

Bollettino del Laboratorio di

FONETICA SPERIMENTALE

«Arturo Genre»

dell'Università di Torino



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

Pubblicazione semestrale

COMITATO SCIENTIFICO

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

Direttore scientifico: ANTONIO ROMANO

COMITATO EDITORIALE

- | | |
|---|---|
| VALENTINA COLONNA – Dip. di Linguistica Gen. e Teoria della Lett. – Università di Granada | DANIELA MEREU – Dipartimento di Lingue e L.S. e C.M. – Università di Torino |
| VALENTINA DE IACOVO – Dipartimento di Lingue e L.S. e C.M. – Università di Torino | ANTONIO ROMANO – Dipartimento di Lingue e L.S. e C.M. – Università di Torino |
| BIANCA MARIA DE PAOLIS – Laboratorio di Fonetica Sperimentale «Arturo Genre» | MAURO UBERTI – Comitato scientifico Laboratorio di Fonetica Sperimentale «Arturo Genre» |

Direttore responsabile: MAURO UBERTI

DIREZIONE E REDAZIONE

Laboratorio di Fonetica Sperimentale «Arturo Genre» dell'Università di Torino
Via Sant'Ottavio n. 20, 10124 Torino - E-MAIL: lfsag.unito@gmail.com
<http://www.lfsag.unito.it/ricerca/phonews/index.html>
Registrazione del Tribunale Ordinario di Torino n° 33 del 22 maggio 2018
Stampato in proprio.

Immagine di copertina: registratori a bobine (multitraccia Revox e magnetofoni Castelli S4000R e LESA Renas C2) in dotazione al LFSAG (foto a cura di Federico Lo Iacono).

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

N. 12 – Dicembre 2023

SOMMARIO

ANTONIO ROMANO,	Discreto e continuo nelle scienze del parlato Presentazione del n. 12 del Bollettino LFSAG	1
CRISTIANA BRIGNONE,	First results of a lexical analysis and a prosodic parsing of English and Spanish speech in weather forecast news . . .	7
ANTONIO ROMANO,	Differenze fonetiche generali tra voci femminili e voci maschili	33
ANTONIO ROMANO,	Contributo acustico preliminare alla descrizione fonetica della debuccalizzazione camuna della <i>s</i>	53
PHONEWS	Un programma di eventi dedicati agli studi sulla voce (a cura di ANTONIO ROMANO, BIANCA DE PAOLIS e VALENTINA COLONNA)	67
	– Summer school «Acoustic Analysis of Speech and Voice» (a cura di ANTONIO ROMANO, VALENTINA DE IACOVO e MARCO GAMBÀ)	67
	– <i>Singing voice analysis and experimental approaches</i>	68
	– <i>Acoustic analysis and benefits for phonetics and phonology</i>	69
	– <i>Speech technologies and scripting</i>	72
	– <i>Introduction to statistical analysis of acoustic measures of speech</i> .	71
	– <i>Phonetic parameters for L2 learner speech and for dysarthric speech</i>	72
	– Rassegna degli eventi salienti a cui ha partecipato il LFSAG (a cura di B. DE PAOLIS)	73
	– Il progetto «Voices of Spanish Poets» (di VALENTINA COLONNA)	77
Norme editoriali	79

Discreto e continuo nelle scienze del parlato

Presentazione del Bollettino LFSAG n. 12

Antonio Romano

Le note che riassumo in questa presentazione si pongono opportunamente sulla scia di considerazioni già proposte nel n. 11 del Bollettino LFSAG, quasi a conferma di alcune piste di riflessione intuite nei lavori di quel numero.

Nel corso della Scuola estiva «Acoustic analysis of Speech and Voice» (organizzata in settembre dal sottoscritto, Valentina De Iacovo e Marco Gamba nel quadro dell'*EAA Summer School for Young Acousticians* associata alla *Forum Acusticum Conference 2023* al Politecnico di Torino) si sono infatti susseguiti interventi che a più riprese hanno portato a riflettere sui cambiamenti che algoritmi e ambienti 'intelligenti' di elaborazione delle informazioni stanno producendo nel campo delle scienze del parlato.

Tra le attività del laboratorio del II semestre 2023, oltre a una massiccia partecipazione all'*ICPhS* e all'organizzazione della suddetta scuola estiva (di cui daremo maggiormente conto nella sezione *PhoneWS*), rammentiamo anche il contributo all'organizzazione del *LVI Congresso internazionale della Società di Linguistica Italiana* (Torino, 14-16 settembre 2023) dal titolo «Continuo e discreto nelle scienze del linguaggio».

Ed è proprio questo il tema centrale e trasversale per gli articoli selezionati per questo numero che si apre con Cristiana Brignone «First results of a lexical analysis and a prosodic parsing of English and Spanish speech in weather forecast news». In quest'articolo l'autrice, neolaureata in Lingue e Letterature Straniere dell'Università di Torino con una Tesi preparata presso il LFSAG, propone un'analisi linguistica di una selezione di trasmissioni televisive dedicate alle previsioni del tempo. Al confronto tra le strategie sintattiche e lessicali più diffuse in questo genere testuale mediatico in spagnolo e inglese (americano) associa un'interessante discussione sulla possibilità di segmentare unità delimitate da pause (che non corrispondono alle frasi 'logiche', ben formate, delimitate dall'interpunzione tradizionale) e unità enunciative in unità intonative. Si tratta in molti casi proprio di un problema di separazione di una successione che può apparire senza soluzione di continuità. Il *parsing* prosodico, come proposto in diversi lavori anche internazionali, sembra voler infatti ignorare i fenomeni di sandhi intonativo che portano alla fusione di unità a contatto¹.

¹ Da ormai un decennio, a lezione, mi capita di discutere delle distinzioni possibili in

Un altro esempio di difficile separazione di un continuum è offerto nel mio articolo «Differenze fonetiche generali tra voci femminili e voci maschili»: sebbene il sesso sia assegnato (anagraficamente) alla nascita sulla base di evidenti conformazioni anatomiche, nel corso della sua crescita nella società l'individuo definisce (generalmente in modo inconsapevole) una voce che, pur risentendo di vincoli biologici, si colloca lungo un continuum dipendente da qualità personali. A una distinta classificazione in voci di tipo femminile e maschile possono quindi corrispondere disposizioni diverse sulle quali si può intervenire esternamente modificando volontariamente alcuni degli assetti fonarticolatori sulla base di scelte individuali. L'articolo proposto – di cui si descriverà meglio la collocazione nella sezione *PhoneWS* – si dilunga su alcuni di questi temi, precludendo ai materiali discussi nel corso di una relazione tenuta da me e Valentina De Iacovo al corso di formazione «La voce parlata

italiano tra (almeno) tre modi diversi di acconsentire ringraziando: «Sì grazie.» (che può essere scritto così proprio in virtù delle sue modalità di enunciazione) è diverso da «Sì, grazie.» e da «Sì. Grazie.». In quest'ultimo caso si potrebbe parlare di due unità enunciative discrete (assertive, conclusive), mentre nel caso di «Sì, grazie.» si realizza una successione, ancora facilmente analizzabile, di due unità intonative (una continuativa e una dichiarativa terminale). La soluzione che alcuni faticherebbero ad

ammettere è invece la prima esemplificata. A garantire la fusione avvenuta in «Sì grazie.», nelle varietà di italiano che dispongono sistematicamente del RF, c'è invece (come sappiamo sin da Camilli 1913) proprio la lunghezza della consonante iniziale di «grazie» che sancisce la saldatura tra le due presunte unità ancora separabili negli altri due casi (e d'altra parte i più avveduti si saranno accorti che l'univerbazione garantisce la continuità di espressioni già persino lemmatizzate come «Sissignore/a»).

e cantata della persona transgender. Dalla ri-abilitazione alla chirurgia di ri-assegnazione della voce» (Torino, 2-3 dicembre 2023).
Segue poi un articolo in cui io stesso sintetizzo i risultati di alcune ricerche condotte sugli esiti debuccalizzanti di *s* in alcuni dialetti lombardi orientali (tra il bergamasco e il bresciano). In particolare, in questo contributo preliminare discuto di una prima valutazione acustica dei tassofoni individuati nel dialetto di Esine (BS) da Vittorio Volpi, ex-allievo di Gaetano Berruto, il quale gli ha trasmesso l'interesse per una visione razionale del patrimonio di informazioni che sopravvivono nei dialetti e nella variazione, in questo caso, fonetica.

Ma è proprio da alcune suggestive considerazioni che G. Berruto ha presentato nella sua relazione su invito al «LVI Congresso internazionale SLI di Torino» (v. sopra) che partono le riflessioni che propongo in questo editoriale, in riferimento agli altri articoli contenuti in questo numero e ad alcuni

ammettere è invece la prima esemplificata. A garantire la fusione avvenuta in «Sì grazie.», nelle varietà di italiano che dispongono sistematicamente del RF, c'è invece (come sappiamo sin da Camilli 1913) proprio la lunghezza della consonante iniziale di «grazie» che sancisce la saldatura tra le due presunte unità ancora separabili negli altri due casi (e d'altra parte i più avveduti si saranno accorti che l'univerbazione garantisce la continuità di espressioni già persino lemmatizzate come «Sissignore/a»).

dati presentati nel numero precedente.

Sebbene la relazione fosse incentrata su «Il continuo e il discreto, le varietà di lingua e la sociolinguistica recente», il suo percorso di rivalutazione razionale dell'evoluzione delle scienze linguistiche negli ultimi decenni ha evidenziato alcuni snodi evolutivi (inter)disciplinari in un avvincente succedersi di problemi teorici e soluzioni offerte da vari paradigmi sperimen-

tamente deducibili dall'osservazione empirica P_α .

Se invece si dispone di n enuncianti con le caratteristiche β , i vincoli strutturali A, B, C e D grazie ai quali queste realizzazioni si manifestano discendendo con curve con angoli non molto appariscenti, da P_β si ha l'impressione che non ci siano elementi discreti (gradini o spigoli) a guidare la discesa (Fig. 1).

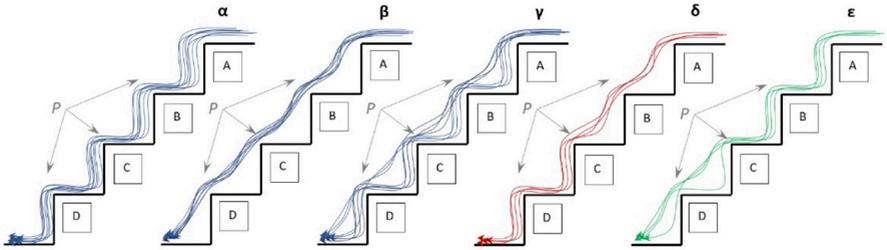


Fig. 1

mentali, soffermandosi in particolare in rappresentazioni che ho provato a riassumere visivamente negli schemi che riproduco in figura 1.

Sia una lingua in cui esista una struttura (un sistema di relazioni, una frase, un'organizzazione di contrasti...) i cui elementi, discreti, sono i gradini (o spigoli, punti di cambiamento direzionale) A, B, C e D.

Se si dispone di n realizzazioni (suoni, enuncianti...), rappresentate dalle linee direzionate discendenti da destra verso sinistra, con le caratteristiche α , i vincoli posti (o le classi individuate) dai gradini A, B, C e D sono immedia-

S'immagini invece la condizione γ : l'insieme delle n realizzazioni osservato da P_γ non lascia apparire i vincoli strutturali A, B, C e D, ma potendo scorporare gli enuncianti di tipo δ (che lasciano apparire le proprietà dei gradini C e D) e studiare separatamente le realizzazioni di tipo ϵ (che mostrano l'esistenza dei gradini A e B) è possibile rianalizzare γ e descrivere i 4 vincoli che guidano complessivamente la discesa (anche se alcune realizzazioni γ sono scivolote sui primi due e altre hanno smussato le proprietà degli ultimi due).

Con il metodo dell'osservazione separata, discriminando le realizzazioni

esperite, emergono le proprietà discrete che guidano la loro produzione.

Ecco però che anche nelle condizioni di osservazione P_{β} vengono messi a disposizione strumenti che sono in grado di rilevare i minimi movimenti che caratterizzano la discesa seguita dalle realizzazioni osservate. Si nota quindi delle curve attenuate e l'osservatore può discernere gli elementi discreti che sottostanno all'esecuzione linguistica. Tuttavia accade che lascino apparire anche sgranature locali (livelli intermedi): l'osservatore deve decidere quali siano i gradini più rilevanti e quali quelli accidentali.

Un esempio concreto di queste condizioni può essere dato dallo studio dei sistemi vocalici. Dall'ascolto delle produzioni del parlante *reg3* (studiate nell'articolo di Romano & Robaldo 2023) è infatti possibile confermare per le parlate dell'Alta Langa i rilievi

di una lunga tradizione di studi che ha ragionato su un sistema di quattro gradi di apertura per le vocali anteriori non arrotondate. L'osservazione della dispersione dei valori acustici consente di individuare facilmente i centroidi delle quattro classi di allofoni che individuano i fonemi /i/, /e/, /ɛ/ e /a/ (fig. 2).

Non sorprenderà che, come spesso accade trattando con dati di parlato, si presentino alcune sporadiche sovrapposizioni: alcuni allofoni di /i/ e /e/, sulla base delle due sole coordinate F_1 e F_2 , si attestano infatti in una stessa area di confine (la separazione tra le due classi può essere tuttavia accertata da un'indagine statistica su queste dispersioni). Allo stesso modo, un accavallamento tra le dispersioni di valori caratterizza in generale la separazione di classi tra alte e medio-alte (anche /y/ vs. /ø/ e /u/ vs. /o/)².

È però, in particolare, la separazione tra le aree di esistenza di /ø/ e di un presunto /ə/ che pone le maggiori difficoltà: neanche gli strumenti statistici aiutano in questo caso, data la non-significatività della debole differenziazione complessiva.

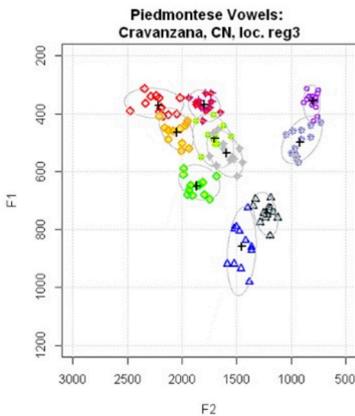


Fig. 2

² L'articolo citato dettaglia anche le criticità che si presentano nella separazione /u/ ~ /o/ per le produzioni di un'altra parlante e, in particolare, in condizioni di nasalizzazione vocalica (giustificando le incertezze nelle trascrizioni offerte dalle fonti atlantistiche per i dialetti di quest'area).

In queste condizioni, oltre che affidarsi alla misurazione di altre variabili acustiche³, la discretezza rimane sostenuta solo da un dato storico-evolutivo ed è come se a un unico (arci)fonema, risultante da una neutralizzazione di opposizioni in corso, corrispondessero ancora due classi di varianti combinatorie ancora debolmente contrastate (forse anche solo per via di una lunghezza distintiva).

Ecco, dunque, un interessante campo di applicazione di una lucida visione dei progressi metodologici che ancora aspettano alcuni campi della ricerca linguistica e nei quali dovremo imparare ad accogliere e contenere i contributi dei nuovi strumenti. Disporre di metodi sperimentali in grado di portare maggiori conoscenze sui fenomeni che spiegano la variazione linguistica

(come nella condizione β dello schema sopra) è senz'altro un ottimo risultato. Ma non dimentichiamo che oltre al risultato, allo studioso interessa anche il percorso seguito per raggiungerlo, nel corso del quale si presuppone si sia prodotto un arricchimento di conoscenze personali. (1) Scrivere un articolo con uno stato dell'arte affidato a sistemi 'intelligenti', (2) definire un protocollo sperimentale mutuandolo da altri lavori e (3) analizzare e discutere risultati ottenuti attraverso procedure automatiche di estrazione e analisi statistica dei dati (con ricorso magari a sistemi dei quali non si ha un pieno controllo), non sono attività in grado di configurare un'esperienza del ricercatore tale da garantire un suo reale progresso personale e, tutto sommato, neanche collettivo.

³ I tentativi di D. Robaldo di ricorrere a misure di F_3 per rafforzare il contrasto potrebbero rivelarsi produttive ricorrendo al concetto di F_2 efficace (cioè F_2' , basata su valori di F_2 ponderati da quelli di F_3 che influiscono sulla percezione del grado di apertura).

First results of a lexical analysis and a prosodic parsing of English and Spanish speech in weather forecast news

Cristiana Brignone

Introduction

The weather forecast is a ubiquitous and necessary part of daily life. It informs individuals and communities of upcoming weather patterns and helps them prepare accordingly. The presentation of weather information to the public is done through various media forms, such as radio and television, where presenters face the challenge of delivering meteorological information, thus attributable to an inherently scientific discipline, in an objective and rational manner, while also resorting to non-scientific and attention-grabbing features of speech. In the field of Linguistics and Media Studies, numerous studies have been conducted on the language and prosody of TV news broadcasts (Bell, 1991; Conboy 2007; Busà 2014), but little attention has been devoted to analyzing the peculiarities of weather-forecasting spoken reports. From a linguistic and prosodic point of view, it can be stated that there are several sim-

ilarities between these two forms of speech. Primarily, the pragmatic use of language to achieve specific purposes, by employing recognizable linguistic features, such as a technical jargon and a straightforward, declarative language structuring; prosody-wise, in both forms of speech presenters modulates their voice according to specific prosodic features, such as a higher speaking rate and recognizable intonation patterns, in order to communicate the information being delivered appropriately (Montgomery 2007). Still, the linguistic alignment of these two types of speech does not preclude the possibility of studying the language of weather forecasts in greater depth, whether these are considered as a subtype of the macro-genre of news bulletins, or as a genre in its own right (Moore Mauroux 2016; Romano 2018).

In this paper I summarize some key findings on this topic from an MA thesis I defended last July at

the University of Turin, based on experimental research conducted at LFSAG. For reasons of space, the article will not provide a comprehensive theoretical overview of the existing literature on media language and broadcast news speech. Nor will it be able to expand on how prosody is used in speech analysis (for a review of previous experimental research on television news speech styles see Brignone 2023). Instead, it will focus on the study conducted on some American and Spanish weather reports, first clarifying the materials and methods used in analyzing their lexical features, and then detailing the results of the analysis of the general differences in timing and prosodic parsing.

In particular, since this style of discourse requires reviewing the events of an entire day in different geographic locations and forces utterances to be connected in a limited time with rapid transitions, clear boundaries between intonational units are not realized. Specific sandhi phenomena and focal accents contribute to making prosodic parsing much less easy than in other styles of media discourse, and one of the results of the work was

precisely to illustrate under what conditions these «fusions» occur.

1. Models of reference for the analysis of weather forecasts

Before looking at data collection and analysis procedures, it seems necessary to briefly focus on two studies of weather forecasts, which served as models of reference for conducting my research.

The first study was conducted by S. Moore Mauroux in 2016 and focused on a corpus of UK weather forecasts recordings, which were analyzed by using acoustic analysis software PRAAT. The data collected were then interpreted through a lexical and a prosodic approach.

On the lexical front, text analysis revealed that the forecasts in question were characterized by a «varied register» made up of specialized terms, such as *pressure*, *system(s)*, *front(s)* among the most frequent, and «compound word formations» (multi-word units), like *low/high pressure system*, *cold/warm front*. Other than expectable references to weather events, such as *snow*, *cloud*, *temperature(s)*, other widespread lexical phenomena include geographical references to parts of the UK, most of times

accompanied by «directional or descriptive adjectives» like *southern/northern*, as well as a considerable number of time-related expressions denoting different parts of the day, and prepositions providing specific information about time and location. For instance, the term *through* appears more than 50 times, in expressions like *through the day/night*, and *across* is the tenth most common word in the corpus. Furthermore, there is a noticeable abundance of verbs that depict «movement, change or influence». Despite the use of a specialized jargon, in the weather reports analyzed there is evidence of more informal terms and expressions belonging to an oral register employed in conversational contexts, such as *struggle*, *chilly*, or set expressions like *we're not out of the woods yet*, as well as fuzzy expressions, namely *fairly* and *pretty*. One last characteristic that emerged from the analysis concerns the employment of collocations built through the pairing of one or more weather-related words that occur together repeatedly, thus creating fixed and predictable associations, *chilly temperatures*, or *outbreaks of rain* among the others (Moore Mauroux 2016: 5-6).

On the prosodic front, the focus of the study lay on intonation patterns, which were investigated in terms of pitch movement and nucleus in order to explore their roles in «discourse structure, topic change and highlighting». The examples illustrated by the author show that, when shifting to a new topic, intonational phrases tend to be subjected to a «pitch reset» process, meaning that the first element of the phrase, which is usually a connector like *and* or *but*, is pronounced on a higher pitch than the words coming before it. The initial high pitch then gradually decreases throughout the intonational phrase mirroring the anticipated decline and gradually decelerating towards the conclusion. Throughout the corpus, pitch modulation was also used to highlight certain pieces of information, specifically events, locations and tendencies, and also to create contrast or continuation between significant facts. Significant pitch variations were also associated with the nucleus, that is the central element of the phrase. The pattern illustrated shows a descending pitch on the nucleus beginning from a higher point, emphasizing its significance, whereas

the second nucleus experiences only a slight decline, suggesting it is considered to be less important than the preceding information. Even though many examples are provided in the article, findings clearly revealed that the primary prosodic indicator was a sustained high pitch, employed to signal both a transition and draw attention to a specific item within a given intonation unit¹. The shift in topic was often accompanied by a pitch reset, where the preceding intonation

unit concludes with a low pitch that contrasts with the subsequent intonation unit's initial high pitch. Moreover, inside intonation units themselves, it was common to observe a slower pace and/or a gradual pitch decline toward the end. Lastly, elevated pitch consistently served as a tool for emphasis, by highlighting various aspects such as the information itself, its scope, and degree of relevance. Similarly, high pitch was used to signify ongoing or shifting trends, as well as to provide commentary on the information presented (Moore Mauroux 2016: 8-11).

The second study, conducted by A. Romano in 2018, compared the speech of British weather forecasts (BBC) and the one of Italian weather forecasts (RAI), through the analysis of 38 audio files, with a total duration of about 22 minutes.

As far as lexical and syntagmatic composition is concerned, the

¹ Much knowledge in this field comes from the results of experiments presented in Gasser et al. (2019). They rely on the use of a corpus made up of sentences originally uttered by radio newscasters and rerecorded by readers with no prior knowledge of broadcast news delivery and instructed to read in a spontaneous and conversational way, and then compared to the original samples in order to identify the main prosodic differences between the two styles (Similar experiments have been conducted in Italy well before by Giannini & Pettorino 1999). A perception-based task was then designed, allowing participants to establish whether the audio recordings used and produced for the production experiment were traceable to news talk or to everyday speech. As regards the production experiment, data showed no significant discrepancies in pitch range and in the total amount of pitch accents, but a higher number of intonation phrases were found in newscasters' recordings (Gasser et al. 2019: 14-15). The results of the production experiment showed that participants were able to establish that the audio clips had been uttered by a newscaster with an average accuracy of 57.83%, and they

showed a higher ability to correctly identify newscasters rather than dismissing non-newscasters from the corpus, which might be due to the fact that recognizing newscasters' voices was an easier task than identifying others, and that other features that were not present in this study, such as discourse structure and lexical content could have been more effective indicators for the identification of newscaster speech (Gasser et al. 2019: 19).

major differences between the two types of weather forecast include a higher frequency of the names of the regions making up the UK territory, among others *England* with 32 occurrences and *Scotland* with 25; whereas, in the highest ranks, RAI data provided evidence of the name of only one region, *Sicilia*, with 6 occurrences. In line with Moore Mauroux's study, the BBC samples are rich in cardinal references (entries of nouns and adjectives referring to the four cardinal points are almost 100) and prepositional phrases constructed with *across* (26), *through* (19) etc.; RAI samples show a limited presence of cardinal directions, due to a preference for more detailed toponyms, linked to mountains, seas, or specific national areas, without preventing the use of more generic terms, such as *regioni* (regions), *parte* (part) and *area* (id.), which also occur in similar number in BBC weather reports. Regarding verb forms, in English, evolution and variation of weather events are primarily signaled by verbs like *see/seeing* (25), *go* and its different forms (18), and *get* (9); whereas in Italian, by *tendere* (tend to) and its different forms (17), *andare* (go) (10), and

arrivare (get) (8). Generally, there is correspondence between the names of the different atmospheric conditions (*rain/s* and *pioggia/e*, etc.), except for *snow*, which occurs 30 times in BBC forecasts, and *neve/nevicate* appearing at far higher ranks, and *sleet*, with no counterpart, as is the case with *temporale/i* (thunderstorm) in RAI forecasts.

On the prosodic front, momentary interruptions in the regular flow of speech are generally attributable to breathing needs or in any case to short and silent pauses, as hesitations and filled pauses are very rare. The number of short pauses is practically equal in the two groups of recordings, whereas long pauses, though rarely performed overall, are more numerous in BBC forecasts. As regards speech tempo, speech rate and fluency rate are higher in Italian weather forecasts (6,56 syll/s and 5,78 syll/s), and have a remarkable difference in numerical value; this incongruence is less marked in English (5,06 syll/s and 4,60 syll/s) due to a more balanced proportion between the expected syllables and those actually articulated. For what concerns intonation, the transcripts of the recordings were

segmented into utterance units and intonation units, in order to highlight potential discrepancies in how speech was structured and enunciated. Results showed that for both BBC and RAI, though more distinctly in English, rhythmic and intonational patterns tend to deviate from the logical and syntactical structure; this is a finding further reinforced by the fact that pauses and breathing-in moments do not recur with regularity, resulting in a performance misaligned with respect to phrasal structures and overwhelmed by discontinuous pausal interruptions. More specifically, Italian utterance units are characterized by a higher number of syllables (24 vs. 16 syllables) and, on average, a longer duration (3,75 vs. 3,24². Besides, BBC's utterance units consist almost entirely of one or three intonation units, whereas

in the RAI data, there is a predominance of utterance units containing four or five intonation units.

In conclusion, the Italian corpus is defined by a higher degree of formality, not only as a result of lexical and syntactic choices, but also of its prosodic structure, with longer utterances and less variable intonation units³; whereas BBC announcers opt for lexical, phrasal mechanisms attributable to a more colloquial style of speaking and prosodic modalities, leading to the general perception of a more effective and pressing rhythm, even though speech rate is higher in the Italian weather forecasts.

2. Methodology of the study

This section will be devoted to the description of data collection, as well as the instruments and procedures employed to conduct the analysis.

² Particularly for Italian, it seemed useful to take into account the diachronic changes reported by Giannini (2004) who was able to observe how, over the course of about thirty years, radio and television speech data showed a higher articulation rate and, consequently, an increase in the percentage of spoken material at the expense of the percentage of silence. This decrease is made even more evident by a consistent lengthening of the utterances, resulting from the practice of limiting «empty» moments to strictly necessary breathing pauses

and of prolonging the articulatory segments when possible. The final perception is that of a faster speech due to increased density and concentration of the verbal flow, as confirmed by a 50% increase in fluency (Giannini, 2004: 59).

³ This does not contradict what Giannini (2004; see previous footnote) observed: a lower silence percentage with fewer and shorter breaks, higher fluency index and speaking rate, indeed corresponds to relatively less intonational variation (see below).

2.1. Data collection and corpus composition

Since the aim of the present thesis is to explore how weather is linguistically reported in Spain and the USA, the material under investigation was divided into two different corpora: one consisting of American English weather forecasts and the other of Spanish ones. The data for this study were obtained online from two news sources: CNN, a major American news network, and RTVE, which is the Spanish national public broadcasting corporation responsible for producing radio and television programs and for operating several channels. The process of data collection covered a time span of about five months, from October 2022 to February 2023, even though the two corpora include four weather forecasts going back to August and September 2022 as well, which were selected at a later date in order to have an overview of the reports that could somehow match seasonal changes. The resulting corpora contains 16 weather-forecast wave (.wav) files, with 8 from CNN and 8 from RTVE. Every recording includes the full content of the weather bulletin, each of which is two to three minutes long in its integral form.

As can be seen from the figures summarized in the Appendix, the total gross time analyzed is 18 minutes and 49 sec of CNN speech and 17 minutes and 36 sec of RTVE speech, divided between eight American and 8 Spanish weather presenters respectively. The choice fell on certain weather bulletins according to two main criteria, the gender of the speakers and the duration of the forecasts. The first criterion allowed for a good variability in terms of gender-related acoustic features, making it easier to achieve a more comprehensive understanding of the results, while ensuring gender balance and avoiding bias; the second criterion resulted in recordings that were not exaggeratedly divergent in terms of time length, so as to ensure standardization and control of the amount of content, and to facilitate analysis, making it easier to compare the different features of the recordings.

2.2. Methods and tools for data analysis

Once the material was collected, all 18 speech samples were transcribed into text format in order to obtain overall visual feedback of the flow of speech and to facilitate

the subsequent transcription on speech-analysis software PRAAT. Afterwards, the actual analysis took place on two fronts, lexical and prosodic.

On the lexical front, weather forecast texts of the same language were merged into a single *MS Word* file (one for Spanish and one for American English) that was then screened out by using Antconc, which is a free and open-source corpus analysis software designed to explore language data through a range of text analysis tools (Anthony 2022). In the case of our research, *AntConc's Word Tool* made it possible to generate a word-frequency list equipped with multiple markers, including the rank and the total number of occurrences for each word, among others; it also enabled an efficient identification of co-occurring terms, common word patterns and expressions, through the *Cluster Tool*, the *N-Gram Tool*, and the *Collocate Tool*. Although this freeware allowed for a thorough level of text analysis, it was still combined with the use of the *Find* function on *MS Word*, as it facilitated the distinction procedure between tokens and types, the highlighting of larger portions

of texts, as well as enabling further proof regarding the presence of set expressions and collocations considered to be characteristic of the weather forecast genre.

On the prosodic front, the analysis was carried out with PRAAT, which is a widely-used software equipped with a range of tools that allow users to perform many types of acoustic analyses on audio data (Boersma & Weenink 2023). Consequently, the 18 wav-format recordings were uploaded one at a time onto the software, listened to and processed according to different modalities, depending on what elements were to be visualized.

Once uploaded, each audio file was listened to and processed according to a two-pronged approach: on the one hand, since a specific computer and acoustic tool was used, it seemed right to take on an experimental approach, especially with regard to more objective measurements, such as pause duration and different speaking rates; this, however, did not prevent to complement the analytical modality by resorting to a perceptual approach when this was needed or more suitable, e.g., for the detection of breathing or silent pauses, and

pitch patterns, always taking into account the risk of making choices based on evaluative arbitrariness or impressionistic assessments. Before illustrating the phenomena considered, some methodological and terminological clarifications are required. Firstly, as anticipated above, speech was segmented and annotated into three levels, or tiers. The first one, labelled UTT (utterance tier), entailing a subdivision into interpausal units, intended as a stretch of speech preceded and followed by silence; since there was only one speaker, the boundaries (/ /) were established either by breathing pauses (<bp> = <inspiration>), short pauses (<sp>) or, though rarely, by long pauses (<lp>). Care was taken to detect, thus not to consider as pause, the holding phase of plosive consonants, and to distinguish short from long pauses by calculating the duration of the two following syllables: if this value was greater than the duration of the interruption, a short pause was established; conversely, if the value was lower, then the pause was labeled as a long one (De Iacovo et al. 2020). The second tier, labelled INT (intonation tier), was created to divide the pho-

netic content of the utterances into intonation units, the boundaries of which were established according to auditory discernment by evaluating audible pitch changes and, if in doubt, by observing the pitch curve or, in cases of correspondence between perceived melodic movements and syntax, the semantic content of the utterances. The third and last tier, labelled SYLL (syllable tier), contains the syllabic scansion of the intonation units. The number of syllables was determined by following the traditional criteria of syllable division and by deciding, as a general rule, to count only the ones actually pronounced by the speaker. So called filled pauses were regularly transcribed and taken into account into the syllabic count, both those uttered due to momentarily poor speech planning and the ones linked to articulation phenomena (e.g., schwa realizations). On the contrary, syllables pronounced in an elongated manner were not counted because, algebraically speaking, it seemed reasonable to consider them as single even though they took up more time than expected in a given utterance context, and also because very few cases were found. Turning

now to the content of the analysis, by taking as a reference the studies already addressed before, the focus was placed on the rates at which speech occurred, the number of pauses made by different speakers and the intonational structure, with the ultimate aim of comparing the ways weather is reported in the US and in Spain; these, then, are the phenomena considered:

- speech rate: calculated by dividing the number of syllables by the total duration of the speech, excluding moments of noise and conversation between the weather presenter and other professionals, both found at the beginning and end of some recordings;
- articulation rate: obtained by dividing the number of syllables by the duration of all the utterance units, thus avoiding pauses;
- number of pauses and silence percentage within the whole speech;
- number of utterance units and number of intonation units they are composed of;
- syllabic composition of the intonation units;
- considerations regarding the most recurrent intonational patterns;

3. Results and discussion

3.1. Analysis of lexical findings

As far as lexical composition is concerned, the two corpora observed exhibit the following features. Focusing on the first 100 entries, in CNN texts, the content words with 10 or more occurrences are weather, found at rank 19 and with 27 occurrences, followed by *degrees* (17), *storm* (17) at rank 29, *temperatures* (15) ranked 34, *snow* (13) in position number 39 and *rain* at rank 65. Other words related to weather or, more specifically, to atmospheric events occupy lower positions with very few occurrences. In RTVE texts, the highest-ranked content words are *temperaturas* (24), *día* (22), *mañana* (22), *nubes* (19), *grados* (18), *precipitaciones* (15) and many others among the ones with more than 10 occurrences, thus indicating a different composition and distribution of the words carrying semantic content if compared to the texts in American English.

Terms belonging to the specialized jargon of weather forecasting like *inches*, *system*, *front(al)* and *pressure* do not stand out within the corpus, since they are found at rank 39, 45, 128, 146, with 13, 12, 6, and 5 occurrences respectively, and are

usually found in multi-word formations such as *low-pressure system*, *storm system*, *cold front*, *low/high pressure*. Spanish counterparts, *sistema* (4), *frente* (5), *presiones* (1), are scarce, as can be seen from the number of their occurrences.

Direct references to geographical areas are sporadic as well, names of US states and cities can be found but with less than 10 occurrences (*Texas* is the highest one with 8). Similarly, indirect geographical references do not occupy high positions; cardinal points, adjectives or compounds including them are relatively low. For example, *North* reaches 17 occurrences in units like *northern*, *Northeast* and *Northwest*; we also have eastern (8) and *Midwest* (7). More general references like *area(s)*, *region(s)*, *part(s)* are detectable with, respectively, 12, 8, 6 occurrences are usually paired up with frequent prepositions like *across* (34). For what concerns Spain, names of geographical areas are not highly significant as well: within the first 100 entries, *Cantábrico*, *Canarias*, *Mediterráneo* occur 10 times, *Galicia* (9), *Baleares* (8); though, some cardinal points are slightly more frequent: *norte* (16), *sur(este/oeste)* (9), *este* (7). Gen-

eral references like *área(s)*, *región(es)*, *parte(s)* are, however very rare.

Temporal deixis is linked to words like *day(s)* occurring 20 times, *week-end* (11), *afternoon* (10), whereas the days of the week have, on average 3 occurrences. In RTVE's samples, *día* (22) and *mañana* (22) and *tarde* (15) rank high positions, week days occur 2 times on average.

Geographical and temporal items are usually paired up with frequent prepositions, like *across* (34), *through* (26) in English; in Spanish similar phrases are built up with, *durante* (16), *a lo largo de* (9).

In the context of verb forms, the evolution of atmospheric events is marked by verbs like *see/ing* (14), *go/ing/es* (11), *expect/ed* (10), *look/ing* (8), *come* (4), *pick (up)* (4). Focusing on Spanish, *ir (irán/va/van)* (33), *ver/vemos* (22), *dejar (dejarse/dejará/dejarán)* (13), *esperar (espera/mos)* (12).

Finally, it is worth pointing out that the American weather forecasts, from a lexical and syntactic point of view, were perceived by the author as intentionally more informal and spontaneous than the Spanish ones. In fact, constructions belonging to a more colloquial type of speech were found; among the most recurrent

ones, *keep an eye on, we could pick up...* and *let's time things out for you*; playful expressions were found in special occasions, with the intention of entertaining, such as *it will have impacts there for trick or treaters* in the Halloween forecast (see Brignone 2023: App. 3), and somehow creating a bond with the viewer, like *Many of us traveling home at the end of the weekend* in the Thanksgiving bulletin (see Brignone 2023: App. 6).

3.2. General timing

For a general representation of timing (speech rate and pause) we rely on a number of measures and assessments summarized below.

As shown in Figures 1, 2 and 3,

Spanish forecasters have a higher speaking rate than the American ones, both if taken individually and if averaged, and with a fairly considerable difference (RTVE: 6,34 syll/s, CNN: 5,12 syll/s); the same observations can be made if pauses are not counted, regarding the articulation rate (RTVE: 6,94 syll/s, CNN: 5,67 syll/s).

As regards pauses, it can be observed that, as expected, the largely prevalent pauses are the ones taken for physiological reasons (breathing pause <bp>), which are present in almost equal proportions in the American and in the Spanish weather bulletins (see Table below). Immediate-

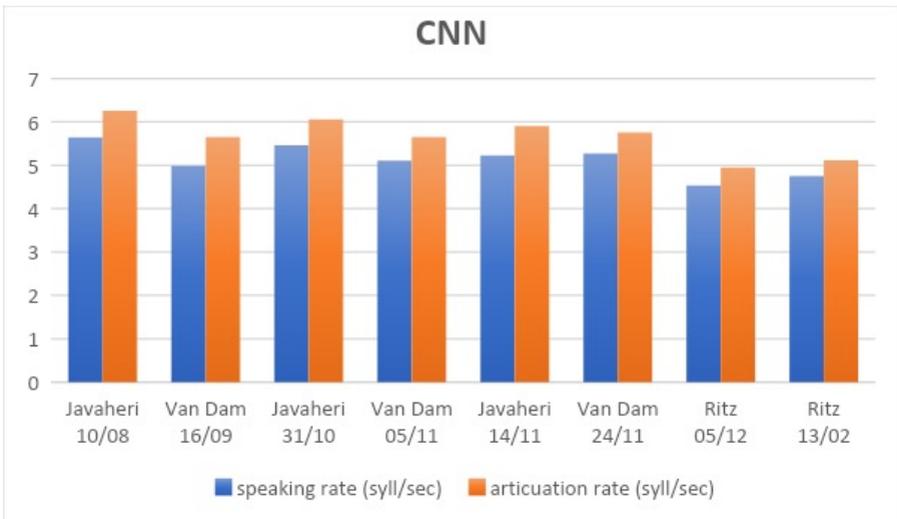


Figure 1 – Histogram: speaking rate and articulation rate of CNN's data.

ly following in numerosity, are short pauses, which are slightly more numerous in the US data (CNN: 67 vs. RTVE: 52). Long pauses are very sporadic, as are pauses containing hesitations in

the form of mumbling (as already explained, so-called full pauses – *uh*, *uhm* – were considered as (non-)lexical entities and, therefore, regularly transcribed as parts of speech).

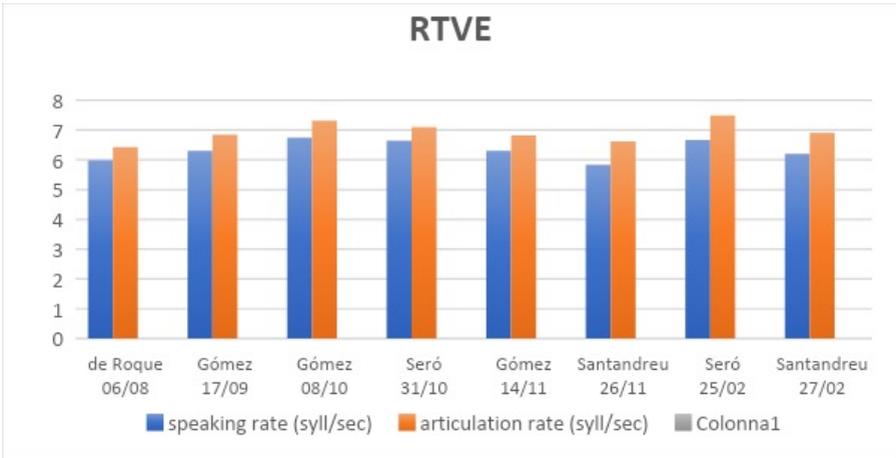


Figure 2 – Histogram: speaking rate and articulation rate of RTVE’s data.

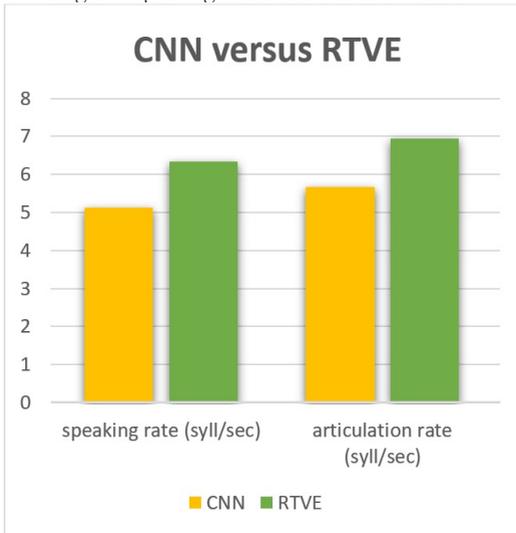


Figure 3 – Histogram: average speaking rate and average articulation rate of CNN’s data and RTVE’s data.

Table – Types and number of pauses in the two corpora.

Type of pause	CNN	RTVE
<bp>	293	299
<sp>	67	52
<lp>	3	3
<sp hesitation>	3	1
<lp hesitation>	2	2

Ultimately, despite the differences just described, the time occupied by pauses and the time occupied by actual articulation are almost identical in both corpora: in the entire utterance, the percentage of silence is 10%, that of speaking is 90%.

It is worth making a few observations about the position of pauses within the speech, to understand how they contribute to its structure. As regards CNN weather speech, slightly more than a third of the breathing pauses are placed in the spots where, in written texts, a full stop would be placed, signaling a change in topic (see Brignone 2023: Appendix 6):

(1) *Don't drive over any flooded roads if some of that ponding does occur. <bp> Here's a look for the day on Sunday. / Many of us traveling home at the end of the weekend. <bp> And we do have*

a larger storm system that will be impacting the East Coast.

In RTVE's samples, the same phenomenon is present but to a lesser extent, in fact affecting about one-sixth of the total cases (see Brignone 2023: Appendix 13):

(2) *Como les decía, durante las próximas jornadas <bp> el tiempo seguirá revuelto, / sobre todo mañana con lluvias en prácticamente toda España. <bp> Durante la jornada del miércoles un nuevo frente <bp> entrará por el oeste del país, <bp> y el jueves parece que las lluvias <bp> también se extenderán / a prácticamente todo el territorio.*

The same occurs for short pauses; in American English, this affects about a third of the occurrences, as with breathing pauses (see Brignone 2023: Appendix 4):

(3) *You can see the Winter Weather alerts in place for so many / Americans at the moment, <bp> with uh lots of snowfall / with the series of storms moving through <bp> and, again, the heavy rainfall impacting places like Seattle <bp> as well as Portland. <sp> Here's a quick look at your temperatures for today.*

As far as Spanish is concerned, short pauses placed after a full stop to signal topic change are, on the contrary, quite uncommon, affecting only 10% of the total occurrences.

Moreover, in the CNN data <bp> + *uh* associations are very common: out of 26 occurrences of *uh*, the association appears 17 times, albeit concentrated in recordings of the same speaker. This leads to claiming that it is more likely a trait of the speaker's personal style rather than a general trend (see Brignone 2023: App. 2 and 6):

(4) *So <bp> uh we are going to look for some replenishing rains for that area, <bp> uh not necessarily a drought buster, but certainly, we'll take what we can get. / The rest of the country right now, <bp> looking nice / along the eastern seaboard, <bp> uh the Midwest OK as well. We do have se- / the potential for some <bp> uh heavier rainfall across the upper Midwest, particularly into Minnesota.*

(5) *The Pacific Northwest also anticipating <bp> uh some wet weather with high-elevation snowfall <bp> uh for places like Seattle to Boise to Billings as well. <bp>*

Temperatures for the day today, upper fifties / to around sixty degrees, depending on where you're located. <bp> And then taking you through the course of the weekend, <bp> uh that is where we anticipate / temperatures to remain / into the upper fifties to around sixty degrees for <bp> many of the East Coast cities.

In the rest of the cases and in both corpora, the different positions of the pauses are quite variable, which makes it possible to presume that there is no precise correspondence with a syntactic structuring of the discourse.

Through the calculation of the two speech rates, the laying-out of speech composition in terms of silence and speaking time, and the exemplification of the most recurrent positions of the pauses, the intention was to determine which of the two groups of speakers reported the weather faster and to understand how the scanning of pauses affected the structuring of speech and thus a listener's perception. In this respect, pauses, precisely because they are not always present in expected positions, contribute to the formation of a perceptually-irregular speech, a factor that will be further

developed by also considering the number of intonation units, their syllabic composition, and the pitch movements that characterize them.

3.3. Types of units and prosodic construction

Regarding the prosodic organization of the texts in American English and Spanish, our analysis and labeling yielded a quantitative sketch as follows.

Through the histogram of Figure 4 an attempt was made to relate two layers of speech structuring: the first is that of utterance units, and the second concerns the sub-level of intonation units.

As can be seen from the graph, most utterances are formed by 1, 2, and some of 3 intonation units, in both CNN and RTVE samples. The major differences are as follows: in the CNN data, 70% of utterances (270 out of 383) consist of a single intonation unit, whereas in the RTVE data, the percentage amounts to 40% (166 out of 369); in addition, Spanish utterances are segmented into 2 INTs in more than 40% of cases (120 out of 369) and into 3 INTs with a frequency of about 15%, whereas in American English, scanning into 2 or 3 units is more sporadic, in the former case 20% of UTs are affected, in the latter case less

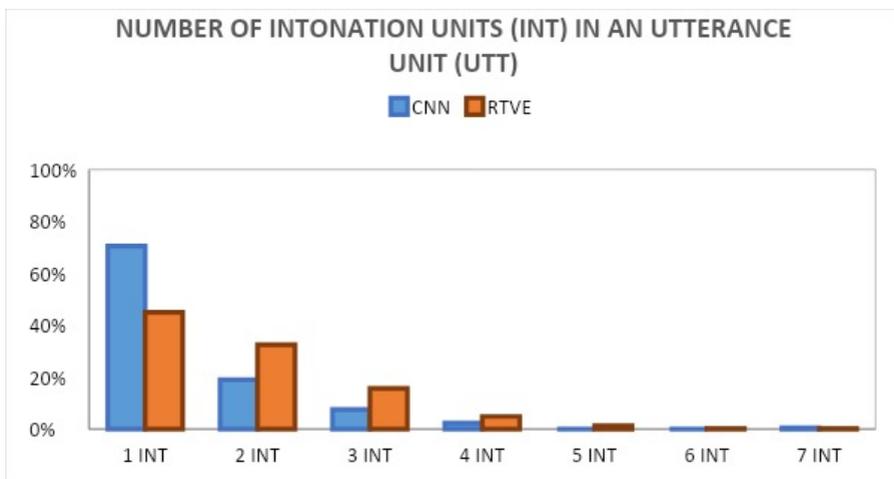


Figure 4 – Histogram: Percentage of utterance units formed by 1 to 7 intonation units in the two corpora.

than 10%. As the number of INTs goes up, the percentages decrease, with very few occurrences in the case of 5, 6, 7 INTs in Spanish data, almost none in American English. In general, it can be argued that the distribution of the different INTs is more spread out in Spanish texts, on the contrary, English texts predispose a more exponential decrease since the UTs are formed in the vast majority of cases by a single INT, and in a few cases by 2 or more INTs. Moreover, it is important to note that, contrary to expectations, the average intonation unit in American English is shorter than in Spanish (15,6 syllables in American English and 17,7 in Spanish). In this case, although Spanish intonation units tend to be slightly longer in terms of syllables, the overall speaking and articulation rates are higher in Spanish. This could indicate that a better spread-out distribution of utterances in more intonational units, and the fact that Spanish forecasters may pronounce individual syllables faster or speak with less pause between them might contribute to an overall higher speech rate.

3.4. *Recurrent intonational patterns*

Given the difficulties that will be highlighted in the following paragraphs, a systematic quantitative study was not conducted to bring out indices of differentiation between the two samples. However, as pointed out in Romano (2018) for Italian and British English, it immediately became apparent how the most recurrent phenomena in the analyzed speech data are not common to the two corpora.

3.4.1. *Continuation rise in Spanish*

From the analysis of pitch contour (blue line in the screenshots of the PRAAT graphic window), rising intonational patterns are frequently observed, thus establishing intonational boundaries by scanning the subdivision of utterances into continuative intonational units.

An example of an expected parsing is in Figure 5, where three rising pitch contours, albeit with different degrees of slope, correspond to four intonation units.

Another example is shown in Figure 6, where it can be seen, in particular, that the first part of the third interpausal unit (highlighted) is characterized by a clear rising intonational movement that makes it

possible to establish an intonation-boundary, clearly visible on the spectrogram, as well as audibly per-

ceptible. However, within this unit, the first word – which has been separated from the preceding ele-

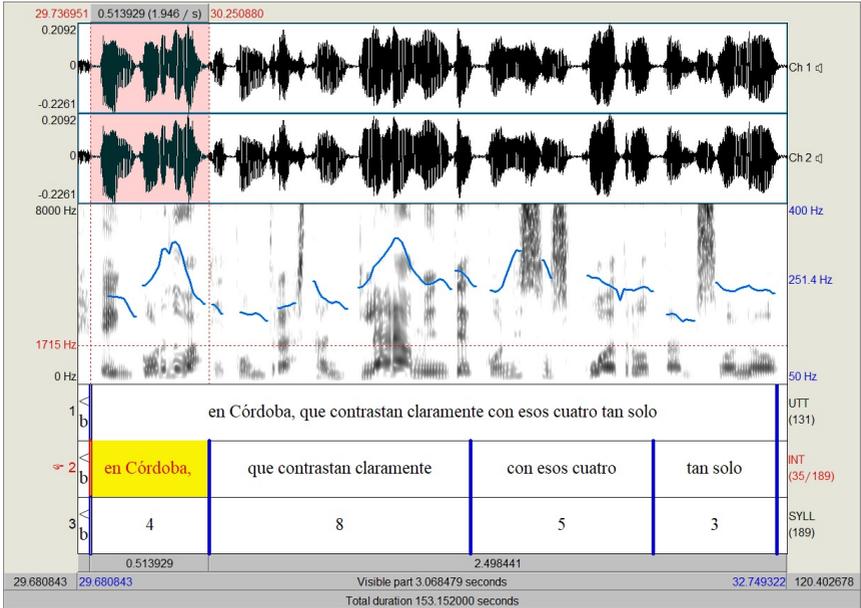


Figure 5 – Nùria Seró 25/02: rising pitch contours in RTVE data

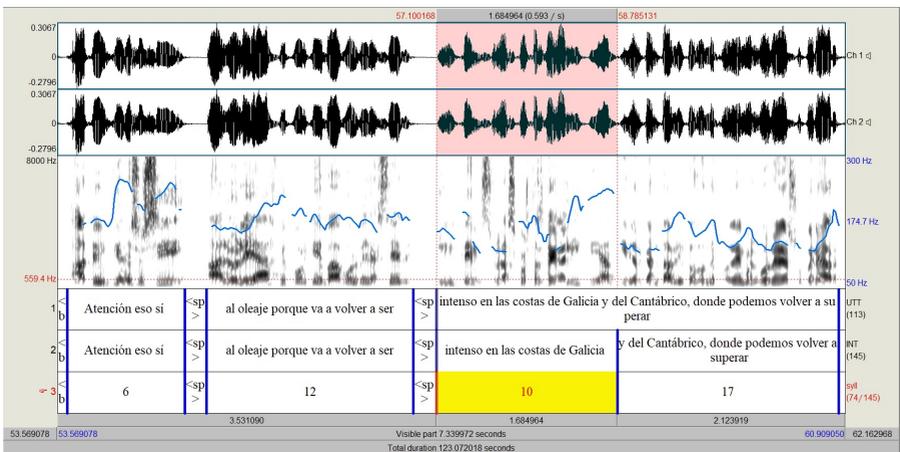


Figure 6 – Marc Santandreu 26/11: rising pitch contours in RTVE data

ments of the syntactic phrase on which it depends – is realized with a specific intonation contour that is well connected to the following pattern.

The same phenomenon is found in the unexpected fusion between *Cantábrico*, on which a residual continuation function is perceived, and the following unit.

3.4.2. *Downstepping and non-terminal units in English*

Rising intonational patterns do apply to American forecasters as well, but to a lesser extent, as intonational boundaries are signaled by pitch phenomena that are quite variable depending on their position and, therefore, not recurrent enough to establish a general trend. Below are some examples of pitch profiles variation.

In these examples, the highlighted utterance is characterized by a pitch that goes downward, underscoring a depletion of energy as one approaches the end of the unit (a phenomenon also reported in Moore Maroux 2016 and well documented in more general intonation studies). However, focusing on the sub-level of the intonation units, it is clearly visible that they

downstep as well. On a perceptual level, this phenomenon gives rise to an initial higher pitch segment followed by units collapsing without clear boundaries (on a lower pitch) (Figures 7 and 8).

Very long utterances depend on such a general attitude to connect intonation units (and sub-units) without any clear boundary marker. A progressively downstepping intonational pattern is associated to a depletion of energy, but it is regularly marked by more or less pronounced pitch accents confined to smaller units. Pitch movement is here associated with the nuclei of the utterance, as a way for prosodic prominence to pointing out significant parts of the information. In Figure 8, H* pitch accents are aligned with *pressure*, *mid-Atlantic*, *snowfall*, *early* and *morning*, thus emphasizing all the essential elements for understanding the content of the message.

In Figure 8, the first intonational unit has a rising pitch that creates a suspended intonational movement, which is then followed by a second closing unit: the pitch drops sharply to signal the end of the information produced up to that point. However, despite the end of the

subject signaled by a full stop, the utterance does not yet end, but has a third intonational unit in which the new piece of information is already introduced, and will be developed in the next utterance.

3.4.3. Unit connection in Spanish

As for the Spanish data, it is necessary to report some examples, in which a similar intonation sandhi phenomenon may be illustrated at even major textual boundaries. It was difficult to establish intonational boundaries, which were discernible by ear, but not clearly placeable because of articulatory phenomena involving contiguous words and arising precisely on those supposed boundaries. Some of them are illustrated below. Figure 9 shows an example of a vowel encounter between two items – *repite* and *En* – at an expected terminal intonation boundary. The highlighted segment identifies the two vowels, instead of being pronounced separately, collide and merge into one, thus hindering the boundary between two audible intonational movements, signaled at the level of punctuation by the full stop, which could have been useful for clearer speech scansion, also in

view of the following topic change.

In Figure 10, the perceived change in intonation between *lluvias* and *ya* does not correspond to a separation at the segment level, since the speaker does not completely realize the two separate sounds corresponding to the voiceless fricative /s/ and a /ʎ/-type approximant: the continuation profile signalled by the rising-falling pitch movement on the first syllable of the word *lluvias* dissolves into the initial profile of the following parenthesis unit (Figures 9 and 10).

With these examples I tried to show how intonation parsing interacts with segmental phono-syntactic phenomena and affect the prosodic patterning of utterances: intonation boundaries are many and varied within the Spanish data. In any case, since these phenomena are subject to perceptual and subjective interpretation, I chose, as a general rule, to establish only one intonational unit, rather than two.

From a prosodic standpoint, the analysis of speech rates showed that both CNN and RTVE weather presenters tend to speak fast, but, with an almost equal number of pauses, the latter are characterized by higher speaking and articulation

rates, thus leading to the perception that Spanish is a faster language. Moreover, the acoustic analysis conducted through Praat demonstrated that American forecasts feature utterances largely formed by a single intonation unit. This holds true for Spanish forecasts as well, but with a more equally spread-out distribution, as, in comparison, they display a higher number of utterances including multiple intonation units. The more melodically-cadenced proceeding of the latter is not to be found, however, in the syllabic composition, since

Spanish intonational units consist, on average, of a higher number of syllables than American English ones. Nevertheless, although Spanish intonation units tend to be slightly longer in terms of syllables, the overall speaking and articulation rates are higher in Spanish. This could indicate that a better spread-out distribution of utterances in more intonational units, and the fact that Spanish forecasters may pronounce individual syllables faster or speak with less pause between them might contribute to an overall higher speech rate.

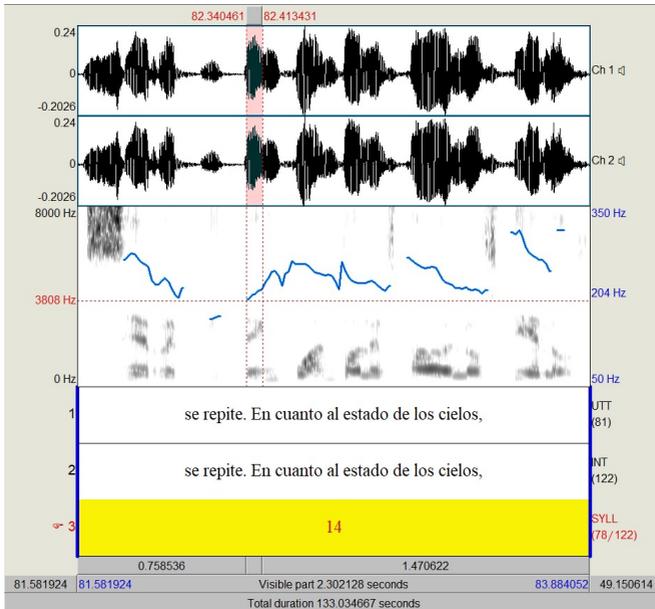


Figure 9 – Ana de Roque 06/08: Intonation sandhi phenomenon avoiding the realization of a boundary.

Focusing on intonation, Spanish forecasters modulate speech and, consequently, information through a higher number of final rising intonational contours, whereas intonational boundaries in CNN texts are signaled by more variable pitch phenomena. Finally, the acoustic analysis highlighted varied phono-syntactic phenomena involving contiguous linguistic items in close proximity to intonational boundaries, thus making it difficult to establish clearly-distinguished intonation units, especially in RTVE weather forecasts.

Conclusions

As a result of the analysis of a corpus of recordings related to the weather forecasts of TV networks CNN and RTVE in the period August 2022-February 2023, we were able to detect, with the same time of organization of the broadcast and under conditions of minimal textual genre variability, interesting differences in lexical choices and spatio-temporal references (only partly attributable to the two different territories involved).

Even more interesting proved to be the study of the different modes

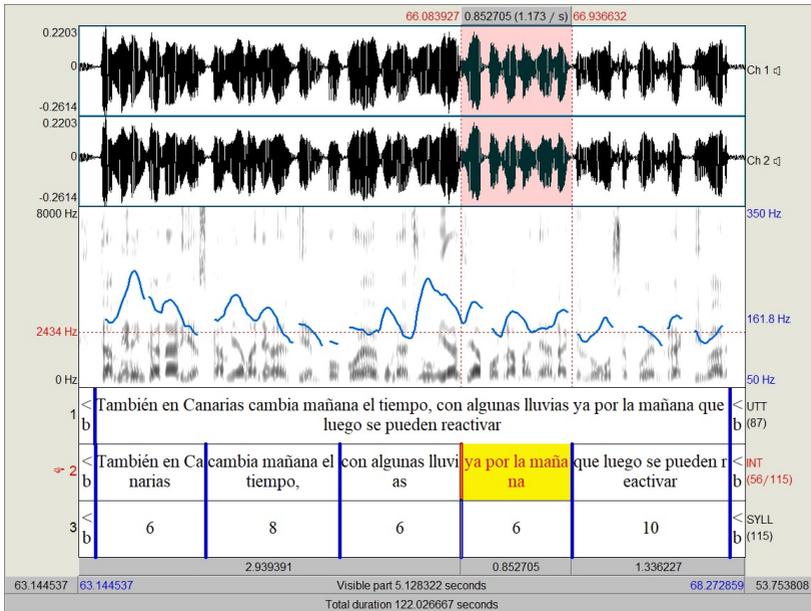


Figure 10 – Andrés Gómez 17/09: Difficult placement of the intonational boundary.

of enunciation adopted by the studied speakers. While the two groups of announcers did not presented relevant differences in terms of fluency, the recorded eloquence speeds suggest that Spanish speakers are generally faster. This is reflected in interesting prosodic contact fusion phenomena between units, making it difficult to measure an index of utterance construction in terms of intonation units along the lines of the study conducted by Romano (2018). While there is a frequent use of continuation units as a typical strategy used in Spanish (with a dominance therefore of non-terminal rising contours), a connection strategy by American speakers is also asserted that leads them to articulate longer and less articulate utterance units. Within the regular downstepping that occurs in these cases, there are thus lower-level units signaled by pitch accents that recur with greater regularity (cf. Moore Mauroux 2016).

References

- Anthony L. (2022). *AntConc* (Version 3.2.1) [Computer Software: <http://www.antlab.sci.waseda.ac.jp>].
- Bell A. (1991). *The language of the news media*. Oxford: Blackwell.
- Brignone C. (2023). «Reporting the weather in two different cultural spaces: a linguistic analysis of American and Spanish weather forecasts». MA Thesis, Dept. of Foreign Languages, University of Turin, *manuscript*.
- Boersma P. & Weenink D. (2023). *PRAAT: doing phonetics by computer* (Version 6.3.10) [Computer Software: <http://www.praat.org>].
- Busà M.G. (2013). *Introducing the Language of the News - A Student's Guide*. London: Routledge.
- Conboy M. (2007). *The Language of the News*. London: Routledge.
- De Iacovo V., Colonna V. & Romano A. (2020). «La pausa». *Bollettino del LFSAG*, 5, 41-48 [www.lfsagunito.it/ricerca/phonews/05/05_5.pdf].
- Gasser E., Byron A., Napoli D. J. & Zhou Z. L. (2019). «Production, perception, and communicative goals of American newscaster speech». *Language in Society*, 48(2) [doi: 10.1017/S0047404518001392].
- Giannini A. (2004). «Analisi acustica del parlato televisivo», In P. Cosi (ed.), *Misura dei Parametri, aspetti tecnologici ed implicazioni nei modelli linguistici* (Atti del I Convegno Nazionale AISV - Associazione

Italiana di Scienze della Voce, Padova, 2-4 Dicembre 2004), Padova: ISTC/EDK ed., 49-61.

Giannini A. & Pettorino M. (1999). «I cambiamenti dell'italiano radiofonico negli ultimi 50 anni: aspetti ritmico-prosodici e segmentali». In: R. Delmonte & A. Bristot (eds.), *Aspetti computazionali in fonetica, linguistica e didattica delle lingue: modelli e algoritmi*, Roma: Esagrafica, 65-81.

Montgomery M. (2007). *The Discourse of Broadcast News: A Linguistic Approach*. London: Routledge.

Moore Mauroux S. (2016). «Patterns and Variation in the Weather Forecast: Can Prosodic Features be Predicted Too?», *Anglophonia* [<https://doi.org/10.4000/anglophonia.755>].

Romano A. (2018). «Indici di connessione del parlato mediatico nelle previsioni del tempo nazionali in inglese e italiano». In F. Bermejo Calleja & P. Katelhön (eds.), *Lingua parlata. Un confronto fra l'italiano e alcune lingue europee*, Berlin: Peter Lang, 57-91.

APPENDIX

Table 1 – Corpus of CNN's weather forecasts.

SPEAKER	DATE	DURATION [s]
Pedram Javaheri	August 10, 2022	121,173
Derek Van Dam	September 16, 2022	164,779
Pedram Javaheri	October 31, 2022	109,888
Derek Van Dam	November 5, 2022	127,381
Pedram Javaheri	November 14, 2022	141,227
Derek Van Dam	November 24, 2022	168,277
Britley Ritz	December 5, 2022	169,259
Britley Ritz	February 13, 2023	127,381

Table 2 – Corpus of RTVE’s weather forecasts.

SPEAKER	DATE	DURATION [s]
Ana de Roque	August 6, 2022	133,035
Andrés Gómez	September 17, 2022	122,023
Andrés Gómez	October 8, 2022	125,312
Nùria Seró	October 31, 2022	140,715
Andrés Gómez	November 14, 2022	132,523
Marc Santandreu	November 26, 2022	123,072
Nùria Seró	February 25, 2023	153,152
Marc Santandreu	February 27, 2023	126,823

Differenze fonetiche generali tra voci femminili e voci maschili*

Antonio Romano

Laboratorio di Fonetica Sperimentale "Arturo Genre", Università di Torino

1. Introduzione

Sappiamo dai manuali di fonetica e da testi specialistici, anche recenti, che variabili come l'altezza o il timbro di una voce dipendono da caratteristiche anatomiche come la lunghezza delle pliche vocali o le dimensioni delle cavità sopraglottidali (contributi vari in Schindler O. 2009; Henrich Bernardoni 2021).

Non ci sorprende quindi che a laringe con cartilagini tiroidee ben pronunciate (a cui corrispondono spazi laringei di maggiore estensione) corrispondano voci generalmente più gravi, cioè con valori di frequenza fondamentale (f_0) più bassi.

È evidente quindi che, al di là dei condizionamenti culturali e delle abitudini personali, voci di individui che dipendono da fattori biologici / ormonali di tipo maschile tendono ad avere estensioni vocali in intervalli di frequenza più bassi di individui biologicamente femminili¹.

Per comprendere questo punto, occorre approfondire alcune semplici caratteristiche della produzione dei suoni linguistici alla luce di: 1) la teoria sorgente-filtro (interazione tra caratteristiche laringee della voce ed effetti di risonanza dovuti all'articolazione dei suoni nel condotto vocale, Fant 1960), 2) l'osservazione delle modalità di sviluppo della curva di f_0 nel corso degli enunciati (intonazione di frase), 3) lo studio contrastivo tra modelli diversi di dispersione statistica dei valori individuali di f_0 (ispezio-

¹ Sappiamo, sin da Bralley *et alii* (1978) (ma v. anche interventi di G. Motta e D. Cossu), che la somministrazione di testosterone determina una maggiore crescita della laringe (con allungamento della glottide) provocando effetti di abbassamento generale di f_0 nei transessuali gino-androidi (ai quali è d'uso riferirsi come parlanti *AFAB = assigned female at birth*). Da questi stessi studi si apprende che l'opposto non è facile da conseguire nel caso dei transessuali andro-ginoidi (*AMAB = assigned male at birth*) somministrando estrogeni. Ma, al di là dei trattamenti endocrinologici, una compensazione si può ottenere con un'adeguata riabilitazione vocale, consistente in attività che non devono prevedere **solo** un innalzamento generale di f_0 (cfr. Panzeri & Falciati 2003; per altri spazi linguistici v. Spencer 1988).

* Il presente contributo riassume alcuni contenuti presentati nel corso del mio intervento tenuto al corso di formazione «La voce parlata e cantata della persona transgender» organizzato a Torino da *Voce-InForma* nei giorni 2 e 3 dicembre 2023.

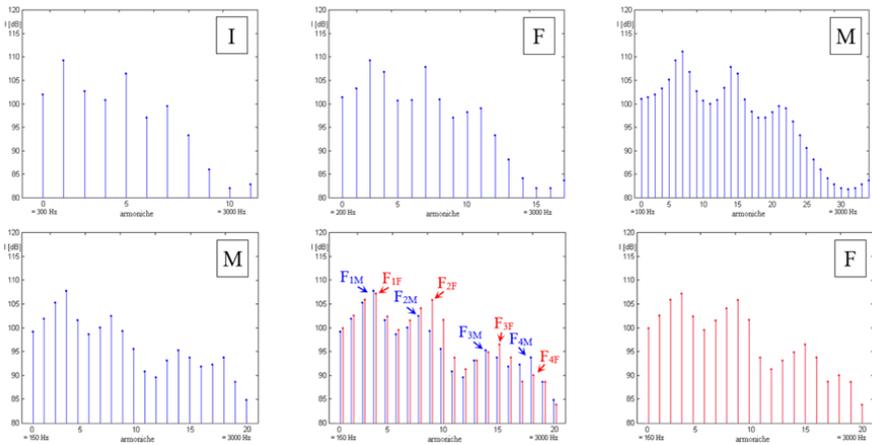


Fig. 1. In alto: strutture armoniche fittizie per profili timbrici simili di voci di tonalità diversa (di tipo infantile, I con $f_0 = 300$ Hz, femminile, F con $f_0 = 200$ Hz, maschile, M con $f_0 = 100$ Hz). In basso: profili timbrici diversi per voci maschili (M, a sinistra) e femminili (F, a destra) a parità di f_0 ($= 150$ Hz) (al centro: confronto tra i due profili M e F ravvicinati).

ne di curve di *densità di probabilità*, *ddf*).

Riguardo al primo punto (v. §3), ho avuto occasione di chiarire, con gli esempi riportati in Fig. 1 (in alto, cfr. Ferrero *et alii* 1979: 77), la minore definizione dello spettro acustico (decomposizione armonica di Fourier) di suoni prodotti da voci infantili (I, in alto a sinistra), femminili (F, in alto al centro) o, comunque, più acute, rispetto a voci maschili (M) caratterizzate da valori di f_0 più bassi e conseguente *pattern* armonico più fitto (in alto a destra). D'altra parte, come si vede dai grafici della stessa figura (in basso), anche a parità di f_0 e di qualità timbrica percepita, voci di tipo M o F presentano schemi formantici lievemente diversi (con valori di F_2 o F_3 , ma spesso anche delle altre formanti,

generalmente più alti nel caso di queste ultime, v. §3) (Fig. 1).

Per le voci maschili ciò vale prevalentemente nelle posizioni degli enuncianti in cui raggiungono valori più gravi, ma in queste posizioni non si può escludere che anche alcune voci femminili raggiungano limiti inferiori simili. A questo proposito vale la pena ricordare che non sono per nulla normativi campioni (anche di diverse centinaia di parlanti) se basati sulle caratteristiche di pronuncia di vocali isolate (sostenute per pochi secondi, per la realizzazione delle quali si determinano assetti in molti casi innaturali nel corso di produzioni decontestualizzate². Si tratta di dati li-

² A questo riguardo si vedano i confronti tra voci di diversa età nel campione di Romano & De Iacovo (2021), v. dopo. Le differenze che

mitati, che non tengono conto degli intervalli esplorati nel corso di compiti linguistici più generali e che, riferendosi esclusivamente a specifiche popolazioni, non sono estendibili ad altri.

Non soddisfano neanche misurazioni di estensioni basate su pochi dati, senza considerare la variazione interna agli enunciati, così come restano generiche e abusive considerazioni del tipo «le voci degli uomini presentano un maggior numero di toni discendenti» (citata da altri lavori in Panzeri & Falciani 2003: 213)³.

In un manuale di prosodia tra i più influenti pubblicati negli ultimi anni in Italia si trova una definizione di

appaiono in questo campione di voci italiane presentano effetti poco sistematici su media e dispersione dei valori: i dati vocali dei parlanti considerati meriterebbero inoltre approfondimenti anche in termini di *jitter* e *shimmer*, cioè di leggera variazione di frequenza o di ampiezza in cicli glottidali successivi.

³ In queste condizioni, nel caso dell'inglese americano (dell'area di Toronto) ad es. si registra che «Older females had a lower f_0 [...] than younger females, but younger and older males did not differ. Shimmer increased with age for males, but neither jitter nor shimmer increased with age for females, whereas noise measures were similar for both ages» (Goy *et alii* 2013). Su simili affermazioni, insieme ad altre piuttosto malferme, si basano lavori come quello di Gucciardo (2007) e Gucciardo *et alii* (2010), i quali – al di là di ottime intuizioni e buoni riferimenti culturali – propongono argomenti spesso esposti con formulazioni impressionistiche e senza riferimenti autorevoli.

estensione tonale come *pitch range*, riferito ai valori minimi e massimi «di un contorno intonativo» (Soriano 2006: 42), laddove per definire la distribuzione dei valori di f_0 in produzioni con diversi livelli di attivazione si notano differenze non trascurabili (Romano 2000) ed è quindi indispensabile riferirsi a un *pitch range* ottenuto mediante valutazioni che si riferiscano a più enunciati e a diverse condizioni di produzione⁴.

D'altra parte, è molto limitante riportare solo indicativamente valori medi, minimi e massimi di f_0 per distinguere voci maschili, femminili o infantili (riferendosi, come nell'esempio ideale proposto sopra, e come fanno alcuni manuali, a valori medi fittizi di tipo 100, 200 e 300 Hz oppure a valori di media e deviazione standard, $dev.st = \sigma$). Come mostrerò, infatti, sulla scia di ricerche passate in rassegna in lavori precedenti, i valori di f_0 si disperdono concentrandosi in alcuni intervalli di maggiore «comodità» della voce, ovvero in intervalli di altezza relativa nei quali occorre soffermarsi nel rispetto di vincoli accentuali, intonativi ed espressivi (cfr. Patterson & Ladd 1999). E ancora, come

⁴ L'intervallo di variazione di f_0 nelle voci è studiato in numerosi lavori in cui, oltre che come *pitch range* e *pitch span*, è indicato anche come *pitch compass* (v. Jassem 1971). Sulla distinzione tra «estensione» e «tessitura» v. Romano *et alii* (2012).

mostrano esempi di voci maschili di estensione molto acuta e voci femminili particolarmente gravi (si pensi ad es. a figure televisive note, come il giornalista Mario Giordano e la conduttrice Maria De Filippi), non è neanche scontato che individui biologicamente di un determinato sesso presentino voci attribuibili inequivocabilmente a quel sesso.

2. Maggiore variabilità nell'altezza delle voci femminili (ed effetto del sesso dell'ascoltatore)

Un parametro che contribuisce a influenzare l'identificazione del sesso di un parlante è anche il sesso dell'ascoltatore (Weston *et alii* 2015)⁵. Ad es. è più facile discriminare tra voci maschili in termini di identificazione (tanto per le donne quanto per gli uomini), mentre nel caso di voci ambigue le

ascoltatrici tendono più spesso a identificarle con voci maschili⁶.

Tuttavia, conoscendo le non linearità della percezione uditiva (che si riflettono sulle scale musicali di tipo logaritmico diffuse in molte culture), è intuitivo attendersi che sensazioni di altezza di scarti melodici simili in una scala più grave corrispondano a variazioni di maggiore estensione in un *range* più alto (si vedano in Fig. 2 le diverse modulazioni di due voci che producono enunciati corrispondenti a uno stesso contenuto frasale con grado di attivazione simile).

Nell'esempio di figura, al di là di due diverse modalità di accentazione che si manifestano con picchi di f_0 più «accentuati» nel caso della voce femminile, le due voci percorrono lo spazio della loro tessitura con variazioni di estensione diversa. Fatte salve le mancate misure (errori di misurazione dell'algoritmo), la voce maschile in questo caso risulta compresa tra circa 103 e 183 Hz, mentre quella femminile (ignorando le false rilevazioni in coda all'ultimo vocoide) si estende da 176 a 307 Hz.

Partendo da queste considerazioni (e dalle limitazioni che prefigurano) proviamo a vedere come sia possibile classificare rigorosamente queste voci anche solo in riferimento alla loro variazione

⁵ A questo riguardo, un altro aspetto rilevato in Joassin *et alii* (2011), riguarda l'associazione tra voci e volti che costituisce un elemento dell'identificazione: «Human social interactions are shaped by our ability to recognise people. Faces and voices are known to be some of the key features that enable us to identify individual people, and they are rich in information such as gender, age, and body size, that lead to a unique identity for a person. A large body of neuropsychological and neuroimaging research has already determined the various brain regions responsible for face recognition and voice recognition separately, but exactly how our brain goes about combining the two different types of information (visual and auditory) is still unknown».

⁶ Un esperimento simile, limitato al parlato letto, ma basato su indici acustici segmentali, è quello di Munson *et alii* (2006).

di f_0 in questi due enunciati (v. Fig. 3).

Ora, è vero che la f_0 media del primo è a 146 Hz (con $\sigma = 19$ Hz) e che l'estensione del suo intervallo di variazione è di 80 Hz (da 103 a 183) ed è pur vero che la f_0 media della seconda è a 241 Hz (con $\sigma = 41$ Hz) e che il suo *pitch range* è di 131 Hz (da 176 a 307). Tuttavia rileviamo come le due voci si siano soffermate più tempo intorno a certi valori che nel primo caso sono stati a 132, 154 (soprattutto) e 178 Hz mentre nel secondo caso si sono dispersi mag-

giormente, concentrandosi attorno a 180, 208, 246, 272 (soprattutto) e 306 Hz. Certo, si tratta di pochi valori, ma ciò che emerge dall'analisi quantitativa di un gran numero di produzioni di uno stesso locutore (come ho potuto mostrare sin da Romano 2000) si caratterizza comunque per una dispersione quasi mai riconducibile a una *ddp* gaussiana e quindi non rappresentabile con valori quali la media, la deviazione standard (σ) o, ancor meno, il minimo e il massimo (o l'intervallo tra questi).

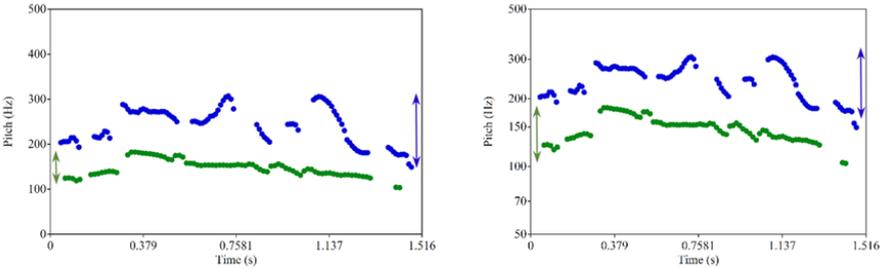


Fig. 2. Profili melodici simili per voci di estensione diversa rappresentati su scale diverse. A sinistra, scala lineare: la voce di tipo femminile (blu) si presenta con un'estensione apparentemente maggiore di quella di tipo maschile (verde). A destra, scala logaritmica: il range di variazione di f_0 nei due casi appare congruente (le variazioni melodiche della voce di tipo maschile risaltano in modo più evidente e in accordo con le sensazioni di altezza che producono).

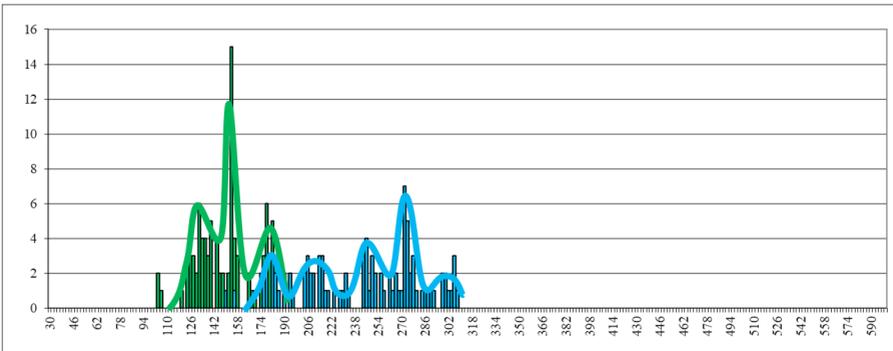


Fig. 3. Istogrammi di densità di probabilità con passo 2 Hz per le due voci di Fig. 2.

Nell'articolo Hudson *et alii* (2007) le f_0 medie di 100 parlanti maschi britannici tra i 18 e 25 anni sono state analizzate statisticamente, mostrando una distribuzione tra i 90 e i 150 Hz. Si tratta di una distribuzione di medie (100 parlanti -> **100** valori) e non di una *ddp* cumulativa.

Nell'articolo di Romano & De Iacovo (2021), invece, di due gruppi di 58 parlanti maschi (29 di un gruppo più giovane, 18-25, e 29 di uno più anziano, 25-50) si sono valutate le dispersioni di valori di f_0 di ciascuno di essi collasate in *ddp* cumulative di più di **150000** valori. La prima ha presentato una f_0 media di 141 Hz ($\sigma = 44$ Hz) e la seconda una f_0 media di 126 Hz ($\sigma = 44$ Hz). Si tratta di *ddp* simili a gaussiane ma piuttosto eccentriche (la moda è di 15-20 Hz più bassa). Si riporta in Fig. 4 la somma delle due *ddp* (con l'aggiunta di valori accuratamente rilevati tra 76 e 80 Hz). Si tratta di circa di **312000** valori con

media a 134 Hz e moda a 114 Hz.

Le voci femminili di quest'esperimento, rispettivamente quelle di 78 e 35 parlanti (più di **400000** misurazioni nel primo caso), presentano invece f_0 media 209 Hz ($\sigma = 49$ Hz) e f_0 media 202 Hz ($\sigma = 52$ Hz). Tuttavia la forma delle *ddp* non somiglia a una gaussiana, quanto piuttosto alla sovrapposizione di due «campane» pseudo-gaussiane: una più bassa, con moda a 80 Hz, e una più alta, con moda a 198 Hz, nel primo caso e una più bassa e un po' meno distinguibile, con moda sempre attorno a 80 Hz, e un'altra con moda a circa 180 Hz. I minimi delle due campane più alte si situano tra 100 e 130 Hz mentre i massimi, raggiunti con code molto smorzate, si trovano attorno a 380 Hz. Non si tratta quindi di estensioni valutabili rilevando i minimi e i massimi di singole produzioni che non avrebbero fatto apparire la composizionalità delle due *ddp*. Sebbene la campana più grave sia caratterizzata da

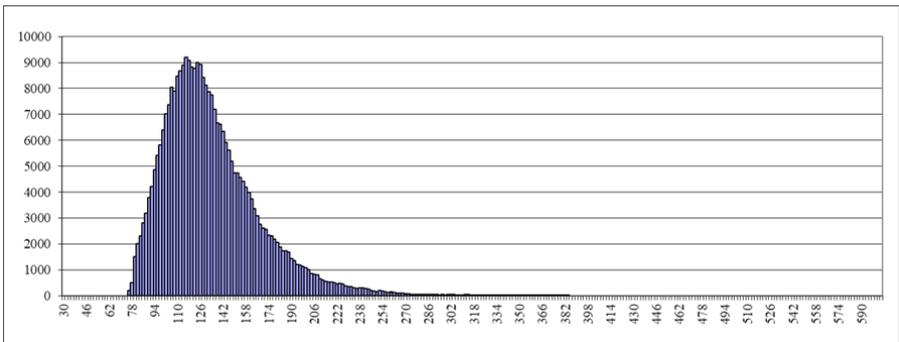


Fig. 4. Istogramma di densità di probabilità con passo 2 Hz per i due gruppi di voci maschili di Romano & De Iacovo (2021) [dati unificati e verificati manualmente tra 76 e 80 Hz].

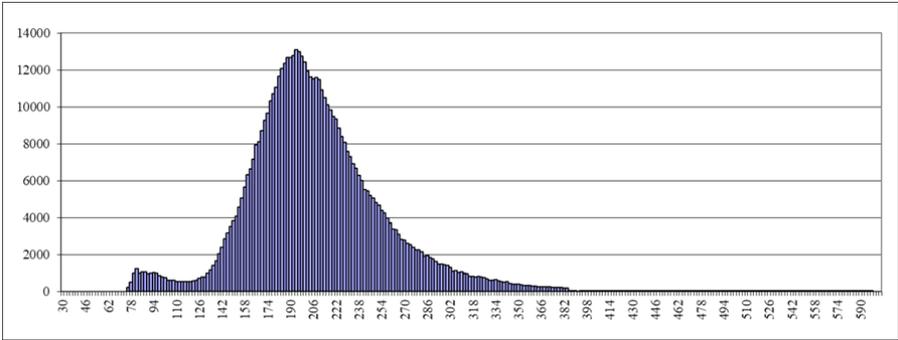


Fig. 5. Istogramma di densità di probabilità con passo 2 Hz per le voci femminili di Romano & De Iacovo (2021) [dati unificati e verificati manualmente tra 76 e 80 Hz].

un numero inferiore di valori, le due campane sembrano rispondere al ricorso da parte di questi gruppi di parlanti a due distinti meccanismi di produzione della voce: uno (f_0 di Roubeau *et alii* 2009 e Henrich Bernardoni 2021), usato meno frequentemente, in un registro di voce cricchiata e un altro (M1), ampiamente esplorato – ma, appunto, con minimi tra 100 e 130 – nel registro di voce modale. Anche in questo caso le due *ddp* sono state unificate e i valori tra 76 e 80 Hz verificati manualmente. Si riporta quindi in Fig. 5 la *ddp* risultante, basata su circa di **584000** valori,

con media a 207 Hz e moda a 192 Hz.

Dal confronto diretto tra le due *ddp* risultanti si può osservare come le voci ricoprano in buona misura lo stesso spazio di variazione di f_0 nelle basse frequenze (v. Fig. 6). Per questi parlanti di italiano, non esiste un modo netto di separare le due *ddp*: tra i 75 Hz e i 265 entrambe le voci coesistono, anche se con diversa probabilità. Fissando una soglia arbitraria di 1000 valori, l'intervallo di maggior ricoprimento si potrebbe circoscrivere tra i 130 e i 200 Hz, con equiprobabilità a circa 155 Hz: intorno

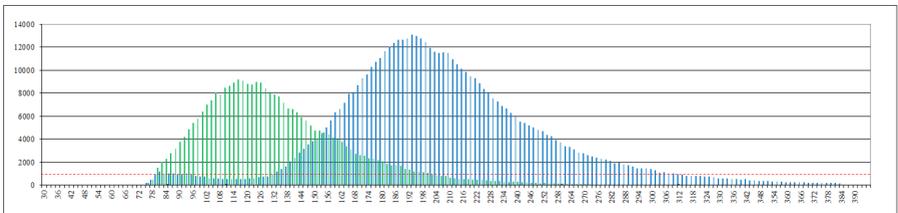


Fig. 6. Istogrammi di densità di probabilità con passo 2 Hz per le voci maschili e femminili di Romano & De Iacovo (2021).

a questi valori (per i quali sono state misurate migliaia di valori, > 4000!) una voce ha le stesse probabilità di essere considerata maschile o femminile (fatti salvi gli altri fattori anticipati)⁷.

3. Caratteristiche timbriche di voci di tipo maschile e femminile

Nonostante le numerose omissioni teoriche e i difetti procedurali che affliggono il loro studio, Panzeri & Falciati (2003) giungono a conclusioni

⁷ Osservo tuttavia che anche tra gli 80 e i 100 Hz la soglia di 1000 valori è raggiunta: a 85 Hz le voci di tipo femminile hanno la metà delle probabilità di muoversi in questo *range* rispetto a quelle di tipo maschile (riconquistando però l'equiprobabilità nei rari casi in cui la voce si localizza sotto gli 80 Hz, cioè nel registro di cricchiato, meccanismo M0). Alla luce di tutto ciò, non deve quindi sorprendere se un valore medio di 155 Hz, inferiore a 160 Hz, faccia «già identificare come femminile la voce di un transessuale *male-to-female*»: uno scarto tra un valore medio di f_0 e una soglia già di per sé non è significativo senza indicazione della varianza o, appunto, dell'estensione. In questo caso, inoltre, il valore assunto come indice presenta uno scarto risibile rispetto a una soglia arbitraria (per di più attribuita a Wolfe *et alii* 1990 che, in riferimento all'inglese, cioè per una comunità di parlanti ben precisa, stabiliscono un limite a 155 Hz!). Sorprende invece che un test di attribuzione di genere, condotto su 8 voci italiane e isolando solo produzioni di /a/ (Panzeri & Falciati 2003: 215), abbia potuto confermare la validità di una simile soglia. Approfito per segnalare che la terminologia riguardo a questi temi è quanto mai soggetta a rapida obsolescenza (basti pensare al cambiamento di titolo che si è reso necessario per la terza edizione, 2019, di Adler *et alii* 2006).

verificabili. Partendo da una premessa sostanzialmente corretta, conforme a una visione semplicistica della teoria-sorgente-filtro, «è stato dimostrato che quando viene eliminata la differenza della frequenza fondamentale, le altre formanti [sic] aiuterebbero l'ascoltatore a percepire la mascolinità o la femminilità della voce» (p. 214). Si tratta evidentemente di un campo esplorato da numerosi autori e con varie metodologie. Si vedano, ad es., Halberstam & Raphael (2004), e riferimenti citati, che ribadiscono l'importanza dell'interazione tra f_0 e l'identificazione dei *pattern* vocalici (e viceversa), e Camargo *et alii* (2012), che indicano la pendenza spettrale come altro fattore dirimente.

D'altra parte l'interazione è nota anche da studi di diversi ricercatori italiani (Barillari *et alii* 1995, Ferrero *et alii* 1995, Maisano 1996, Calamai 2003) che hanno fornito i materiali di riferimento, partendo da Fant (1978) e fonti citate (v. in particolare §7, p. 81 «Male-female difference», v. Fig. 7a), oltre che da Ferrari-Disner (1980), tra i primi a proporre metodi di valutazione oggettivi anche in termini sensoriali.

Oltre a differenze acustiche specifiche, gli effetti percettivi di F_2' (cioè dei valori efficaci di F_2 pesati da altre componenti energetiche presenti nello spettro acustico) sono stati studiati in Ferrero *et alii* (1996) e, nell'ottica di lavori di normalizzazione, in Pätzold &

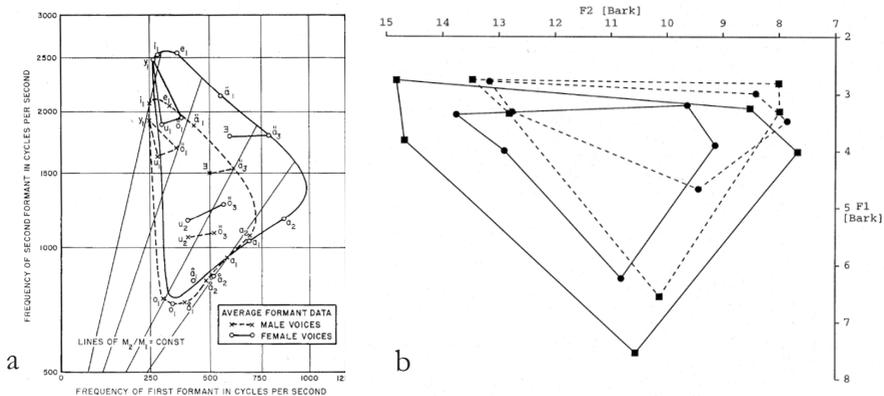


Fig. 7. Differenze tra dispersioni di valori formantici (F_1 e F_2) per uomini e donne di lingua svedese (6a) e tedesca (6b) [tratterispettivamente da Fant (1978) e Pätzold & Simpson (1997)].

Simpson (1997) (v. Fig 7b)⁸. Il metodo di confronto tra voci sulla base della superficie occupata dai poligoni definiti dalla dispersione dei valori formantici è stato poi studiato per l'italiano (fiorentino e pisano) da Calamai (2004) e Calderone & Calamai (2007)⁹. La valutazione dello spazio articolatorio si fa

in questi casi considerando la superficie dello spazio acustico-articolatorio dei diversi parlanti e valutando un coefficiente che stabilisce la relazione di proporzione: i poligoni di tipo maschi-

⁸ Citando Peterson & Barney (1952), Ferrero *et alii* (1979: 78) sottolinea come si trovino differenze per i valori formantici di una «stessa vocale» pronunciata da un uomo, da una donna o da un bambino «Le differenze uomo-donna per F_1 variano tra il 5% e il 30%, per F_2 tra il 10% e 25% e per F_3 tra il 10% e il 20%». Riguardo a F_2 , già in Ferrero *et alii* (1979: 124) si definisce una «seconda formante percettiva calcolata [...] a partire dai valori delle prime quattro formanti acustiche». In riferimento a Bladon & Fant (1978) si aggiunge poi che questi «hanno fornito una definizione più perfezionata di F_2 nell'ambito di un modello a due formanti delle vocali cardinali» (pp. 124-125).

⁹ L'importanza di queste conoscenze sfugge a molti ricercatori di ambito medico-logopedico che non accedono alla lettura fondamentale del campo fonetico e/o non si riferiscono alle ricerche che sono state condotte in Italia per decenni facendo luce su tutti questi aspetti e contribuendo a definire una disciplina. In mancanza di questi riferimenti, lavori come quelli menzionati sopra potrebbero presentare risultati fuorvianti o comunque non rappresentativi dello stato di conoscenze raggiunto in un determinato periodo all'interno di una comunità scientifica (alla quale si approssimano acriticamente, ripartendo da conoscenze relative ad altre epoche e ad altri spazi linguistici, mentre invece, nell'estrema semplicità del loro esperimento, passando in rassegna diversi lavori, Sulpizio *et alii* 2015 hanno messo in evidenza l'essenzialità della dipendenza dalla lingua).

le, essendo associati a condotti vocali stimati mediamente di lunghezza 17,5 cm (Giannini & Pettorino 1992), hanno minore estensione (generalmente in entrambe le dimensioni definite da F_1 e F_2) rispetto a quelli di tipo femminile. In questi casi infatti le ridotte dimensioni antropometriche definiscono una minore lunghezza dello spazio risonanziale (15,5 cm in stime medie piuttosto grossolane) a cui corrisponde (essendo la frequenza inversamente proporzionale alla lunghezza d'onda) un «poligono formantico» più esteso.

Esempi di confronto tra sistemi vocalici di parlanti ben differenziati (per l'area barese) sono proposti anche in Romano & Manco (2004). I valori della frequenza delle prime due formanti discussi in questo studio (v. Fig. 8) sono compatibili con quelli presentati da Ferrero *et alii* (1979)¹⁰.

¹⁰ I grafici di riferimento per voci maschili e femminili si trovano alle pp. 122-123 (v. Appendice D), insieme a molte considerazioni che – ignorando gli studi di Fant e quello di Pätzold & Simpson (1997) – devono essere sfuggite a Panzeri & Falciani (2003). Pur dedicando a F. Ferrero il loro articolo (e avendo inserito tra i riferimenti proprio questo manuale) non sembrano infatti aver preso in considerazione i passaggi che li avrebbero scoraggiati a basare le loro valutazioni su F_2 e F_3 (e non su F_1 o F_2). In particolare, in Ferrero *et alii* (1979) si dice che il diagramma di esistenza F_3 – F_2 «dimostra l'inadeguatezza di F_3 a differenziare tra loro, altrettanto nettamente di F_1 – F_2 , le sette vocali italiane» (pp. 123-124). Vi si legge anche: «è stato dimostrato che le formanti superiori alle prime due aggiungono poco» (nella definizione del timbro).

L'area dei poligoni definiti dai centroidi delle dispersioni di valori presentate dai 4 speaker considerati è di circa 1089000 Hz² e 1024000 Hz² per i due locutori (M) e di circa 1474000 Hz² e 1317000 Hz² per le due locutrici (F). Si hanno quindi coefficienti compresi tra 1,3 e 1,4 per passare da una dispersione tipica di queste voci maschili a quella delle voci di tipo femminile che riproducono contrasti simili (si vedano le aree di esistenza dei fonemi)¹¹.

Come si vede dai grafici di Fig. 9, tuttavia, l'espansione deve interessare diversamente i vari timbri, dato che ad es. per le realizzazioni di /u/ i valori delle formanti misurati per i locutori e le locutrici risultano molto simili, mentre è soprattutto per /i/ e per /a/ che si creano delle divergenze acusticamente molto rilevanti.

¹¹ Per offrire un altro riferimento, rispetto ai valori qui presentati, riassumo in Appendice I quelli prototipici desumibili dai grafici di Ferrero *et alii* (1979: 122-123). Con quei valori, le superfici dei poligoni valgono rispettivamente circa 790000 Hz² (M) vs. 1076000 Hz² e il valore del coefficiente di conversione spaziale M → F risulta 1,36 (compatibile con quanto visto sopra). L'applicazione di questo coefficiente richiede però di lavorare a partire da diagrammi di esistenza simili, mentre naturalmente questo non è facile da conseguire in un Paese come l'Italia caratterizzato da forte variazione dialettale (in termini geografici). I valori rilevati per F_2 di /a/ in Panzeri & Falciani (2003) sembrano quindi corrispondere ai valori medi che si riscontrano in voci tipizzate come M o F e sono in accordo con i valori riportati anche negli studi citati sopra (apparsi per la verità in anni seguenti). Alla

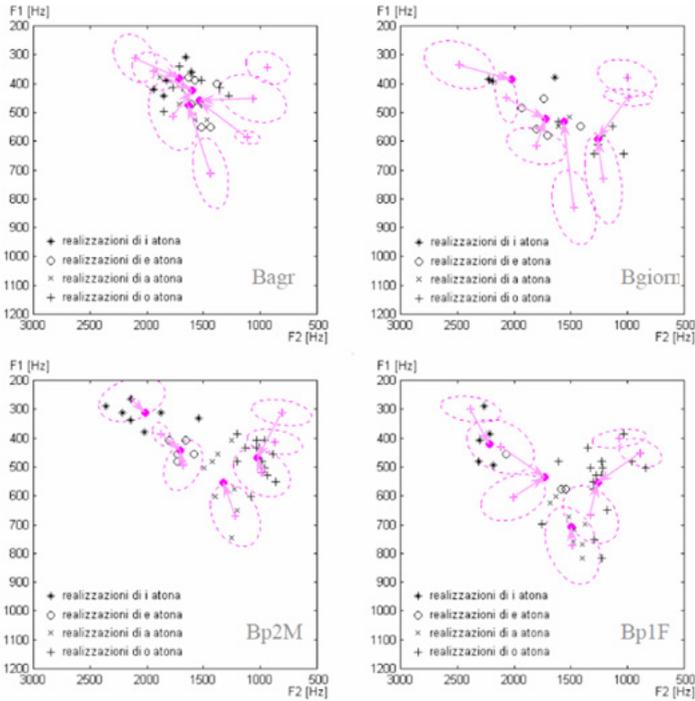


Fig. 8. Dispersioni di valori formantici (F_1 e F_2) per uomini e donne baresi (un agronomo e una giornalista, in alto, e due studenti, M e F, in basso) [tratte da Romano & Manco 2004].

Il lavoro di mappatura che ristabilisce, a un livello solo teorico, le somiglianze che emergono in termini

differenza di F_2 per /a/ tra M (ca. 1250 Hz) e F (ca. 1450) rispondono mediamente bene i valori medi ricavabili dai valori di F_2 tabulati per i gruppi di controllo ($F_{2M} = 1241$ Hz e $F_{2F} = 1503$). Tuttavia la dispersione di valori per le voci maschili (σ per $F_{2M} = 226$ Hz vs. σ per $F_{2F} = 145$ Hz) differisce notevolmente da quella osservata negli studi citati sopra (v. anche le voci maschili di Fig. 7, con F_{2M} media = 1332 Hz e $\sigma = 138$ vs. F_{2F} media = 1478 con $\sigma = 149$) e testimonia di una variazione che interessa potenzialmente anche rilevanti differenze nelle rese fonetiche (alcuni speaker di controllo hanno valori di F_2

perceptivi tra sistema di contrasti realizzato da voci maschili e femminili deve quindi prevedere una pondera-

per /a/ che fanno pensare a rese di tipo [a]. L'uniformità nella variazione diatonica perseguita invece negli studi riassunti nella rassegna di Calamai (2004) (o nei dati riportati nei diagrammi di Fig. 7), oltre a limitare le differenze tra M e F per una maggiore convergenza intra-comunità si legano anche a una variazione più contenuta nei valori medi nelle realizzazioni di questa vocale rappresentata da allofoni simili, laddove una pratica riabilitativa dovrebbe poter considerare anche le specificità geografica del campione di riferimento, assicurando anche il passaggio da un allofono all'altro (v. dopo).

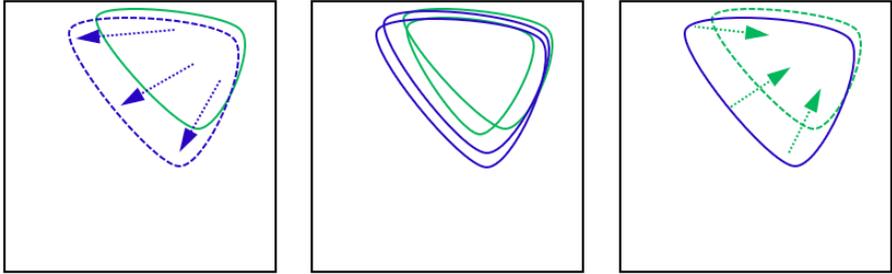


Fig. 9. Spazi vocalici idealizzati di voci maschili (M, verde) e femminili (F, blu) (in riferimento a quelli definiti da locutori e locutrici di Fig. 7). Il passaggio M>F dovrebbe prevedere un'estensione dello spazio vocalico (come nel grafico a sinistra), mentre il passaggio F>M dovrebbe prevedere una contrazione (come nel grafico a destra).

zione di questo coefficiente (della frazione che vogliamo restituire alle due dimensioni).

Ovviamente questo può ottenersi, diversamente, attraverso una mappatura su un piano formantico «deformato» in modo da far coincidere meglio le aree per mezzo di misure indirette (sottraendo costanti e introducendo coefficienti).

Diverse soluzioni erano già state proposte sin dagli anni '80, al punto che per rispecchiare «meglio il modo in cui suonano all'orecchio», Uguzzoni (1992) suggeriva di lavorare con scale di rappresentazione (in bark) in cui si riportano i valori di *spread* (cioè $F_2 - F_1$) vs. *flat* ($F_2 + F_1$), valutando anche, in un'ottica di valutazione congiunta tra valori risonanziali e componenti armoniche, « $F_1 - F_0$ » (Uguzzoni 1992: 32-33).

Tuttavia, come hanno mostrato Ferrero *et alii* (1995: 263, sulla base dei valori della tabella di p. 261 riprodotta

in Appendice II), «la normalizzazione “ $F_2 - F_1$ in Bark” è utile per le vocali posteriori /u, o, ɔ/ e la centrale /a/ per le quali la differenza tra VM e VF non è significativa ($p < 005$). Non è così per le vocali anteriori /i, e, ε/ per le quali la distanza tra F_2 e F_1 nelle vocali femminili è significativamente maggiore rispetto a quelle maschili»¹².

Ecco quindi che sono soprattutto le vocali anteriori che, nel passaggio andro-ginoide, dovrebbero essere articolate in modo da aumentare le dimensioni dello spazio di articolazione

¹² Sempre al livello teorico, per una rappresentazione maggiormente compatibile con valutazioni sensoriali, Ferrero *et alii* (1995: 263, v. ultimo diagramma in Appendice II) propongono di basarsi su diagrammi che abbiano in ascissa la differenza $F_2 - F_1$, in bark, e in ordinata $F_1 - F_0$, in bark. A questo riguardo, si noti che una delle formule usate per il passaggio da Hz a bark è oggi quella di Traunmüller: $f_{\text{bark}} = [(26,81 * f_{\text{Hz}}) / (1960 + f_{\text{Hz}})] - 0,53$.

(v. Fig. 9 a sinistra), e nel passaggio gino-androide dovrebbe invece ricentrarsi in uno spazio acustico contratto con valori di formanti più bassi (v. Fig. 9 a destra)

Come sia possibile perseguire questi due opposti scopi resta una materia complessa. L'eventuale obiettivo potrebbe essere raggiunto adottando metodi di compensazione articolatoria che risultano idealmente più facilmente risolvibili nel caso $M > F$, per via del fatto che un condotto vocale anatomicamente più lungo può essere sfruttato in una sua dimensione ridotta. Risulta invece oggettivamente più difficile nel caso $F > M$ abbassare le formanti aumentando la lunghezza del condotto, anche se resta vero che una disposizione più avanzata nell'articolazione dei suoni anteriori, così come una maggiore protrusione labiale o una maggiore apertura ottenuta con l'abbassamento della mandibola (nonché della radice della lingua e della laringe), potrebbero offrire un *setting* articolatorio maggiormente orientato verso questi risultati. In questa condizione però abbiamo già visto che i trattamenti farmacologici sortiscono effetti migliori sulla laringe. E, infatti, per tornare ai limiti di una teoria sorgente-filtro banalizzata a una separazione totale dei contributi di laringe e cavità sopraglottidali, giocherebbe (e gioca) un ruolo fondamentale un lavoro sull'assetto del-

la laringe, condizionata in modo da produrre e muovere armoniche tali da stimolare le risonanze rispettivamente più acute o più gravi nell'inviluppo risonanziale del condotto vocale (v. Fig. 1) con dinamiche tali da assicurare uno spazzolamento (*sweeping*) di un *range* di f_0 corrispondente a quello del modello di voce desiderato (v. Fig. 6) purché compatibile con le proprie condizioni anatomiche.

4. Conclusioni

Con il breve intervento riassunto in questo contributo ho inteso ristabilire un ordine di priorità in merito alla consultazione di studi scientifici precedenti sulle caratteristiche vocali di parlanti maschi e femmine di varietà di italiano. Lo studio non pretende di essere esaustivo, ma, sulla base di un certo numero di esempi e riflessioni, ha mostrato come non sia prioritaria la ricerca di metodi di studio che – ignorando lo stato dell'arte in materia – ripartano da zero e fondino la loro conoscenza su metodi sviluppati per le comunità linguistiche maggiormente rappresentate nelle riviste internazionali.

I dati riportati e le discussioni scaturite dall'osservazione di questi mostrano come spesso – purtroppo ancora in tempi recenti – questi condizionamenti pregiudichino ricerche condotte in ambito medico-logopedico, nel quale è particolarmente

sconsigliabile che i gruppi di ricerca che lavorano su questi temi sensibili non partano da un'adeguata formazione interdisciplinare. Questa si è già rivelata imprescindibile sin dai primi lavori in questo campo avviati in Italia in seno al *Gruppo di Fonetica Sperimentale* dell'*Associazione Italiana di Acustica* e proseguiti poi anche all'interno dell'*Associazione Italiana di Scienze della Voce* e, in parte, del *Gruppo di Studio sulla Comunicazione Parlata* della *Società di Linguistica Italiana*.

Il quadro qui offerto si è soffermato dapprima su valutazioni di misurazioni di f_0 eseguite su parlanti italiani con voci di tipo maschile e femminile appartenenti a un campione di dimensioni mai raggiunte finora. Sulla base dei dati discussi, ha permesso di concludere considerando con maggiore cautela le reali estensioni dei valori assunti da questa variabile prima di incoraggiare gli interessati a spingere la propria voce su valori molto spesso solo stereotipati sulla base di intuizioni non verificate.

In un secondo tempo, l'articolo si è concentrato sul confronto dei timbri oggettivi e delle sensazioni da questi generate in funzione di altri fattori, vincoli strutturali e limitazioni statistiche. Alla discussione dei valori delle formanti studiate da vari autori per voci di tipo maschile e femminile nei sistemi di varie comunità italiane (Firenze, Padova, Bari e Lecce) si è aggiunta una

valutazione delle possibilità di normalizzazione applicate in Italia alla fine degli anni '90. L'obiettivo, oltre che tener conto le varie scale di conversione, utili ai fini di un confronto tra le distinte dispersioni di valori determinate da caratteristiche anatomiche, è ovviamente soprattutto quello di riportare la qualità acustica dei timbri alle possibilità plastiche del condotto vocale.

Rispetto alle ricerche condotte alcuni decenni fa, in riferimento a quelle di studiosi interessati ad altri spazi linguistici, con diversi condizionamenti culturali e socio-fonetici, e grazie alla disponibilità di risultati ottenuti recentemente in lavori sperimentali condotti proficuamente in Italia da specialisti del settore, il ventaglio di soluzioni proponibili è cambiato considerevolmente. Le conoscenze oggi disponibili grazie alla facilità di circolazione delle informazioni e la maggiore diffusione di strumenti che agevolano le misurazioni e le valutazioni statistiche possono infatti rivelarsi utili per comprendere meglio la materia, aiutando a predisporre metodologie efficaci per un *counselling* professionale.

Appendice bibliografica

Adler R.K., Hirsch S. & Mordaunt M./ Pickering J. (2006). *Voice and Communication Therapy for the Transgender/ Transsexual Client*. San Diego: Plural (3^a ed. 2019, > *Voice and Communication Therapy for the Transgender/ Gender Diverse Client*).

- Barillari U., Giannini A., Pettorino M. & Siani M. (1995). «Indici acustici e movimenti articolatori nella voce maschile e femminile». *Studi Italiani di Linguistica Teorica e Applicata*, XXIV/2, 265-276.
- Calamai S. (2003). «Vocali d'Italia. Una prima rassegna». In: P. Cosi, E. Magno Caldognetto, A. Zamboni (a cura di), *Voce Canto Parlato. Studi in onore di Franco Ferrero*, Padova: Unipress, 49-57.
- Calamai S. (2004). «Vocali fiorentine e vocali pisane a confronto». In: F. Albano-Leoni, F. Cutugno, M. Pettorino & R. Savy (a cura di), *Il Parlato Italiano* (Atti del Convegno Naz., Napoli, 13-15 febbraio 2003), Napoli: D'Auria (CD-ROM).
- Calderone B. & Calamai S. (2007). «Lo spazio delle vocali». In: V. Giordani, V. Bruseghini & P. Cosi (a cura di), *Scienze Vocali e del Linguaggio - Metodologie di Valutazione e Risorse Linguistiche* (Atti del III Convegno Nazionale AISV - Associazione Italiana di Scienze della Voce, ITC-IRST Povo - Trento, 29 nov. - 1 dic. 2006), Torriana (RN): EDK, 77-89.
- Camargo Z., Madureira S., Pessoa A. & Rusilo L. (2012). «Voice Quality and Gender: Some Insights on Correlations between Perceptual and Acoustic Dimensions». *Proc. Speech Prosody 2012* (Shanghai, 22-25 maggio 2012), 115-118.
- Fant G. (1960). *Acoustic Theory of Speech Production*. L'Aja: Mouton.
- Fant G. (1978). «Vowel perception and specification». *Rivista Italiana di Acustica*, II(2), 69-87.
- Ferrari-Disner S. (1980). «Evaluation of vowel normalization procedures». *Journal of the Acoustical Society of America*, 67, 253-261.
- Ferrero F., Genre A., Boë L.J. & Contini M. (1979). *Nozioni di fonetica acustica*. Torino: Omega.
- Ferrero F., Magno Caldognetto E. & Cosi P. (1995). «Le vocali al femminile». In: G. Marcato (a cura di), *Dialettologia al femminile. Atti del Conv. Intl. di Studi* (Sappada, 26-30 giugno 1995), Padova: Cleup, 413-436.
- Ferrero F., Magno Caldognetto E. & Cosi P. (1996). «Sui piani formantici acustici e uditivi delle vocali di uomo, donna e bambino». In: A. Peretti & P. Simonetti (a cura di), *Atti del XXIV Convegno Naz. AIA* (Trento, 12-14 giugno 1996), 169-178.
- Giannini A. & Pettorino M. (1992). *La fonetica sperimentale*. Napoli: Ed. Scientifiche Italiane.
- Goy H., Fernandes D.N., Pichora-Fuller M.K. & van Lieshout P. (2013). «Normative Voice Data for Younger and Older Adults». *Journal of Voice*, 27 (5), 545-555.
- Gucciardo A.G. (2007). *Voce e sessualità*. Torino: Omega.
- Gucciardo A.G., Merendi E., Paolillo N.P., Fussi F. (2010). «Sulla voce del transgender /Transessuale. Alcune osservazioni di sesso-vocologia generale

e artistica». In: F. Fussi (a cura di), *La voce del cantante*, VI, Torino: Omega, 503-509.

Halberstam B. & Raphael L.J. (2004). «Vowel normalization: the role of fundamental frequency and upper formants». *Journal of Phonetics*, 32, 423-434.

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

Hudson T., de Jong G., McDougall K., Harrison Ph. & Nolan F. (2007). «F0 Statistics for 100 Young Male Speakers of Standard Southern British English». *Proc. of ICPHS XVI* (Saarbrücken, 6-10 agosto 2007), 1809-1812.

Jassem W. (1971). «Pitch and compass of the speaking voice». *Journal of the International Phonetic Association*, 1, 59-68.

Joassin F., Pesenti M., Muraige P., Verreckett E., Bruyer R. & Campanella S. (2011). «Cross-modal interactions between human faces and voices involved in person recognition». *Cortex*, 47, 3, 367-376.

Kachel S., Radtke A., Skuk V.G., Zäske R., Simpson A.P. & Stefens M.C. (2018). «Investigating the common set of acoustic parameters in sexual orientation groups: A voice averaging approach». *PLOS ONE*, 13(12), e0208686.

Maisano L. (1996). «Sistemi vocalici maschili e femminili a confronto». In: A. Peretti & P. Simonetti (a cura di), *Atti del XXIV Convegno Naz. ALA* (Trento, 12-14 giugno 1996), 183-186.

Munson B., McDonald E.C., De-Boe N.L. & White A.R. (2006). «The acoustic and perceptual bases of judgments of women and men's sexual orientation from read speech». *Journal of Phonetics*, 34(2), 202-240.

Oates J. & Dacakis G. (1983). «Speech Pathology Considerations in the Management of Transsexualism: A Review». *British Journal of Disorders of Communication*, 18(3), 139-151.

Panzeri M. & Falciati L. (2003) «Percezione della voce di parlanti transessuali: Requisiti acustici e percettivi per l'identificazione del genere». In: P. Cosi, E. Magno Caldognetto, A. Zamboni (a cura di), *Voce Canto Parlato. Studi in onore di Franco Ferrero*, Padova: Unipress, 213-218.

Patterson D. & Ladd R. (1999). «Pitch range modeling». *ICPhS99*, 1169-1172.

Pätzold M. & Simpson A.P. (1997). «Acoustic analysis of German vowels in the Kiel Corpus of Read Speech». *Arbeitsberichte des Instituts für Phonetik und digitale Sprachverarbeitung Universität Kiel*, 32, 215-247.

Ralph C. Bralley, Glen L. Bull, Cheryl Harris Gore, Milton T. Edgerton (1978). «Evaluation of vocal pitch in male transsexuals». *Journal of Communication Disorders*, 11(5), 443-449

Romano A. (2000). «Statistiche di frequenza fondamentale per uno stesso locutore in diverse condizioni di produzione». *Atti del 28° Convegno Nazionale dell'Associazione Italiana di*

Acustica (Trani, 10-13 giugno 2000), 249-252.

Romano A. & De Iacovo V. (2021). «Statistiche di f0 per 200 parlanti di italiano». *Bollettino LFSAG*, 8, 21-33 [http://www.lfsag.unito.it/ricerca/phonews/08/8_2.pdf]

Romano A. & Manco F. (2004). «Incidenza di Fenomeni di Riduzione Vocalica nel Parlato Spontaneo a Bari e a Lecce». In: F. Albano Leoni, F. Cutugno, M. Pettorino & R. Savy (a cura di), *Il Parlato Italiano* (Atti del Convegno Naz., Napoli, 13-15 Febbraio 2003), Napoli: D'Auria (CD-ROM).

Romano A., Cesari U., Mignano M., Schindler O. & Venero I. (2012). «Voice Quality» / «La qualità della voce». In: A. Paoloni & M. Falcone (a cura di), *La voce nelle applicazioni* (Atti dell'VIII Convegno dell'Associazione Italiana Scienze della Voce, Roma, 25-27 gennaio 2012), Roma: Bulzoni, 75 (art. int. CD 35 pp.).

Roubeau B., Henrich N. & Castellenigo M. (2009). «Laryngeal Vibratory Mechanisms: The Notion of Vocal Register Revisited». *Journal of Voice*, 23(4), 425-438.

Schindler O. (a cura di) (2009). *La voce: fisiologia, patologia clinica e terapia*. Padova: Piccin.

Spencer L. (1988). «Speech characteristics of male-to-female transsexuals: A perceptual and acoustic study». *Folia Phoniatrica*, 40(1), 31-42.

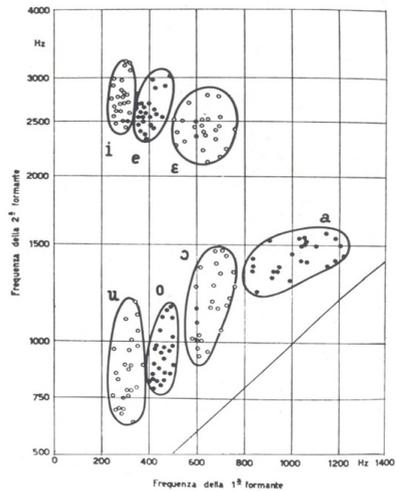
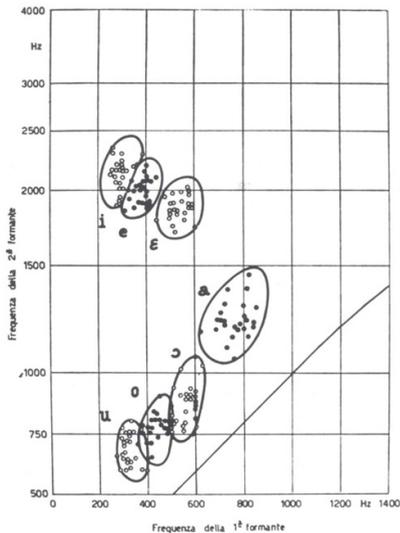
Sulpizio S., Fasoli F., Maass A., Paladino M.P., Vespignani F., Eyssel F. & Benteler D. (2015). «The Sound of Voice: Voice-Based Categorization of Speakers' Sexual Orientation within and across Languages», *PLOS ONE*, 10(7), e0128882.

Uguzzoni A. (1992). «Rappresentazione e classificazione delle vocali in termini uditivi». *Atti delle 3^e Giornate di Studio del GFS* (Padova, 19-20 nov. 1992), 27-35.

Wolfe V.I., Ratusnik D.L., Smith F.H. & Northrop, G. (1990). «Intonation and fundamental frequency in male-to-female transsexuals». *Journal of Speech & Hearing Disorders*, 55(1), 43-50.

Weston Ph.S.J. *et alii* (2015). «Discrimination of voice gender in the human auditory cortex». *NeuroImage*, 105, 208-214.

Appendice I



Diagrammi F1-F2 riprodotti da Ferrero *et alii* (1979: 122-123) (a loro volta ripresi da quelli pubblicati alle pp. 11 e 12 di F. Ferrero «Caratteristiche acustiche di fonemi vocalici italiani», *Parole e Metodi*, 3, 1972, pp. 9-31). A sinistra grafico con le aree di esistenza dei 7 fonemi definita sulla base delle realizzazioni di 25 informatori maschili. A destra le aree di esistenza ottenute partendo dalle realizzazioni di 25 soggetti femminili (v. valori approssimativi nella tabella sotto).

	i	e	ε	a	ɔ	o	u
F1 (M)	300±95	400±70	550±95	770±145	550±65	450±80	320±70
F2 (M)	2200±270	2000±250	1850±250	1250±225	900±180	750±150	650±140
F1 (F)	300±40	410±55	620±100	1050±220	650±120	450±70	300±50
F2 (F)	2880±350	2700±360	2450±400	1450±200	1250±270	950±220	850±270

Appendice II

Tabella di p. 261 di Ferrero *et alii* (1995) relativa a suoni prodotti da 20 informatori padovani (10 M e 10 F)¹³:

	i	e	ε	a	ɔ	o	u
F1 (M)	251±31	351±34	526±34	716±85	514±42	383±46	258±28
F2 (M)	2071±98	1953±76	1755±110	1208±83	847±71	788±80	747±52
F1 (F)	290±34	420±37	681±77	956±78	681±39	470±59	321±38
F2 (F)	2630±96	2498±134	2086±101	1507±82	1065±108	854±85	747±52

¹³ I valori di deviazione indicati risultano più contenuti rispetto a quelli della tabella sopra perché basati su ellissi di equiprobabilità al 75%.

Contributo acustico preliminare alla descrizione fonetica della debuccalizzazione camuna di S

Antonio Romano, LFSAG

(con la collaborazione di Vittorio Volpi)

1. Introduzione

Sebbene siano diversi i processi evolutivi che portano allo sviluppo di suoni di tipo [h], nello spazio romanzo sono note diverse aree in cui alla presenza di suoni di questo tipo si associa a una reputazione di vero e proprio tratto-bandiera. Le comunità nelle quali il processo si è concluso sono facilmente individuate per questa loro caratteristica soprattutto da parte di comunità confinanti in cui il suono originario da cui proviene presenta esiti intermedi o oscillanti oppure si conserva integro (Contini 1975).

Ad es. tanto in castigliano quanto in gascone o nei dialetti calabresi centrali si possono trovare esiti di una laringalizzazione di lat. F (es. sp. *bijo* ['ixo] < FILIU, gsc. di Momas *bòrt* [hòrt] < FORTE, cal. di Catanzaro *hazzu* ['hatsu] < FACIÒ, *fhuju* ['fhuju] < FUGIÒ) e in questi casi anche la notazione grafica ricorre a <h>, <h̃> o <fh> (come ad es. in Calabria, cfr. Rohlfs 1966: §154, cfr. Trumper 1997, v. dopo)¹.

Esiti laringali si hanno invece in Sardegna per /k/ (Contini 1975) che può avere anche rese occlusive glottidali.

Sono infine ben noti i casi del processo storico per cui in francese, così come osservabile in molti casi nel cosiddetto spagnolo atlantico, una s preconsonantica (di coda sillabica o originariamente anche iniziale) è andata incontro a una *debuccalizzazione*, manifestandosi come fenomeno di aspirazione e giungendo persino alla cancellazione, in seguito a varie conseguenze sui suoni contingui (SCHOLA > fr. *école*, BĒST(I)A > fr. *bête*, MAGISTRU > fr. *maître*, Grammont 1946: 346; MĀ(GI)S > sp. and. *mah*, PISCATU > sp. and. *pebkado*, *pek^bado* o *pek^kado*, SPĒRAT > sp. and. *ehpera*, *ep^bera* o *ep^kera*, Navarro Tomás 1939: 185; cfr. O'Neill 2009).

Si parla in questi casi di *debuccalizzazione* perché il luogo d'articolazione laringale è anche quello deputato alla generazione della voce e, in un modello di produzione che identifichi i contributi del filtro acustico che ne definisce il timbro nelle cavità buccali (e nasali) e quelli della sorgente, cioè proprio della voce, nella laringe, il mutamento del luogo di articolazione di un suono verso questa sede corrisponde a una

¹ Rese simili si hanno per il Poitevin-Saintongeais dove /h/ corrisponde alle posizioni in cui in francese si ha /ʒ/ ed è per questo spesso trascritto <jh> (es. *najbe* [nah] 'nuoto', cfr. fr. *nage* [naz], v. Boula de Mareüil et al. in c. di p.).

perdita di *buccalità* del suono (v. §3)².

Tuttavia, al di là del processo generale, una notevole variabilità di condizioni può prevedere anche un'evoluzione di *s* intervocalico, o in qualsiasi altra posizione, verso suoni di tipo laringale, con articolazioni secondarie e sviluppi di difficile classificazione anche per i suoni contigui. In alcuni *patois* francoprovenzali, come ad es. nel frp. di Fénis (Favre 2000)³ si hanno ad es. esiti diversi in casi come /la 'tsaʔə/ «la chasse» 'la caccia', /be'ʔũŋ/

² Non ha evidentemente molto senso parlare di "spirantizzazione di *s*" (Massariello Merzagora 1988: 91 e 92); semmai potrebbe trattarsi di "aspirazione di *s*". Il fenomeno generale, oggi noto anche come *unmouthing*, è molto studiato nelle trattazioni elaborate nell'ambito di varie teorizzazioni fonologiche (per una rassegna v. Molinu, 1997) e – soprattutto nel caso degli esiti di *s* – comporta modalità di riorganizzazione segmentale in conseguenza di sfasamenti nella coordinazione oro-laringea (tra gli altri, Mairano et al. 2009, Parrell 2012). Il processo descritto per queste lingue trova numerosi contributi anche nella riflessione glottologica e nella linguistica storica per via delle numerose corrispondenze tra /s/ latino e /h/ greco (es. *super-* vs. *hyper-*, *serpente(m)* vs. *herpetón*, *septe(m)* vs. *heptá...*; v., tra gli altri, Beekes 2011). Sono molto interessanti anche i casi di restituzione con *f* di *s* originario, attraverso *h*, come mostrano – proprio per l'area qui indagata – anche gli esempi di Bernini (1987: 219-225): *fumnà* 'seminare' in Val Cavallina (Grone), *nuf* 'noce' in Franciacorta (Brione).

³ Ascoli (1878: 97) esemplifica il fenomeno in riferimento ai *patois* di Saint-Rémy, dove registra 'ò 'sale', e quello di Fénis, appunto, per il quale riporta 'ain 'sano' (cfr. fr. *sain*).

«jumeau» 'gemello' etc. o, appunto, costrittivi di tipo [h] come in /əna 'kɔʔa 'kɔʔe/ → [əna 'kuha 'kɔʔe] «une rose rose» 'una rosa rosa', o ancora di laringalizzazione transitoria come in /la 'meʔa/ → [la 'meɣa] «la messe» 'la messa' o /l e'liʔə/ → [l e'li:ɦə] «l'église» 'la chiesa'⁴.

Tra le altre aree in cui sono tradizionalmente noti sviluppi di tipo *s > h* in tutte le posizioni, ci riferiamo qui a quella generalmente indicata come bergamasca da Guarnerio (1918: 518) e Rohlfs (1966: §165 e §308)⁵. Da fonti più specifiche (v., tra gli altri, Massariello Merzagora 1988) risulta invece evidente l'estensione dell'area al bresciano (il cui sistema è sufficientemente noto sin da Beltrami 1884)⁶.

⁴ Come si vede da alcuni esempi, in corrispondenza di suoni di solito percepiti come occlusivi (colpi di glottide), nei diversi contesti, si possono manifestare incisivi effetti sulla qualità dei suoni vocalici confinanti, con rese che includono l'attivazione di meccanismi fonatori non modal (Contini et al. 2005).

⁵ Cfr. anche Lurati (1988: 493). Basta un riferimento ai dati *AIS* e *ALI* per avere un colpo d'occhio sulla diffusione areale di simili fenomeni (cfr. anche i casi retoromanzi descritti sin dai Ascoli 1873). Ad es., in *AIS* con *s > h* si trova il punto 247, Monasterolo (in Valle Cavallina).

⁶ Per una dettagliata analisi del diasistema v. vari contributi in Sanga (1987). La delimitazione areale del fenomeno non rientra tra gli obiettivi del presente lavoro. Resta tuttavia urgente una verifica della stabilità di questi trattamenti nell'area in questione tenendo conto delle delimitazioni proposte in Bonfadini (1997) e delle

In particolare, benché si abbia una buona descrizione del fenomeno nella Media Val Trompia (Massariello Merzagora 1988: 91; cfr. altri riferimenti nelle nn. prec.), i nostri dati riguardano la diffusione e gli esiti del processo nei dialetti lombardi camuni per i quali questi fenomeni sono «facilmente documentabili» (Ibid., 92), da Capo di Ponte (media Val Camonica) a tutta la Frianciocorta, slargandosi in corrispondenza del lago d'Iseo (Brescia) a Nord del quale, ancora in bassa valle, si situa Esine, il cui dialetto è considerato in questo studio.

2. Contesto e dati

Tra il 2019 e il 2023, grazie ad alcuni studenti lombardi che seguono l'insegnamento di Linguistica Generale del corso Magistrale a Torino, ho avuto l'opportunità di raccogliere campioni di parlato dialettale da località dell'Isola bergamasca, della Valtellina e della Val Camonica (v.

sottili distinzioni sociofonetiche delineate *ante litteram* da Bonfadini (1995). D'altra parte rese interferite sono oggi attestate persino nelle località bergamasche e bresciane più esposte all'urbanesimo o lungo le maggiori direttrici commerciali e turistiche (come ho potuto verificare in un'analisi sommaria di registrazioni di giovani parlanti di Sotto il Monte e in occasione di una mia incursione a Ome, Gardone Val Trompia e Sarezzo). Cfr. Canepari (2004: 260) "Ormai, [h] per /s/ è solo del contado, nonostante il tradizionale *Bèrghem de hura* e *Bèrghem de bóta*, citatissimo anche dai cittadini?".

Archivio vocale «tramontane»). Alla prima occasione fui immediatamente colpito dalle rese laringali che uno dei parlanti registrati, uno studente poco più che ventenne di Pian Camuno (BS), riservava regolarmente a porzioni di enunciato in cui ci saremmo aspettati suoni di tipo [s].

Non fui sorpreso da questi interessanti fenomeni. Da un lato perché risultano ben descritti in opere fondamentali della dialettologia italiana (da vari contributi apparsi sin dai primi numeri di *AGI*, v. tra gli altri, Ascoli 1878; cfr. Trumper 1977, Sanga 1987) e dall'altro perché nei primi anni 2000 avevo già avuto occasione di lavorare sui suoni laringali che risultano da trattamenti di s in alcune località della Valle d'Aosta (oltre che come esito di altri suoni originari in dialetti dei domini retoromanzo e sardo; v. Contini et al. 2005).

Un altro evento fortunato che ha condotto al presente studio risulta dall'esperienza di collaborazione con lo studioso Vittorio Volpi, di Esine, ma a lungo bibliotecario a Iseo, sempre nella provincia di Brescia, il quale, oltre a fornirmi materiali di ottima qualità per uno studio acustico preliminare, con gli strumenti di un dialettologo esperto e di un fonetista provetto, aveva riflettuto a lungo sulle varianti combinatorie che qui saranno passate in rassegna.

3. Obiettivo e risorse tecniche

Trascurando di dettagliare le realizzazioni che si presentano in contesti di evoluzione di una variante sonora (un'approssimante dentale a più vasta diffusione areale, di solito indicata con simboli di tipo δ , v. Bonfadini 1987: 369-371), ci concentriamo qui sugli esiti di s non-sonora e cioè, come anticipato, quelli di tipo h^7 .

L'analisi spettrografica di [h] presenta formanti turbolente, specialmente in posizione intervocalica (Quilis 1999: 269), energia di alta frequenza e aperiodicità nella forma d'onda (Ladefoged 2005: 58) causata da una costrizione glottidale (*glottal closure or stricture*): «In place of a true stop, a very compressed form of creaky voice or some less extreme form of stiff phonation may be superimposed on the vocalic stream» (Ladefoged & Maddieson 1996: 75).

Ora, nei dati sui quali da tempo esercito attività di osservazione acustica, quello che emerge in modo più appariscente è la particolare dinamica che, in contesti di sonorità, determina la comparsa di

suoni approssimanti e mormorata e, in contesti di non-sonorità, produce fenomeni di frizione corrispondenti a strutture spettrali non sempre ben definite, ma spesso dipendenti dal contesto timbrico circostante.

All'obiettivo di discernere queste varianti fornisce l'utile supporto di una serie esempi descrittivi il già citato studio di Contini *et al.* (2005).

Sulla scia di quello, propongo qui un'esemplificazione delle caratteristiche acustiche dei diversi tassofoni presenti nelle produzioni dello stesso Vittorio Volpi.

4. Osservazione acustica delle varianti

Cominciamo con quello che potremmo considerare l'allofono naturale, /h/ > [h], precisando tuttavia che le caratteristiche dei suoni di tipo [h] iniziali, come vedremo, dipendono fortemente dalle proprietà spettrali del suono seguente (ad es. si hanno: [ɛ̥ɛ] in *bèc* 'secco', con [ɛ̥ ≈ ç], [ɪɪ] in *btinc* 'rigido', con [ɪ̥ ≈ ç], [ɻ̥ɻ] in *btüggjà* 'studiare', con [ɻ̥ ≈ ç^w])⁸.

⁷ In realtà, come osserva V. Volpi, anche la sorda di *déb* 'dieci' [deç] (v. §4.3, ad es., in un contesto fonosintattico di sonorizzazione, segue le sorti della sonora: *déb agn* 'dieci anni' > [de^h δaɲ], con [ð] = [ð̥] (al punto da incoraggiare una grafia con <ð> o con <ð̥>: *am pi de ded'agn* [am^hpi δe δe^hðaɲ] 'un bambino di dieci anni').

⁸ Sulle differenze acustiche tra suoni di tipo [ç], [ç] e [x] v. De Sio & Romano (2003). In particolare mentre [ç] tende ad assumere una struttura spettrale di tipo [i] indipendentemente dal contesto, [ç] presenta un colore spettrale negoziato con le proprietà timbriche dei suoni sonori circostanti, mantenendo una sua maggiore 'identità', di tipo [ç], in contesti di non-sonorità e in assenza di articolazioni secondarie.

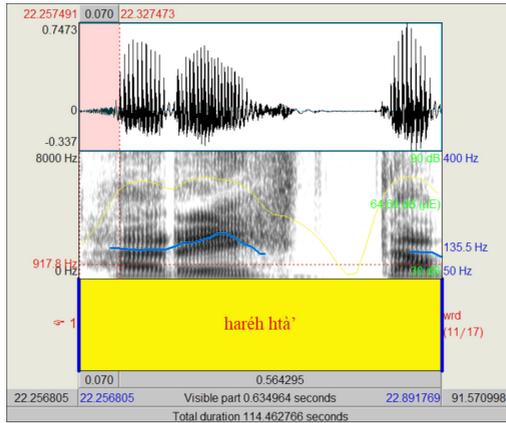


Fig. 1. Esempio spettrografico di $S > [h]$ (la registrazione integrale include altri due contesti interessanti: uno visibile nello stesso spettrogramma, $V_a b \# b t V$, dove si ha $/(V_a)h/ > [ç]$ e $[ç h(t)] > [ç(t)]$, per questo tassofono v. 4.3; $V_a b \# b V$, dove si ha assimilazione totale regressiva, v. anche 4.2).

4.2. $s > [h]$ in posizione intervocalica etc. (v. dopo): (*al ghe piàh*) *htà hentà' ddo bo* (*la teràha*) [(al ge pjax) htà hen'ta d:ò ho la t'rahò] 'gli piace star seduto sul balcone' (v. Fig. 2, per questa realizzazione cfr. Contini et al. 2005)

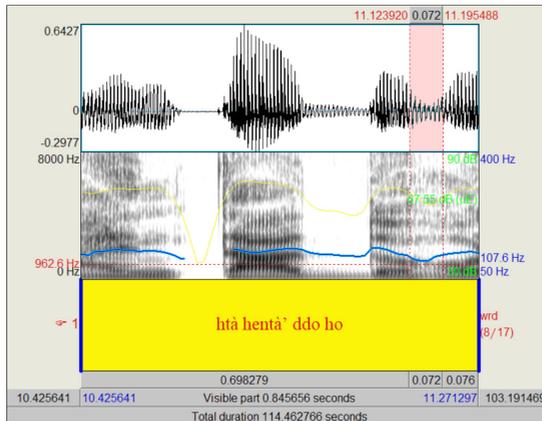


Fig. 2. Esempio spettrografico di $S > [h]$ (la registrazione integrale include altri tre contesti interessanti: lo stesso tassofono in contesto di suoni vocalici anteriori; $(V_a b \#) b t V$, qui non visibile, v. dopo; $V_a b \# d V$, dove si ha assimilazione totale regressiva, v. anche 4.1).

4.3. $s > [ç]$ finale dopo vocale anteriore ($V_a b\#$): *an hmüb* [an'hmyç] ‘uno smusso’ (con $F_2 \approx 2000$ Hz) (v. Fig. 3).

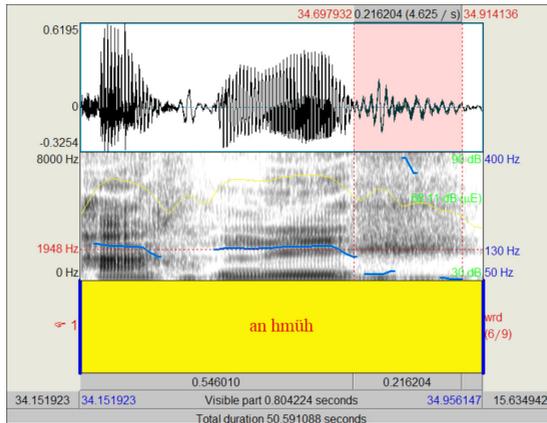


Fig. 3. Esempio spettrografico di $S > [ç]$ (l'esempio include un altro contesto interessante: $n\#bC_p$, dove si ha $/h/ > [h]$, v. 4.10).

4.4. $s > [x]$ finale dopo vocale posteriore ($V_p b\#$): (*i nób*) *htrah* [(i₁nox) ^htrax] ‘i nostri stracci’ ($F_2 \approx 1700$ Hz) (cfr. finale dopo vocale anteriore ($V_a b\#$): *he 'l fòh htàt atènto*) (v. Fig. 4).

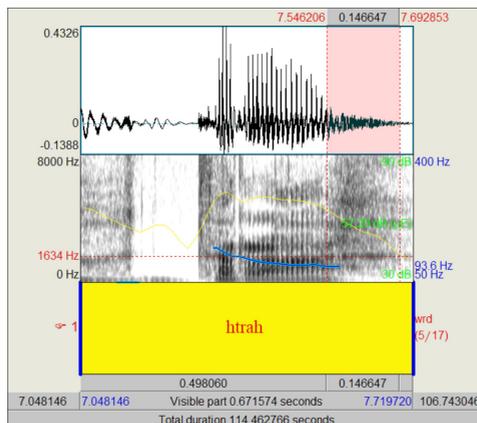


Fig. 4. Esempio spettrografico di $S > [x]$ (la registrazione integrale include un altro contesto interessante qui non visibile: ($V_p b\#$)*htrV*, dove si ha $/(V_p)h/ > [x]$ e $[x h(t)] > [x(t)]$, cfr. 4.1).

4.5. Trattamenti di /h/ in contesti fonosintattici, $Vb\#bV$: *al ghe piàh htà* (*bentà' ddo bo la teràha*) 'gli piace star (...)' (v. 4.2), /h/ > [h] (con $F_2 \approx 1600$ Hz) (v. Fig. 5).

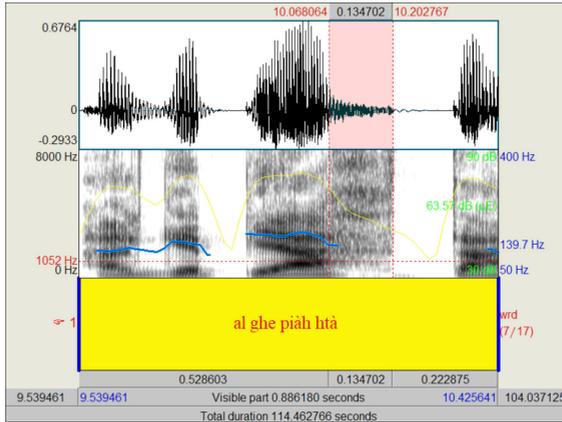


Fig. 5. Esempio spettrografico di realizzazione in contesto fonosintattico $Vb\#bV$: /-h/ + /h-/ > [h].

4.6. Trattamenti di /h/ in contesti di assimilazione in fonosintassi, intervocalica: *tròp héc > trò' héc* / *tròh 'hek* / [trɔx 'hek], con (inizialmente $F_2 \approx 1200$ Hz, ma poi fasi $(\text{ɔ})x + \xi(\text{ɛ})$) (v. Fig. 6).

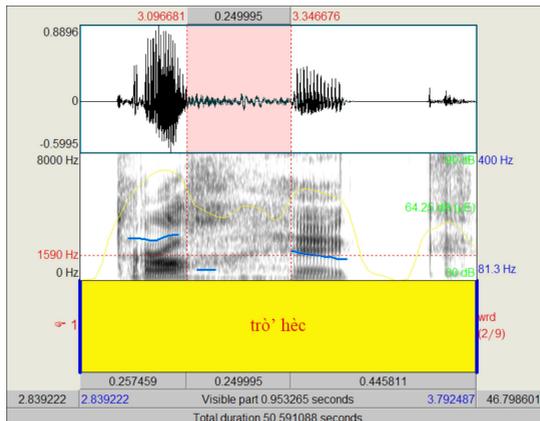


Fig. 6. Esempio spettrografico di realizzazione a confine di parola, in posizione intervocalica (sul piano fonologico è un caso di assimilazione regressiva, come tutti gli altri, ma le fasi iniziali e finali differiscono e tendono a dipendere dagli altri suoni contigui, con emergenza di tassofoni diversi).

4.7. Trattamenti di /h/ in contesti di assimilazione regressiva, postvocalica: *tròp htrèt* > *trò' htrèt* [trɔ'htret] ‘troppo stretto’ (con $F_2 \approx 1520$ Hz, perché parte da F_2 di [ɔ(p)]) (cfr. *tùt htrùt* > *tù' htrùt*, con $F_2 \approx 1600-1750$ Hz, perché parte da vocale anteriore ed evolve verso [t] che esplode verso 1750 Hz) (v. Fig. 7).

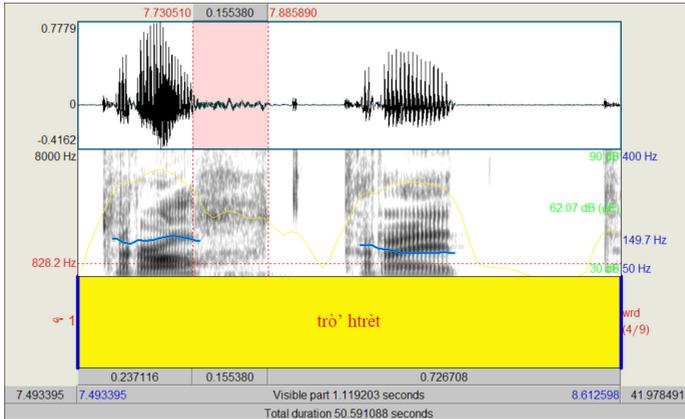


Fig. 7. Esempio spettrografico di realizzazione a confine di parola, con assimilazione in posizione postvocalica.

4.8. Trattamenti di /h/ in contesti fonosintattici di parziale (blanda) assimilazione: *l#ht* (*pal htórt*), /h/ > [h] (con $F_2 \approx 1800$ Hz); cfr. *m#ht* (*fa-m htrobà*), *r#ht* (*fèr htórt*) (*ht* è l'unico in cui la resa può essere di tipo [s], v. *htórt htortènt*) (v. Fig. 8).

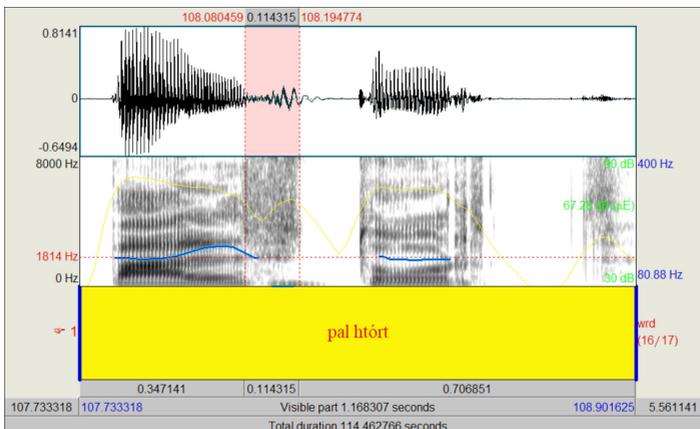


Fig. 8. Esempio spettrografico di realizzazione a confine di parola, con tracce di assimilazione dopo /l/.

4.9. Trattamenti di /h/ in contesti fonosintattici di assimilazione: $f\#b(C) > [h]$ (*fif hparà-ga a chè?*) (con $F_2 \approx 2200 > 1600$ Hz, da [i] a [p]) (v. Fig. 9). L'esempio, in realtà già risultante da un primo trattamento di assordimento (*fif < fl-v*), mi dà modo di accennare ai trattamenti di tipo $F > h$ che pure possono interessare l'area in questione (cfr. §1)⁹ (nota a pag. 62).

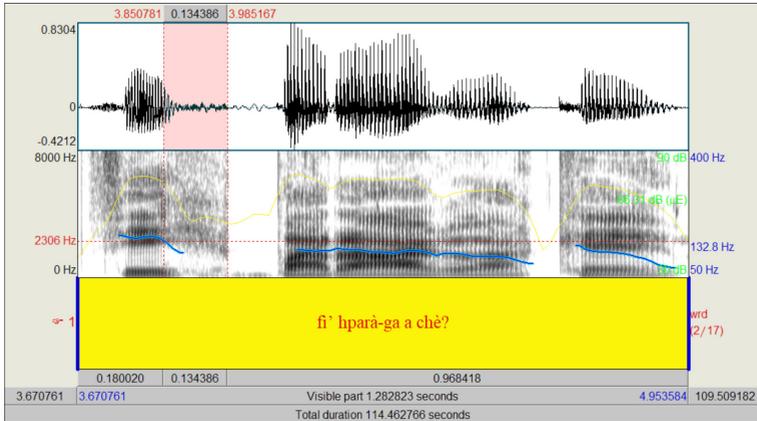


Fig. 9. Esempio spettrografico di realizzazione a confine di parola, con assimilazione totale di /f/.

4.10. Contesti *nbt, nbk, nbh, nbm, nbn, nbv* > rese narinali $[\tilde{n}]$ ($[\tilde{n}]$ $[\tilde{n}]$ $[\tilde{n}]$... $[\tilde{n}]^w$).
 Es. *nbn*: *an hnèta* $[an^h n\tilde{e}.t\tilde{a}]$ ‘noi ripuliamo’ > /h/ > $[\tilde{n}]$ (v. Fig. 10).

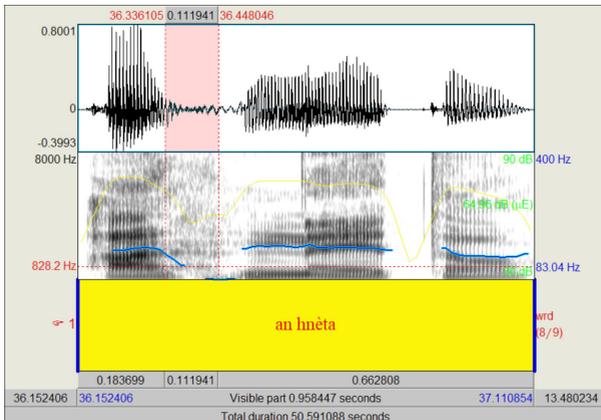


Fig. 10. Esempio spettrografico di realizzazione a confine di parola, dopo (e prima di) nasale.

Sono questi i contesti più difficili da analizzare e verosimilmente anche più interessanti da approfondire, alla luce di realizzazioni simili descritte per il birmano (e il tibetano), dove però i fenomeni di coarticolazione che si determinano in questi contesti hanno conseguenze diverse sul sistema fonologico e definiscono condizioni evolutive diverse (v. Dantsuji 1984; cfr. Nishida 1975)¹⁰.

Conclusioni

Con questo studio acustico preliminare ho inteso predisporre alcune indicazioni bibliografiche utili per lo studio dei fenomeni di debuccalizzazione di *s* nei dialetti camuni e, allargando l'orizzonte geografico, ai dialetti lombardi orientali (bergamaschi e bresciani) noti per queste interessanti caratteristiche.

La disponibilità di dati sonori registrati sulla base di un corpus di frasi, organizzato secondo principi distribuzionali e criteri sistematici dallo studioso Vittorio Volpi, mi ha permesso

di illustrare dettagliatamente vari contesti di realizzazione dell'esito laringale di *s* nel dialetto di Esine.

Dall'osservazione spettrografica di una selezione di dieci di questi contesti (v. §4) ho concluso a favore dell'esistenza di alcuni tassofoni, sistematici in questi dati, che meritano ora di essere verificati in un campione di registrazioni più esteso e maggiormente rappresentativo delle condizioni di variazione diatopica (e sociofonetica) descritte nella letteratura specifica di quest'area dialettale.

Tra le realizzazioni più interessanti emerse nel corso di questa disamina sottolineo tuttavia quelle individuate nell'osservazione dell'ultimo contesto analizzato: quello cioè in cui una resa debuccalizzata si localizzi in prossimità di suoni nasali che, per via di specifiche condizioni di coordinazione orolaringea, causano fenomeni di assimilazione bidirezionale. Questi ultimi, in funzione del grado di coarticolazione, producono esiti narinali ancora poco descritti persino nella letteratura internazionale e che rappresentano forse – come giustamente rilevano le preziose osservazioni dello studioso Volpi – uno dei più intriganti fenomeni fonetici presenti nelle parlate di un territorio noto per le sue suggestive testimonianze preistoriche e ancora tutto da scoprire per specifici fatti linguistici (pensando anche alle suggestive considerazioni di Benozzo 2004).

⁹ Nei dati di *AIS* 238 (Borno) troviamo esempi del tipo *hil de hèr* 'fil di ferro'. A Esine si hanno invece alternanze imprevedibili del tipo *fibbià / bibbià* 'fischiare'. Ringrazio ancora una volta Vittorio Volpi per queste utilissime segnalazioni.

¹⁰ Oltre a un aggiornamento bibliografico sugli studi sperimentali in quest'ambito, occorrerà aggiungere maggiori riflessioni derivanti dall'applicazione di tecniche di analisi acustica basate sulla definizione di indici di coarticolazione e di soglie classificatorie conseguenti come quella proposta in Romano et al. (2005) sulla scorta di una letteratura generale che meriterebbe anch'essa di essere aggiornata.

Bibliografia

- AIS - Jaberg K. & Jud J. (1928-1940). *Sprach- und Sachatlas Italiens und der Südschweiz*. Zofingen: Universität Zürich - Ringier [trad. it.: *AIS. Atlante linguistico ed etnografico dell'Italia e della Svizzera meridionale*, vol. I, a cura di G. Sanga, Milano: Unicopli, 1987].
- ALI - Bartoli M., Terracini B., Vidossi G., Grassi C., Genre A. & Mas-sobrio L. (1995-), *Atlante Linguistico Italiano*, Istituto dell'ALI, Università di Torino (materiali inediti) - Roma: Istituto Poligrafico e Zecca dello Stato (10 volumi pubblicati).
- Archivio vocale «tramontane» - Romano A. & De Iacovo V. (2019). «La base di dati «Tramontane»: dati di parlato su lingue, dialetti, etnoletti e interletti del laboratorio di fonetica sperimentale “Arturo Genre”». In: D. Piccardi, F. Ardolino, S. Calamai (a cura di), *Gli archivi sonori al crocevia tra scienze fonetiche, informatica umanistica e patrimonio digitale* (Studi AISV 6), Milano: Officinaventuno, 49-57.
- Ascoli G.I. (1873). «Saggi ladini», *Archivio Glottologico Italiano*, 1, 1-537.
- Ascoli G.I. (1878). «Schizzi franco-provenzali», *Archivio Glottologico Italiano*, 3, 61-120.
- Beekes R.S.P. (2011). *Comparative Indo-European Linguistics: An Introduction*. Amsterdam: Benjamins (1^a ed. 1995).
- Beltrami A. (1884). Del consonantismo e degli accidenti fonetici generali del dialetto odierno di Brescia. *Commentari dell'Ateneo Bresciano*, 205-211.
- Benozzo F. (2004). «Alcune considerazioni sull'aspirazione di /s/ nei dialetti lombardi orientali: per un approfondimento alpino della *Paleolithic Continuity Theory*», *Quaderni di Semantica*, 25, 243-253.
- Bernini G. (1987). «Il dialetto della Media Val Seriana e della Valle Gandino». In: G. Sanga (a cura di), II vol., 197-287.
- Bonfadini G. (1995). «I sistemi consonantici dei dialetti alto-italiani: il caso dell'Alta Val Camonica». In: E. Banfi, G. Bonfadini & P. Cordin (a cura di), *Italia settentrionale crocevia di idiomi romanzzi*, Tubingen: Max Niemeyer, 25-41.
- Bonfadini G. (1997). «Profilo dialettale della valti bresciana». In: C. Boroni (a cura di), *Itinerari culturali nel bresciano. La montagna*, Venezia: Corbo & Fiore 380-407.
- Boula de Mareuil Ph., Evrard M., François A. & Romano A. (2024). «Computer modelling of innovations relative to Latin in contemporary Romance dialects». Com. presentata al *53rd Symposium on Romance Languages* (Paris, 26-30 June 2023) (in c. di p. in *Selected papers from LSRL53*, a cura di L. Brunetti, I. Chitoran, A. Mardale, M. Russo & Sh. Ulfsbjorninn).
- Canepari L. (2004). *Manuale di Fonetica*. Monaco: Lincom.
- Contini M., Carpitelli E. & Romano A. (2005). «Des occlusives glottales dans l'espace roman». *Estudios Ofre-*

cidos a Antonio Quilis, vol. I, Madrid: Consejo Superior de Inv. Científicas / Univ. Nac. de Ed. a Distancia / Univ. de Valladolid, 127-145.

Contini M. (1975). «*f* > *h* et *k* > *h* et le problème des laryngales du sarde», *Bulletin de l'Institut Phonétique de Grenoble*, IV, 27-66.

Dantsuji M. (1984). «A Study on Voiceless Nasals in Burmese». *Studia phonologica*, 18, 1-14.

De Sio L. & Romano A. (2003). «Caratteristiche acustiche delle fricative dello svedese: uno studio delle transizioni formantiche in contesto VCV». In: G. Marotta & N. Nocchi (a cura di), *La Coarticolazione* (Atti delle XIII Giornate di Studio del GFS, Pisa, 26-28 novembre 2002), Pisa: ETS, 321-329.

Favre S. (2000). «Le *fun-ubèn*: un patois singulier», In: AA.VV., *Fénis. Une communauté au fil de l'histoire*, Quart: Commune de Fénis et Musumeci, 429-441.

Grammont M. (1946). *Traité de Phonétique*. Paris: Delagrave (1^e éd. 1933).

Guarnerio P.E. (1918). *Fonologia Romanza*. Milano: Hoepli.

Henriksen N. & Harper S.K. (2016). «Investigating lenition patterns in south-central Peninsular Spanish /sp st sk/ clusters», *Journal of the International Phonetic Association*, 46(3), 287-310.

Ladefoged P. & Maddieson I. (1996). *The Sounds of the World's Languages*. Oxford: Blackwell.

Ladefoged P. (2005). *Vowels and Consonants*. Malden (Mass.): Blackwell.

Lurati O. (1988). «Italienisch: Areal-linguistik III. Lombardei und Tessin». In G. Holtus *et alii* (a cura di), *Lexicon der Romanistischen Linguistik*, IV, 485-516.

Mairano P., Romano A., Stevens M. & Hajek J. (2009). «La preaspirazione nelle lingue del nord Europa e in italiano: dati di base per un confronto tra islandese e senese». In: L. Romito, V. Galatà & R. Lio (a cura di), *La fonetica sperimentale: metodi e applicazioni* (Atti del IV Convegno Nazionale AISV - Associazione Italiana di Scienze della Voce, Cosenza, 3-5 Dic. 2007), Torriana (RN): EDK, 179-195.

Massariello Merzagora G. (1988). *Lombardia (Profilo dei Dialetti Italiani)*. Pisa: Pacini.

Molinu L. (1997). «L'alternance /k/ - [ʔ] dans les parlers de la Barbagia d'Ollolai. Une approche géophonologique non-linéaire». *Géolinguistique*, 7, 133-157.

Navarro Tomás T. (1939). «Dédoublement de phonèmes dans le dialecte andalou». *Travaux du cercle linguistique de Prague*, 8, 184-186.

Nishida T. (1975). «On the Development of Tones in Tibetan». *Acta Asiatica*, 25, 43-55.

O'Neill P. (2009). «S-Aspiration and Occlusives in Andalusian Spanish: Phonetics or Phonology?», in O. Parker Jones & E. Payne (a cura di), *Oxford University Working Papers in Linguistics, Philology and Phonetics*, vol. 12, 73-85.

Parrell B. (2012). «The role of gestural phasing in Western Andalusian Spanish aspiration». *Journal of Phonetics*, 40, 37-45.

Quilis A. (1999). *Tratado de fonología y fonética españolas*. Madrid: Gredos (1^a ed. 1993).

Romano A., Mancini F., Zovato E. (2005). «Nasali eterosillabiche in italiano e spagnolo: l'energia di banda come parametro discriminante nella classificazione dei nessi NC». In: P. Cosi (a cura di), *La misura dei parametri: Aspetti tecnologici ed implicazioni nei modelli linguistici* (Atti del I Convegno Nazionale AISV - Associazione Italiana di Scienze della Voce, Padova, 2-4

Dicembre 2004), Padova: ISTC/EDK ed., 101-133.

Sanga G. (a cura di) (1987). *Lingua e dialetti di Bergamo e delle Valli* (3 voll.). Bergamo: Lubrina.

Trumper J. (1977). «Ricostruzione nell'Italia Settentrionale: sistemi consonantici. Considerazioni sociolinguistiche nella diacronia». In: R. Simone & U. Vignuzzi (a cura di), *Problemi della ricostruzione linguistica* (Atti del Convegno Internazionale di studi della SLI, Pavia, 1-2 ottobre 1975), Roma: Bulzoni, 259-310.

Trumper J. (1997). «Calabria and Southern Basilicata». In: M. Maiden & M. Parry (a cura di), *The Dialects of Italy*, London: Routledge, 355-364.

PhoneWS

Un programma di eventi dedicati agli studi sulla voce

(a cura di ANTONIO ROMANO, BIANCA DE PAOLIS E VALENTINA COLONNA)

Questa ricca rubrica di informazioni sulle attività del LFSAG nel periodo luglio-dicembre 2023 si articola in tre sezioni:

1. Relazione sulla *Summer School* «Acoustic analysis of Speech and Voice» (A. Romano);
2. Rassegna degli eventi salienti a cui ha partecipato il LFSAG (B. De Paolis);
3. Il progetto «Voices of Spanish Poets» (V. Colonna)¹.

1. *Summer school* «Acoustic Analysis of Speech and Voice»

(org. Antonio Romano, Valentina De Iacovo, Marco Gamba)

FORUM ACUSTICUM - Torino, Cittadella Politecnica, via Borsellino 53/N – 8-10 settembre 2023



¹ I nomi dei componenti attuali del LFSAG presenti in questo testo saranno abbreviati come segue: AR = Antonio Romano; VDI = Valentina De Iacovo; DM = Daniela Mereu; VC = Valentina Colonna; BDP = Bianca M. De Paolis; AA = Anna Anastaseni; FLI = Federico Lo Iacono.

Nell'ambito del congresso internazionale *Forum Acusticum 2023*, svoltosi presso il Politecnico di Torino e organizzato da Arianna Astolfi, il laboratorio ha contribuito all'organizzazione e all'erogazione di borse per i partecipanti di uno dei dieci *hot topic* della scuola estiva «EAA Summer School for Young Acousticians» coordinata da Louena Shtrepi (Politecnico di Torino) e Iván Herrero-Durá (Bertrandt Group, Germania).

Le lezioni si sono svolte dal venerdì 8 (13:30) alla domenica 10 settembre 2023 (13:00) prima dell'inaugurazione del *Forum Acusticum Conference (FA2023)* ed è stata frequentata da 16 dei 17 studenti iscritti. Il programma delle lezioni è stato presentato presso il Dipartimento di Lingue e L.S. e C.M. (Saletta riunioni, 3° piano, Compl. Aldo Moro), l'8 settembre 2023 alle ore 11:00.

Hanno partecipato i seguenti docenti (in ordine alfabetico, con attività indicate dal titolo associato):

Nathalie Henrich Bernardoni (Univ. di Grenoble-Alpes, Francia): *Singing voice analysis and experimental approaches*;

Catia Cucchiarini (Univ. di Nijmegen, Paesi Bassi): *Phonetic parameters for L2 learner speech*;

Barbara Gili Fivela (Univ. del Salento): *Acoustic analysis and benefits for phonetics and phonology*;

Antonio Origlia (Univ. di Napoli Federico II): *Speech technologies and scripting*;

Paolo Mairano (Univ. di Lille, Francia): *Introduction to statistical analysis of acoustic measures of speech*;

Helmer Strik (Univ. di Nijmegen, Paesi Bassi): *Phonetic parameters for dysarthric speech*.

Grazie ai sostenitori (aziende o associazioni partner: Translated, AISV e PerVoice) sono state assegnate 11 borse a studenti provenienti da tre Paesi europei oltre l'Italia (Spagna, Francia e Germania).

Un riassunto delle attività svolte è ricostruibile partendo dalle relazioni svolte dai sette frequentanti che hanno conseguito i crediti formativi accordati dal FA2023.

Riproduco qui una selezione dei testi consegnati (con minimi adattamenti e traduzioni) per ciascuno degli interventi seguiti.

1. *Singing voice analysis and experimental approaches* - Nathalie Henrich Bernardoni (Univ. Grenoble-Alpes, Francia):

La lezione di Nathalie Henrich Bernardoni si è focalizzata su come si produce la voce cantata, dalla respirazione alla fonazione, all'articolazione. Alle proprietà articolatorie sono state associate le caratteristiche acustiche ri-

levabili soprattutto nel corso della produzione dei suoni vocalici, studiati con metodi di indagine sperimentale più o meno invasiva e il ricorso a software in grado di isolare i contributi acustici (distinguendo le risonanze orali dalle formanti rilevabili sullo spettro).

L'intervento si è concentrato sulla produzione vocale partendo dal modello sorgente-filtro (Fant, 1960) e distinguendo meccanismi di voce modale, voce finta e falsetto. A una rassegna di strumenti di misurazione utili ai fini dell'analisi dei meccanismi laringei ha associato la disamina di alcuni casi in cui fonazione e articolazione portano a particolari forme di distribuzione dell'energia vocale (voci bulgare leggere *leka* e pesanti *peshka*).

Sono stati descritti alcuni approcci sperimentali per lo studio delle diverse «componenti» delle pliche vocali che entrano in vibrazione a causa dei meccanismi aerodinamici e della tensione muscolare (vibrazione dei legamenti o della mucosa). Sono state illustrate modalità di rilevamento con metodi MRI, con metodi EGG che consentono la misurazione della *muscle thickness* e con metodi grafici che delimitano automaticamente le dimensioni e la forma della glottide (oltre che la forma d'onda glottale, la velocità e i quozienti di apertura/chiusura) [tradotto dall'inglese].

Queste lezioni hanno offerto le basi della *source-filter theory* e alcuni esempi per mostrare le relazioni tra i valori

delle formanti F_1 - F_2 e le armoniche che caratterizzano la voce [tradotto dall'inglese].

2. *Acoustic analysis and benefits for phonetics and phonology* - Barbara Gili Fivela (Univ. del Salento)

La lezione di Barbara Gili Fivela ha affrontato il tema delle transizioni formantiche e sulla loro correlazione con la posizione e il grado di sollevamento della lingua. È stato analizzato in particolare uno dei primi esperimenti condotti per valutare gli effetti di questi movimenti sull'identificazione dei suoni consonantici, partendo dalle traiettorie delle formanti nella fase iniziale di produzione del vocoide seguente. Anche la variazione melodica è stata discussa in relazione a esperimenti di *gating* (percezione categoriale; ad es. continuazione vs. domanda).

L'intervento di Barbara Gili Fivela si è basato sull'analisi acustica dello spettrogramma, fornendo gli strumenti principali per l'osservazione delle formanti (in diversi contesti) in relazione ai parametri articolatori. Ha fatto riferimento a EPG ed EMA nel caso di parlato patologico (disartria) e ha offerto un quadro più completo introducendo elementi di percezione categoriale rilevati nell'ambito degli studi prosodici.

Barbara Gili Fivela ha proposto un programma in cui ha discusso dell'importanza degli indici acustici sia in termini segmentali (studio delle transi-

zioni formantiche per la decodifica del parlato) sia in riferimento alla dimensione soprasegmentale (studio dell'allineamento dei *pitch accent*) nel quale le categorie fonologiche sono più sfumate e dipendono da variabili stilistiche (e geografiche).

Nella lezione di Barbara Gili Fivela abbiamo imparato a portare attenzione a due ambiti di definizione del parlato in termini di produzione e percezione: il piano fonetico (con riferimento alla notazione delle differenze tra i suoni, alle loro proprietà acustiche e alle nostre capacità uditive) e il piano fonologico (che risente del posizionamento di soglie più o meno definite in spazi di variazione continui). In questo contesto, l'analisi spettrografica aiuta a collegare la variazione acustica con gli assetti e i movimenti nell'articolazione dei suoni [tradotto dall'inglese].

La docente ha proposto l'osservazione di parlanti fisiologici e patologici mostrando l'utilità di strumentazioni di misura e le possibilità dell'analisi dei dati acustici e articolatori. In particolare ha offerto esempi relativi a fenomeni di affricazioni studiati nelle produzioni di parlanti toscani e di variazione melodica nella realizzazione di schemi intonativi in diverse varietà.

Nell'analisi del parlato, oltre alla necessità di distinguere le voci dei soggetti (uomo/donna/bambino, sani/patologici...) si deve tenere conto di

variazioni di pronuncia importanti come le sfumature che fanno percepire differenze dialettali (ad es. *verso* > *ver(t)so* o *penso* > *pen(t)so*, in alcune aree della Toscana).

La docente ha fornito le basi per distinguere fonetica da fonologia e, tra gli esempi analizzati, ha discusso delle modalità di opposizione tra scempie e geminate [tradotto dall'inglese].

3. *Speech technologies and scripting* - Antonio Origlia (Univ. di Napoli Federico II)

La lezione di Antonio Origlia si è caratterizzato per un orientamento più informatico, mostrando come gli strumenti del *machine learning* e dell'IA possano ricoprire un ruolo importante nella ricerca. Nelle premesse è emersa l'importanza della comunicazione tra studiosi di campi diversi: tra chi raccoglie i dati, chi li analizza e chi li interpreta deve esserci collaborazione. In funzione del tipo di dati si sceglie il modello di analisi e si definiscono i limiti di generalizzabilità dei risultati ottenuti, innescando in molti casi processi circolari (fino a casi nei quali può essere necessario ricominciare dall'organizzazione dei dati e ridefinire le modalità di analisi).

L'intervento è stato incentrato sulla statistica applicata alla fonetica (*data-driven processing*) e ha mostrato la necessità dell'interdisciplinarietà a contatto tra informatica e linguistica. L'invito

rivolto ai ricercatori di questi campi è quello di seguire il modello di studioso definito da Renata Savy, in grado di congiungere competenze in termini di analisi rigorosa dei dati e di capacità interpretativa. È stato illustrato l'ambiente di analisi offerto da OpenSmile e sono stati considerati esempi di interazione con software disponibili in vari ambienti di programmazione e/o *scripting* (Orange, Python).

Antonio Origlia ha mostrato l'utilità dell'introduzione nell'analisi delle variabili acustiche che caratterizzano il parlato di strumenti di trattamento dei dati operanti in ambienti di sviluppo e simulazione come OpenSmile, Orange e PyCharm. Un'applicazione esemplare è quella relativa alla sillabazione fonetica, quella cioè in grado di definire la posizione dei nuclei di sonorità rilevabili negli enunciati (nei quali non sempre intensità e rilievi melodici sono sufficienti per individuare le sillabe, soprattutto quelle non accentate). Il caso studiato ha permesso di approfondire come si può migliorare l'estrazione di informazioni dal dato e definire un algoritmo progressivamente più adatto alla loro elaborazione. In particolare, senza avventurarsi nella ricerca di indici acustici responsabili dei confini che si stabiliscono fonologicamente tra le sillabe, la procedura illustrata ha ottimizzato i suoi risultati grazie allo studio dell'entropia.

Con Antonio Origlia abbiamo appreso come si possono adattare all'analisi del parlato tecniche di *machine learning* (*supervised and unsupervised*). Dopo una parte teorica iniziale, la lezione si è concentrata su esempi in cui può essere molto utile basarsi sugli script disponibili in varie librerie (es. Python, Parselmouth) [tradotto dall'inglese].

Antonio Origlia ha illustrato varie tecniche di analisi applicate allo studio della variabili acustiche responsabili della prominente sillabica. Presentando vari ambienti di elaborazione delle informazioni ha analizzato i limiti dello scripting in PRAAT e ha introdotto principi di programmazione in Python.

Concentrandosi praticamente sul problema dell'individuazione delle sillabe fonetiche (e della prominente), il docente ha mostrato l'importanza della distribuzione dell'energia spettrale (ad es. nell'opposizione tra suoni orali e suoni nasali). Ricordando il prezioso lavoro svolto dalla Prof.ssa Savy, ha inoltre mostrato come la collaborazione tra studiosi di diversa formazione possa portare beneficio a un campo, come quello dello studio dell'intonazione, in cui sono necessarie conoscenze interdisciplinari.

La lezione di Antonio Origlia si è concentrata sulle tecniche di *machine learning* (*supervised and unsupervised*) per l'estrazione e l'analisi di informazioni contenute nel parlato [tradotto dall'inglese].

4. *Introduction to statistical analysis of acoustic measures of speech* - Paolo Mairano (Univ. di Lille, Francia)

L'intervento di Paolo Mairano ha illustrato le possibilità offerte dall'ambiente R per l'analisi statistica e il trattamento grafico-numerico di dati relativi alle caratteristiche del parlato [tradotto dall'inglese]

[...] Gli esempi illustrati hanno mostrato le possibilità di analisi attraverso modelli statistici che consentono di analizzare più variabili, valutandone solo alcune – di volta in volta – rispetto ad altre considerate «fattori».

Paolo Mairano ha mostrato alcuni utili esperimenti condotti in R-studio su dati di parlato e variabili foneticamente rilevabili [tradotto dall'inglese]

Paolo Mairano ha analizzato diversi casi di studio conducendo analisi di dati linguistici con R-studio. Ad es. ha trattato un semplice esperimento sulle differenze di f_0 in voci maschili e femminili. Ha poi illustrato le possibilità di analisi offerte nel caso di misure di durata delle geminate, analizzate in un campione di parlanti di diverse regioni. Infine, ha discusso dei risultati che si possono ottenere, con interessanti rappresentazioni grafiche, nel caso di misure relative alla formanti vocaliche (triangoli vocalici).

L'intervento è stato incentrato sull'uso di R (R-studio) per l'analisi statistica di dati relativi a esempi concreti (in particolare formanti vocaliche, durate

delle geminate...). È stata una relazione con esempi pratici svolti insieme al gruppo di apprendenti in aula. L'interesse di questi ambienti di analisi dei dati è legato alla possibilità di verificare ipotesi iniziali sui sistemi fonologici verificabili non solo grazie a grafici accattivanti (ottenuti con la *library ggplot*, tra le altre) ma anche grazie alla possibilità di effettuare test statistici (es. *t-test*, *ANOVA* etc.).

[...] Tra gli esempi illustrati, in particolare, si è mostrato molto utile quello relativo all'analisi della lunghezza consonantica e al modo in cui scempie e geminate contrastano nelle diverse posizioni (ad es. in sillabe accentate e non).

5. *Phonetic parameters for L2 learner speech and for dysarthric speech* - Catia Cucchiari & Helmer Strik (Univ. di Nijmegen, Paesi Bassi)

Nell'intervento congiunto condotto da Helmer Strik e Catia Cucchiari sono state analizzate situazioni in cui l'uso di misurazioni di variabili acustiche è d'aiuto nel caso della stima di errori di pronuncia di apprendenti stranieri di una lingua (es. nederlandese da parte di ispanofoni) e nell'analisi del parlato nel caso di patologie del linguaggio (ad es. variabili che oppongono occlusive a fricative). È stato illustrato il lavoro presentato dal loro gruppo di ricerca al convegno *Interspeech* [tradotto dall'inglese]

È stato un bellissimo intervento in cui si è trattato dell'importanza di misure acustiche nello studio dell'apprendimento di una L2 e in contesti generali in cui si fanno studi psicolinguistici e/o di acquisizione del linguaggio. Sono stati analizzati i timbri vocalici (formanti), le differenze statistiche nelle misure di variazione dell'energia acustica (consonanti) ed è stata valutata l'importanza dei tempi di silenzio (pause).

Trattando di analisi di parlato patologico (disartrico) e di fattori che influenzano la pronuncia di apprendenti stranieri, Catia & Helmer hanno mostrato le loro *pipeline* di analisi e i principali risultati di alcuni esperimenti da loro presentati in varie occasioni [tradotto dall'inglese]

L'intervento di Strik e Cucchiari ha esplorato le potenzialità dell'analisi fonetica nello studio delle disartrie e di alcuni disturbi dell'apprendimento, in particolare le difficoltà di pronuncia nelle lingue straniere. Sono stati analizzati alcuni indici acustici che contraddistinguono la tipica *fluency* di alcuni parlanti disartrici e apprendenti di L2, principale ragione per cui anche uditivamente percepiamo un parlato che si distingue da quello dei parlanti fisiologici e madrelingua.

Gli interventi di Catia Cucchiari e Helmer Strik sono stati incentrati sull'utilità di misurazione condotte su grandi quantità di dati di parlato nel-

lo studio di apprendenti L2 e pazienti disartrici e sulle possibilità offerte da questi nell'indirizzo di strategie correttive e terapie adeguate, rispettivamente. Sono stati esposti studi preliminari, condotti per comprendere e quantificare le caratteristiche acustiche di un parlato disartrico (*severity, intelligibility*), e studi specifici mirati alla selezione di variabili particolarmente indicative in condizioni determinate (studi longitudinali). Nel caso di parlato di apprendenti L2 è stata presentata una tecnica di analisi basata su un sistema di ASR, valido nella rilevazione degli errori (*error detection*, come nel caso delle differenze tra /x/ e /k/), ma anche delle disfluenze.

Questi ultimi due interventi sono stati svolti da un punto di vista teorico e pratico e hanno trattato di un'analisi acustica del parlato in grado di discriminare riflessi di vario tipo delle modalità di coordinazione di attività di produzione (respirazione, fonazione, articolazione) finalizzate alla creazione di enunciati o frammenti di enunciato con caratteristiche speciali nei disturbi del parlato o dell'apprendimento. Dallo studio di indici in grado di prevenire condizioni di demenza, si è passati all'analisi delle differenze oggettive tra i suoni di lingue diverse e delle difficoltà di alcuni parlanti ad articolare determinati suoni. L'impatto della prima lingua sulle lingue straniere apprese successivamente è stato studiato sul

piano delle possibilità di convergenza (da un sistema di partenza a un sistema d'arrivo) osservate su un diagramma di dispersione dei timbri vocalici.

(resoconto stilato a partire dalle annotazioni dei frequentanti
a cura di Antonio Romano,
26 settembre 2023)

2. Rassegna degli eventi salienti a cui ha partecipato il LFSAG

(a cura di B. De Paolis)

Il semestre che è appena trascorso è stato particolarmente fecondo in termini di nuove acquisizioni e partecipazione a eventi.

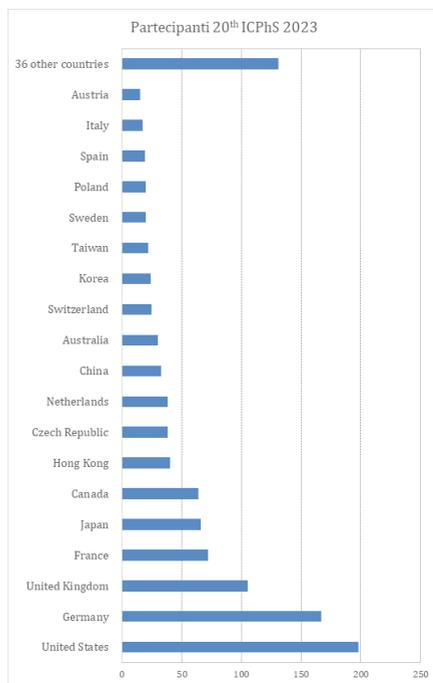
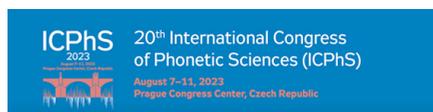
Comincia con la consegna a cura della *Mute-labs* (mute-labs.com) di una cabina insonorizzata modello *QUAT-TRO*, acquistata grazie a un finanziamento accordato al Dip. di Lingue e LS e CM sui fondi del progetto IRI-DIHUM del centro interdipartimentale DISH (ref. Cristina Trinchero) e installata dalla ditta *Acoustic Lab* (acousticlab.com) presso una nuova saletta concessa dal CLA al LFSAG.

Successivamente, durante un periodo di difficile accesso ai locali (per via di lavori di ristrutturazione in corso nell'atrio di Palazzo Nuovo), si sono svolte diverse attività a distanza. Tra queste ricordiamo soprattutto, a fine luglio: 1) il concorso di dottorato DH c/o UNIGE (25 luglio 2023) dal quale è risultato vincitore Federico Lo Iacono (il quale ha poi preso servizio

dal 1° novembre, divenendo presenza costante del laboratorio nel corso delle attività intraprese successivamente); 2) il seminario svolto su invito da Antonio Romano all'interno del *Ciclo de Webinar's da línea temática FCT «Variação Linguística»* (Portogallo, 28 luglio 2023). L'intervento, dal titolo «Les voix d'AMPER: une procédure pour l'analyse de l'extension vocale», ha generato nuovi interessi riguardo ai materiali del progetto AMPER che ancora promettono di offrire un utile supporto agli studi su variazione prosodica e voce, una volta completato l'allestimento della Base di Dati BD-AMPER. Collegata a questa ripresa di lavori, in vista dell'allestimento definitivo dei materiali, è stata poi la visita della collega portoghese Lurdes de Castro Moutinho, che ha trascorso un periodo di ricerca presso il LFSAG tra il 15 e 20 dicembre 2024.

Un'occasione di grande impegno e notevole coinvolgimento del laboratorio è stata poi quella offerta dal 20° International Congress of Phonetic Sciences - *ICPhS*, svoltosi a Praga nei giorni 7-11 agosto 2023. Il laboratorio ha partecipato con due comunicazioni orali (derivanti da una collaborazione tra VC e AR e tra BDP, BA e AR) e due poster (derivanti da una collaborazione tra VC e AR e tra AA e AR), a cui si aggiunge la presentazione di VDI (in collaborazione con Paolo Mairano e John Hajek).

Purtroppo, come si vede dal grafico che riassume la presenza al convegno da parte di ricercatori dei vari Paesi, l'Italia era relativamente poco presente (anche se in modo più consistente rispetto al penultimo *ICPhS* che si era svolto a Melbourne nel 2019, v. notizia nel Bollettino n. 4, 2019) e il gruppo torinese ha contribuito notevolmente a rappresentarla, partecipando a diverse sessioni in tutti i giorni del convegno.



A settembre, vari membri del laboratorio (BA, BDP, VC, AA, FLI) hanno partecipato al *Forum Acusticum 2023 EAA SUMMER SCHOOL* presso il Politecnico di Torino, coinvolti nella Scuola estiva «Acoustic analysis of Speech and Voice», organizzata da Antonio Romano, Valentina De Iacovo e Marco Gamba nel contesto dell'*EAA Summer School for Young Acousticians* (v. sezione 1 di questo *PhoneWS*).



Nello stesso mese, l'Università di Torino ha ospitato il *LVI Congresso della Società di Linguistica Italiana*, dal tema «Continuo e discreto nelle scienze del linguaggio», per cui AR e VDI sono stati parte del comitato di organizzazione. Durante il congresso, sono stati presentati due interventi da parte di membri del laboratorio: AA ha presentato un poster dal titolo «Interconnessioni tra comportamenti grafici, rese fonetiche e sistema fonologico: il caso delle palatali approssimanti e laterali in un task di scrittura sotto dettatura», mentre BDP ha presentato una comunicazione orale intitolata «Il contrasto come oggetto continuo: uno studio sulla realizzazione del focus da parte di parlanti nativi e non nativi di francese e italiano». Il mese di settem-

bre si è concluso con la relazione su invito di DM al workshop «Uso e riuso di corpora di italiano parlato e dialogico», organizzato dalla Libera Università di Bolzano; DM ha presentato (con S. Dal Negro) un intervento dal titolo «Cioè nell'italiano parlato: riduzione fonetica e distribuzione funzionale».

Nel mese di ottobre, BDP è stata invitata, insieme a Fabián Santiago, a condurre un workshop presso l'Università di Grenoble, nell'ambito delle *Journées d'études du réseau d'acquisition L2* dal titolo «Che cos'è Polytonia? Trascrivere la prosodie en langue seconde» e AR ha partecipato al Festival dell'Innovazione di Settimo T.se «CONTATTI – Linguaggi che uniscono» con un intervento sulle caratteristiche fonetiche del cosiddetto «parlare in corsivo» (v. contributi di AR e BDP, AA e VDI nel Bollettino n. 10, 2022).



La partecipazione al workshop «Publish or Perish? Zur Persistenz webbasierter Forschung - Abschluss-tagung von Verba Alpina» (24-25 ottobre 2023 Schloss Fürstenried, München) ha visto la presentazione di una relazione da parte di AR dal titolo: «Archivi vocali e Open Science: ipertrofia di analisi quantitative e derive di una linguistica liquida», mentre infine,

il 28-29 ottobre, Ph. Boula de Mareuil & AR hanno partecipato al convegno «Il patrimonio linguistico storico della Liguria», organizzato a Varazze in ricordo di Fiorenzo Toso, curando un intervento dal titolo «Cartografia di innovazioni rispetto al latino attraverso un atlante sonoro dell'Europa romanza».

L'8 novembre, VDI, in collaborazione con il Consolato del Vietnam a Torino, ha organizzato presso il Dipartimento di Lingue e LS e CM la «Giornata Viet-Nam-Italia» nel corso del quale AR è intervenuto con una relazione dal titolo «Sounds and rhythms of the Vietnamese language (Tiếng Việt)».

Nello stesso mese, VDI ha anche partecipato al convegno «L2APR: L2

Accent And Pronunciation Research: Acquisition, Teaching, Attitudes» organizzato dall'Università Ca' Foscari, con un intervento in collaborazione con Paolo Mairano dal titolo: «Effetti della variazione intonativa regionale sulla percezione di apprendenti di italiano L2». Nello stesso periodo, DM presentava presso l'Università di Aveiro una relazione plenaria su invito, dal titolo «Una prospettiva ecologica allo studio delle lingue di minoranza: il sardo e le sue varietà», in occasione delle «III Jornadas em Línguas Minoritárias».

Infine, a dicembre si sono svolti diversi eventi: la «Giornata In ricordo di Giovanni Ronco» presso il Dipartimento di Lingue e Letterature Straniere, Lingue e Culture Moderne il 1° dicembre, durante la quale AR ha tenuto un intervento dal titolo «Ricordo di Giovanni attraverso le tesi/dissertazione dei suoi studenti». Inoltre, VDI e AR hanno partecipato al corso di formazione «La voce parlata e cantata della persona transgender. Dalla riabilitazione alla chirurgia di riassegnazione della voce» a Torino il 2-3 dicembre, con un intervento dal titolo «Storia della voce legata all'identità di genere e stereotipi vocali» in collaborazione con altri.

Si segnala, a conclusione dell'anno 2023, la partecipazione al convegno «It's NOT only Rock 'n'Roll – On air», organizzato dai colleghi L. Bellone, L. Bonato e E. Madrussan e patro-



cinato dal LFSAG. L'evento si è svolto presso l'Auditorium Complesso Aldo Moro dal 13 al 15 dicembre, con un intervento di AR dal titolo: «Cantare in una lingua inventata. Glossolalia e vocalità».



VOICES OF SPANISH POETS
Vocal Archive and Experimental Study on Poetry Reading



3. Il progetto «Voices of Spanish Poets»: un finanziamento MSCA (Marie Skłodowska-Curie Actions) per lo studio delle voci dei poeti (di Valentina Colonna)

Il 13 febbraio 2023, ho ricevuto la comunicazione di essere stata selezionata per una Marie Skłodowska-Curie Postdoctoral Fellowship, con il progetto Voices of Spanish Poets: Vocal Archive and Experimental Study on Poetry Reading (VSP) proposto con l'host spagnolo Università di Granada (UGR) e i supervisors Antonio Pamies Bertrán, Francisca Serrano, ricevendo una valutazione di 100/100.

Nell'ambito delle Marie Skłodowska-Curie Actions (MSCA), questa Postdoctoral Fellowship è finanziata dall'Unione Europea (UE) sotto il programma di ricerca e innovazione Horizon Europe, con il grant agreement Marie Skłodowska-Curie No. 1011109465. Il finanziamento ricevuto corrisponde a € 165.312,96.

Si tratta di una fellowship di tipo «European» di durata di due anni, con data di inizio 01/05/2023, che si svolge presso l'UGR. Le strutture coinvolte sono, rispettivamente, il Dipartimento de Lingüística General y Teoría de la Literatura e il Centro de Investigación Mente, Cerebro y Comportamiento (CIMCYC).

Il progetto di ricerca, di cui formano parte la Principal Investigator (PI), Valentina Colonna, e i due *supervisor* (Antonio Pamies Bertrán, Francisca Serrano), è finalizzato allo studio interdisciplinare di tipo sperimentale, principalmente fonetico e psicolinguistico/cognitivo/psicofisiologico, della voce dei poeti spagnoli. I risultati confluiranno in un archivio vocale dinamico online.

VSP trova le sue radici nel precedente progetto dottorale della PI Voices

of Italian Poets (VIP), sviluppato presso il LFSAG grazie al dottorato in Digital Humanities (Linguistica, Linguistica applicata e Onomastica), svolto presso le Università di Genova-Università di Torino e concluso nel marzo 2021 sotto la supervisione di Antonio Romano.

Si riporta di seguito la presentazione introduttiva a VSP, pubblicata sulla pagina CORDIS dell'UE, dove sono disponibili ulteriori informazioni sul progetto (<https://doi.org/10.3030/101109465>):

«I lettori del tesoro costituito dalla poesia spagnola non hanno potuto apprezzare in modo approfondito i versi scritti attraverso la loro resa vocale. Con il sostegno del programma

di azioni Marie Skłodowska-Curie, il progetto VSP compirà un pionieristico passo in avanti negli studi umanistici digitali e nell'espressione artistica proponendosi di ridare vita alle voci registrate dei poeti spagnoli moderni. Attraverso una fusione innovativa di scienze umane digitali e metodologie interdisciplinari, la missione di VSP è quella di colmare il divario tra teoria e pratica, svelando l'intricata essenza melodica contenuta nei versi poetici. Grazie alle sue innovative strategie interdisciplinari, il progetto non promette solamente di ridefinire le tecniche di insegnamento e di esecuzione, ma funge anche da trampolino di lancio per l'avviamento di iniziative di collaborazione ed esplorazione multilingue.»

Norme editoriali

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

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

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

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

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

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

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

Es.:

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

Figure e tabelle devono essere corredate da una didascalia (nella quale, se l'immagine o i dati non sono frutto di un lavoro originale, si rinvia alla fonte). Le immagini, rigorosamente in bianco e nero, non devono essere inserite nel testo ma allegate separatamente.



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

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

¹ Scaricabile gratis, e.g., dal sito <https://www.wfonts.com/font/garamond>.

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

Riferimenti bibliografici (in fondo al testo).

Esempi:

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

Bertinetto P.M. & Magno Caldognet-

to E. (1993). Ritmo e intonazione, In A.A. Sobrero (a cura di) (1993a), 141-192.

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

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

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

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

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

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

Sitografia

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

Il registratore magnetico

Da quando disponiamo di registratori digitali e microfoni di buona qualità, inclusi negli smartphone persino di bassa gamma, e/o ricorriamo a una vasta gamma di dispositivi di digitalizzazione audio che passano attraverso l'elaborazione di sistemi ADC, ci sembra lontanissima l'epoca delle registrazioni su nastro magnetico.

«Il registratore magnetico (o magnetofono) è un dispositivo che consente di immagazzinare un segnale elettrico (per esempio il segnale acustico trasformato in tensione elettrica da un microfono) sotto forma di variazioni dello stato magnetico di particelle di ossido di ferro depositate su un lato di un supporto plastico, chiamato nastro magnetico [...] La corretta comprensione del principio di funzionamento del registratore presuppone dunque la conoscenza di alcune nozioni relative al fenomeno della “magnetizzazione residua”, che qui non si ritiene necessario richiamare.

[... U]n registratore magnetico consta delle seguenti unità fondamentali:

- un sistema meccanico per il trasporto del nastro;
- tre testine magnetiche, rispettivamente per registrare, riprodurre e cancellare il segnale;
- un insieme di circuiti elettrici attraverso i quali passa il segnale prima di essere registrato;
- un insieme di circuiti elettrici attraverso i quali passa il segnale preventivamente registrato, per essere riprodotto;
- un oscillatore di alta frequenza, necessario per espletare le funzioni di registrazione e cancellazione del nastro».

[Da: Ferrero F, Genre A., Boë L.J. & Contini M. (1979). *Nozioni di fonetica acustica*. Torino: Omega, p. 87]

