

ACQUISIZIONE DI FENOMENI TEMPORALI E RITMICI DELL'ITALIANO

Analisi di produzioni di apprendenti anglofoni di italiano L2

Paolo MAIRANO, Marta MOIS, Valentina DE IACOVO, Antonio ROMANO¹

ABSTRACT • Acquisition of temporality and rhythm of the Italian language: Analysis of productions of English speaking learners of Italian as L2. Despite the conspicuous number of studies on the acquisition of L2 segmental patterns, research on L2 prosody is scarce and mostly deals with L2 English. In this paper we characterise the rhythmic patterns of L1 English learners of L2 Italian. We analysed read speech of 24 learners and compared them with data of L1 Italian and L1 English and compute several rhythm metrics. The results partially support the hypothesis that learners produce intermediate rhythmic patterns between the L1 and the target language. However, non-normalized rhythm metrics illustrate a different pattern, whereby L1 English learners of Italian seem to overshoot some rhythm characteristics. Such results are explained on the basis of the interference of speech rate and point to the need to use normalized metrics for studying rhythmic patterns of L2 speech. Finally, we did not find significant effects of variables of L2 acquisition (years of study, length of stay in Italy, competence in other Romance languages) on values of rhythm metrics.

KEYWORDS • L2 Italian; Rhythm Metrics; L2 Prosody; Pronunciation; Acquisition

1. Introduzione

Gli aspetti segmentali dell'acquisizione della fonetica godono di una letteratura cospicua e modelli espliciti (PAM di Best 1995, SLM di Flege 1995, teoria della marca di Eckman 1977). Invece, gli studi sull'acquisizione degli aspetti prosodici di una L2 sono molto più rari. Inoltre, la maggioranza della ricerca si è concentrata sull'acquisizione dell'inglese L2, mentre sono più rari gli studi che trattano altre L2. In questo articolo ci proponiamo di contribuire a colmare questa mancanza, caratterizzando gli aspetti prosodici del parlato di apprendenti anglofoni di italiano L2. In particolare, analizzeremo le caratteristiche ritmiche del parlato letto prodotto da 24 apprendenti, confrontandole con dati simili prodotti da inglesi nativi e italiani nativi.²

¹ L'articolo è il risultato di una collaborazione tra gli autori tanto sul piano sperimentale quanto su quello redazionale. Per scopi concorsuali precisiamo la seguente ripartizione del lavoro finale presentato. Sono da attribuirsi all'autore PM la conduzione dei test statistici e la redazione dei §§ 1.1, 2.3 e 3; all'autrice MM la raccolta dei dati di apprendenti inglesi, una prima annotazione dei materiali e la redazione del § 2.2; all'autrice VD la revisione delle annotazioni e la redazione del § 2.1; all'autore AR la definizione del protocollo di annotazione e la redazione dei §§ 1.2, 1.3 e 2.4.

² Sebbene sia un'assunzione di comodo quella di separare nettamente i due piani di strutturazione, ritmico e intonativo, del parlato, precisiamo che in quest'articolo il riferimento esclusivo alle metriche ritmiche e altre variabili dell'organizzazione temporale non prescinde da questo secondo livello organizzativo in

1.1. *Acquisizione di aspetti prosodici nella L2*

Come accennato sopra, la letteratura sull'acquisizione della fonetica e fonologia delle L2 si è soprattutto occupata di studiare aspetti segmentali. Più nello specifico, i vari modelli esistenti di acquisizione della fonologia L2 (per es. PAM e SLM) si sono occupati soprattutto di stabilire la difficoltà di produzione e percezione di determinati fonemi od opposizioni fonologiche da parte di parlanti non-nativi. Generalmente, questi studi tendono a convergere verso l'idea che i parlanti non-nativi riescano più facilmente a stabilire categorie fonologiche per suoni la cui realizzazione acustica è più lontana rispetto alle categorie esistenti nella loro L1; viceversa, categorie fonologiche realizzate nella L2 attraverso suoni vicini a categorie della L1 risentiranno maggiormente dell'influenza della L1. Mentre le ipotesi che scaturiscono da questi modelli sono state testate in numerosissimi studi (cf. Colantoni, Steele & Escudero 2015) per una panoramica, l'acquisizione delle caratteristiche prosodiche delle L2 non ha goduto della stessa fortuna: molto meno frequenti sono gli studi che se ne sono occupati, e al momento l'unico modello specifico che emetta predizioni sull'acquisizione della prosodia è quello recentissimo di Mennen (2015), che però si occupa esclusivamente di fenomeni intonativi.

Ultimamente però, l'acquisizione di fenomeni prosodici ha cominciato a destare l'interesse della comunità scientifica, come dimostrato dal sorgere di conferenze e giornate di studio dedicate interamente a questo tema, come *Methodological Perspectives on Second Language Prosody*, organizzato proprio in Italia (Busà & Stella, 2012) e il *Workshop on Second Language Prosody*, ormai alla 3ª edizione (a cura di Foltz, Cooper & Levendon nel 2016). Molti studi si sono concentrati sull'acquisizione dell'intonazione delle L2, studiando l'acquisizione di specifiche configurazioni intonative in produzione e/o percezione (cf. Santiago, Mairano & Delais-Roussarie, 2014), o mantenendosi su un piano globale e rivelando per esempio che l'escursione di f_0 tende a essere inferiore in parlanti di una L2 (Mennen, Schaeffler & Dicke, 2014). Anche le caratteristiche ritmiche e temporali non sono state trascurate, come dimostrato dagli studi di White & Mattys (2007), Tortel & Hirst (2010), De Meo (2012), Pettorino, De Meo & Vitale (2012), Vitale, Pellegrino, De Meo & Rasulo (2015), De Meo, Vitale & Pellegrino (2016), Tarasi & Romito (2016).

1.2. *Studi sul ritmo*

Come noto, e come dimostrato in diversi studi riassunti – tra gli altri – in [ANONIMI] (2011), un'etichetta di tipo “stress-timed” è tradizionalmente applicata a lingue la cui organizzazione temporale sembra dominata da *pattern* accentuali che sacrificano le sillabe non prominenti a favore di quelle prominenti, mentre un'etichetta di tipo “syllable-timed” è usata per quelle lingue la cui regolarità temporale sembra dominata da una certa rigidità sillabica che induce intervalli di prominenza di lunghezza variabile.

Questa distinta tipologia di lingue nelle quali il piede o la sillaba sembrano affermarsi maggiormente nell'organizzazione del parlato è generalmente basata su intuizioni e criteri di osservazione diversi proposti da linguisti di varia formazione nelle fasi del progresso scientifico in quest'ambito nel corso del '900. La difficoltà di valutare nel parlato spontaneo la presenza di

quanto, come segnalato e discusso da Bertinetto e Bertini (2010) l'efficacia delle misurazioni dipende, oltre che dalla velocità d'eloquio, dalla connessione del discorso e dalla lunghezza delle catene foniche. Nel caso di parlanti non nativi (cfr. anche De Meo 2012, Prieto *et aliae* 2012, Soriano 2012) la frammentazione è tale da produrre unità intonative brevi e spesso interrotte, causando forti riduzioni sui numeri di segmenti utili ai fini della valutazione del ritmo.

una regolarità dominante di un tipo o dell'altro ha indotto infatti un certo scetticismo rispetto al fenomeno in sé e alle procedure sperimentali per la sua misurazione.

Tuttavia, se – da un lato – nel primo decennio del XXI sec., alcuni autori ancora si chiedono se il ritmo linguistico sia una variabile fonologica o la conseguenza fonetica di altri eventi fonologici, la definizione di uno spazio ritmico pluridimensionale e di due poli “stress-based” (STB) e “syllable-based” (SYB) verso i quali si orienterebbero i tipi linguistici, in considerazione di una diversa disposizione a controllare variabilmente una di queste dimensioni, è avvenuta sulla base dell'affermarsi di modelli di valutazione delle cosiddette “metriche ritmiche” (a partire da Ramus *et alii* 1999 e Grabe & Low 2002; cfr. Bertinetto & Bertini 2010). Il principale vantaggio offerto dalle metriche ritmiche (v. § 2.3) consiste nel permettere di quantificare le variabili ritmico-temporali, e quindi nel permettere confronti quantitativi tra varietà o comunità linguistiche, nonché semplicemente tra diversi parlanti.

1.3. Obiettivo di questo studio

All'affermarsi di questi nuovi paradigmi, discussi tra gli altri da Kohler (2009) e Tarasi & Romito (2016), sono corrisposti numerosi tentativi di applicazione dei nuovi metodi di misurazione dapprima alle lingue più svariate (tra cui gli stessi Ramus, Nesport & Mehler 1999 e Grabe & Low 2002), poi alle varietà regionali e dialettali (Giordano & D'Anna 2010; Leemann, Dellwo, Kolly & Schmid 2012) e alle varietà di apprendimento. Riguardo a queste ultime, White & Mattys (2007) hanno mostrato che il parlato di apprendenti anglofoni di spagnolo L2 e di apprendenti ispanofoni di inglese L2 tendevano a presentare caratteristiche ritmiche a cavallo tra le due lingue. Questo sembrava mettere in luce un processo di “calibrazione” delle proprietà ritmiche durante l'acquisizione. Un risultato simile è stato ottenuto da Tortel & Hirst (2012), su dati di apprendenti francofoni di inglese L2: in questo caso gli autori hanno anche riscontrato differenze tra due diversi gruppi di apprendenti (adulti non specialisti vs. studenti universitari di inglese al 2° anno). In questo caso, l'analisi statistica condotta sulle metriche ritmiche permettevano di discriminare i due gruppi di apprendenti e il gruppo di controllo di inglesi nativi con un tasso del 70%. Kinoshita & Sheppard (2011) hanno invece confermato che i risultati delle metriche su un gruppo di apprendenti coreani di giapponese L2 sono correlati con i giudizi uditivi di nativi esperti. Pellegrino (2017) riporta che, dopo una fase di training basata su auto-imitazione (stimoli manipolati a partire da registrazioni di loro stessi), apprendenti nipponofoni di italiano L2 imitavano meglio le caratteristiche ritmiche del parlato, rispetto alle caratteristiche intonative. Pellegrino, He & Dellwo (2017) hanno mostrato che metriche ‘combinata’ (durata e f_0) erano in grado di discriminare le produzioni di sinofoni nativi e apprendenti italiani di cinese.

Tra gli studi di ricercatori che si sono dedicati all'analisi delle proprietà ritmiche di diverse lingue attraverso misurazioni sul parlato di apprendenti stranieri, non si annoverano casi, a nostra conoscenza, di ricerche condotte su apprendenti anglofoni di italiano L2 e che dispongano di una base di confronto su dati di un campione di parlanti nativi delle due lingue in questione. In questo contributo ci proponiamo quindi non solo di analizzare le caratteristiche ritmiche di apprendenti anglofoni di italiano L2, ma anche di confrontarle direttamente con dati di italiano L1 e inglese L1 elicitati con lo stesso paradigma sperimentale. Siccome l'italiano appartiene alle lingue *syllable-timed* o *syllable-based* (Bertinetto 1977), mentre l'inglese appartiene alle lingue *stress-timed* o *stress-based* (Abercrombie 1967), ci pare ragionevole ipotizzare che gli apprendenti anglofoni mostreranno proprietà ritmiche che si discostano da quelle della loro L1 (inglese) e che procedono in direzione della L2 (italiano), senza però raggiungere i valori riscontrati per italiani nativi.

2. Dati e metodologia

2.1. Partecipanti

Per questo studio sono state registrate produzioni orali di parlato letto per 24 apprendenti inglesi (14 F, 10 M; età mediana = 19 anni, range = 18-46 anni) iscritti ai corsi di italiano lingua straniera dell'Università di Strathclyde (Glasgow). Data la penuria di studenti disponibili, sono stati registrati apprendenti con caratteristiche piuttosto eterogenee. Al momento della registrazione, gli studenti frequentavano corsi di livello diverso; 18 di essi aveva studiato italiano per meno di 2 anni, ma 19 di essi aveva passato 1 o più anni in Italia e 7 di loro hanno dichiarato di parlare italiano con almeno 1 familiare. Inoltre, 15 di essi hanno dichiarato di parlare un'altra lingua romanza oltre all'italiano (6 hanno dichiarato di parlare il francese, 2 lo spagnolo, 7 francese e spagnolo).

Al fine di analizzare le proprietà temporali e ritmiche delle produzioni degli apprendenti (d'ora in avanti: gruppo AP), i risultati sono stati confrontati con dati prodotti sullo stesso testo e in condizioni analoghe da (i) parlanti nativi di italiano registrati in italiano L1 (d'ora in avanti, gruppo IT) e (ii) parlanti nativi di inglese registrati in inglese L1 (d'ora in avanti, gruppo EN). I dati dei gruppi IT ed EN provengono da Mairano & Romano (2011) e pertanto non possono essere considerati come veri e propri gruppi di controllo, ma piuttosto come punti di riferimento per l'italiano L1 e l'inglese L1. Per questo studio non è infatti stato possibile registrare specifici gruppi di controllo con variabili sociolinguistiche corrispondenti al gruppo sperimentale. Il gruppo IT era costituito da 14 parlanti di italiano standard, mentre i 4 partecipanti del gruppo EN sono parlanti standard di inglese britannico, americano, australiano e neozelandese rispettivamente e sono stati pubblicati nelle illustrazioni IPA di queste quattro varietà (rispettivamente, Roach 2004, Ladefoged 1989, Cox & Palethorpe 2007, Bauer *et alii* 2007). Rimandiamo a Mairano & Romano (2011) per maggiori dettagli sui parlanti dei gruppi EN e IT.

2.2. Metodologia

Le registrazioni del gruppo AP sono avvenute presso l'Università sopra menzionata, in un'aula non insonorizzata ed è stato utilizzato un registratore Tascam DR-40. Il testo su cui si basa la registrazione di parlato per tutti e 3 i gruppi è offerto dalle versioni inglesi e italiana del racconto esopico de "la tramontana e il sole".

Un giorno il vento di tramontana e il sole discutevano su chi dei due fosse il più forte, quando videro arrivare un passante con addosso un mantello.

Allora decisero che il più forte sarebbe stato quello dei due che fosse riuscito a fargli togliere il mantello.

Il vento per primo cominciò a soffiare con tutta la sua forza, ma più soffiava e più il passante si stringeva nel mantello, finché alla fine il vento non si diede per vinto.

Il sole a sua volta incominciò a splendere e subito il calore obbligò l'uomo a togliersi il mantello.

Così la tramontana fu costretta a riconoscere che il sole era più forte di lei.

Ti è piaciuta la storiella?

Vuoi che te la racconti di nuovo?

I partecipanti del gruppo IT sono stati registrati per la maggior parte nella cabina silente del laboratorio di fonetica di [ANONIMIZZATO], mentre i dati del gruppo EN sono stati realizzati in modalità analoghe dai vari autori delle illustrazioni IPA.

La segmentazione è stata inizialmente operata a un livello di distinzione tra catene foniche ininterrotte e pause, interruzioni o ispirazioni. L'annotazione è avvenuta secondo i criteri in uso nel laboratorio ispirati a quelli dei progetti nazionali API e CLIPS. Successivamente, all'interno degli intervalli delimitati è stata disposta un'etichettatura di tipo CV degli intervalli delimitati su un livello di analisi segmentale (nessi consonantici e incontri vocalici) secondo le indicazioni necessarie per l'analisi col *software* Correlatore (Mairano & Romano 2010).

2.3. Metriche

Al fine di fornire una descrizione delle caratteristiche ritmico-temporali del parlato di apprendenti anglofoni di italiano L2, abbiamo utilizzato le metriche ritmiche più utilizzate negli studi degli ultimi 20 anni. Come già menzionato nell'introduzione, esse misurano il grado di variabilità degli intervalli consonantici e vocalici nella catena di parlato. In particolare, le metriche utilizzate per questo studio sono le seguenti:

- *Vperc* (percentuale vocalica): indica in che percentuale il parlato è costituito da segmenti vocalico vs consonantici. Maggiori valori di *Vperc* sono associati in letteratura a un maggiore grado di isosillabicità (Ramus *et alii* 1999).
- *Vdev* (deviazione standard degli intervalli vocalici, spesso chiamata): indica il grado di variabilità delle durate degli intervalli vocalici. Maggiori valori di *Vdev* sono associati in letteratura a un più alto grado di isoaccentualità (Ramus *et alii* 1999).
- *Cdev* (deviazione standard degli intervalli consonantici): come sopra, ma consonantico.
- *varcoV* (coefficiente di varianza degli intervalli vocalici): come *Vdev*, ma include un coefficiente per normalizzare rispetto a diverse velocità d'eloquio tra i partecipanti (Dellwo & Wagner 2003).
- *varcoC* (coefficiente di varianza degli intervalli consonantici): come sopra, ma consonantico.
- *Vnpvi* (indice normalizzato di variabilità a coppie degli intervalli vocalici): indica il grado di variabilità delle durate degli intervalli vocalici adiacenti e include un coefficiente di normalizzazione rispetto a diverse velocità d'eloquio tra partecipanti. Maggiori valori di *Vnpvi* sono associati in letteratura a un più alto grado di isoaccentualità (Grabe & Low 2002).
- *Crpvi* (indice non-normalizzato di variabilità a coppie degli intervalli consonantici): come sopra, ma consonantico e senza coefficiente di normalizzazione, che viene considerato come superfluo nel caso delle consonanti. Maggiori valori di *Crpvi* sono associati in letteratura a un più alto grado di isoaccentualità.

I valori di queste metriche sono stati calcolati per le produzioni di ogni partecipante tramite il *software open-source* Correlatore (Mairano & Romano 2009). Le analisi statistiche sono state condotte su R (R Core Team 2017).

2.4. Catene foniche e pause

Uno studio preliminare è stato condotto sulle catene foniche e sulle pause, dato che i soggetti si sono differenziati sensibilmente su questo piano (si sono avute 39 e 40 unità ininterrotte per i parlanti 8 e 13, rispettivamente, vs. 89 unità per gli apprendenti 10 e 17).

Ad es. l'apprendente 8 produce a un certo punto “che il più forte sarebbe stato quello dei due” così come 13 dice “che il più forte sarebbe stato <pb> quello de<ee>i du<uu>e”, mentre la parlante 10, in corrispondenza, produce un passaggio incerto e frammentato “*ce +?il <pb> più forte<ee> <pb> sarebbe<ee> <pl> stato quello d+ <pb> *di<ii> due”.

Un confronto preliminare può essere affidato alla Fig. 1 nella quale si può osservare come a valori mediamente più alti e relativamente variabili delle durate delle catene foniche (es. 1, 8, 9 e 13) corrisponde di solito un numero minore delle stesse, mentre ai valori più contenuti (e generalmente uniformi: 7, 10, 11, 14, 15, 17...) corrispondono numeri elevati di catene foniche, risultando in una lettura estremamente frammentata (tra 76 e 89 intervalli, cfr. §3).

In Figura 2 si può osservare una durata piuttosto uniforme delle pause brevi <pb> (circa 30 ms) e buona corrispondenza tra pause lunghe <pl> (di circa 1 s, in media) e frontiere tra i blocchi testuali. I soggetti 5, 6, 8, 12, 13, 16, 20, 22, che ne realizzano meno di tre, non ne fanno o le inseriscono tra il testo principale e le due domande finali, o tra la premessa e il nucleo del racconto, o tra questo e la morale conclusiva.

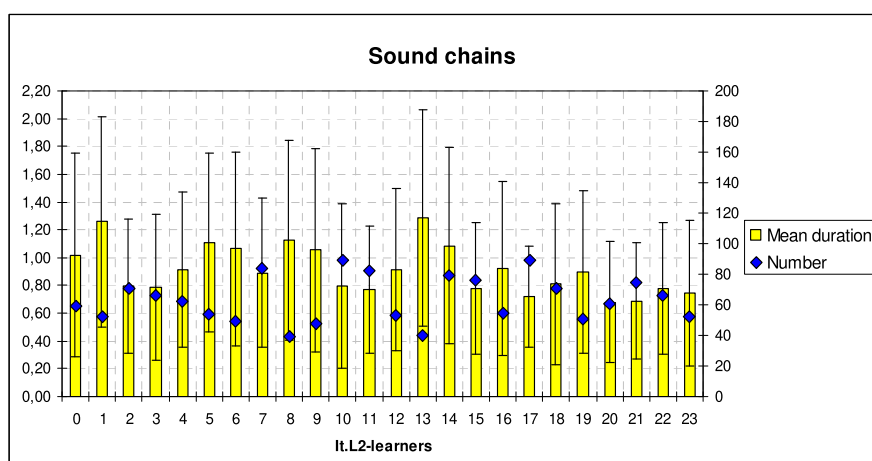


Figura 1: istogramma di variazione nel numero e nella durata delle catene foniche prodotte per lo stesso testo dai 24 soggetti del gruppo AP.

La separazione tra pause brevi (generalmente sotto i 60 ms) e pause lunghe³ (tra i 60 ms e i 3 s ca., nei casi in cui sono presenti in modo più significativo: 1, 2, 3, 9, 10, 11, 14, 15, 17 e 18) consente quindi di approfondire aspetti interessanti della segmentazione del testo in lettura. Ci limitiamo qui a dire che la correlazione tra l'occorrenza di pause lunghe e il livello generale dell'apprendente di cui si discute al §3, meriterebbe ancora una certa attenzione. Tuttavia, sarà interessante sottolineare qui la presenza di posizioni in cui un'interruzione è più ricorrente e altre in cui invece si mantiene una certa coesione fonica. Una frontiera prosodica terminale è offerta ad es. da un'interpunzione forte (soprattutto se accompagnata da a capo).

³ La velocità d'eloquio percepita (generalmente piuttosto variabile per lo stesso locutore) è servita come base per l'etichettatura manuale della pause (la revisione delle etichettature da parte di un secondo operatore esperto esclude il rischio di confusione tra casi di dialefe con glottidalizzazione e/o fasi di tenuta di occlusive sorde e pause brevi reali). Come si vede dai grafici, non sono state invece messe in discussione alcune scelte del primo operatore che non ha mancato di annotare come brevi pause di durata maggiore di 60 ms, così come, sporadicamente, ha percepito (e in qualche caso erroneamente etichettato) come lunghe pause di durata inferiore. La soglia è stata introdotta a posteriori su basi puramente impressionistiche (ma i grafici confermano una distinta distribuzione).

È così che tra i sintagmi mantenuti coesi sistematicamente dai soggetti osservati troviamo: “Un giorno”, “più forte” (2/3), “a soffiare”, “per vinto”, “(con) tutta la sua forza”, “il sole” (2/3), “di lei”, “di nuovo”, “a splendere”, “a sua (volta)”. Non stupisce anche la coesione fonica di “non si” e “Così la”, seguiti rispettivamente da parole che richiedono una preanalisi endofasica, data la complessità o l'infrequenza (“diede” e “tramontana”). Invece; la mancanza di pausa o scansione prosodica in corrispondenza di “due” e “fosse” in “quello dei due che fosse riuscito” ci fa supporre che i nostri partecipanti (tutti di livello non superiore a B1) non conoscessero la parola “fosse”. A un livello di coesione alto, ma non universale, notiamo infine “e subito”, “il mantello” e “la storiella”, mentre sono altrettanto sistematiche le pause dopo “Allora” e “decisero”.

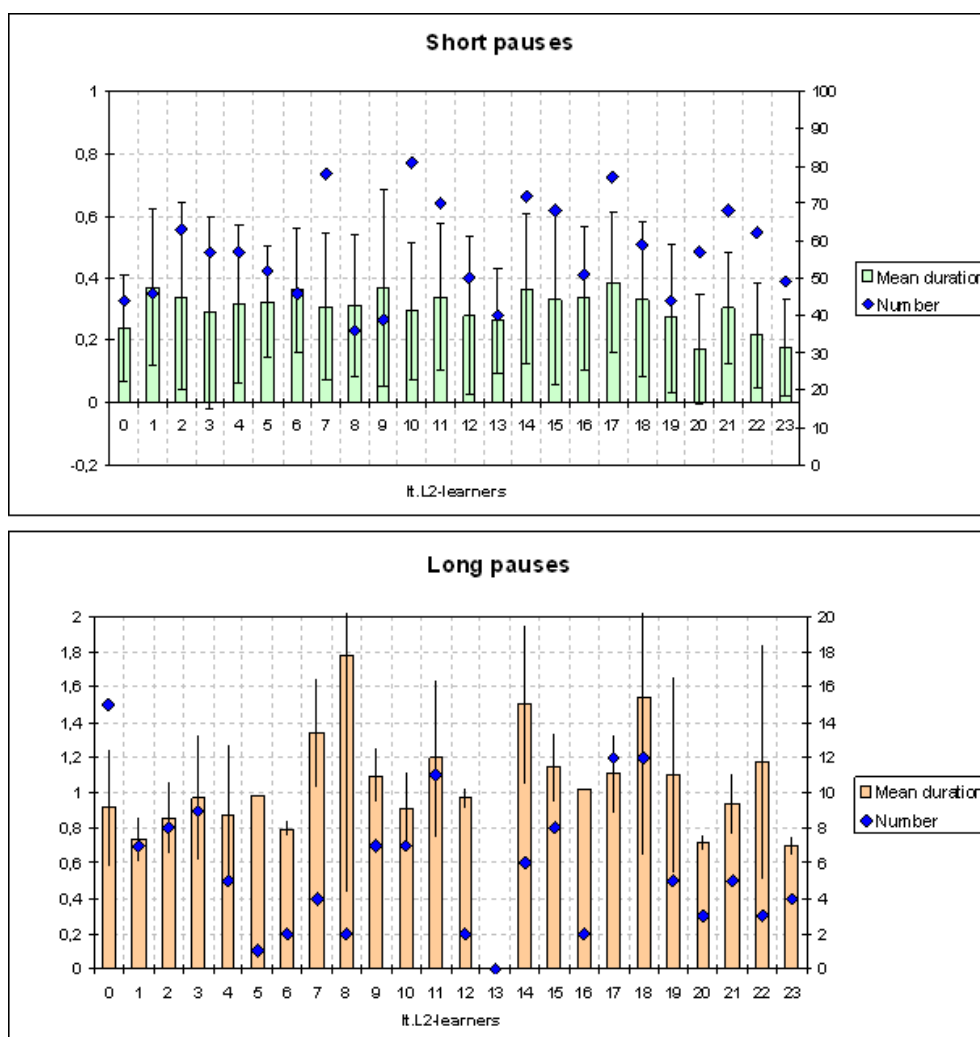


Figura 2: istogrammi di variazione nel numero e nella durata delle pause prodotte per lo stesso testo dai 24 soggetti del gruppo AP.

3. Risultati

3.1. Metriche ritmiche

I risultati ottenuti per ognuna delle metriche ritmiche sono illustrati tramite grafici di densità della distribuzione in figura 1. Questi grafici mostrano bene quali di queste metriche permettono una differenziazione dei gruppi sulla base delle proprietà ritmico-temporali: i tre gruppi sembrano infatti distribuirsi su valori diversi in relazione a V_{perc} , V_{npvi} e (in minor misura) V_{dev} . Invece, le curve di densità dei 3 gruppi mostrano un alto grado di *overlap* per le altre metriche, che quindi risultano avere un potere di classificazione meno elevato per i nostri dati. Come suggerito da uno dei revisori anonimi, questo risultato potrebbe derivare dalla grande variabilità dei dati analizzati, che include apprendenti dalle caratteristiche molto eterogenee.

Abbiamo poi sottoposto questi risultati a un'analisi statistica per verificare la significatività delle differenze riscontrate tra i diversi gruppi; per ogni metrica, l'analisi è stata condotta tramite test dei ranghi con segno di Wilcoxon. Siccome la nostra variabile di gruppo ha 3 livelli (AP, EN, IT), per stabilire i valori di significatività statistica abbiamo usato test a coppie con correzione di Bonferroni-Holm per confronti multipli.

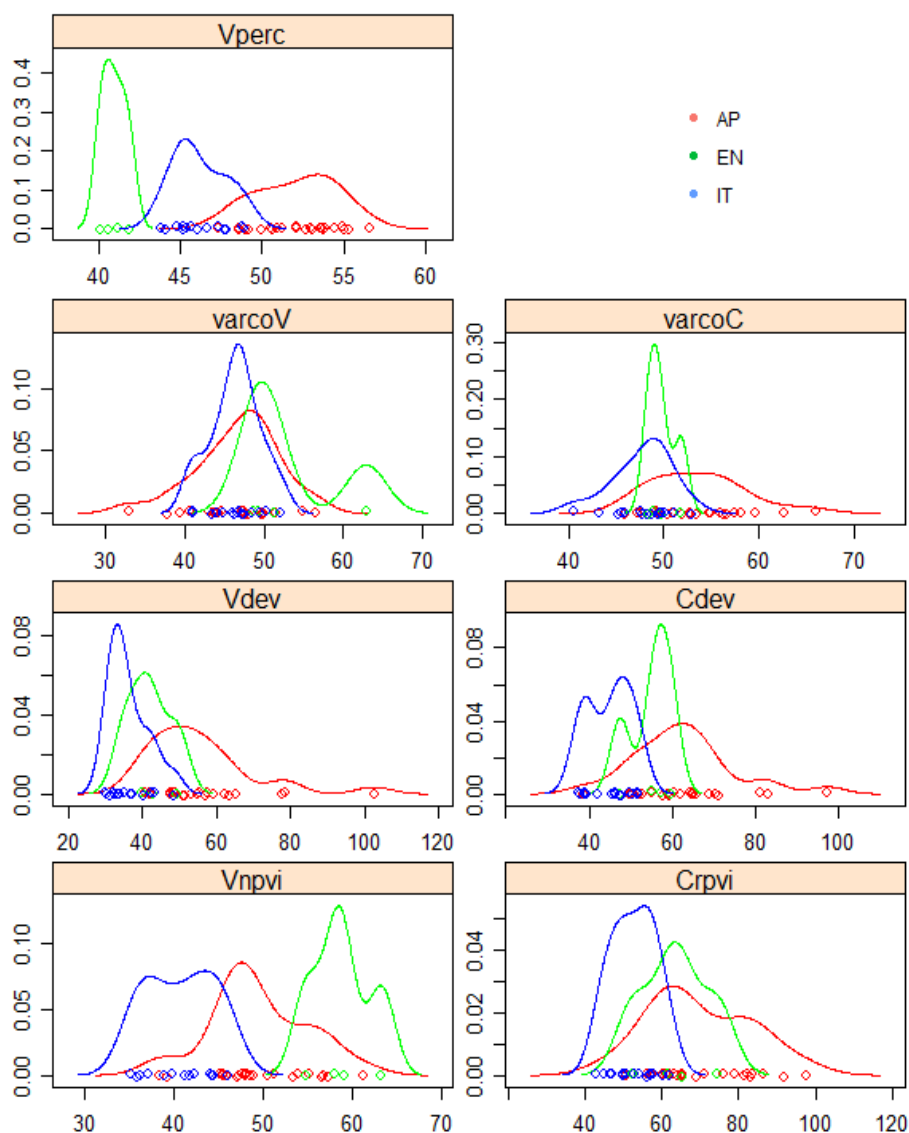


Figura 3: grafici di densità della distribuzione per le 7 metriche ritmiche considerate e i 3 gruppi (AP in rosso, EN in verde, IT in blu).

Secondo l'esito dell'analisi, tutte le differenze tra i valori di Vperc e Vnpvi sono significative: le produzioni del gruppo AP mostrano una percentuale vocalica (Vperc) più alta del gruppo EN ($p < 0.001$) e del gruppo IT ($p < 0.001$), i quali differiscono significativamente tra loro ($p < 0.001$). I valori più alti di Vperc nel caso del gruppo AP si spiegano con i maggiori tempi di articolazione, legati a un maggiore esitazione (che si realizza naturalmente sulle vocali) e a una velocità d'eloquio complessivamente ridotta: è infatti risaputo che i parlanti di una L2 tendono a parlare più lentamente rispetto a parlanti nativi (Trofimovich & Baker 2006). Tali risultati di Vperc ricordano inoltre quelli ottenuti da White & Mattys (2007): in questo studio, apprendenti anglofoni di spagnolo L2 e apprendenti ispanofoni di inglese L2 mostravano entrambi valori più alti di Vperc in relazione ai rispettivi gruppi di controllo nativi. Per quanto riguarda la variabilità degli intervalli vocalici (normalizzata per velocità d'eloquio), il gruppo AP si situa con valori medi a metà strada, significativamente più bassi rispetto al gruppo EN (p

< 0.01) ma significativamente più alti rispetto al gruppo IT ($p < 0.001$); naturalmente, anche i due gruppi di controllo IT ed EN differiscono in maniera statisticamente significativa ($p < 0.001$).

Passando all'analisi delle altre metriche, nessuno dei tre gruppi differisce in maniera statisticamente significativa per i valori di $Vdev$ e $varcoV$; notiamo solo una tendenza marginalmente significativa del gruppo EN a distanziarsi da AP ($p = 0.05$ per $Vdev$) e da IT ($p = 0.06$ per $varcoV$). Per quanto riguarda le metriche consonantiche, notiamo sorprendentemente che il gruppo AP presenta costantemente valori medi più alti di variabilità degli intervalli consonantici rispetto agli altri gruppi. Tuttavia, le differenze sono statisticamente significative solo tra i gruppi AP e IT ($p < 0.001$ per tutte le metriche consonantiche), ma mai tra AP e EN. Questo ci porta quindi a pensare che il gruppo AP tenda a comportarsi in maniera simile al gruppo EN per quanto riguarda i fenomeni di durata consonantica.

3.2. Regressione logistica multinomiale

Avendo stabilito che $Vperc$ e $Vnpvi$ sono i parametri che, presi singolarmente, sono maggiormente in grado di differenziare le produzioni dei 3 gruppi presi in esame, passiamo a un'analisi multiparametrica che permetta una visione più globale. La figura 2 offre quattro rappresentazioni bidimensionali con coppie di metriche secondo gli abbinamenti suggeriti in letteratura: rispettivamente, $Cdev$ e $Vperc$ (Ramus *et alii* 1999), $Cdev$ e $Vdev$ (ibid.), $varcoC$ e $varcoV$ (Dellwo & Wagner 2003), $Crpvi$ e $Vnpvi$ (Grabe & Low 2002). Una prima osservazione sembra suggerire che il grafico $Cdev / Vdev$ offra la ripartizione in gruppi maggiormente distinta, con poca sovrapposizione dei punti relativi ai 3 gruppi. Tuttavia, un'analisi più attenta rivela che la separazione dei gruppi è dovuta essenzialmente ai valori di $Vperc$ (asse delle ordinate), che abbiamo visto essere la metrica con maggior potere discriminante rispetto a questi dati.

Risulta interessante notare che il gruppo AP si colloca su valori estremi (ovvero oltre IT ed EN) per le metriche che non normalizzano la velocità d'eloquio, ma su valori intermedi tra IT ed EN per le metriche che normalizzano la velocità d'eloquio. Ci pare in effetti ragionevole pensare che gli alti valori di $Vperc$ e $Vdev$ del gruppo AP siano dovuti alle differenze di eloquio già evocate in merito a questo gruppo. La questione della relazione tra velocità d'eloquio e ritmo è molto delicata, soprattutto per dati di L2 (cf. Gut 2012). Ci pare comunque appropriato separare questi due concetti; non per 'eliminare' dall'analisi la velocità d'eloquio, che sicuramente è un dato temporale molto importante, bensì per analizzare questi due fenomeni separatamente.

Al fine di stabilire quali siano le variabili più rilevanti per una classificazione dei parametri ritmici dei nostri tre gruppi, abbiamo proceduto a un'analisi tramite regressione logistica multinomiale: il fattore gruppo (AP, EN, IT) costituiva la variabile dipendente del modello, mentre ogni metrica presa in esame costituiva una variabile indipendente, secondo la seguente formula: $group \sim Vperc + Cdev + Vdev + Crpvi + Vnpvi + varcoV + varcoC$. La procedura di eliminazione progressiva degli effetti non significativi ha portato il modello a convergere sulla seguente formula ottimale: $group \sim Vdev + Vnpvi + varcoV$, con un valore AIC di 16.00 e una devianza residua dello 0.0003. Questo risultato indica che le 3 metriche sono le uniche necessarie per modellare statisticamente in maniera ottimale le differenze ritmiche tra i tre gruppi presi in esame.

In considerazione del fatto che le metriche non normalizzate potrebbero semplicemente riflettere le già evocate differenze di velocità d'eloquio (anch'essa una variabile temporale, strettamente collegata al ritmo) più che differenze di altri fenomeni ritmico-temporali, abbiamo provato anche a ripetere l'analisi includendo solo le metriche che offrivano una

normalizzazione. Partendo dalla formula $group \sim Vnpvi + varcoV + varcoC$, il processo di ottimizzazione del modello converge proprio sulla formula iniziale, confermando quindi l'utilità di tutte e tre le variabili per la classificazione dei gruppi. Tuttavia, avendo normalizzato (e quindi di fatto eliminato dall'analisi) gli effetti della velocità d'eloquio, il modello è decisamente meno potente, come dimostrato dal valore di AIC 33.00 e da una devianza residua di 17.90. Questo conferma quindi che una grossa parte delle differenze visibili nei grafici è un prodotto della velocità d'eloquio (come facilmente intuibile nel caso di una L2 e come già discusso sopra), più che delle proprietà temporali inter-segmentali che le metriche ritmiche vorrebbero catturare.

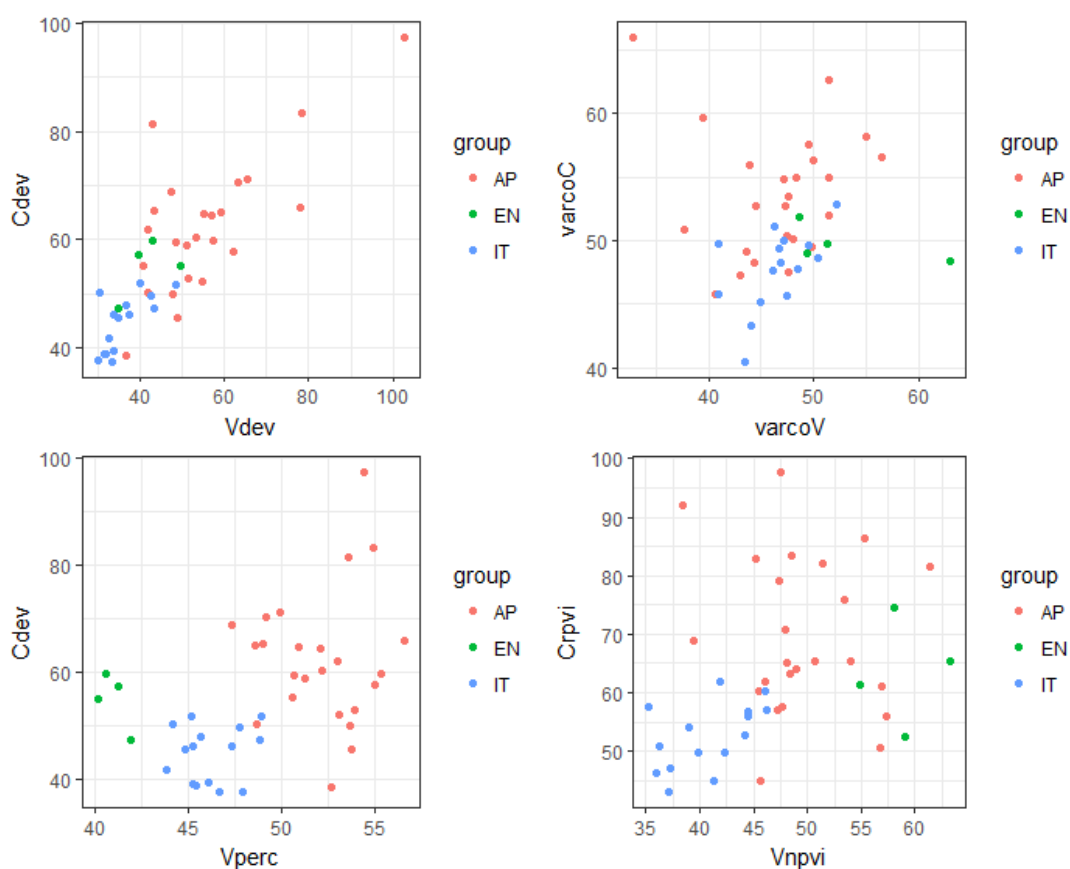


Figura 4: Grafici di Cdev / Vdev, Cdev / Vperc, varcoC / varcoV, Crpvi / Vnpvi per i locutori dei 3 gruppi presi in esame.

3.3. Analisi degli effetti delle variabili di acquisizione

Infine, abbiamo concentrato la nostra analisi sui potenziali effetti delle variabili di acquisizione della L2 sulle metriche ritmiche. Come precedentemente menzionato, i partecipanti avevano compilato un questionario con informazioni anagrafiche e relative al loro percorso di apprendimento dell'italiano L2. Questo ci ha permesso di valutare la portata degli effetti delle seguenti variabili:

- gli anni di studio dell'italiano L2 (range: 0.5 - 10)
- gli anni passati in Italia (range: 0 - 6)

- l'utilizzo dell'italiano con i familiari
- la competenza in un'altra lingua romanza (francese e/o spagnolo)

Come già menzionato, i parlanti registrati avevano caratteristiche molto eterogenee rispetto alle quattro variabili analizzate; il controllo di tutte queste variabili richiederebbe un numero molto più consistente di registrazioni e di selezione dei partecipanti, comportando fondi e risorse che non erano disponibili per questo contributo. Abbiamo comunque cercato di misurare eventuali correlazioni, tenendo comunque presente che la significatività statistica di tale analisi non può essere che debole, di fronte a tale eterogeneità.

In una prima analisi, abbiamo calcolato la correlazione tra ognuna delle metriche e gli anni di studio dell'italiano L2 e di permanenza in Italia. I risultati sono riportati nella tabella 1 e mostrano valori di correlazione piuttosto bassi. In generale, notiamo comunque che i valori delle metriche tendono ad abbassarsi (e quindi ad avvicinarsi ai valori del gruppo IT) con il crescere degli anni di studio e degli anni passati in Italia.

	Vperc	Cdev	Vdev	varcoC	varcoV	Crpvi	Vnpvi
Anni di studio di italiano L2	-0.27	-0.30	-0.24	-0.14	0.04	-0.35	0.01
Anni passati in Italia	-0.23	-0.07	-0.22	-0.16	-0.33	-0.14	-0.37

Tabella 1. Tabella di correlazione tra i valori delle metriche ritmiche per il gruppo AP e gli anni di studio di italiano L2 e di permanenza in Italia.

Per testare la validità statistica di queste correlazioni, e per valutare l'effetto delle altre variabili (l'utilizzo dell'italiano con i familiari e la competenza in altre lingue romanze), abbiamo poi costruito dei modelli lineari per ognuna delle metriche ritmiche, in cui le variabili indipendenti erano i fattori menzionati sopra e includendo alcune interazioni, secondo la seguente formula: $Vperc \sim \text{anni studio di italiano} * \text{anni passati in Italia} + \text{parla francese} * \text{parla spagnolo} + \text{parla italiano con familiari}$. La variabile dipendente è Vperc nell'esempio, altri modelli equivalenti sono stati valutati per tutte le metriche prese in esame. I risultati hanno mostrato che nessuna di queste variabili sembra avere un effetto statisticamente significativo sui valori di nessuna metrica ritmica per i nostri dati. L'unica eccezione è l'effetto della variabile "parla italiano con familiari" su Vdev ($p < 0.05$). Inoltre, gli anni di studio di italiano L2, gli anni passati in Italia e la conoscenza dello spagnolo sembrano avere effetti marginalmente significativi su Vperc ($p = 0.06$) sui dati dei nostri partecipanti. Risultati simili sono stati trovati da Gut (2009) e Sarmah, Gogoi & Wiltshire (2009) su dati più controllati, in cui le variabili di apprendimento (in particolare l'esposizione alla L1) non hanno effetti significativi sui valori delle metriche ritmiche. Tuttavia, il nostro studio utilizza variabili sovrapposte, ovvero i partecipanti che parlano italiano con i familiari sono parzialmente gli stessi che parlano un'altra lingua romanza, etc. Queste rende l'analisi poco affidabile, e lascia quindi le porte aperte a una verifica più controllata dell'effetto di tale variabili sui valori delle metriche ritmiche.

4. Conclusioni

In questo contributo, abbiamo dato una caratterizzazione ritmica di apprendenti italo-foni di italiano L2. La nostra ipotesi di partenza postulava che il gruppo di apprendenti anglo-foni di italiano L2 avrebbero mostrato caratteristiche ritmiche intermedie tra la L1 e la L2, come tipico delle interlingue. Questa ipotesi è effettivamente confermata da alcune delle metriche ritmiche prese in esame, in particolare quelle che offrono una normalizzazione rispetto alla velocità d'eloquio, come Vnpvi e varcoV. I risultati di queste metriche mostrano infatti valori più bassi per il gruppo di italiani nativi, più alti per il gruppo di inglesi nativi (coerentemente con i valori riportati in letteratura da Ramus *et alii*, 1999, e Grabe & Low, 2002) e intermedi per il gruppo di

apprendenti. Questo sembra quindi indicare un minor controllo delle durate vocaliche da parte del gruppo di apprendenti di italiano L2, che probabilmente sfruttano processi di riduzione vocalica propri della L1, causando quindi più alti valori di V_{npvi} e $varcoV$ rispetto ai parlanti nativi di italiano. Il quadro offerto dalle metriche ritmiche a livello consonantico è invece meno definito, in quanto gli apprendenti mostrano una grossa variabilità e le differenze rispetto al gruppo EN non sono significative: limitatamente agli intervalli consonantici, sembrerebbe quindi che gli apprendenti riproducano in italiano L2 le caratteristiche temporali tipiche della loro L1.

I risultati di V_{perc} meritano invece un discorso a parte. Il quadro offerto da questa metrica non riflette la nostra ipotesi iniziale. Il gruppo di apprendenti non mostra infatti valori di V_{perc} intermedi tra quelli tipici dell'inglese nativo (valori bassi) e dell'italiano nativo (valori alti), bensì addirittura più alti rispetto alla lingua target. Una configurazione simile era stata trovata da White & Mattys (2007): parlanti inglesi di spagnolo L2 mostravano una V_{perc} più alta rispetto agli stessi parlanti nativi di spagnolo. Gli autori lo descrivono come un caso di *overshooting*, quindi una specie di ipercorrezione a livello prosodico. Nel nostro caso, sembrerebbe che tali valori di V_{perc} siano dovuti semplicemente all'interferenza della velocità d'eloquio sui risultati delle metriche senza normalizzazione. Gli indizi che puntano verso questa interpretazione sono due: da un lato, anche V_{dev} (che non include nessuna normalizzazione rispetto alla velocità d'eloquio) offre una ripartizione dei gruppi in cui gli apprendenti non si situano in posizione intermedia tra la L1 e la lingua target; dall'altro, abbiamo verificato che i modelli di regressione logistica multinomiale che includevano metriche non normalizzate per la velocità d'eloquio avevano un potere predittivo significativamente maggiore, indicando quindi un'importante interferenza della stessa. Crediamo che questo risultato sia significativo dal punto di vista metodologico ed evidenzia l'importanza di scindere (o, per lo meno, essere consapevoli del-) le varie componenti segmentali e globali che influiscono sulle misure ritmiche, in particolare su dati di L2 vs L1, una problematica già sollevata da Gut (2012). Gut suggerisce che un confronto tra parlanti L1 vs L2 sia possibile solo se viene chiesto ai locutori L1 di parlare a una velocità d'eloquio simile a quella dei locutori L2 con cui dovranno essere confrontati. Sebbene il nostro punto di vista sia forse meno netto, rimane certo che gli studi che confrontano gruppi di locutori di L1 e L2 devono tenere in conto possibili differenze sistematiche in termini di velocità d'eloquio, che hanno conseguenze sui valori delle metriche ritmiche non normalizzate.

Infine, l'analisi delle variabili di acquisizione in relazione alle caratteristiche ritmiche del gruppo di apprendenti non ha rivelato effetti significativi per i dati disponibili, tranne eccezioni marginali. È possibile che le caratteristiche ritmiche siano piuttosto legate ad altre variabili più difficilmente misurabili attraverso un questionario, come l'attitudine o la motivazione. Oppure, la mancanza di correlazioni significative è semplicemente da ascrivere all'eterogeneità dei dati e al fatto che si tratti di dati con variabili sovrapposte, ovvero non controllate singolarmente. È comunque da rilevare che alcune correlazioni sono descrittivamente presenti (anche se non significative) e che la loro direzione va nel senso delle nostre predizioni (ovvero i valori delle metriche ritmiche si avvicinano ai valori dei nativi con il crescere degli anni di studio e con gli anni passati in Italia). Probabilmente il campione di dati non è sufficientemente ampio e controllato per far emergere questo tipo di correlazioni, aprendo quindi la strada a futuri studi che esploreranno più nel dettaglio e su un campione più ampio e controllato l'effetto delle variabili di acquisizione sulle caratteristiche ritmiche del parlato di apprendenti.

BIBLIOGRAFIA

- Abercrombie, D. (1967) *Elements of General Phonetics*. Edinburgh University Press.
- Bauer, L., Warren, P., Bardsley, D., Kennedy, M. & Major, G. (2007). New Zealand English. *JIPA*, 37(1), 97-102.
- Bertinetto, P. M. (1977). Syllabic blood: ovvero l'italiano come lingua ad isocronismo sillabico. *Studi di grammatica italiana*, 6, 69-96.
- Bertinetto, P.M. & Bertini, C. (2010) Towards a unified predictive model of Natural Language Rhythm. In Russo, M. (ed.) *Prosodic Universals. Comparative Studies in Rhythmic Modeling and Rhythm Typology* (pp. 43-77), Roma: Aracne.
- Best, C. T. (1995) A Direct Realist View of Cross-Language Speech Perception. In W. Strange (ed.) *Speech perception and linguistic experience: Issues in cross-language research* (pp. 171-204), Timonium, MD, York Press.
- Busà, M.G., & Stella, A. (2012) *Methodological Perspectives on Second Language Prosody*. Padova: CLEUP.
- Colantoni, L., Steele, J. & Escudero, P. (2015). *Second language speech*. Cambridge University Press.
- Cox, F. & Palethorpe S. (2007) Australian English. *JIPA*, 37 (3), 341-350.
- De Meo A. (2012) How credible is a non-native speaker? Prosody and surroundings. In Busà & Stella (2012), 3-9.
- De Meo, A., Vitale, M. & Pellegrino, E. (2016). Tecnologia della voce e miglioramento della pronuncia in una L2: imitazione e autoimitazione a confronto. Uno studio su sinofoni apprendenti di italiano L2. In: F. Bianchi & P. Leoni (eds) *Linguaggio e apprendimento linguistico. Metodi e strumenti tecnologici* (pp. 13-25), Studi AltLA, Milano: Officinaventuno.
- Dellwo, V. & Wagner, P. (2003) Relations between language rhythm and speech rate. *Proc. of the 15th International Congress of Phonetics Sciences* (Barcelona, Spain), pp. 471-474.
- Eckman, F. (1977), Markedness and the Contrastive Analysis Hypothesis, *Language Learning*, 27 (2), 315-330.
- Flege, J.E. (1995). Second language speech learning: Theory, findings, and problems. In Strange, W. (ed.), *Speech perception and linguistic experience: Theoretical and methodological issues in cross-language speech research* (pp. 233-277), Timonium, MD, York Press.
- Grabe, E. & Low, E.L. (2002) Durational variability in speech and the rhythm class hypothesis. In Gussenhoven, C. & Warner, N. (eds.) *Papers in Laboratory Phonology 7*, Berlin: Mouton de Gruyter, 515-546.
- Giordano, R., & D'Anna, L. (2010). A comparison of rhythm metrics in different speaking styles and in fifteen regional varieties of Italian. In *Proc. of Speech Prosody 2010*, 11-14 May, Chicago (US).
- Gut, U. (2012). Rhythm in L2 speech. *Speech and Language Technology*, 14/15, 83-94.
- Gut, U. (2009). *Non-native Speech: A Corpus-based Analysis of Phonological and Phonetic Properties of L2 English and German*. Frankfurt: Peter Lang.
- Kinoshita, N., & Sheppard, C. (2011). Validating acoustic measures of speech rhythm for second language acquisition. In *Proc. of the 17th International Congress of Phonetic Sciences* (pp. 1086-1089), 17-21 August, Hong Kong (China).
- Kohler, K.J. (2009) Rhythm in speech and language - A new research paradigm. *Phonetica*, 66(1), 20-45.
- Ladefoged, P. (1989) American English. *JIPA*, 19(2), 77-80 (v. anche *IPA 1999. Handbook of the International Phonetic Association*, Cambridge: Cambridge University Press, 41-44).
- Leemann, A., Dellwo, V., Kolly, M. J., & Schmid, S. (2012). Rhythmic variability in Swiss German dialects. In *Proc. Speech Prosody 2012*, 22-25 May, Shanghai (China).
- Mairano, P. & Romano, A. (2011) Rhythm metrics for 21 languages. *Proc. of XVII ICPHS (International Congress of Phonetic Sciences)*, Hong Kong (China), 17-21 August 2011.
- Mairano, P. & Romano, A. (2010) Un confronto tra diverse metriche ritmiche usando Correlatore. In Schmid, S., Schwarzenbach, M. & Studer, D. (eds.) *La dimensione temporale del parlato* (pp. 79-100), Torriana: EDK Editore.

- Mennen, I. (2015). Beyond segments: Towards a L2 intonation learning theory. In E. Delais-Roussarie, M. Avanzi & S. Herment (Eds.) *Prosody and language in contact* (pp. 171-188). Springer, Berlin, Heidelberg.
- Mennen, I., Schaeffler, F., & Dickie, C. (2014). Second language acquisition of pitch range in German learners of English. *Studies in Second Language Acquisition*, 36(2), 303-329.
- Pellegrino, E. (2017). Can second language suprasegmentals be learned? A study on Japanese learners of Italian as foreign language. *The Journal of the Acoustical Society of America*, 141(5), 3519-3519.
- Pellegrino, E., He, L., Dellwo, V. (2017). Computation of L2 speech rhythm based on duration and fundamental frequency. In: J. Trouvain, I. Steiner, B. Mobius (eds). *Elektronische Sprachsignalverarbeitung 2017*. Dresden: TUD press, 246-253.
- Pettorino, M., De Meo, A., & Vitale, M. (2012). La competenza prosodico-intonativa nell'italiano L2. Analisi e sintesi del segnale fonico di cinesi, vietnamiti e giapponesi. In *La linguistica educativa. Proceedings of the 44th International Congress of Italian Linguistics Society SLI* (pp. 329-342).
- Prieto, P., Vanrell, M.M., Astruc, Ll., Payne, E. & Post, B. (2012) Phonotactic and phrasal properties of speech rhythm. Evidence from Catalan, English, and Spanish. *Speech Communication*, 54, 681-702.
- R Core Team (2017) *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. Retrieved from: <https://www.R-project.org/>.
- Ramus, F., Nespore, M. & Mehler, J. (1999) Correlates of linguistic rhythm in the speech signal. *Cognition*, 73(3), 265-292.
- Roach, P. (2004). British English: Received pronunciation. *JIPA*, 34(2). 239-245.
- Santiago, F., Mairano, P., & Delais-Roussarie, E. (2014). Non-native perception of final boundary tones in French interrogatives. In *Proceedings of the 7th International Conference on Speech Prosody* (pp. 563-567).
- Sarmah, P., Gogoi D. V. & Wiltshire C. (2009). Thai English. Rhythm and vowels. *English World-Wide*, 30, 196-217.
- Sorianello, P. (2012) Prosodic features in native and non-native speech segmentation. In De Meo, A. & Pettorino, M. (eds.) *Prosodic and Rhythmic Aspects of L2 Acquisition* (pp. 15-35), Cambridge: Cambridge Scholar Pub.
- Tarasi, A. & Romito, L. (2016) *Un nuovo paradigma nello studio del ritmo delle lingue naturali*, Saarbrücken: Edizioni Accademiche Italiane.
- Tortel, A. & Hirst, D. (2010) Rhythm metrics and the production of English L1/L2. In Hasegawa-Johnson, M. (ed.), *Proceedings of Speech Prosody 2010*, Chicago, 100959:1-4.
- Trofimovich, P., & Baker, W. (2006) Learning second language suprasegmentals: Effect of L2 experience on prosody and fluency characteristics of L2 speech. *Studies in second language acquisition*, 28(1), 1-30.
- Vitale, M., Pellegrino, E., De Meo, A., & Rasulo, M. (2015). Misurare la competenza prosodica. Le richieste in italiano e in inglese lingue straniere. In Chini M. (Ed.) *Il parlato in italiano L2: aspetti pragmatici e prosodici* (pp. 59-72), Milano: Francoangeli.
- White, L., & Mattys, S. L. (2007). Calibrating rhythm: First language and second language studies. *Journal of Phonetics*, 35(4), 501-522.

PAOLO MAIRANO • is currently *maître de conférence* (Assistant Professor) at the University of Lille (France). After obtaining a PhD in linguistics from the University of Turin, he also worked as research fellow at the Universities of Grenoble (France) and Warwick (UK), as teaching fellow at the University of Rouen (France), and as speech scientist at Nuance Communications Inc. (Italy). His research interests range from phonetic and phonological aspects of L2 acquisition to language typology and speech technologies.

E-MAIL • paolo.mairano@univ-lille.fr

MARTA MOIS • studies English and Russian at the University of Turin. In 2015 she got a Bachelor's degree in Modern Languages with a thesis in Linguistics with the title "A corpus-based

study of syntactic transfer by Italian learners of English”. During her Master years, she took part to the Erasmus Traineeship project and worked as language assistant at the Department of Italian studies of the University of Strathclyde (Glasgow), where she started her researches for the Master degree.

E-MAIL • marta.mois@edu.unito.it

VALENTINA DE IACOVO • is finishing a PhD in Digital Humanities in collaboration with the Universities of Genoa and Turin, working specifically on the Italian section of the International AMPER “Atlas Multimédia Prosodique de l’Espace Roman” project; she is also organising the LFSAG speech archives and defining a guidelines definition for the prosodic annotation.

E-MAIL • vdeiacov@unito.it

ANTONIO ROMANO • is Professor (habil. as Full Professor) at the Dept. of Foreign Languages and Literatures of the University of Turin and scientific director of the Laboratory of Experimental Phonetics “Arturo Genre” and Director of the Master in Translation for Cinema, TV and Multimedia. He is also co-ordinator of AMPER Atlas Multimédia Prosodique de l’Espace Roman (with M. Contini) and director of the Italian committee (AMPER-ITA).

E-MAIL • antonio.romano@unito.it