PAROLA”, ma raggiunge valori ben più elevati in presenza di un contesto (Figura 7).

Figura 7 - Punteggi medi per modalità di presentazione dello stimolo,  
struttura sillabica “+1”



Non è però chiaro in quale misura tale ipotesi si possa generalizzare ad altri elementi bersaglio intuitivamente più trasparenti. Anche nel caso dello stesso gruppo “+1”, inoltre, il cospicuo incremento dei punteggi in presenza di un contesto non implica necessariamente che la soluzione sia stata suggerita dal contesto in maniera indipendente dall’elemento

bersaglio: al contrario, il contesto potrebbe aver ridotto l'incertezza del partecipante rispetto a un numero limitato di opzioni verosimili, indirizzandolo in maniera più netta verso la soluzione divenuta più plausibile alla luce della nuova cornice semantica e sintattica (Muikku-Werner 2014). Tutte queste considerazioni sono destinate a restare solo speculative sulla base dei dati a disposizione, ma potranno essere utilmente riprese dalla ricerca futura.

### 3.3 *Criteri di valutazione*

Strettamente legata al ruolo del contesto è anche la definizione dei criteri di valutazione delle risposte. Alcune infatti risultano semanticamente corrette, ma etimologicamente irrelate alla parola bersaglio, es. *baco* > 'bruco', *dardo* > 'freccia', *studio* > 'ufficio'. Nel presente studio si è scelto di applicare un criterio restrittivo, per cui nella condizione "FRASE" solo le risposte etimologicamente prossime alla parola bersaglio sono state considerate corrette: per le altre, per quanto semanticamente accettabili, permane infatti il dubbio che siano state intuite piuttosto che riconosciute (nella condizione "PAROLA" naturalmente il problema non si pone). Ci sono però alcuni controsensi, qui esemplificati con l'elemento bersaglio ROCENT, da *docente*, i quali sembrerebbero suggerire che la scelta di un sinonimo anziché della parola più prossima possa avvenire anche in presenza di un'inferenza corretta (3).

- (3) a. *maestro*.  
b. *insegnante (docente)*.

In entrambi i casi l'elemento bersaglio era presentato nella condizione "PAROLA", cioè *in assenza* di contesto. Le due risposte sembrano suggerire che la parola italiana più prossima all'elemento bersaglio, i.e. *docente*, per qualche motivo non fosse sentita come adeguata, e che il rispondente abbia perciò ritenuto opportuno sostituirla con un sinonimo, specificando in alcuni casi l'elemento lessicale che ha effettivamente guidato l'inferenza.

Si potrebbe poi obiettare che in un contesto comunicativo reale, difficilmente l'efficacia della comunicazione è affidata alle sole coppie minime, come invece in parte si simula qui (le pseudo-parole con struttura sillabica "—" differiscono dalla controparte italiana per un solo segmento). Albano Leoni (2009) sostiene che talvolta un'opposizione tra fonemi può effettivamente veicolare una differenza di significato altrimenti irre-

cuperabile, come nei sintagmi francesi *une glace à la menthe* [yn glas a la māt] ‘un gelato alla menta’ e *une glace à l'amande* [yn glas a la mād] ‘un gelato alla mandorla’. Si tratta però di casi del tutto marginali: l'autore argomenta infatti convincentemente che la trasmissione dell'informazione si basa normalmente su un'immagine acustica olistica.

#### *4. Conclusioni: discreto e continuo nel contesto dell'intercomprensione*

La lezione che è possibile trarre dallo studio è che per affrontare il problema del riconoscimento di parole nel contesto dell'intercomprensione è indispensabile adottare un approccio sperimentale al tempo stesso rigoroso e quanto più semplice possibile: in assenza di grandi coorti di informanti, infatti, non è auspicabile prevedere numerose combinazioni di variabili, come invece è almeno in parte il caso del presente studio. Tale conclusione è resa evidente da un lato dalla complessa interazione tra variabili discrete (struttura sillabica, sillaba contenente la consonante manipolata, parametro articolatorio manipolato); dall'altro, dal fatto che una variabile continua come la distanza interlinguistica, universalmente indicata come la più determinante per spiegare il riconoscimento di parole etimologicamente imparentate, è in realtà la conseguenza di tali manipolazioni di variabili discrete. Trattandosi di valori normalizzati, inoltre, la relazione è resa più complessa dalla sua dipendenza dalla lunghezza delle parole considerate, la quale a sua volta è fortemente correlata all'ampiezza del vicinato fonologico, cioè all'insieme delle parole in assoluto più simili a quella in esame. In conclusione, dunque, si suggerisce qui che la ricerca dovrà primariamente occuparsi di sciogliere sperimentalmente tali complesse relazioni, concentrandosi sul ruolo di una specifica variabile solo una volta che tutte le altre siano state controllate o neutralizzate. Tale approccio permetterà forse di descrivere con maggiore precisione il sostrato cognitivo comune, sul quale in ogni caso insiste evidentemente una notevole variabilità individuale (Vanhouve & Berthele 2017).

#### *Riferimenti bibliografici*

- Albano Leoni, Federico. 2009. *Dei suoni e dei sensi: il volto fonico delle parole*. Bologna: Il Mulino.
- Berthele, Raphael. 2011. On abduction in receptive multilingualism. Evidence from cognate guessing tasks. *Applied Linguistics Review* 2. 191–220.

- Comesaña, Montserrat & Soares, Ana P. & Marçet, Ana & Perea, Manuel. 2016. On the nature of consonant/vowel differences in letter position coding: evidence from developing and adult readers. *British Journal of Psychology* 107(4). 651–674.
- Gabry, Jonah & Češnovar, Rok. 2022. cmdstanr: R Interface to “CmdStan.” (<https://mc-stan.org/cmdstanr/>) (Consultato il 29.11.2024.)
- Gooskens, Charlotte. 2007. The contribution of linguistic factors to the intelligibility of closely related languages. *Journal of Multilingual and Multicultural Development* 28(6). 445–467.
- Gooskens, Charlotte & Heeringa, Wilbert & Beijering, Karin. 2008. Phonetic and lexical predictors of intelligibility. *International Journal of Humanities and Arts Computing* 2(1–2). 63–81.
- Gooskens, Charlotte & van Heuven, Vincent. 2019. How well can intelligibility of closely related languages in Europe be predicted by linguistic and non-linguistic variables? *Linguistic Approaches to Bilingualism* 10(3). 351–379.
- Goslin, Jeremy & Galluzzi, Claudia & Romani, Cristina. 2014. PhonItalia: a phonological lexicon for Italian. *Behavior Research Methods* 46(3). 872–886.
- Heeringa, Wilbert. 2004. *Measuring dialect pronunciation differences using Levenshtein distance*. Rijksuniversiteit Groningen: Groningen (Tesi di dottorato.)
- Heinz, Christof. 2009. Semantische Disambiguierung von false friends in slavischen L3: die Rolle des Kontexts. *Zeitschrift für Slawistik* 54(2). 145–166.
- Jágrová, Klára & Avgustinova, Tania & Stenger, Irina & Fischer, Andrea. 2019. Language models, surprisal and fantasy in Slavic intercomprehension. *Computer Speech & Language* 53. 242–275.
- Levenshtein, Vladimir. 1966. Binary codes capable of correcting deletions, insertions and reversals. *Soviet Physics Doklady* 10(8). 707–710.
- Marx, Nicole & Mehlhorn, Grit. 2010. Pushing the positive: encouraging phonological transfer from L2 to L3. *International Journal of Multilingualism* 7(1). 4–18.
- Möller, Robert & Zeevaert, Ludger. 2015. Investigating word recognition in intercomprehension: methods and findings. *Linguistics* 53(2). 313–352.
- Muikku-Werner, Pirkko. 2014. Co-text and receptive multilingualism – Finnish students comprehending Estonian. *Eesti ja soome-ugri keeleteaduse ajakiri. Journal of Estonian and Finno-Ugric Linguistics* 5(3). 99–113.

- Nerbonne, John & Heeringa, Wilbert. 2009. Measuring dialect differences. In Peter Auer & Schmidt, Jürgen Erich (a cura di), *Language and Space*, 550–567. Berlin: Mouton de Gruyter.
- PhonItalia = <http://> (consultato il ....).
- R Core team. 2021. R: A language and environment for statistical computing. Vienna: R Foundation for Statistical Computing. (<http://www.R-project.org/>) (Consultato il 29.11.2024.)
- Rastle, Kathleen & Brysbaert, Marc. 2006. Masked phonological priming effects in English: are they real? Do they matter? *Cognitive Psychology* 53(2). 97–145.
- Schmandt, Silvana & Nazzi, Thierry & New, Boris. 2022. Consonant, vowel and lexical neighbourhood processing during word recognition: new evidence using the sandwich priming technique. *Language, Cognition and Neuroscience* 37(9). 1115–1130.
- Stan Development Team. 2022. Stan modeling language users guide and reference manual. (<https://mc-stan.org>) (Consultato il 29.11.2024.)
- Stenger, Irina & Jagrova, Klara & Avgustinova, Tania. 2020. The INCOMSLAV platform: experimental website with integrated methods for measuring linguistic distances and asymmetries in receptive multilingualism. In *Proceedings of the LREC 2020 workshop “Citizen linguistics in language resource development”*, 40–48. Marseille: European Language Resources Association.
- Vanhouve, Jan & Berthele, Raphael. 2017. Interactions between formal distance and participant-related variables in receptive multilingualism. *International Review of Applied Linguistics in Language Teaching* 55(1). 23–40.

## Appendice

La Tabella 4 presenta in ordine decrescente di valore assoluto della media i coefficienti stimati dal modello statistico. Le interazioni sono sintetizzate secondo il seguente schema: modalità di presentazione, numero di sillabe della pseudo-parola rispetto alla controparte italiana, sillaba contenente la consonante manipolata, parametro articolatorio manipolato.

Tabella 4 - *Coefficienti del modello statistico*

variable	mean	sd	q2.5	q97.5
dist_parola	-2.71	0.72	-4.09	-1.26

variable	mean	sd	q2.5	q97.5
<i>pred</i>	2.5	0.57	1.39	3.64
<i>FRASE_uguale_2_grado</i>	1.97	0.45	1.09	2.85
<i>PAROLA_meno1_1_modo</i>	-1.89	0.53	-2.99	-0.89
<i>PAROLA_piul1_1_luogo</i>	-1.85	0.48	-2.82	-0.93
<i>PAROLA_meno1_2_modo</i>	-1.65	0.5	-2.62	-0.67
<i>FRASE_uguale_1_grado</i>	1.63	0.51	0.67	2.65
<i>PAROLA_uguale_1_modo</i>	-1.57	0.44	-2.42	-0.74
<i>PAROLA_uguale_2_modo</i>	-1.55	0.41	-2.36	-0.75
<i>PAROLA_piul1_1_modo</i>	-1.42	0.48	-2.37	-0.47
<i>PAROLA_uguale_2_luogo</i>	-1.39	0.4	-2.2	-0.62
<i>FRASE_uguale_1_luogo</i>	1.25	0.48	0.31	2.2
<i>FRASE_meno1_2_grado</i>	1.22	0.5	0.22	2.19
<i>FRASE_piul1_1_grado</i>	1.1	0.46	0.19	2.02
<i>FRASE_meno1_1_grado</i>	1.05	0.43	0.18	1.88
<i>PAROLA_piul1_1_grado</i>	-0.99	0.47	-1.89	-0.04
<i>FRASE_uguale_2_luogo</i>	0.98	0.4	0.19	1.75
<i>PAROLA_uguale_2_grado</i>	0.87	0.41	0.02	1.66
<i>fqTotL_parola</i>	0.59	0.23	0.12	1.03
<i>Phon_N_parola</i>	-0.35	0.09	-0.54	-0.17
<i>PAROLA_uguale_1_grado</i>	-0.85	0.44	-1.71	0
<i>FRASE_meno1_2_luogo</i>	0.76	0.47	-0.18	1.66
<i>FRASE_uguale_1_modo</i>	0.74	0.46	-0.18	1.65
<i>FRASE_uguale_2_modo</i>	0.67	0.4	-0.13	1.44
<i>FRASE_piul1_2_modo</i>	-0.58	0.49	-1.5	0.36
<i>FRASE_meno1_2_modo</i>	0.56	0.47	-0.34	1.48
<i>PAROLA_piul1_2_modo</i>	-0.55	0.69	-1.95	0.74
<i>PAROLA_piul1_2_luogo</i>	0.46	0.62	-0.77	1.68
<i>FRASE_piul1_2_grado</i>	-0.46	0.49	-1.44	0.5

variable	mean	sd	q2.5	q97.5
<i>Phon_N_MFreq_parola</i>	-0.42	0.38	-1.17	0.33
<i>PAROLA_piul_2_grado</i>	0.41	0.6	-0.76	1.56
<i>PAROLA_meno1_2_grado</i>	0.4	0.46	-0.48	1.3
<i>dist_frase</i>	0.38	0.75	-1.08	1.84
<i>FRASE_piul_1_modo</i>	0.37	0.46	-0.52	1.27
<i>FRASE_meno1_1_luogo</i>	0.35	0.43	-0.49	1.19
<i>FRASE_piul_2_luogo</i>	-0.32	0.49	-1.29	0.65
<i>PAROLA_uguale_1_luogo</i>	-0.31	0.44	-1.2	0.55
<i>PAROLA_meno1_2_luogo</i>	-0.28	0.46	-1.18	0.61
<i>PAROLA_meno1_1_luogo</i>	-0.23	0.46	-1.15	0.69
<i>FRASE_meno1_1_modo</i>	-0.17	0.44	-1.05	0.68
<i>FRASE_piul_1_luogo</i>	0.15	0.45	-0.74	1.05
<i>Phon_N_frase</i>	-0.15	0.09	-0.33	0.03
<i>PAROLA_meno1_1_grado</i>	0.08	0.45	-0.79	0.96
<i>Phon_N_MFreq_frase</i>	0.02	0.38	-0.72	0.77
<i>fqTotL_frase</i>	0.01	0.23	-0.46	0.45

La Tabella 5 presenta in ordine decrescente di valore assoluto della media i contrasti post-hoc tra i coefficienti relativi all’interazione compresa nel modello statistico (solo contrasti che con il 95% di probabilità non contengono il valore 0). Le interazioni sono sintetizzate secondo il seguente schema: modalità di presentazione, numero di sillabe della pseudo-parola rispetto alla controparte italiana, sillaba contenente la consonante manipolata, parametro articolatorio manipolato. Le righe colorate in grigio corrispondono ai contrasti riguardanti la sola condizione “PAROLA”.

Tabella 5 - *Contrasti post-hoc (solo significativi al 95%)*

rango	par1	par2	par1 - par2
1	<i>PAROLA_uguale_2_modo</i>	<i>PAROLA_uguale_2_grado</i>	-2.42 (-2.89, -1.96)

range	par1	par2	par1 - par2
2	PAROLA_piu1_2_luogo	PAROLA_piu1_1_luogo	2.31 (0.83, 3.76)
3	PAROLA_uguale_2_grado	PAROLA_uguale_2_luogo	2.26 (1.83, 2.68)
4	PAROLA_meno1_2_grado	PAROLA_meno1_2_modo	2.05 (1.36, 2.79)
5	PAROLA_meno1_1_grado	PAROLA_meno1_1_modo	1.97 (1.09, 2.96)
6	PAROLA_piu1_2_luogo	PAROLA_uguale_2_luogo	1.85 (0.39, 3.24)
7	PAROLA_uguale_2_grado	PAROLA_uguale_1_grado	1.72 (0.57, 2.85)
8	PAROLA_meno1_1_luogo	PAROLA_meno1_1_modo	1.66 (0.74, 2.65)
9	PAROLA_meno1_1_luogo	PAROLA_piu1_1_luogo	1.62 (0.38, 2.82)
10	PAROLA_uguale_1_luogo	PAROLA_piu1_1_luogo	1.53 (0.31, 2.8)
11	PAROLA_meno1_2_modo	PAROLA_meno1_2_luogo	-1.37 (-2.11, -0.65)
12	PAROLA_uguale_1_luogo	PAROLA_uguale_1_modo	1.26 (0.77, 1.76)
13	PAROLA_piu1_1_grado	PAROLA_piu1_1_luogo	0.86 (0.22, 1.49)
14	PAROLA_uguale_1_modo	PAROLA_uguale_1_grado	-0.73 (-1.2, -0.25)
15	PAROLA_meno1_2_grado	PAROLA_meno1_2_luogo	0.68 (0.1, 1.29)
16	PAROLA_uguale_1_luogo	PAROLA_uguale_1_grado	0.53 (0.04, 1.05)
17	PAROLA_uguale_1_grado	FRASE_uguale_1_grado	-2.47 (-3.3, -1.7)
18	FRASE_uguale_2_grado	FRASE_piu1_2_grado	2.44 (1.13, 3.75)
19	PAROLA_uguale_2_luogo	FRASE_uguale_2_luogo	-2.37 (-2.86, -1.89)
20	PAROLA_uguale_1_modo	FRASE_uguale_1_modo	-2.31 (-2.97, -1.67)
21	PAROLA_uguale_2_modo	FRASE_uguale_2_modo	-2.22 (-2.73, -1.72)
22	PAROLA_meno1_2_modo	FRASE_meno1_2_modo	-2.21 (-3.13, -1.3)
23	PAROLA_piu1_1_grado	FRASE_piu1_1_grado	-2.09 (-3.04, -1.16)
24	PAROLA_piu1_1_luogo	FRASE_piu1_1_luogo	-1.99 (-2.93, -1.06)
25	PAROLA_piu1_1_modo	FRASE_piu1_1_modo	-1.79 (-2.7, -0.87)
26	PAROLA_meno1_1_modo	FRASE_meno1_1_modo	-1.72 (-2.73, -0.78)
27	FRASE_meno1_2_grado	FRASE_piu1_2_grado	1.68 (0.33, 3.03)
28	PAROLA_uguale_1_luogo	FRASE_uguale_1_luogo	-1.57 (-2.33, -0.84)
29	FRASE_piu1_1_grado	FRASE_piu1_2_grado	1.56 (0.3, 2.87)

<i>range</i>	<i>par1</i>	<i>par2</i>	<i>par1 - par2</i>
30	<i>FRASE_uguale_2_modo</i>	<i>FRASE_uguale_2_grado</i>	-1.31 (-1.88, -0.74)
31	<i>FRASE_più1_2_luogo</i>	<i>FRASE_uguale_2_luogo</i>	-1.31 (-2.54, -0.03)
32	<i>FRASE_meno1_1_grado</i>	<i>FRASE_meno1_1_modo</i>	1.22 (0.72, 1.73)
33	<i>PAROLA_uguale_2_grado</i>	<i>FRASE_uguale_2_grado</i>	-1.1 (-1.77, -0.45)
34	<i>PAROLA_meno1_2_luogo</i>	<i>FRASE_meno1_2_luogo</i>	-1.05 (-1.88, -0.25)
35	<i>FRASE_uguale_2_grado</i>	<i>FRASE_uguale_2_luogo</i>	0.99 (0.42, 1.58)
36	<i>PAROLA_meno1_1_grado</i>	<i>FRASE_meno1_1_grado</i>	-0.97 (-1.7, -0.25)
37	<i>FRASE_più1_1_grado</i>	<i>FRASE_più1_1_luogo</i>	0.95 (0.55, 1.36)
38	<i>FRASE_uguale_1_modo</i>	<i>FRASE_uguale_1_grado</i>	-0.89 (-1.75, -0.07)
39	<i>FRASE_più1_1_grado</i>	<i>FRASE_più1_1_modo</i>	0.73 (0.31, 1.16)
40	<i>FRASE_meno1_1_luogo</i>	<i>FRASE_meno1_1_grado</i>	-0.7 (-1.19, -0.2)
41	<i>FRASE_meno1_1_luogo</i>	<i>FRASE_meno1_1_modo</i>	0.53 (0.02, 1.06)

La Tabella 6 presenta gli elementi bersaglio dello studio. Le variabili *fqtotl*, *phon\_n* e *phon\_n\_mfreq* (si veda la sezione 2.1) si riferiscono alla forma di parola italiana assunta come modello e hanno perciò identico valore per tutte le pseudo-parole da essa derivate. Le variabili *tipi* e *pred* indicano rispettivamente la dispersione in tipi delle risposte e l'accuratezza media.

Tabella 6 - *Elementi bersaglio.*

<i>italiano</i>	<i>pseudo-parola</i>	<i>para- metro</i>	<i>sillaba modi- ficata</i>	<i>numero sillabe</i>	<i>distan- za</i>	<i>fqtotl</i>	<i>phon _n</i>	<i>phon _n</i>	<i>tipi</i>	<i>pred</i>
<i>asta</i>	<i>ASDA</i>	<i>grado</i>	2	=	0,25	4,34	8	2,93	0.18	0.27
<i>asta</i>	<i>ASPA</i>	<i>luogo</i>	2	=	0,25	4,34	8	2,93	0.18	0.27
<i>asta</i>	<i>ASSA</i>	<i>modo</i>	2	=	0,25	4,34	8	2,93	0.18	0.27
<i>astuta</i>	<i>ASDUTA</i>	<i>grado</i>	2	=	0,17	1,79	3	0,90	0.09	0.00
<i>astuta</i>	<i>ASCUTA</i>	<i>luogo</i>	2	=	0,17	1,79	3	0,90	0.09	0.00
<i>astuta</i>	<i>ASSUTA</i>	<i>modo</i>	2	=	0,17	1,79	3	0,90	0.09	0.00
<i>baco</i>	<i>PAC</i>	<i>grado</i>	1	-1	0,50	1,61	13	1,51	0.64	0.73

italiano	pseudo-parola	para- metro	sillaba modi- ficata	numero sillabe	distan- za	fqtotl _n	phon _n	phon mfreq	tipi	pred
baco	DAC	luogo	1	-1	0,50	1,61	13	1,51	0,64	0,73
baco	MAC	modo	1	-1	0,50	1,61	13	1,51	0,64	0,73
basco	PASAC	grado	1	1	0,60	0,69	11	2,51	0,64	0,00
basco	DASAC	luogo	1	1	0,60	0,69	11	2,51	0,64	0,00
basco	MASAC	modo	1	1	0,60	0,69	11	2,51	0,64	0,00
contadino	CONDADIN	grado	2	-1	0,22	3,64	3	3,04	0,00	0,00
contadino	CONCADIN	luogo	2	-1	0,22	3,64	3	3,04	0,00	0,00
contadino	CONSADIN	modo	2	-1	0,22	3,64	3	3,04	0,00	0,00
contante	CONDANT	grado	2	-1	0,25	1,79	5	3,29	0,55	0,55
contante	CONCANT	luogo	2	-1	0,25	1,79	5	3,29	0,55	0,55
contante	CONSANT	modo	2	-1	0,25	1,79	5	3,29	0,55	0,55
crostata	CROSDATA	grado	2	=	0,13	1,61	0	0,00	0,36	0,00
crostata	CROSPATA	luogo	2	=	0,13	1,61	0	0,00	0,36	0,00
crostata	CROSSATA	modo	2	=	0,13	1,61	0	0,00	0,36	0,00
dado	TAD	grado	1	-1	0,50	1,79	13	2,17	0,00	0,00
dado	GAD	luogo	1	-1	0,50	1,79	13	2,17	0,00	0,00
dado	NAD	modo	1	-1	0,50	1,79	13	2,17	0,00	0,00
dardo	DARAT	grado	2	1	0,40	0,00	11	2,29	0,09	0,00
dardo	DARAB	luogo	2	1	0,40	0,00	11	2,29	0,09	0,00
dardo	DARAR	modo	2	1	0,40	0,00	11	2,29	0,09	0,00
docente	TOCENT	grado	1	-1	0,29	4,16	4	2,95	0,09	0,00
docente	GOCENT	luogo	1	-1	0,29	4,16	4	2,95	0,09	0,00
docente	ROCENT	modo	1	-1	0,29	4,16	4	2,95	0,09	0,00
dogana	TOGANÀ	grado	1	=	0,17	1,10	1	0,69	0,00	0,00
dogana	BOGANÀ	luogo	1	=	0,17	1,10	1	0,69	0,00	0,00
dogana	ROGANÀ	modo	1	=	0,17	1,10	1	0,69	0,00	0,00
domanda	TOMANDA	grado	1	=	0,14	5,83	5	2,93	0,09	0,09

italiano	pseudo-parola	para- metro	sillaba modi- ficata	numero sillabe	distan- za	fqtotl _n	phon _n	phon mfreq	tipi	pred
<i>domanda</i>	<i>BOMANDA</i>	<i>luogo</i>	<i>1</i>	<i>=</i>	<i>0,14</i>	<i>5,83</i>	<i>5</i>	<i>2,93</i>	<i>0.09</i>	<i>0.09</i>
<i>domanda</i>	<i>ROMANDA</i>	<i>modo</i>	<i>1</i>	<i>=</i>	<i>0,14</i>	<i>5,83</i>	<i>5</i>	<i>2,93</i>	<i>0.09</i>	<i>0.09</i>
<i>dondolo</i>	<i>DONOTOL</i>	<i>grado</i>	<i>2</i>	<i>1</i>	<i>0,38</i>	<i>1,79</i>	<i>4</i>	<i>1,14</i>	<i>0.09</i>	<i>0.00</i>
<i>dondolo</i>	<i>DONOBOLE</i>	<i>luogo</i>	<i>2</i>	<i>1</i>	<i>0,38</i>	<i>1,79</i>	<i>4</i>	<i>1,14</i>	<i>0.09</i>	<i>0.00</i>
<i>dondolo</i>	<i>DONOROL</i>	<i>modo</i>	<i>2</i>	<i>1</i>	<i>0,38</i>	<i>1,79</i>	<i>4</i>	<i>1,14</i>	<i>0.09</i>	<i>0.00</i>
<i>morbida</i>	<i>MORPIDA</i>	<i>grado</i>	<i>2</i>	<i>=</i>	<i>0,14</i>	<i>3,64</i>	<i>3</i>	<i>3,06</i>	<i>0.36</i>	<i>0.18</i>
<i>morbida</i>	<i>MORDIDA</i>	<i>luogo</i>	<i>2</i>	<i>=</i>	<i>0,14</i>	<i>3,64</i>	<i>3</i>	<i>3,06</i>	<i>0.36</i>	<i>0.18</i>
<i>morbida</i>	<i>MORMIDA</i>	<i>modo</i>	<i>2</i>	<i>=</i>	<i>0,14</i>	<i>3,64</i>	<i>3</i>	<i>3,06</i>	<i>0.36</i>	<i>0.18</i>
<i>resto</i>	<i>RESD</i>	<i>grado</i>	<i>2</i>	<i>-1</i>	<i>0,40</i>	<i>6,44</i>	<i>13</i>	<i>2,90</i>	<i>0.00</i>	<i>0.09</i>
<i>resto</i>	<i>RESC</i>	<i>luogo</i>	<i>2</i>	<i>-1</i>	<i>0,40</i>	<i>6,44</i>	<i>13</i>	<i>2,90</i>	<i>0.00</i>	<i>0.09</i>
<i>resto</i>	<i>RESS</i>	<i>modo</i>	<i>2</i>	<i>-1</i>	<i>0,40</i>	<i>6,44</i>	<i>13</i>	<i>2,90</i>	<i>0.00</i>	<i>0.09</i>
<i>sbieco</i>	<i>SABIEC</i>	<i>grado</i>	<i>1</i>	<i>1</i>	<i>0,57</i>	<i>1,39</i>	<i>0</i>	<i>0,00</i>	<i>0.27</i>	<i>0.00</i>
<i>sbieco</i>	<i>VABIEC</i>	<i>luogo</i>	<i>1</i>	<i>1</i>	<i>0,71</i>	<i>1,39</i>	<i>0</i>	<i>0,00</i>	<i>0.27</i>	<i>0.00</i>
<i>sbieco</i>	<i>NABIEC</i>	<i>modo</i>	<i>1</i>	<i>1</i>	<i>0,71</i>	<i>1,39</i>	<i>0</i>	<i>0,00</i>	<i>0.27</i>	<i>0.00</i>
<i>sguardo</i>	<i>SAGUARD</i>	<i>grado</i>	<i>1</i>	<i>1</i>	<i>0,25</i>	<i>5,81</i>	<i>1</i>	<i>3,78</i>	<i>0.27</i>	<i>0.27</i>
<i>sguardo</i>	<i>VAGUARD</i>	<i>luogo</i>	<i>1</i>	<i>1</i>	<i>0,38</i>	<i>5,81</i>	<i>1</i>	<i>3,78</i>	<i>0.27</i>	<i>0.27</i>
<i>sguardo</i>	<i>NAGUARD</i>	<i>modo</i>	<i>1</i>	<i>1</i>	<i>0,38</i>	<i>5,81</i>	<i>1</i>	<i>3,78</i>	<i>0.27</i>	<i>0.27</i>
<i>studio</i>	<i>ASDUDI</i>	<i>grado</i>	<i>1</i>	<i>1</i>	<i>0,57</i>	<i>6,16</i>	<i>3</i>	<i>3,77</i>	<i>0.18</i>	<i>0.00</i>
<i>studio</i>	<i>ASCUDI</i>	<i>luogo</i>	<i>1</i>	<i>1</i>	<i>0,57</i>	<i>6,16</i>	<i>3</i>	<i>3,77</i>	<i>0.18</i>	<i>0.00</i>
<i>studio</i>	<i>ASSUDI</i>	<i>modo</i>	<i>1</i>	<i>1</i>	<i>0,57</i>	<i>6,16</i>	<i>3</i>	<i>3,77</i>	<i>0.18</i>	<i>0.00</i>
<i>vasta</i>	<i>VASDA</i>	<i>grado</i>	<i>2</i>	<i>=</i>	<i>0,20</i>	<i>4,04</i>	<i>12</i>	<i>3,36</i>	<i>0.50</i>	<i>0.33</i>
<i>vasta</i>	<i>VASPA</i>	<i>luogo</i>	<i>2</i>	<i>=</i>	<i>0,20</i>	<i>4,04</i>	<i>12</i>	<i>3,36</i>	<i>0.50</i>	<i>0.33</i>
<i>vasta</i>	<i>VASSA</i>	<i>modo</i>	<i>2</i>	<i>=</i>	<i>0,20</i>	<i>4,04</i>	<i>12</i>	<i>3,36</i>	<i>0.50</i>	<i>0.33</i>