

ALESSANDRO VIETTI, DANIELA MEREU¹

Sistemi vocalici in contatto nell'italiano di Bolzano: un'analisi esplorativa *corpus-based*

Vowel systems in contact in Bolzano Italian:
a *corpus-based* exploratory analysis

The aim of our paper is to illustrate an exploratory analysis on the vowel system in Italian spoken in Bolzano by both native and non-native speakers (i.e. Tyrolean German speakers). Data for the analysis come from DIA (Dialogic ItAlian) corpus, a speech corpus of spontaneous Italian spoken in Bolzano and based on dialogic interactions. After a short illustration of the studies on Italian vowels and a phonetic description of the Italian speech variety of Bolzano, we provide some information on the construction of DIA corpus. The central part of the contribution is devoted to the discussion of the specific methodological problems that arise in the acoustic analysis of vowels extracted from speech corpora. To address these problems, we carry out a spectral analysis based on DCT coefficients and we test the hypothesis that there are two vowel sub-systems using linear mixed models.

Keywords: Italian vowels; spontaneous speech; spectral analysis (DCT coefficients).

1. Introduzione

L'obiettivo principale di questo contributo è di presentare un'indagine esplorativa sulle vocali dell'italiano di Bolzano. Il suo carattere esplorativo si spiega con la necessità di definire una procedura di analisi che possa essere valida per lo studio delle vocali estratte da corpora di parlato spontaneo. Il lavoro che qui si presenta si basa infatti sul DIA-Dialogical ItAlian, un corpus di parlato spontaneo della varietà di italiano bolzanino, costruito su un campione di parlanti nativi e non nativi di italiano, ovvero tedescofoni. Prima di entrare nel vivo delle questioni metodologiche e procedurali dell'analisi, forniremo una descrizione della varietà oggetto di indagine, sulla base di una revisione della bibliografia esistente sull'argomento.

Inizieremo, quindi, con alcune note di sintesi sugli studi vocalici per l'italiano (par. 2), in seguito tratteremo una breve rassegna sull'italiano di Bolzano, dando una particolare attenzione alle vocali (par. 3), per poi concentrarci sulla metodologia adottata per la costruzione del corpus di riferimento e per la preparazione dei dati da analizzare (par. 4); la sezione centrale del lavoro si focalizzerà invece sull'a-

¹ Nel suo complesso, il lavoro è il frutto di una collaborazione tra i due autori; tuttavia, per fini accademici, segnaliamo che Alessandro Vietti è responsabile delle sezioni 5, 5.1, 5.2, 5.3, 6, mentre Daniela Mereu è responsabile delle sezioni 1, 2, 3, 4 (4.1, 4.2).

nalisi esplorativa delle vocali (par. 5), chiarendo dapprima la natura dei problemi metodologici legati all'analisi acustica di grandi quantità di dati spontanei (par. 5.1) e, successivamente, illustrando i risultati di un'analisi spettrale condotta con l'impiego di modelli lineari misti (par. 5.2). All'analisi seguiranno alcune osservazioni conclusive (par. 6).

2. *Le vocali dell'italiano*

Gli studi di fonetica sperimentale sulle vocali in italiano vantano una lunga tradizione, a partire dai primi lavori di Ferrero (1968; 1972), che si sono concentrati sulle vocali dell'italiano, senza fare distinzione tra varietà regionali diverse². Da questi lavori pionieristici, le ricerche sulle vocali sono state in seguito affinate secondo diverse prospettive. Una prima linea di indagine ha riguardato lo studio degli spazi vocalici in alcuni dialetti e nelle rispettive varietà locali di italiano (es. Mendicino & Romito, 1991; Romito & Lorenzi, 1997). Parallelamente, sono state sviluppate anche delle riflessioni di carattere metodologico sulla misurazione dei segmenti vocalici, come dimostrano i contributi scientifici presentati al primo convegno italiano su questo tema, sfociato nel volume a cura di Trumper e Romito (1993).

Progressivamente, la ricerca sulle vocali è stata ampliata sia in prospettiva diatopica, attraverso il confronto tra diversi punti linguistici di una stessa varietà e in relazione a diverse varietà di italiano regionale, sia in prospettiva sociolinguistica, con una particolare predilezione verso la dimensione diafasica di variabilità. Questo incremento di studi è stato reso possibile grazie anche alla costruzione di basi di dati di parlato, diatopicamente e stilisticamente diversificate, come il corpus AVIP (Bertinetto, 2001), API (Crocco, Savy, Cutugno, 2002), il DIVA (Albano Leoni, Cutugno & Savy, 1995; 1998) e il CLIPS (es. Albano Leoni, 2006; 2007), sui quali si sono basati molti dei lavori che hanno preso in esame le vocali.

Le principali direttrici di studio e di ricerca seguite hanno riguardato: a) la descrizione di singoli sistemi vocalici, con il confronto tra il sistema tonico e quello atono, es. Nocchi & Calamai (2009), per il parlato toscano in tre punti linguistici; b) l'analisi contrastiva dei sistemi vocalici in chiave diatopica – per es. Clemente, Savy & Calamai (2003), in cui vengono analizzati i sistemi vocalici di sette varietà diverse, riferibili alle città di Palermo, Napoli, Lecce, Pisa, Roma, Milano, Torino; c) l'analisi stilistica delle vocali, in funzione di diversi stili di parlato (lettura di singole parole, lettura di frasi, parlato semi-spontaneo), con una particolare attenzione verso i fenomeni di riduzione e centralizzazione – es. Savy & Cutugno (1996), che si focalizza su diversi fenomeni di riduzione, strutturale e non strutturale, in differenti varietà regionali (Lombardia, Toscana, Lazio e Campania), e in diversi stili di parlato, anche se solo per la varietà campana; Lo Prejato, Clemente, Savy (2003),

² In Calamai (2003a) è presente una dettagliata rassegna dei lavori condotti fino a quel momento sulle caratteristiche acustiche dei sistemi vocalici dell'italiano.

sulla riduzione vocalica nella varietà napoletana; Calamai & Sorianello (2003), sul sistema vocalico romano; Calamai (2003b), sul sistema vocalico pisano.

Prima di entrare nel vivo dell'italiano parlato a Bolzano e, in particolar modo, del suo sistema vocalico, sarà opportuno fornire un breve quadro del sistema vocalico dell'italiano standard, in relazione ai diversi italiani regionali.

Il sistema linguistico dell'italiano standard è composto da sette fonemi vocalici (/i, e, ε, a, ɔ, o, u/), ma non tutte le varietà di italiano mostrano opposizione fonologica tra vocali medio-alte e medio-basse in sillaba tonica. Inoltre, anche laddove esiste tale distinzione, questa si caratterizza per un peso funzionale molto basso.

Sintetizzando³, i diversi sistemi vocalici dell'italiano possono essere ricondotti a tre diversi tipi (Mioni, 2001b; Vietti, 2019): a) sistemi con sette fonemi vocalici, in cui la distribuzione delle vocali medio-alte e medio-basse può anche non coincidere con quella dello standard (oltre alla gran parte dell'Italia centrale, della Campania, della Basilicata e del Molise, in questo primo tipo rientrano anche Veneto e Trentino, di particolare interesse per il nostro discorso); b) sistemi in cui sono presenti sette foni, ma solo cinque fonemi vocalici, perché la distribuzione delle vocali medie è condizionata dal contesto fonetico e quindi non crea opposizioni fonologiche (Lombardia, Piemonte, Emilia-Romagna, Friuli e, probabilmente, Liguria; anche l'italiano regionale di Sardegna mostra una situazione analoga, per effetto della metaforia); c) sistemi vocalici con cinque foni e cinque fonemi vocalici, in cui le vocali medie mostrano valori intermedi tra il punto medio-alto e quello medio-basso (Italia meridionale estrema e Venezia Giulia).

3. *L'italiano di Bolzano con particolare riguardo per le vocali*

Sebbene l'analisi che verrà presentata al par. 5 abbia carattere esplorativo e sia finalizzata quindi più all'elaborazione e alla definizione di una procedura adatta allo studio sulle vocali nei corpora di parlato spontaneo che non alla ricerca di risultati sulla varietà (o sulle varietà) oggetto di indagine, ci sembra possa essere utile al lettore fornire dapprima una descrizione delle caratteristiche fonetiche proprie delle varietà di italiano presenti a Bolzano, ovvero la varietà dei parlanti nativi e la varietà dei parlanti tedescofoni. Il profilo linguistico qui presentato metterà in luce la posizione rilevante che le vocali occupano ai fini della caratterizzazione di queste due varietà.

Il repertorio linguistico della città di Bolzano viene tipicamente definito come un bilinguismo sociale di tipo bicomunitario, in quanto sono presenti due comunità linguistiche prevalentemente monolingui (una italiana e una tirolese) (cfr. Mioni, 1990b). Tuttavia, nel corso degli ultimi decenni la proporzione di parlanti bilingui, in particolare nella comunità tedesca tirolese, è in costante aumento nel contesto ur-

³ Per un'illustrazione esaustiva dei diversi sistemi vocalici dell'italiano, si rimanda a Schmid (1999), Mioni (2001b), Vietti (2019).

bano, al punto che la pratica bilingue e la presenza di varietà di contatto rappresenta la norma (Vietti 2011).

L'italiano parlato a Bolzano è da considerare un caso piuttosto singolare nel panorama sociolinguistico italiano, per il fatto che il suo sviluppo non si configura come l'esito di un processo storico di convergenza dei dialetti verso l'italiano, analogamente a quanto è avvenuto nel caso delle altre varietà regionali di italiano, ma piuttosto come il risultato di una trasformazione repentina del contesto sociolinguistico, che ha comportato un processo di livellamento e koinizzazione⁴.

Per la sua singolarità, l'italiano bolzanino è stato, a più riprese, oggetto di attenzione da parte soprattutto di Mioni (1990a, 2001a), che ha analizzato questa varietà a partire da dati provenienti da informatori riconducibili alla seconda generazione di italofoeni nati a Bolzano (o nella provincia). Sulla base di questi dati, e grazie anche al confronto con un campione di informanti padovani, Mioni (1990a: 201) ha definito l'italiano di Bolzano come una varietà di italiano di Nord-Est più avanzato di quello veneto verso la standardizzazione, attribuendo questa maggiore standardizzazione alla particolare storia dell'immigrazione di italofoeni, avvenuta nel periodo successivo alla Prima guerra mondiale verso la città di Bolzano. Nei primi anni, infatti, nonostante la presenza di un'importante componente di dirigenti lombardi ed emiliano-romagnoli, la parte più consistente della popolazione era costituita per lo più da trentini e veneti, mentre l'arrivo degli italiani provenienti dalle regioni centro-meridionali è da ricondurre al periodo del secondo dopoguerra, ovvero solo quando ormai il modello di riferimento locale per la pronuncia dell'italiano si era già formato e consolidato.

La pronuncia prevalente sarebbe quindi quella veneto-trentina ma con un grado maggiore di standardizzazione (Mioni, 1990a; 2001a). Dal confronto tra il campione padovano e i due campioni bolzanini, uno italofono e l'altro germanofono, Mioni (1990a) individua alcuni fenomeni che contraddistinguono queste tre varietà. Considerato il fine del presente lavoro, menzioneremo solamente i tratti caratteristici delle due varietà di italiano parlate a Bolzano.

L'italiano bolzanino del gruppo di italofoeni si caratterizzerebbe per:

- l'opposizione fonologica tra le vocali medie /e/ ed /ɛ/, in riferimento però a parole diverse rispetto a quelle previste dallo standard (questo tratto accomuna l'italiano di Bolzano a quello di Padova);
- la realizzazione più standardizzata rispetto a Padova delle vocali medie posteriori /o/ e /ɔ/;
- la realizzazione dei fonemi /ʃ/ e /ʎ/ con una durata non lunga analogamente a quanto si registra per Padova, ma con una frequenza maggiore di varianti standard rispetto alla varietà di italiano padovano;

⁴ Per una dimostrazione dell'applicazione al caso bolzanino dei principi di koinizzazione e dei fattori macro-sociolinguistici così come formulati da Kerswill, Williams (2000), si veda Vietti (2017).

- una distribuzione simile a quella padovana dei fonemi /ts/ e /dz/⁵, con tracce sporadiche di realizzazioni fricative dentali [s, z], da ricondurre all'influenza di famiglie di origine trentina o emiliana.

Verso un quadro dell'italiano bolzanino come varietà più standardizzata rispetto all'italiano regionale veneto e trentino, soprattutto in riferimento alla produzione delle vocali medie, sembra convergere anche la descrizione di Canepari (1999). L'autore sostiene infatti che negli item lessicali in cui l'italiano di Bolzano diverge rispetto alle altre due varietà nord-orientali, la differenza risiederebbe in una pronuncia maggiormente standardizzata per l'italiano altoatesino: es. la parola *nebbia* realizzata come [ˈnebbja], *bosco* come [ˈbɔsko], *nome* come [ˈnɔme] (Canepari, 1999). Tuttavia, anche per l'Alto Adige, secondo Canepari (1999), i fonemi /e, ɛ/ e /o, ɔ/ sarebbero dei semifonemi, ovvero delle “acquisizioni precarie e capricciose” che mostrano grandi oscillazioni, motivo per cui possono essere solamente descritte – con riferimento al sistema fonologico standard – le differenze nella loro distribuzione lessicale dei diversi timbri (Canepari, 1999).

Passando all'italiano parlato dai tedescofoni, Mioni (1990a) individua un elenco di tratti, riconducibili quasi totalmente al contatto con il tirolese e alle differenze presenti nei due inventari, per i quali questa varietà si contraddistingue. I fenomeni in questione sono:

- la differente modalità di distinzione tra consonanti sorde e sonore;
- la durata consonantica, inesistente in tedesco;
- la durata vocalica, con la tendenza a pronunciare come lunghe tutte le vocali in posizione finale di sillaba;
- la realizzazione di /r/, che nel dialetto tirolese è prodotta come un'uvulare⁶;
- la realizzazione delle affricate /tʃ/ e /dʒ/, in cui il componente occlusivo mostra un distacco rispetto a quello fricativo;
- la velarizzazione di /l/ in fine di sillaba;
- la frequente realizzazione delle sibilanti /s, z/ come dentali o pre-dentali;
- una maggiore standardizzazione delle vocali medie in posizione tonica, sia in sillaba aperta che chiusa.

La realizzazione delle vocali medie toniche come medio-alte [e, o] in sillaba aperta e come medio-basse in sillaba chiusa [ɛ, ɔ], per effetto del contatto col tirolese, spiega il trattamento molto standardizzato di /e/ in sillaba aperta, ma non fornisce una spiegazione del rendimento altrettanto buono di queste vocali in sillaba chiusa. Dal confronto con i dati padovani, Mioni (1990a) infatti, oltre all'esito standardizzato in sillaba aperta, osserva anche una maggiore standardizzazione in sillaba chiusa tra i parlanti tedescofoni di Bolzano rispetto ai parlanti padovani.

A conclusione del suo studio, l'autore sostiene che il modello di riferimento per l'apprendimento della fonologia italiana da parte dei tedescofoni è costituito dalla fonologia degli italo-foni altoatesini.

⁵ Per un'analisi sociofonetica di questi fenomeni si rimanda a Meluzzi (2020).

⁶ Un'analisi acustica di /r/ in tirolese è illustrata in Galatà, Spreafico, Vietti, Kaland (2016).

4. Metodologia

4.1 Il campione dei parlanti e i dati

I dati presi in considerazione in questa sede derivano dal corpus DIA-Dialogic ItAlian, un corpus di parlato spontaneo sulla varietà di italiano di Bolzano, proveniente da parlanti italofofoni e tedesconi, che rappresenta la base di dati per il progetto di ricerca NEB – *Networks, Exemplars and Bilinguals. The potential of quantitative social network analysis in sociophonetic research*. Considerato che lo scopo del progetto è l'esplorazione degli effetti delle strutture delle reti sociali personali sul comportamento linguistico di parlanti bilingui, la creazione di un corpus di parlato spontaneo è risultata essere la scelta metodologica più appropriata.

Il campione consta di 40 soggetti (età: 18-65 anni; 14 M, 26 F): 23 parlanti nativi di italiano, 12 parlanti nativi di tirolese (bilingui sequenziali) e 5 bilingui simultanei. Dal punto di vista del grado di istruzione e dell'occupazione lavorativa dei parlanti, il corpus si caratterizza per essere molto eterogeneo⁷.

Per la raccolta dei dati, è stato usato un registratore Zoom H4, con il campionamento a 44.100 Hz e la digitalizzazione a 16-bit; al fine di incrementare ulteriormente la qualità, sono stati usati anche due microfoni ad archetto (Shure SM35), uno per ogni informante. Ogni sessione di registrazione ha previsto tre fasi distinte: 1) un'interazione dialogica in coppia tra persone che si conoscevano bene tra loro (coppie di amici, marito e moglie, parenti, etc.); 2) la somministrazione di un questionario volto all'elicitazione di informazioni sulle reti sociali personali, oltre che ai dati di ordine socio-demografico; 3) la lettura di una lista di frasi. Complessivamente i dati ammontano a 30h 43' 25" di parlato, suddivise nel modo seguente: 9h 49' 32" di parlato spontaneo, 19h 03' 26" di parlato proveniente dalle interviste sulle reti sociali personali e 1h 50' 27" di parlato letto.

Per l'analisi che presenteremo nei prossimi paragrafi, sono stati usati i dati delle interazioni dialogiche, che si sono basate su una lista di argomenti comuni (es. problemi sociali a Bolzano, feste tradizionali locali, situazione politica della città, etc.), somministrata ai parlanti, i quali potevano decidere di prendere spunto o meno da questa lista. In totale, i dialoghi registrati sono 20, della durata di circa 30 minuti ciascuno.

La quantità di parlato spontaneo presa in esame in questa sede è di poco più di 30 minuti. Il campione di riferimento è costituito da 8 parlanti donne, di cui 4 native di italiano e 4 native di tirolese (bilingui sequenziali), di età compresa tra i 47 e i 59 anni. Per ogni parlante sono stati analizzati 4 minuti di parlato spontaneo.

4.2 Trascrizione ortografica e annotazione fonologica

I dati sono stati dapprima trascritti ortograficamente mediante lo strumento dei sottotitoli offerto da YouTube (Cangemi, Fründt, Hanekamp & Grice, 2019) e, in

⁷ Per una descrizione dettagliata della costruzione del corpus, ci permettiamo di rinviare a Mereu, Vietti (2021).

seguito, ricontrollati manualmente e corretti dalla seconda autrice. La trascrizione ortografica ha riguardato non solo il parlato, ma anche tutti i “rumori umani” (es. risate, tosse) e le pause piene. Al momento la trascrizione è stata completata per tutti i dati di parlato spontaneo provenienti dalle interazioni dialogiche (10 ore circa, per un totale di 100.966 tokens).

I file trascritti ortograficamente, accompagnati dai rispettivi file audio, sono stati processati su WebMAUS (cfr. Kisler, Reichel & Schiel, 2017), sfruttando lo strumento dell'allineamento forzato. Dopo una serie di passaggi tramite diversi servizi offerti dalla piattaforma, il risultato ottenuto è stato, per ogni file audio, un file Text. Grid contenente tre *tier*: un *tier* ORT con la trascrizione ortografica parola per parola, un *tier* KAN, indicante la trascrizione fonologica per parola, un *tier* MAU, che riporta la trascrizione fono per fono. Anche le segmentazioni e le annotazioni sono state completamente controllate e corrette manualmente dalla seconda autrice. Oltre a questi tre *tier*, al fine di poter svolgere l'analisi sulle vocali, è stato aggiunto manualmente un ulteriore *tier* (STR), contenente l'indicazione delle vocali toniche e atone.

Per quanto riguarda la correzione apportata alla segmentazione e all'annotazione prodotta in maniera automatica mediante WebMAUS, l'intervento manuale è avvenuto in modo uniforme per tutti i dati, tenendo conto di alcuni criteri. Per la correzione delle segmentazioni, tutte le vocali sono state segmentate nel punto di ampiezza zero, facendo corrispondere l'*onset* al punto di inizio del primo ciclo completo e l'*offset* al punto coincidente con l'ultimo ciclo completo incluso. Quanto all'annotazione, le vocali sono state etichettate seguendo un'annotazione fonologica e quindi facendo riferimento a quanto previsto dallo standard italiano, che ha funzionato come punto di riferimento per lo studio della distribuzione delle vocali. Le occorrenze segmentate ammontano a 7784 (2027 toniche, 5757 atone); tuttavia, l'analisi esplorativa dei dati di cui daremo conto prende in esame solamente le vocali toniche.

I dati trascritti ortograficamente e segmentati e annotati fonologicamente sono stati trasformati in un database interrogabile, mediante il sistema EMU (Winkelmann, Harrington & Jansch, 2017), che, oltre a offrire un'interfaccia online, consente di condurre l'analisi sul corpus attraverso il linguaggio R (pacchetto EmuR).

5. *Analisi esplorativa delle vocali: problemi*

Questa sezione dell'articolo è strutturata in questo modo: prima vengono enunciati i principali aspetti problematici che si presentano quando si debba condurre un'analisi acustica di vocali su un corpus di parlato spontaneo caratterizzato da alta variazione sociale; in seguito, sono avanzate alcune delle possibili soluzioni a questi problemi sotto la forma di un'analisi statistica esplorativa di un sottoinsieme di vocali estratte dal corpus DIA.

5.1 Problemi nell'analisi acustica di corpora

L'analisi acustica di vocali (ma non soltanto) condotta su dati elicitati con procedure di tipo naturalistico e poi raccolti in un corpus pone problemi specifici che sono riconducibili solo in parte alle procedure consuete di analisi di dati acquisiti in contesti con maggiore controllo sperimentale. Volendo riassumere le caratteristiche in pochi tratti, possiamo dire che nei corpora di parlato spontaneo, rispetto a quello di laboratorio, ci si trova di fronte a dati caratterizzati da un forte sbilanciamento nella frequenza relativa dei singoli foni e a una maggiore variazione contestuale, prosodica e individuale.

Nella nostra esperienza, tre ambiti metodologici hanno richiesto una particolare riflessione critica:

- a. il grado di affidabilità della misurazione automatica dei parametri acustici;
- b. l'alta variabilità del dato spontaneo;
- c. la complessità della modellizzazione statistica.

In questo contributo, nessuno di questi tre aspetti può essere trattato se non cursoriamente; la finalità è soltanto quella di sollevare il problema, osservarne alcune delle possibilità di soluzione e tentare una prima proposta di analisi.

Il primo problema riguarda le modalità di parametrizzazione acustica delle vocali quando ci si trovi ad analizzare ampi corpora di dati. La soluzione di default, nella ricerca fonetica acustica, è quella di basarsi su algoritmi di estrazione automatica delle formanti (*formant tracking*). Questi algoritmi sono basati in varia misura sulla tecnica di predizione lineare (*Linear Predictive Coding*). In quest'ottica è necessario scegliere l'algoritmo ottimale, ovvero quello più robusto, o per lo meno, quello sul quale i risultati della ricerca in ambito computazionale sembrano convergere. Due tra le alternative più comuni tra i *formant trackers* sono il metodo Burg (Andersen, 1974) e l'algoritmo Split Levinson (Delsarte & Genin, 1986). Il primo è quello implementato di norma anche in Praat, mentre il secondo è utilizzato, per esempio, all'interno del pacchetto *wrassp* nel sistema di database emuR.

Tuttavia, quando si dispone di grandi basi di dati, che presentano grande variabilità in termini di numero di parlanti e di grado elevato di spontaneità del parlato, la strategia di appoggiarsi ad algoritmi di *formant tracking* non sembra essere la più adatta, a meno di essere disposti ad affrontare una fase di controllo manuale intensivo. Il maggiore punto di debolezza del *formant tracking* è la sua rigidità di fronte alla variabilità delle caratteristiche del parlante e del parlato. È necessario infatti operare un continuo adattamento dei parametri non soltanto ai parlanti (in ragione di età e genere), ma anche in relazione a specifiche vocali e/o a interazioni tra vocali e caratteristiche dei parlanti. Anche in questo senso si possono evidentemente ottimizzare algoritmicamente (con opportune procedure iterative) il numero dei coefficienti dei polinomi di predizione lineare così come il *formant ceiling*⁸, tuttavia questa ope-

⁸ Il valore massimo al di sotto del quale cercare di identificare le formanti (Maximum formant in Praat); questo valore è legato alla lunghezza del tratto vocale e perciò può variare da individuo a individuo, ma in media per voci maschili è di 5000 Hz, mentre per voci femminili è di 5500 Hz.

razione risulta onerosa al crescere dei parlanti e in genere della variabilità del parlato (soltanto a titolo di esempio si possono segnalare le osservazioni metodologiche contenute in Escudero, Boersma, Rauber & Bion, 2009).

L'alta proporzione di errore operata dal *formant tracking* automatico eseguito su corpora si manifesta principalmente nella presenza consistente di *outlier*, ovvero misurazioni di F1 e/o F2 non solo distanti dalle tendenze centrali, ma anche chiaramente poco realistici rispetto alla natura del dato. In un contesto sperimentale, è consuetudine la revisione manuale dei valori formantici eseguiti automaticamente poiché la numerosità del dataset lo consente, mentre quando si opera in un contesto di *corpus phonetics* è necessario riflettere su possibili procedure automatiche alternative all'estrazione di formanti, che svincolino il ricercatore dall'attività di correzione di *outliers* (Harrington, 2010: 215).

La seconda caratteristica di questo tipo di analisi è connessa con la variabilità nella realizzazione delle diverse categorie fonetiche. Un corpus di parlato spontaneo di tipo dialogico comporta infatti un elevato grado di variabilità acustica che pone seri problemi all'individuazione di categorie definite e separabili. Volendo schematizzare, le possibili fonti di variazione presenti nel dato spontaneo (segmentale) si possono articolare secondo queste macro-categorie (Harrington 2010):

- i contesti fonetici e fonologici, e più in generale la variabilità posizionale e strutturale;
- le posizioni prosodiche a cui possono corrispondere le diverse funzioni sintattico-pragmatiche;
- gli stili di elocuzione, che a loro volta determinano i gradi di connessione del parlato e i fenomeni di riduzione;
- le caratteristiche sociali degli individui, per esempio in questo caso specifico la distinzione tra parlanti nativi e non nativi;
- le differenze di tipo biologico tra i parlanti, in primo luogo genere ed età.

Non c'è evidentemente qui spazio per riassumere neppure sinteticamente le diverse fonti di variabilità presenti nel dato acustico (per una schematizzazione v. Vietti, in stampa).

Il nocciolo della questione può essere riassunto così: se in un contesto sperimentale le fonti di variabilità possono essere controllate e manipolate in modo selettivo, nel parlato elicitato in contesti naturalistici queste si presentano tutte congiuntamente. Il compito è perciò quello di trovare il modo di organizzare questa variazione attraverso la scomposizione dello spazio di variazione in dimensioni, con lo scopo di meglio definire gli oggetti-categorie. Più in generale, l'opportunità che lo studio di corpora offre all'analisi acustica è quella di poter osservare il modo in cui la variazione contribuisce alla strutturazione di categorie fonetiche e fonologiche.

La necessità di rendere conto dell'alta variazione dei dati di corpora si traduce sul piano metodologico nella ricerca di modelli statistici adeguati. Un caso particolare in questa direzione è rappresentato dall'introduzione (e impiego sempre più esteso) di modelli lineari a effetti misti, che ha comportato un significativo avanzamento delle nostre possibilità di controllare, quasi in una logica sperimentale, quegli effetti

casuali legati alla ripetizione delle misurazioni su medesimi soggetti e alle possibili dipendenze interne tra osservazioni. In questo modo è ora possibile analizzare le relazioni tra variabili che presentano ripetizioni e in generale una struttura gerarchica o multilivello, tenendo sotto controllo alcune fonti di variabilità che altrimenti eserciterebbero un'indebita influenza sul modello finale.

Oltre a questa specifica innovazione, nel corso degli ultimi decenni le possibilità di analisi quantitativa delle vocali si è espansa enormemente, di pari passo con la necessità di formulare ipotesi via via più complesse sulla natura dinamica delle vocali e sull'interdipendenza di più variabili. Volendo di nuovo un po' semplificare il problema in dicotomie, potremmo distinguere tra (a) analisi statiche e dinamiche, (b) su singoli parametri o su aggregati di parametri, (c) su una singola vocale o su interi sistemi vocalici (a titolo puramente esemplificativo si v. Adank et al., 2004; Calamai, 2004; Fox & Jacewicz, 2017; Gahl & Baayen, 2019).

Considerata la natura esplorativa di questo contributo, non vi è spazio qui per una discussione approfondita delle possibilità di analisi e modellizzazione delle vocali (e dei sistemi vocalici), aspetti che evidentemente non sono soltanto di metodo, ma sono implicati dalle prospettive teoriche che si adottano. A fronte di queste limitazioni di intenti, riteniamo sia bene avere presente il contesto delle possibilità e dei problemi metodologici specifici che l'analisi del parlato spontaneo di un corpus pone, prima di accingersi all'analisi.

È utile anche ricordare il fine ultimo a cui si tenderà in questa e in future indagini, ovvero quello di individuare quale analisi statistica ci permetta di valutare le differenze tra sistemi vocalici in varietà di lingua diverse, o meglio ancora in una logica induttiva, quale metodo ci consenta di individuare delle varietà di lingua a partire dalla distribuzione dei valori acustici dei fonemi in un sistema vocalico in relazione a fattori esterni relativi ai parlanti e alle loro caratteristiche.

5.2 Analisi spettrale

Come è stato già anticipato al par. 4.1, l'analisi che presentiamo qui si basa su un primo sottoinsieme del corpus DIA. Il dataset analizzato contiene 2027 occorrenze vocaliche di 8 parlanti donne, 4 parlanti native dell'italiano e 4 non native con tedesco come prima lingua (in realtà si tratta del tirolese, un dialetto tedesco). Le occorrenze estratte dal corpus riguardano le 7 vocali toniche dell'italiano, ovvero quelle vocali che compaiono in sillaba accentata di parola.

Per ognuno dei tre ambiti metodologici delineati sopra abbiamo sviluppato una possibile soluzione:

- a. per la misurazione automatica dei parametri mostriamo un confronto qualitativo tra analisi per formanti e analisi spettrale basata su un insieme di coefficienti coseno discreti (*DCT coefficients*), poi ulteriormente ridotti attraverso analisi di componenti principali;
- b. per la variabilità del dato abbiamo tenuto in considerazione la variazione contestuale e individuale, trattando questi fattori come effetti casuali in un modello lineare a effetti misti;

- c. per la modellizzazione statistica abbiamo invece provato due modelli statistici relativamente elementari basati su un'osservazione statica nel tempo delle vocali (estrazione dei parametri spettrali in un punto medio) con i singoli parametri acustici come variabili dipendenti e le vocali e i fattori extralinguistici come variabili esplicative.

Il dataset analizzato ($n = 2027$) è il risultato di una prima selezione operata eliminando gli *outlier* di durata ($n = 86$), ovvero escludendo i tokens vocalici con valori inferiori a 20 ms e superiori a 1.5 volte la distanza interquartile (in questo caso > 234 ms).

Il nostro punto di partenza, in modo da avere un riferimento nell'interpretazione dei dati, è dato dall'osservazione dei valori delle prime due formanti, estratti secondo l'algoritmo di *formant tracking* di Split-Levinson per come implementato dalla funzione *forest* in emuR ($n = 1726$). Nelle Figure 1 e 2 sono rappresentate le due possibili varietà di italiano.

Figura 1 - F1 e F2 di parlanti native (ellissi di confidenza 95%)

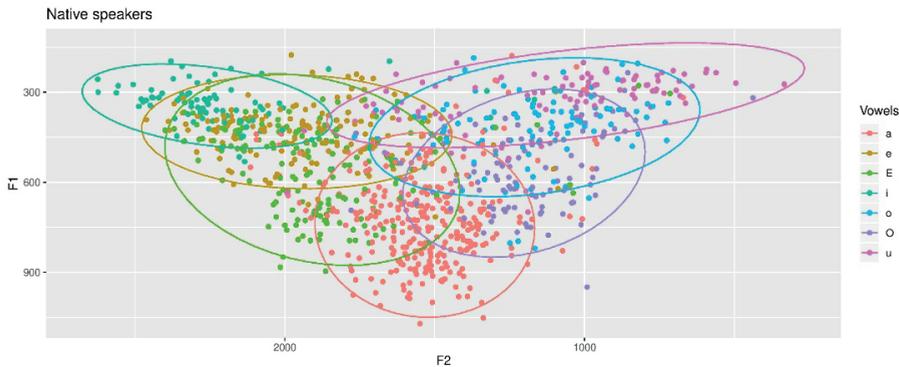
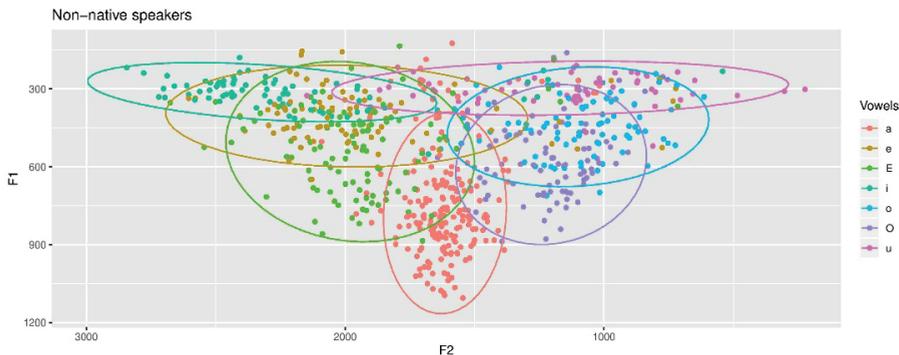


Figura 2 - F1 e F2 di parlanti tedescofone non native (ellissi di confidenza 95%)



Dai grafici si può innanzitutto osservare come, in entrambe le varietà, vi sia una notevole sovrapposizione tra le categorie, caratteristica tipica del parlato spontaneo.

Sebbene raramente univocamente individuabili, degli addensamenti attorno ai valori di tendenza centrale (moda e media) sono ben riconoscibili. Passando a una superficiale osservazione delle differenze tra le due varietà, possiamo notare in primo luogo la differenza nel *range* tra i valori massimi e minimi di F1 e F2 che porta a uno spazio vocalico complessivo più ampio nel caso della varietà di parlanti non nativi. In secondo luogo, si può constatare come la natura della sovrapposizione tra categorie muti in termini di forma e varianza delle ellissi tra le due varietà (si osservi p.e. il caso della /a/ nelle due varietà).

Nell'estrazione dei valori con *formant tracking* emerge un alto numero di occorrenze che presentano un errore nella stima della formante. L'errore più comune è l'assegnazione di un valore zero a F2 (n = 301 tokens), in particolare nel caso di /i/.

L'analisi alternativa che abbiamo deciso di sperimentare è invece un'analisi spettrale che non preveda individuazione e misurazione delle formanti, ma che consideri piuttosto l'informazione spettrale nel suo complesso. L'analisi spettrale ha come obiettivo l'estrazione dei coefficienti della trasformata discreta del coseno (*Discrete Cosine Transform*) nel punto medio della durata del segmento (Zahorian & Jagharghi, 1993). Tutte le operazioni sono state condotte in R utilizzando per lo più le funzioni di analisi del segnale disponibili nel pacchetto EmuR (Winkelmann, n.d.).

La procedura per calcolare i coefficienti è stata la seguente (Harrington 2010: 206-215):

- a. calcolo degli spettri di potenza con trasformata di Fourier veloce (FFT);
- b. trasformazione degli spettri nella scala Bark;
- c. calcolo della seconda trasformata di Fourier eseguita sugli spettri di potenza (*Discrete Cosine Transform*);
- d. estrazione da (c) dei coefficienti DCT che descrivono in modo sintetico le frequenze di risonanza.

L'esito di questo processo è analogo a quello che si applica per ottenere i coefficienti cespstrali trasformati in Mel (*Mel-Frequency Cepstral Coefficients*, Wakita, 1976); la principale differenza risiede nel fatto che la seconda analisi di Fourier viene eseguita usando unicamente funzioni coseno come basi invece che anche funzioni seno. In sintesi, il vantaggio di tale scelta risiede nel fatto che servono meno basi per ricostruire il segnale e, soprattutto, l'analisi può essere condotta utilizzando soltanto numeri reali (e non anche numeri complessi come per la trasformata di Fourier con funzioni seno).

La scomposizione cespstrale del segnale dovrebbe produrre coefficienti indipendenti, tuttavia in Iskarous (2019) si è registrata una certa ridondanza nella codifica delle informazioni relative alle risonanze per i tratti articolatori di altezza e avanzamento delle vocali. Per questo motivo, si è deciso di esaminare l'eventuale presenza di dipendenze tra i coefficienti DCT e la loro riduzione a un numero inferiore di dimensioni attraverso l'analisi delle componenti principali (PCA). L'analisi è stata condotta usando i primi 6 coefficienti DCT, in quanto da un lato sono apparsi sufficienti per ricostruire le caratteristiche salienti dello spettro originale e dall'altro in questo modo la proporzione di varianza cumulativa delle prime tre componenti ri-

sultava ottimale. Le prime due dimensioni (Dim1 = 33.2%, Dim2 = 29.4%) rappresentano infatti il 62.6% della percentuale cumulativa di varianza, mentre sommando le prime tre componenti si raggiunge l'80% (Dim3 = 17.5%). Prima di condurre l'analisi i dati sono stati ricampionati a una frequenza di 16.000 Herz ritenuta più adeguata all'analisi delle vocali.

L'esito di questa analisi è riportato graficamente nelle due figure 3 e 4.

Il risultato che si ottiene dalla PCA è una ridefinizione dei valori dei coefficienti in uno spazio geometrico di dimensioni ridotte, basato sulla struttura delle correlazioni tra le variabili di partenza (in questo caso i sei coefficienti). Si tratta perciò di uno spazio astratto, le cui unità di misura non sono più unità fisiche, ma correlazioni tra variabili. Ciò nondimeno le relazioni tra gli oggetti dovrebbero far emergere un'organizzazione basata sui parametri articolatori.

Da un esame visivo dei grafici risulta infatti che, sebbene ruotata di 100 gradi in senso antiorario rispetto alla consueta rappresentazione, la disposizione dei punti e delle categorie vocaliche appare sensata, ovvero regolata dai parametri articolatori di innalzamento/abbassamento e avanzamento/arretramento. In particolare, la Dimensione 1 (quella che cattura la maggior parte della varianza del nostro dataset) approssima l'altezza del corpo della lingua, mentre la Dimensione 2 coglie in parte la distinzione tra vocali anteriori e posteriori.

Figura 3 - PCA dei DCT coefficients parlanti nativi

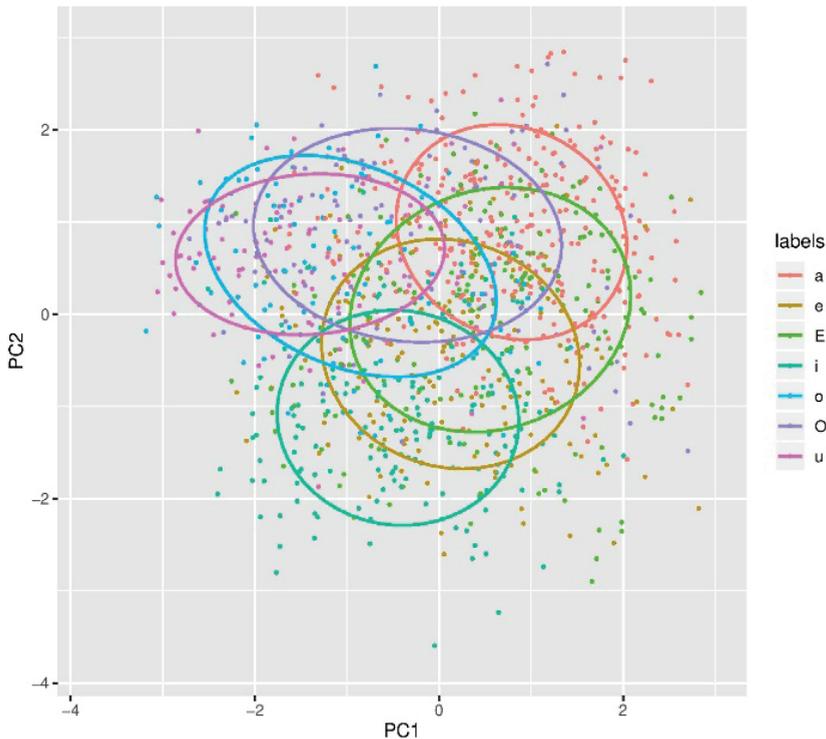
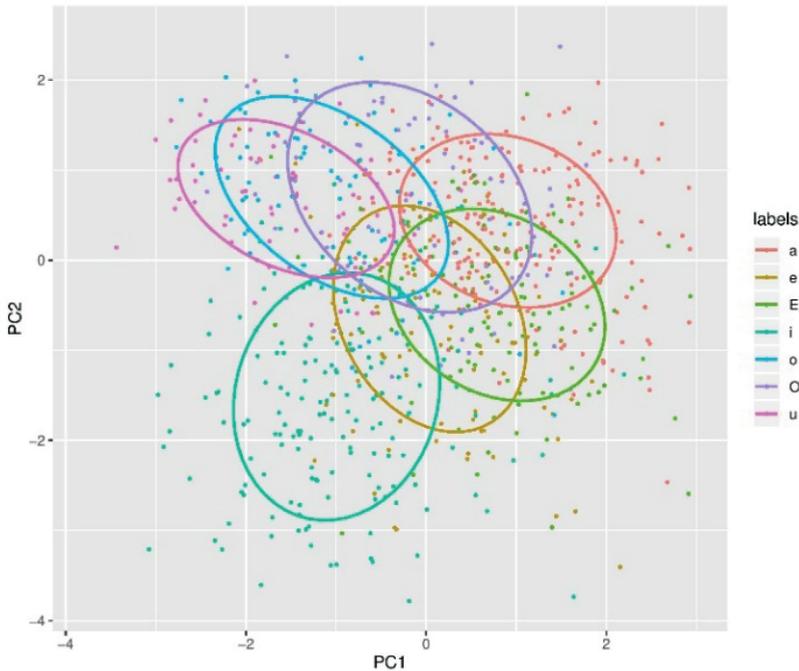


Figura 4 - PCA dei DCT coefficients parlanti tedesco non nativi



In secondo luogo, si può osservare come anche in questo caso l'*overlap* fra le categorie appaia notevole e comunque maggiore rispetto alla rappresentazione per formanti. Tuttavia, sul piano qualitativo possiamo registrare come la tendenza a una migliore discriminazione tra le vocali nella varietà di non nativi sia confermata anche dall'analisi PCA. Come abbiamo già rilevato più in generale, l'alta sovrapposibilità delle categorie non è imputabile né alla precisione della misurazione, né al trattamento quantitativo dei dati; piuttosto è da ricercare nelle condizioni di produzione del parlato che, all'aumentare della spontaneità nella comunicazione, vede diminuire, nel complesso, la definizione delle caratteristiche acustiche. Al piano della distribuzione statistica si può invece senz'altro attribuire la rappresentazione geometrica delle ellissi di dispersione, in quanto come sappiamo, forme e dimensioni (a) sono fortemente influenzate dalla presenza anche di pochi *outlier* e (b) poggiano su un assunto di gaussianità nella distribuzione bivariata tendenzialmente violato (Savy, Clemente, Lo Prejato, 2005).

5.3 Analisi statistica

L'ultima fase dell'analisi prevede l'applicazione di modelli lineari a effetti misti con lo scopo di testare in via preliminare la nostra ipotesi sulla presenza di una differenza tra i due sistemi vocalici di nativi e non nativi, tenendo sotto controllo alcune delle maggiori fonti di variabilità. Con questo obiettivo, abbiamo condotto due analisi parallele utilizzando le due prime dimensioni dell'analisi delle componenti princi-

pali come variabili dipendenti, mentre le categorie vocaliche, il retroterra linguistico dei parlanti e l'interazione tra queste due variabili sono state considerate come variabili esplicative ed effetti fissi nel modello. Sono state poi inserite tre variabili come effetti casuali: il contesto segmentale sinistro, il contesto segmentale destro⁹ e il parlante. Con le prime due variabili si cerca naturalmente di catturare l'aspetto più macroscopico della variazione contestuale connessa con la coarticolazione, e non certo controllare in modo accurato la variazione posizionale e strutturale.

Per il modello relativo alla Dimensione 1 (innalzamento/abbassamento) la differenza tra le diverse vocali è significativa ($F[6, 223.5]$, $p < 0.001$), mentre il retroterra linguistico (italofono vs. tedescofono) non è significativo in quanto effetto principale ($F[1, 0.01]$, $p = 0.95$). Tuttavia, come era prevedibile, c'è una forte interdipendenza tra le varietà di lingua e le vocali ($F[6, 7.44]$, $p < 0.001$). Questo implica che il background linguistico dei parlanti nativi e non nativi non determina due sistemi vocalici significativamente diversi per grado di apertura vocalica, ma sembra piuttosto che vi siano delle interazioni tra vocali all'interno delle due varietà di lingua.

Per la Dimensione 2 i risultati sono analoghi: le vocali sono significativamente distinte per avanzamento/arretramento ($F[6, 238.6]$, $p < 0.001$), il background linguistico dei parlanti non è significativo ($F[1, 1.2]$, $p = 0.92$), mentre vi sono significative interazioni tra le due variabili ($F[6, 7.6]$, $p < 0.001$).

Osservando la componente *random* dei due modelli, è interessante notare come i contesti segmentali sinistro e destro siano caratterizzati da una bassa varianza (attorno a 0.03), mentre la variazione intersoggettiva sia in genere più alta e maggiore nel caso dell'altezza vocalica (Dim. 1, $s^2 = 0.19$) rispetto all'avanzamento (Dim. 2, $s^2 = 0.08$). Questo significa che ci sono maggiori differenze tra parlanti in relazione al parametro articolatorio dell'innalzamento/abbassamento della lingua.

Per comprendere meglio la struttura delle interazioni tra vocali e varietà di lingua è necessario condurre un'analisi post-hoc su tutte le coppie possibili di valori. In sintesi, quest'analisi conferma l'assenza di un effetto principale del background linguistico e mette invece in evidenza come le due varietà siano organizzate in modo diverso al proprio interno lungo le due dimensioni articolatorie. Rispetto al primo aspetto, tutti i test che confrontano la stessa vocale nelle due varietà sono risultati non significativi (p.e. /a/ di nativi e /a/ di non nativi non sono statisticamente diverse). Considerata l'interazione tra vocali e varietà, ci possiamo aspettare che le vocali dei nativi e dei non nativi siano distinte in modo diverso per grado di altezza e avanzamento.

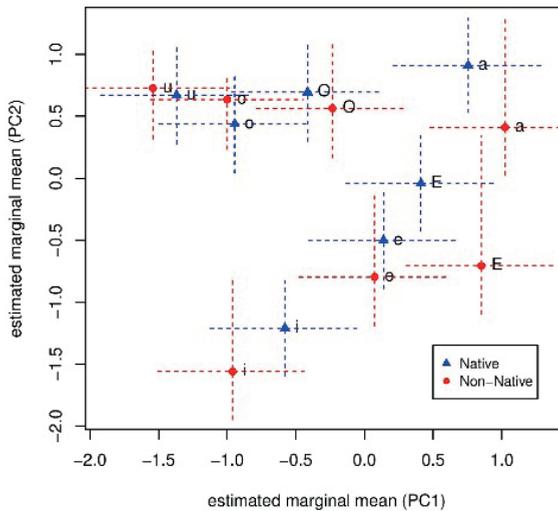
⁹ È stato preso in considerazione il fono precedente e quello seguente alla vocale.

Questo è infatti quanto emerge dall'analisi ed è riassunto nella Tabella 1.

Tabella 1 - *Vocali non distinte statisticamente per Dimensione e Varietà di lingua*

		Suoni non distinti statisticamente
Dimensione 1 (altezza)	Nativi	(u, o), (i, o), (i, ə), (e, ε)
	Non nativi	(i, o), (e, ə), (ε, a)
Dimensione 2 (avanzamento)	Nativi	(u, ə, a)
	Non nativi	(u, o, ə, a), (e, ε)

Figura 5 - *Grafico delle medie stimate dai modelli con relativi intervalli di confidenza*



Per cogliere in modo più chiaro le caratteristiche dei due sistemi vocalici ipotizzati, i valori medi dei diversi parametri così come stimati dai due modelli lineari, insieme ai relativi intervalli di confidenza, sono espressi in forma grafica nella Figura 5.

Il grafico conferma visivamente l'analisi statistica condotta: da un lato, se guardiamo alle posizioni delle diverse categorie (p.e. /o/ nativa vs. /o/ non nativa) appare evidente che non esistono due distinti sistemi vocalici; dall'altro lato, se si osservano le relazioni tra fonemi all'interno dei due sistemi, le configurazioni relative che ne risultano sono ben distinte. Appaiono qui di particolare interesse (a) i diversi rapporti tra le vocali medie nelle due varietà, anche alla luce degli studi condotti da Mioni, e (b) le (possibili) diverse specificazioni articolatorie di /i/ e /a/.

6. Osservazioni conclusive

Dall'analisi statistica emergono infine preziosi elementi di riflessione, sia di metodo che di contenuto, che verranno presi in considerazione in analisi successive. Tra questi vorremmo qui soltanto menzionarne rapidamente tre:

- a. da una rapida osservazione, la distribuzione delle due variabili continue appare decisamente non gaussiana, problema questo che pone in primo luogo seri dubbi sulla significatività dei risultati e in secondo luogo sulla reale struttura interna delle categorie;
- b. considerato il numero di occorrenze non elevato, anche la struttura elementare del modello può correre il rischio di eccedere nella specificazione causando problemi di *overfitting*, convergenza per singolarità e più in generale perdere in capacità predittiva;
- c. lo scarso ruolo della variabilità contestuale, intesa qui come effetto random, è inatteso, perciò andrebbero esaminati con maggiore attenzione i possibili effetti locali di singoli foni precedenti o seguenti in interazione con specifiche vocali;
- d. la variabile legata al retroterra dei parlanti potrebbe cogliere in modo molto approssimativo le differenze nella competenza dei parlanti (in futuro, considerato il contesto di bilinguismo sociale, sarebbe più opportuno utilizzare altre misure di dominanza nell'uso linguistico, come le reti sociali).

Riferimenti bibliografici

ADANK, P., VAN HOUT, R. & SMITS, R. (2004). An acoustic description of the vowels of Northern and Southern Standard Dutch. In *The Journal of the Acoustical Society of America*, 116(3), 1729-1738.

ALBANO LEONI, F. (2006). Il corpus CLIPS: presentazione del progetto. Documento scaricabile al sito <http://www.clips.unina.it> (22.05.2020).

ALBANO LEONI, F. (2007). Un frammento di storia recente della ricerca (linguistica) italiana. Il corpus CLIPS. In *Bollettino d'Italianistica*, 4, 122-130.

ALBANO LEONI, F., CUTUGNO, F. & SAVY, R. (1995). The Vowel System of Italian Connected Speech. In ELENIUS, K. & BRANDERUD, P. (Eds.), *Proceedings of XIIIth International Congress of Phonetic Sciences*, Stockholm 13-19, VIII, 1995, Stockholm, IV, 396-399.

ALBANO LEONI, F., CUTUGNO, F. & SAVY, R. (1998). Il vocalismo dell'italiano televisivo. Analisi acustica di un corpus. In RUFFINO, G. (Ed.), *Atti del XXI Congresso Internazionale di Linguistica e Filologia Romanza*, Centro di studi filologici e linguistici siciliani – Università di Palermo 18-24.IX.1995. Tübingen: Niemeyer, IV: 3-16.

ANDERSEN, N. (1974). On the calculation of filter coefficients for maximum entropy spectral analysis. In *Geophysics*, 39(1), 69-72.

BERTINETTO, P.M. (Ed.) (2001). AVIP – *Archivio di Varietà di Italiano Parlato*. 4 cd-rom. Ufficio Pubblicazioni della Classe di Lettere della Scuola Normale Superiore: Pisa.

- CALAMAI, S. (2003a). Vocali d'Italia. Una prima rassegna. In COSÌ, P., MAGNO CALDOGNETTO, E. & ZAMBONI, A. (Eds.), *Voce canto parlato. Scritti in onore di Franco Ferrero*. Padova: Unipress.
- CALAMAI, S. (2003b). The Pisan Vowel System of Read and Semi-spontaneous Speech. An Exploratory Contribution. In *Quaderni del Laboratorio di Linguistica della Scuola Normale Superiore*, 3, 72-98.
- CALAMAI, S. (2004). *Il vocalismo tonico pisano e livornese. Aspetti storici, percettivi e acustici*. Alessandria: Edizioni dell'Orso.
- CALAMAI, S., SORIANELLO, P. (2003). Aspetti stilistici del vocalismo romano. In *Quaderni del Laboratorio di Linguistica della Scuola Normale Superiore di Pisa*, 4, 27-41.
- CANEPARI, L. (1999). *Manuale di pronuncia italiana*. Bologna: Zanichelli.
- CANGEMI, F., FRÜNDT, J., HANEKAMP, H., GRICE, M. (2019). A semi-automatic workflow for orthographic transcription and syllabic segmentation. In PICCARDI, D., ARDOLINO, F. & CALAMAI, S. (Eds.), *Gli archivi sonori al crocevia tra scienze fonetiche, informatica umanistica e patrimonio digitale. Audio archives at the crossroads of speech sciences, digital humanities and digital heritage*. Milano: Officinaventuno, 419-425.
- CROCCO, C., SAVY, R., CUTUGNO, F. (Eds.) (2002). *API – Archivio del Parlato Italiano*. Dvd. Napoli: Università degli Studi di Napoli "Federico II".
- DELSARTE, P., GENIN, Y. (1986). The split Levinson algorithm. In *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 34(3), 470-478.
- ESCUDERO, P., BOERSMA, P., RAUBER, A.S., & BION, R.A.H. (2009). A cross-dialect acoustic description of vowels: Brazilian and European Portuguese. In *The Journal of the Acoustical Society of America*, 126(3), 1379-1393.
- FERRERO, F.E. (1968). Diagrammi di esistenza delle vocali italiane. In *Alta Frequenza*, 37, 54-58.
- FERRERO, F.E. (1972). Caratteristiche acustiche dei fonemi vocalici italiani. In *Parole e metodi*, 3, 9-31.
- FOX, R.A., JACEWICZ, E. (2017). Reconceptualizing the vowel space in analyzing regional dialect variation and sound change in American English. In *The Journal of the Acoustical Society of America*, 142(1), 444-459.
- GAHL, S., BAAYEN, R.H. (2019). Twenty-eight years of vowels: Tracking phonetic variation through young to middle age adulthood. In *Journal of Phonetics*, 74, 42-54.
- GALATÀ, V., SPREAFICO, L., VIETTI, A. & KALAND, C. (2016). An acoustic analysis of /r/ in Tyrolean. In *INTERSPEECH 2016*, 1002-1006.
- HARRINGTON, J. (2010). *Phonetic analysis of speech corpora*. Malden, MA: Wiley-Blackwell.
- KERSWILL, P., WILLIAMS, A. (2000). Creating a new town koine: children and language change in Milton Keynes. In *Language in Society*, 29, 65-115.
- KISLER, T., REICHEL, U. & SCHIEL, F. (2017). Multilingual processing of speech via web services. In *Computer Speech & Language*, 45, 326-347.
- MELUZZI, C. (2020). *Sociofonetica di una varietà di koinè. Le affricate dentali nell'italiano di Bolzano*. Milano: FrancoAngeli.

- MENDICINO A., ROMITO L., (1991). 'Isocronia' e 'base di articolazione': uno studio su alcune varietà meridionali. In *Quaderni del Dipartimento di Linguistica dell'Università della Calabria, Serie Linguistica*, 3, 49-67.
- MEREU, D., VIETTI, A. (2021). Dialogic ItAlian: the creation of a corpus of Italian spontaneous speech. In *Speech Communication*, 130, 1-14.
- MIONI, A.M. (1990a). La standardizzazione fonetico-fonologica a Padova e Bolzano (stile di lettura). In CORTELAZZO, M., MIONI, A.M. (Eds.), *L'italiano regionale. Atti del XVIII Congresso della SLI*. Roma: Bulzoni, 193-208.
- MIONI, A.M. (1990b). Bilinguismo intra- e intercomunitario in Alto Adige/Südtirol: considerazioni sociolinguistiche. In LANTHALER, F. (Ed.), *Mebr als eine Sprache. Zu einer Sprachstrategie in Südtirol – Più di una lingua. Per un progetto linguistico in Alto Adige*. Merano: Alpha&Beta, 13-36.
- MIONI, A.M. (1994). Vivere senza dialetto?. In LANTHALER, F. (Ed.) *Dialekt und Mehrsprachigkeit / Dialetto e plurilinguismo*. Merano/Meran: Alpha&Beta, 27-47.
- MIONI, A.M. (2001a). L'italiano nelle tre comunità linguistiche tirolesi (con particolare riguardo per la pronuncia). In EGGER, K., LANTHALER, F. (Eds.), *Die deutsche Sprache in Südtirol. Einheitsprache und regionale Vielfalt*. Wien: Folio, 65-76.
- MIONI, A.M. (2001b). *Elementi di fonetica*. Padova: Unipress.
- NOCCHI N., CALAMAI S. (2009). Durata e strutture formantiche nel parlato toscano: Indagini preliminari su un campione di dialoghi semispontanei. In SCHMID, S., SCHWARZENBACH, M. & STUDER, D. (Eds.), *Atti del Convegno AISV 2009 La dimensione temporale del parlato, Zurigo 4-6.II.2009*. Torriana: EDK, 195-223.
- ROMITO L., LORENZI M. (1997). Considerazioni generali sul comportamento di alcune varietà dialettali meridionali e settentrionali rispetto all'accento intensivo. In *Quaderni del Dipartimento di Linguistica dell'Università della Calabria. Serie Linguistica*, 6, 11-34.
- SAVY, R., CLEMENTE, G. & LO PREJATO, M. (2005). Per una caratterizzazione e una misura della riduzione vocalica in italiano. In COSI, P. (Ed.), *Misura dei parametri – aspetti tecnologici e implicazioni nei modelli linguistici, Atti del 1° Convegno Nazionale dell'Associazione Italiana di Scienze della Voce, Padova, 2-4 dicembre 2004*. Torriana (RN): EDK Editore, 135-160.
- SCHMID, S. (1999). *Fonetica e fonologia dell'italiano*. Torino: Paravia Scriptorium.
- TRUMPER, L., ROMITO, L. (Eds.) (1993). *Teoria e sperimentazione: parametri, tratti e segmento. Atti delle II Giornate di Studio del GFS, Calabria, 28-29.XI.1991*. Roma: Esagrafica.
- VIETTI, A. (2011). Comunità alloglotte: tedesca. In SIMONE, R. (Ed.), *Enciclopedia dell'italiano*. Roma: Istituto Enciclopedia Italiana Treccani, 482-485.
- VIETTI, A. (2017). Italian in Bozen/Bolzano: the formation of a "new dialect". In: CERRUTI, M., CROCCO, C. & MARZO, S. (Eds.), *Towards a new standard: theoretical and empirical studies on the restandardization of Italian*. Berlin: De Gruyter, 176-212.
- VIETTI, A. (2019). Phonological variation and change in Italian. In *Oxford Research Encyclopedia of Linguistics*. Oxford: Oxford University Press.
- VIETTI, A. (in stampa). Il ruolo della variabilità acustica nella costruzione del dato linguistico. In BERNINI, G. (Ed.), *Superare l'evanescenza del parlato*.

- WAKITA, H. (1976). Instrumentation for the Study of Speech Acoustics. In LASS, N.J. (Ed.), *Contemporary Issues in Experimental Phonetics*. Academic Press, 3-40.
- WINKELMANN, R. (n.d.). The EMU-SDMS Manual. Retrieved May 24, 2020, from <https://ips-lmu.github.io/The-EMU-SDMS-Manual/>
- WINKELMANN, R., HARRINGTON, J. & JÄNSCH, K. (2017). "EMU-SDMS: Advanced speech database management and analysis in R". In *Computer Speech & Language*, 45, 392-410.
- ZAHORIAN, S.A., JAGHARGHI, A.J. (1993). Spectral-shape features versus formants as acoustic correlates for vowels. In *The Journal of the Acoustical Society of America*, 94(4), 1966-1982.