

GLEND A GURRADO, PATRIZIA SORIANELLO

Sulla percezione del tratto di dominanza/sottomissione in un campione di voci femminili

According to the Frequency Code proposed by Ohala, there is a close relationship among the physical size of many animals (large vs. small), the vocal pitch height of their vocalizations (low vs. high sounds) and their degree of dominance/submissiveness. Also in humans, low-pitch voices are generally associated with high level of dominance, while high-pitched voices normally convey a meaning of submission and politeness. This research aims to explore the possible relationship between female voice pitch and its paralinguistic interpretation. To this purpose, we conducted an auditory test by asking to a group of Italian listeners to evaluate the voices in terms of five semantic scales reproduced on a *Visual Analogue Scale*. The findings show a confusing trend: low-pitch voices increased ratings of dominance and self-confidence than high-pitched ones; nevertheless, the transmission of these paralinguistic meanings is affected by other phonetic aspects, such as formant structure, speech rate and intensity¹.

Key words: Frequency Code, female voice, pitch, dominance/submissiveness dimension, Visual Analogue Scale.

1. *Il dimorfismo della voce*

Il presente lavoro ha inteso indagare la relazione esistente fra il *pitch* della voce e alcuni *markers* della personalità. Punto di partenza della ricerca è il Codice della Frequenza definito in chiave etologica da John Ohala (1983, 1984, 1994). Si tratta di un codice che evidenzia il ruolo svolto dalla voce nella definizione delle relazioni di potere, in uno scambio comunicativo fra emittente e ricevente. Le caratteristiche acustiche della voce, in particolar modo la frequenza fondamentale (f_0), variano sensibilmente a seconda del sesso del parlante. Le differenze anatomiche riscontrabili fra uomini e donne coinvolgono, difatti, anche gli organi preposti alla fonazione, in primo luogo il tratto vocale, lungo in media circa 17-18 cm nell'uomo e 14-14,5 nella donna. La laringe dell'uomo, inoltre, è circa il 50% più larga rispetto a quella della donna, una differenza più marcata nella dimensione antero-posteriore e meno visibile in quella laterale. Ciò comporta la presenza di corde vocali di diversa lunghezza, mediamente 17-18 mm per gli uomini e 13-14 mm per le donne (cfr. Titze, 1989). Tale difformità fa la sua comparsa in età puberale; prima di allora, maschi e femmine possiedono laringi di dimensioni simili. A partire dalla fase adolescenziale, la laringe dell'uomo

¹ Questo studio è il risultato di una collaborazione continua fra le autrici. Tuttavia, GG è responsabile dei §§ 1, 2, 3, 4; PS dei §§ 6, 6.1, 6.2, 6.3, 7.1, 8. I paragrafi 5, 5.1, 5.2, 5.3 e 7 sono comuni, come pure lo svolgimento dell'analisi acustica e del test percettivo.

comincia a evolversi con maggiore possibilità di allungamento, essendo posizionata più in basso nella gola; la voce dell'uomo, sarà per questo tipicamente più grave rispetto a quella della donna. La frequenza fondamentale, da cui dipende il *pitch* della voce, è determinata dalla velocità con cui le pliche vocali, durante il processo di fonazione, compiono il ciclo di adduzione e abduzione al momento del passaggio dell'aria. Le pliche vocali vibreranno più lentamente se grandi e più velocemente se piccole: i suoni gravi sono quindi determinati da una f_0 bassa, mentre i suoni acuti sono il risultato di una f_0 alta. Generalmente, la f_0 media degli uomini adulti si attesta attorno ai 110 Hz, laddove quella delle donne si aggira attorno ai 205 Hz (Titze, 1989; Simpson, 2009). Questi valori sono peraltro influenzati anche dall'età, oltre che dal sesso, essendo fortemente correlati alla produzione ormonale. Infatti, nell'uomo la f_0 si abbassa progressivamente fino a 35 anni, per poi ricominciare a innalzarsi verso i 55 anni di età; nelle donne, al contrario, la f_0 è stazionaria fino al periodo della menopausa, a quel punto comincia a decrescere raggiungendo, intorno ai 70 anni, il suo valore minimo. Dalla grandezza delle pliche vocali dipende anche l'estensione melodica, ossia l'intervallo in frequenza calcolato fra il valore massimo e il valore minimo in Hz, rilevabili nel contorno intonativo dell'enunciato: nel parlato, l'escursione melodica di un maschio adulto è compresa fra gli 80 e i 160 Hz, laddove quella di una donna adulta si attesta fra i 160 e i 320 Hz.

2. *La comunicazione emotiva animale*

I filoni di ricerca a cui è possibile fare riferimento sono ampi e diversificati, nonché saldamente ancorati allo studio della comunicazione nel mondo animale. A tal proposito, il riferimento a Charles Darwin (1872) è imprescindibile. L'autore si interrogò sugli effetti che le emozioni producono sulle caratteristiche acustiche dei suoni emessi dagli animali. In una situazione di particolare coinvolgimento emotivo, gli animali subiscono una contrazione involontaria dei muscoli del petto e della glottide che determina l'emissione di suoni vocali non abituali e visibili cambiamenti fisici. Ad es. nel cane, la rabbia induce versi tipicamente gravi, come il ringhio, e l'innalzamento dei peli della schiena. Le spinte emotive possono essere positive o negative, a seconda che il contesto tranquillizzi o metta in allerta l'animale: nel primo caso l'animale tende a instaurare un rapporto amichevole fondato sulla remissività, mentre nel secondo sarà portato a trasmettere segnali ostili. In questa circostanza, l'intento, perseguito mediante i versi prodotti e i cambiamenti fisici assunti nella fase di pre-attacco, sembrerebbe identificarsi con il desiderio di apparire il più grande e temibile possibile alla presenza del nemico, al fine di incutere timore e imporre la propria supremazia.

Appare evidente lo stretto rapporto esistente fra il *pitch* e lo stato emotivo. Difficoltosa è però l'attribuzione di un verso a una particolare sensazione; molto spesso i cambiamenti sonori possono risultare, all'orecchio umano, quasi impercettibili: un rapporto biunivoco è difficilmente intuibile nei contesti emotivi realizzabili nel mondo animale. Un ruolo fondamentale è svolto dall'abitudine, elemento fuorviante

in alcuni casi, quanto utile in altri: produzione e percezione del suono potrebbero essere fortemente influenzate proprio dal fattore abitudine, che avrebbe favorito la creazione, poi consolidatesi nel tempo, di relazioni fra versi ed emozioni.

Molti versi appaiono distinti sia dal punto di vista acustico che motivazionale, ma spesso la percezione dell'intento soggiacente a essi non è così semplice e immediata. L'analisi delle regole sulla comunicazione animale ha affascinato numerosi studiosi del mondo della scienza e della linguistica. Fra essi, spicca Eugene Morton (1977), il quale, partendo dall'osservazione del comportamento comunicativo di numerosi mammiferi e volatili, ha elaborato le *Motivation-Structural Rules*, regole che influenzano la produzione di versi animali caratterizzati da una specifica struttura acustica. Nelle situazioni comunicative intraspecifiche o interspecifiche, il dato che permette ai soggetti di percepire le sfumature motivazionali che sottostanno all'uso dei suoni emessi è senza dubbio la prossimità: è facile intuire come uno scambio comunicativo realizzato a distanza conduca inevitabilmente alla perdita di sfumature acustiche cruciali ai fini della trasmissione dei cambiamenti motivazionali. Tali intenti possono essere espressi, quindi, solo mediante una gradualità sonora che si inserisce all'interno di uno spettro acustico avente come punti estremi da un lato i suoni a bassa frequenza, utilizzati dagli animali prevalentemente in contesti ostili, e dall'altro i suoni ad alta frequenza, usati al fine di avvicinare il nemico in modo amichevole, predisponendolo positivamente nei propri confronti. Morton sottolinea come i poli di tale spettro siano in realtà usati raramente dagli animali, sono invece i gradi frequenziali mediani a esprimere meglio le mutazioni d'intento.

Gli *endpoints* dello spettro acustico corrispondono a suoni ad alta e a bassa frequenza, valori che rappresentano motivazioni diametralmente opposte. Eugene Morton attribuisce tale correlazione a una giustificazione di tipo fisico-anatomico: il primo strumento utilizzato dagli animali per comunicare l'informazione sulla propria mole è la voce. Fisiologicamente, difatti, un animale caratterizzato da una corporatura ampia produce suoni a bassa frequenza, mentre un animale di dimensioni ridotte produce suoni ad alta frequenza. Di conseguenza, in uno scambio comunicativo, il soggetto ricevente, partendo dai soli versi, potrà ricavare l'informazione relativa alla taglia dell'animale emittente, anche in assenza del contatto visivo, e decidere, in base a questa informazione, se procedere all'attacco o ritirarsi prima che inizi lo scontro. È facile comprendere quanto tale correlazione abbia giocato un ruolo decisivo a livello evolutivistico, visto che, come è noto, un animale grande tenderà sempre a vincere su uno più piccolo e a dettarne l'esclusione dal branco. Tuttavia, in numerose situazioni è proprio l'utilizzo dei suoni ad alta frequenza a determinare la sopravvivenza dell'animale: ad es., al fine di evitare lo scontro, il soggetto, alla stregua di un infante, emette suoni acuti trasmettendo un messaggio di sottomissione e una richiesta di protezione. Appare evidente quanto lo strumento voce permetta agli animali di gestire gli incontri in base alle relazioni di potere da essi impostati, modulando la frequenza dei versi emessi e dando vita a uno scontro verbale in cui le armi in gioco sono la dominanza e la sottomissione.

3. *La voce della dominanza*

Fatta questa premessa, non c'è da stupirsi che la voce sia uno strumento di trasmissione degli intenti, anche con riferimento all'uomo.

Diversi linguisti si sono interrogati sulla correlazione esistente fra le caratteristiche acustiche dei contorni intonativi degli uomini e i messaggi sociali da essi trasmessi, riconducibili a un certo grado di sottomissione o dominanza. In particolar modo, il linguista americano John Ohala (1984) ha indagato il ruolo svolto dalla frequenza fondamentale negli scambi comunicativi fra uomini adulti, elaborando il Codice della Frequenza. Tale modello assume che un contorno intonativo avente una f_0 particolarmente alta, o in progressivo innalzamento, trasmetterà sottomissione e insicurezza; diversamente, un contorno intonativo caratterizzato da una f_0 bassa, o in progressiva discesa, invierà un messaggio di dominanza e sicurezza di sé. Secondo Ohala, tale correlazione ha un carattere innato e non va assunta come risultato dell'evoluzione della comunicazione. È sufficiente menzionare l'anatomia del tratto vocale maschile e femminile che, come già descritto, differenzia anche la struttura acustica dei suoni prodotti che, pertanto, saranno acuti per le donne e gravi per gli uomini. Tale difformità comporta una relazione fra la dimensione corporale e i significati veicolati, dovuta all'associazione del tratto vocale grande a una maggiore robustezza e prestanta fisica (dominanza) e del tratto vocale piccolo a una ridotta fisicità (sottomissione). Questa differenza, pur essendo innata, si è consolidata nel tempo innescando stereotipi sociali di discriminazione sessuale. Si tratta di universali cross-culturali e cross-linguistici, fra i quali rientra, sul piano propriamente linguistico, anche la contrapposizione fra il contorno intonativo ascendente delle frasi interrogative, che trasmette la dipendenza informativa di colui che la produce, e il contorno discendente delle frasi affermative, autosufficienti dal punto di vista informativo poiché complete: una differenza rilevabile nella maggior parte delle lingue.

John Ohala evidenzia che il Codice della Frequenza ha carattere innato o inconscio se connesso al dimorfismo sessuale dell'apparato fonatorio di uomini e donne, carattere affettivo se connesso all'uso consapevole dei significati trasmessi dai diversi livelli di f_0 . Al fine di dimostrare le proprie argomentazioni, Ohala raccolse degli stimoli audio di parlato spontaneo, realizzati da 4 adulti angloamericani (2 femmine e 2 maschi). Gli stimoli sono stati successivamente sintetizzati, per eliminare le informazioni sul sesso e sul contenuto semantico della frase e allo stesso tempo innalzati e abbassati in frequenza. I segnali, abbinati in coppia, sono stati proposti a un gruppo di ascoltatori, i quali hanno dovuto giudicare quale fra i due stimoli della coppia fosse più dominante rispetto all'altro. Il test ha dimostrato che il 92% degli ascoltatori ha percepito come più dominante lo stimolo audio caratterizzato da una frequenza fondamentale media ($f_0\bar{x}$) più bassa; inoltre, si è rilevato che la discesa finale dei livelli di f_0 è il tratto che più contribuisce alla percezione della dominanza.

Gussenhoven (2002) riprende la teoria del Codice della Frequenza, o *Size Code*, definendo altri due codici biologici: il Codice delle Sforzo (*Effort Code*) e il Codice della Produzione (*Production Code*). Il primo riguarda la quantità di energia impie-

gata dal parlante durante la produzione dei suoni; l'impiego di uno sforzo maggiore non solo rende più preciso il movimento articolatorio e più ampia l'escursione melodica, ma è anche la causa principale dei processi di prominenza. Il Codice della Produzione evidenzia invece il ruolo della dinamica respiratoria nella resa del contorno intonativo di un enunciato assertivo, notoriamente caratterizzato da declinazione intonativa. I picchi intonativi sono per questo più alti in frequenza all'inizio dell'enunciato, quando l'energia è massima, mentre il livello frequenziale più basso si riscontra alla fine, in corrispondenza del momento di minima pressione laringea. In relazione al *Size Code*, Gussenhoven sottolinea come la correlazione esistente fra la larghezza della laringe e la velocità di vibrazione delle pliche vocali permetta l'instaurazione delle relazioni di potere fra i parlanti, in base alla percezione di una dominanza fisica e sociale. I tre codici possono essere letti attraverso interpretazioni affettive, che attribuiscono caratteristiche al parlante, e interpretazioni informazionali, che associano caratteristiche al messaggio linguistico. Con riferimento al *Size Code*, le interpretazioni affettive riguardano l'opposizione tra i valori afferenti alla sfera della sottomissione (es. gentilezza, vulnerabilità), veicolati da alte frequenze, e i valori della dominanza (es. sicurezza di sé, aggressività) trasmessi da basse frequenze. Le interpretazioni informazionali evidenziano la contrapposizione tra il senso di incertezza convogliato da un messaggio linguistico prodotto ad alte frequenze e il senso di certezza trasmesso da un messaggio a basse frequenze, nonché il confronto tra frase affermativa e frase interrogativa, caratterizzata l'una da un contorno intonativo discendente, l'altra da un contorno finale ascendente. Poiché tali corrispondenze si ravvisano nella maggior parte delle lingue del mondo, è lecito supporre che gli usi del Codice della Frequenza abbiano subito una sorta di grammaticalizzazione.

4. *Gli studi*

In questa direzione, una serie ampia e diversificata di ricerche svolte in ambito linguistico e psicolinguistico ha indagato il ruolo della voce come vettore di un messaggio pregno di significati 'altri' rispetto a quello linguistico, sconfinando nel campo della trasmissione dei marcatori della personalità del parlante. Oltre agli studi incentrati sull'origine dello stereotipo sociale (Scherer, 1972; Henton, 1989; Jiang, 2011), diverse ricerche hanno preso in esame anche la relazione fra la voce e la percezione della dominanza. Dagli studi non emerge un quadro interpretativo sempre nitido. Apple, Streeter & Krauss (1979) e Scherer (1979) rilevano una correlazione positiva fra il *pitch* e la percezione di alcuni tratti della personalità, giungendo però a risultati fra loro discordanti. Tale relazione sarebbe assente per Aronovitch (1976), mentre Tusing, Dillard (2000) e Collins (2000), basandosi su un test di formulazione di giudizio da parte di ascoltatrici donne su voci maschili, affermano che le voci caratterizzate da un *pitch* più basso trasmettono un maggior grado di dominanza. In modo speculare, Collins, Missing (2003) hanno dimostrato come voci più alte in frequenza denotino sottomissione e subordinazione del parlante. I principi del Codice della Frequenza ohaliano sono confermati anche da uno studio svolto da

Chen, Gussenhoven e Rietveld (2004) sulla lingua inglese e olandese. Gli autori, attraverso la sintesi di specifici contorni intonativi e la somministrazione di un test percettivo, elaborato mediante una *Visual Analogue Scale*, attestano come al decrescere dei livelli di f_0 e di estensione tonale corrisponda un aumento della percezione della dominanza. Fra le due lingue indagate, emergono comunque delle sostanziali differenze percettive, un dato che mette in luce la presenza di una certa connotazione linguo-specifica dei significati paralinguistici.

Rimanendo nell'ambito degli studi sui marcatori della personalità, spiccano le ricerche di Feinberg, Jones, Little, Burt & Perret (2005) e di Puts, Gaulin & Verdolini (2006), inerenti alle influenze che la manipolazione, tramite sintesi, di f_0 e delle formanti determina sul grado di attrazione esercitato sulle donne da maschi adulti. A questo proposito, Feinberg et al. (2005) osservano che negli uomini le frequenze basse trasmettono un senso di mascolinità e di capacità riproduttiva. Gli autori elaborano un test di sintesi vocale; a un gruppo di dodici donne si chiede di giudicare il grado di mascolinità di voci maschili, precedentemente impiegate in 12 *task* linguistici. I risultati dimostrano che le voci caratterizzate dall'abbassamento di f_0 sono più spesso associate a significati di grandezza, mascolinità e maturità rispetto alle voci caratterizzate da una f_0 alta. Parimenti, anche il grado di attrazione risulta positivamente connesso ai valori della frequenza fondamentale. Sempre con riferimento alla percezione di mascolinità, Puts et al. (2006) esaminano la relazione fra *pitch* e dominanza (sociale e fisica), durante un *dating-game*, indagando anche la possibile correlazione fra f_0 e il successo degli uomini con l'altro sesso. Il test ha previsto la registrazione di voci maschili e la formulazione del giudizio di dominanza percepito da parte di maschi adulti. I risultati hanno rivelato che voci gravi sono associate a un più alto grado di dominanza rispetto a voci acute; inoltre, in un contesto competitivo, gli uomini che ritengono di essere fisicamente più dominanti rispetto al proprio concorrente tendono ad abbassare il *pitch* della loro voce. Tuttavia, sebbene gli uomini che hanno dichiarato di avere avuto un maggior numero di partner sessuali siano accomunati da una voce particolarmente bassa in frequenza, la relazione fra f_0 e livello di successo con le donne non è statisticamente significativa.

Rilevante è infine il filone di studi dedicati ai tratti acustici e prosodici della voce carismatica, per lo più incentrati sull'analisi del parlato politico. Fra tutte, merita di essere menzionata la ricerca di Rosenberg, Hirschberg (2009). Lo studio ha preso in esame i discorsi realizzati da politici democratici americani. I segnali audio sono stati sottoposti al giudizio percettivo di otto ascoltatori nativi, ai quali è stato chiesto di valutare il grado di carisma presente nei discorsi. Per il giudizio, gli autori si avvalgono di una scala Likert a 5 punti e dell'ausilio di una gamma di aggettivi, a scelta dell'ascoltatore, da associare a ciascuna voce, in modo da pervenire a una valutazione più affinata della propria percezione di carisma. I risultati provano che livelli alti di $f_0\bar{x}$ conferiscono maggior carisma al parlante, lo stesso vale per la variazione dell'estensione tonale e dell'intensità. Sulla stessa scia, altre ricerche hanno approfondito tale tematica (Signorello, D'Errico, Poggi & Demolin, 2012; Signorello, Demolin, 2013); per meglio definire le varie 'dimensioni del carisma', gli autori han-

no analizzato i discorsi di politici francesi e brasiliani, anche con riferimento contrastivo alla lingua italiana (D'Errico, Signorello & Poggi, 2013).

Da questa breve sintesi è possibile dedurre un aspetto centrale. I campioni impiegati nelle ricerche citate sono per lo più costituiti da voci maschili, mentre i test sono per gran parte composti da un uditorio femminile chiamato a giudicare voci maschili in termini di dominanza sociale. Le lingue maggiormente esplorate in questa direzione sono inglese e angloamericano. Diversamente, la lingua italiana, oggetto di studio sulla variazione delle frequenze formantiche di uomini, donne e bambini e sulle modalità di normalizzazione (Ferrero, Magno Caldognetto & Cosi, 1995; Ferrero, Magno Caldognetto & Cosi, 1996), non è stata indagata in questa specifica direzione. Recentemente, la correlazione tra voce e tratti della personalità è stata presa in esame da Sorianello & Gurrado (in stampa). Partendo dalla letteratura disponibile, con esplicito riferimento ai principi del Codice della Frequenza, le autrici hanno cercato di verificare se il rapporto fra i parametri prosodici e la percezione della personalità trovasse riscontro anche all'interno di un campione di sole voci femminili. La finalità della ricerca, improntata su un corpus di parlato letto prodotto da cinque locutrici italiane, era quella di accertare se voci femminili con caratteristiche timbrico-frequenziali molto diverse tra loro potessero essere associate a quei tratti della personalità afferenti alle sfere della dominanza o della sottomissione. L'analisi acustica e la successiva raccolta di giudizi percettivi, formulati da cinquanta ascoltatori per mezzo di una *Visual Analogue Scale*, producono risultati disomogenei. Non sempre la correlazione fra voce (acuta vs. grave) e tratti della personalità percepiti (sottomissione vs. dominanza) appare infatti verificata, sebbene il test, nel complesso, riveli un certo collegamento tra la presenza di un *pitch* grave e la trasmissione di un senso di dominanza e sicurezza di sé. A discapito dello stereotipo sociale, è stato inoltre rilevato che anche la voce femminile può trasmettere un messaggio di dominanza e sicurezza, se contrassegnata da specifiche caratteristiche acustiche, e se confrontata con una voce più alta in frequenza di pari genere. Tuttavia, come atteso, nella valutazione percettiva delle voci, diversi parametri fonetici come ad es. l'intensità e la velocità elocutiva, e non solo il *pitch*, interagiscono ai fini della trasmissione dei marcatori della personalità. Il quadro che emerge da questo recente studio non è comunque nitido, dato che alcuni fattori potrebbero aver confuso i giudizi. Da un lato il numero elevato di *speakers* potrebbe aver destabilizzato gli uditori, dall'altro lo stile imposto dalla lettura potrebbe aver annullato, o almeno ridotto, alcune caratteristiche peculiari delle voci del campione.

5. *La ricerca*

La presente ricerca trae origine proprio dalle criticità metodologiche rilevate da Sorianello, Gurrado (in stampa). Partendo dai risultati già conseguiti, abbiamo operato una duplice modifica al protocollo di ricerca: da un lato abbiamo ridotto il campione delle locutrici da cinque a tre, lasciando solo le voci più rappresentative, cioè una voce acuta (CI), una voce media (GM) e una grave (RV) (*ultra*), dall'altro

abbiamo ampliato la gamma dei materiali, affiancando, in un'ottica contrastiva, il parlato letto a un corpus di parlato spontaneo.

Lo studio intende esplorare più aspetti. In primo luogo, vuole verificare se un numero più ristretto di voci produca una migliore polarizzazione dei giudizi percettivi in rapporto alla dimensione di 'dominanza/sottomissione'. L'ipotesi di lavoro, fondata sui principi del Codice della Frequenza, è che la voce grave (RV) sia, a una valutazione percettiva, più spesso associata ai tratti della dominanza, mentre la voce acuta (CI) si disponga in senso opposto, classificandosi come voce non dominante. In secondo luogo, sarà possibile verificare se i giudizi degli uditori risentano del genere stilistico (parlato letto vs. parlato spontaneo). Il perseguimento di tali obiettivi è subordinato alla verifica acustica dei materiali che pertanto ha costituito il primo stadio della ricerca.

5.1 I partecipanti

Per dare fondatezza alle nostre ipotesi, la fase preliminare dello studio ha riguardato la selezione dei parlanti. A tal fine, sono state scelte informatrici di sesso femminile, aventi, a un riscontro uditivo, voci molto diverse fra loro e particolarmente marcate in frequenza. Il campione analizzato si compone di tre partecipanti, da ora indicate con le iniziali del loro nome (CI, GM, RV), di pari provenienza ed estrazione socio-culturale. Le informatrici sono nate e vissute a Bari, hanno un'età compresa tra 22 e 25 anni e sono in possesso di laurea. Nel dettaglio, la voce CI è percepita come acuta, la voce RV mostra un timbro vocale particolarmente grave, mentre la voce GM si colloca infine in una posizione intermedia tra le due. Le partecipanti hanno compilato un breve questionario su provenienza, età, grado di istruzione, contenente pure specifiche domande su peso, altezza e principali interessi.

5.2 I materiali

Ogni partecipante è stata coinvolta in due diversi *task* linguistici, la lettura di un testo e la descrizione di una ricetta gastronomica². La registrazione dei materiali è avvenuta in un ambiente insonorizzato con strumenti digitali (WAV. format, 44100 Hz, 32 bit). I soggetti hanno accettato il consenso informato sul trattamento dei loro dati personali e sull'utilizzo dei materiali audio, ai soli fini scientifici.

5.3 Il protocollo

Il disegno sperimentale dello studio prevede due fasi complementari coincidenti la prima con una analisi acustica, la seconda con una analisi percettiva. La prima fase ha un valore preliminare, poiché consente di caratterizzare la produzione vocale delle tre informatrici sul piano spettro-acustico, un dato imprescindibile per poter poi interpretare i risultati del test percettivo e dunque l'associazione di ciascuna voce con la dimensione di 'dominanza/sottomissione'.

² Si tratta di un testo giornalistico, tratto da *La Repubblica* dal titolo *Perché i Bronzi di Riace non andranno all'Expo di Milano?*, pubblicato il 10/04/2014.

Per la verifica acustica delle registrazioni abbiamo considerato i seguenti parametri: 1) valore medio, massimo e minimo della frequenza fondamentale ($f_0\bar{x}$, $f_{0\min}$, $f_{0\max}$), 2) escursione melodica convertita in semitoni (EM, ST), 3) intensità media (INT, dB), 4) velocità di eloquio (VE, sill/sec), 5) computo della *Vowel Space Area* (VSA)³, previa estrazione dei valori in Hz di F1 e F2 delle 3 vocali cardinali toniche /i a u/⁴. Questo parametro, il cui risultato è espresso in Hz^2 , definisce l'area di un poligono irregolare. Si tratta di un indice utilizzato per valutare l'ampiezza dello spazio vocalico, anche in caso di disturbi linguistici, ma è soventemente sfruttato in ambito dialettologico per cogliere, in modo contrastivo, le differenze globali della dimensione tra sistemi vocalici marcati sul piano sociolinguistico, ad es. per stile o diatopia (in questa direzione, Jacewicz, Fax & Salmons, 2007). VSA si calcola applicando la formula riportata in (1).

$$(1) \text{VSA: } [(F1i (F2a-F2u) + F1a (F2u-F2i) + F1u (F2i-F2a)]/2$$

L'analisi è stata condotta mediante il *software Praat* (versione 5.0.16, Boersma, Weenink, 2013). La significatività statistica è stata testata mediante t-test e ANOVA, per le comparazioni multiple è stato impiegato il test *post-hoc* di *Tukey*, la significatività statistica è stata impostata per $p < 0.05$.

6. Risultati

Gli esiti acustici avvalorano l'impressione iniziale, in quanto le informatrici manifestano più tratti fonetici differenzianti che non riguardano solo il *pitch*, ma coinvolgono, in modo globale, anche gli altri parametri esaminati. Di seguito, esporremo i risultati ottenuti per la frequenza fondamentale (6.1), il timbro vocalico (6.2) nonché per l'intensità e la velocità elocutiva (6.3).

6.1 La frequenza fondamentale

Come atteso, il *pitch* si conferma il parametro più robusto ai fini della distinzione delle voci delle tre informatrici, le quali effettivamente presentano valori di f_0 differenziati. Molte delle tendenze rilevate mostrano, inoltre, una buona sistematicità a un confronto tra parlato letto e spontaneo. Per quantificare tali differenze, abbiamo assunto quale riferimento la $f_0\bar{x}$ e abbiamo computato lo scarto tra le singole voci, adottando quale unità di misura il semitono. Tale confronto è sempre significativo. Precisamente, nel parlato letto la distanza tra la $f_0\bar{x}$ di CI e RV, le voci più rappresentative, è di 4,2 ST (dev.st. 0,2), mentre nello spontaneo è pari a 4 ST (dev.st. 1,1). Minore è la distanza tra queste voci e GM, rispettivamente, per la lettura, CI/GM = 2,3 (dev.st. 0,4) e RV/GM = 1,9 (dev.st. 0,7); valori simili si attestano anche nello spontaneo (CI/GM = 2,2, dev.st. 0,5; RV/GM = 1,8, dev.st. 1,6).

³ Le ricerche che hanno applicato VSA sono numerose, fra tutte si vedano gli studi di Higgins, Hodge (2002); Vorperian, Kent (2007); Flipsen, Lee (2012).

⁴ L'estrazione delle formanti è stata effettuata manualmente mediante tecnica FFT, in corrispondenza dello *steady state* delle vocali.

Nel complesso, il comportamento di f_0 è più regolare nel parlato letto che non nello spontaneo, sebbene le differenze siano sempre significative [Lettura, $F(2,33) = 117,099$, $p = .000$; Spontaneo, $F(2,33) = 36,305$, $p = .000$]. Nello spontaneo, malgrado ciò, alcune differenze si riducono: è quanto avviene per f_{0min} e f_{0max} di GM e RV ($p = .88$), si veda la rappresentazione della Figura 1.

Figura 1 - Valori medi dei 3 parametri di f_0 (in Hz) per le tre locutrici nel parlato letto (L) e spontaneo (Sp)

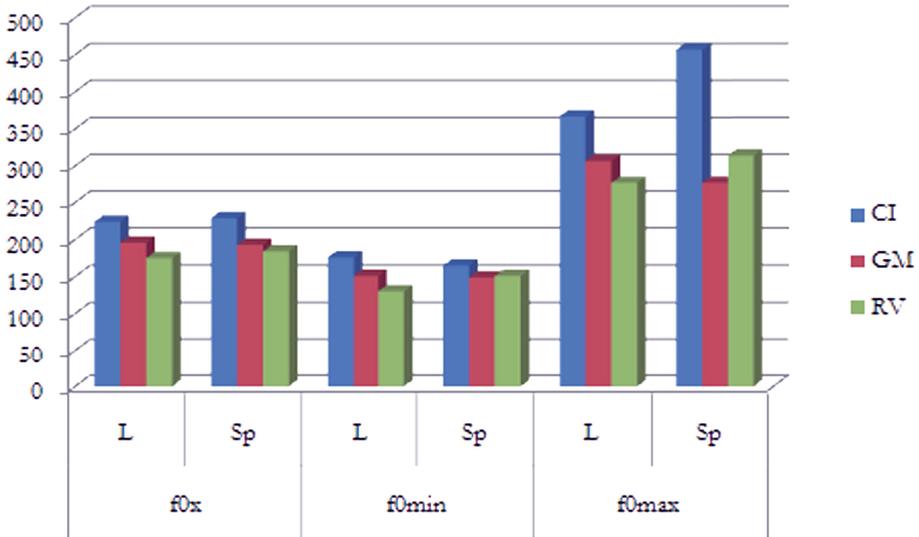
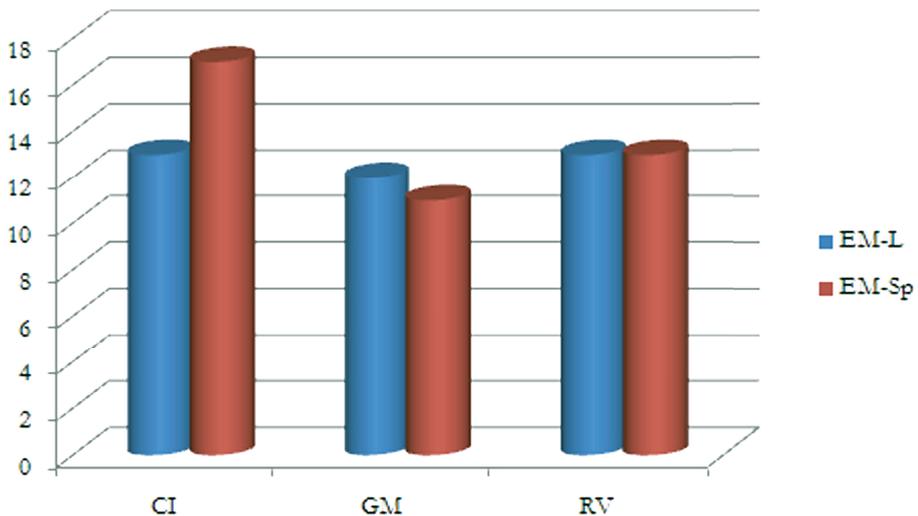


Figura 2 - Valori medi relativi all'Escursione Melodica in semitoni per le tre locutrici nel parlato letto (L) e spontaneo (Sp)



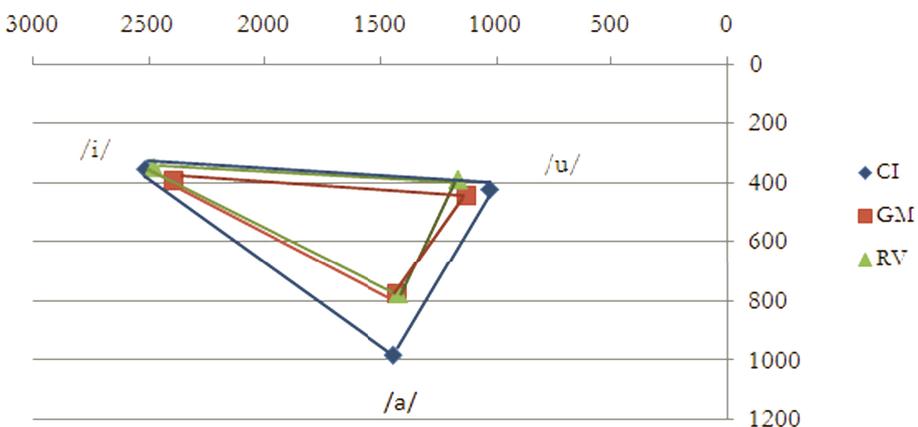
Per quanto concerne l'escursione melodica, nel parlato spontaneo le parlanti presentano valori più bassi rispetto allo stile di lettura, ma anche più variegati, la dispersione intorno alla media è di circa 2 ST per tutte le locutrici (Figura 2). I dati sull'EM, inoltre, essendo piuttosto simili, non acquisiscono uno *status* differenziante ai fini della caratterizzazione delle tre voci; l'analisi della varianza in merito a questo parametro non è, infatti, significativa, fatta eccezione per l'EM di CI estratta per lo spontaneo.

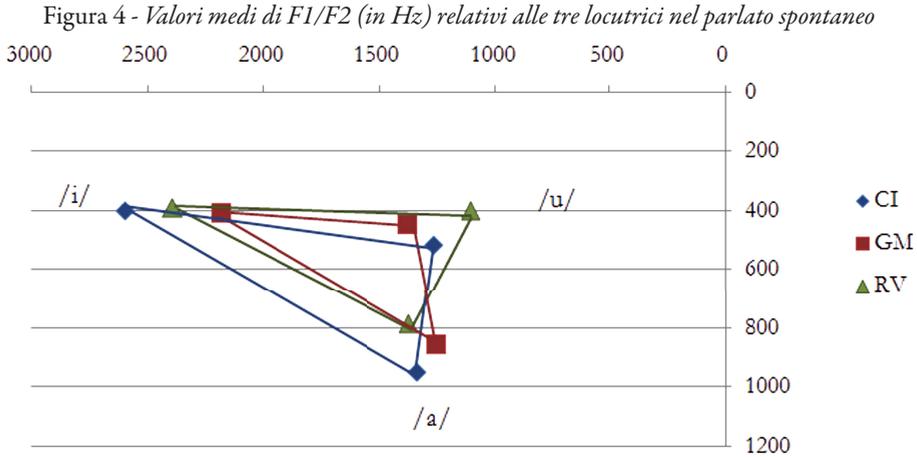
6.2 Il timbro vocalico

Dai risultati ottenuti per le frequenze formantiche si desume che il timbro concorre a differenziare le locutrici. Nei due stili considerati (letto e spontaneo), l'assetto vocalico mostra diversi elementi distintivi che riguardano l'articolazione delle vocali nella loro duplice dimensione: altezza e grado di anteriorità/posteriorità, come si vede dai sistemi vocalici lineari (in Hz) rappresentati nelle Figure 3 e 4.

È tuttavia emblematico osservare che, a parità di stile, le informatrici si comportano diversamente. Nel parlato letto, con riferimento alle parlanti GM e RV, l'aspetto più cospicuo riguarda lo slittamento verso il centro della vocale /a/, una contrazione che riduce sensibilmente l'area vocalica nel suo complesso, con specifica ripercussione sulla dimensione dell'altezza vocalica. Diversamente, CI presenta uno spazio vocalico ampio e ben definito, dalla classica forma triangolare. Tali effetti sono ragionevolmente da ascrivere a differenze anatomiche, come ad esempio la dimensione della cavità orale, la grandezza della mandibola, ma anche a una certa postura articolatoria che potrebbe condizionare il grado di protrusione labiale o determinare la faringalizzazione delle vocali.

Figura 3 - Valori medi di F1/F2 (in Hz) relativi alle tre locutrici nel parlato letto





Nel parlato spontaneo il quadro è meno nitido. Uno stile meno controllato e più espressivo determina i ben noti processi di riduzione e di centralizzazione timbrica. Tuttavia, i parlanti non si comportano allo stesso modo. Al pari del parlato letto, GM presenta un assetto vocalico centralizzato, il posizionamento dei timbri cardinali causa una complessiva riduzione dello spazio. Più contenuta è invece la contrazione del sistema vocalico di RV; in questa parlante sia /i/ che /u/ mantengono una posizione periferica, mentre in CI i timbri più stabili, meno centralizzati, sono /i/ e /a/. Il comportamento vocalico osservato per le tre parlanti potrebbe indurre un effetto timbrico particolare, non indifferente ai fini della successiva valutazione percettiva delle voci.

Per quantificare con maggiore precisione le dinamiche articolatorie delle tre parlanti, e quindi giungere a una stima quantitativa delle differenze vocaliche, abbiamo ritenuto utile calcolare anche l'Area dello Spazio Vocalico (in Hz²) o *Vowel Space Area*, da ora VSA.

I valori medi relativi a VSA sono riportati nella Figura 5. Nella Tabella 1 abbiamo invece calcolato per le tre locutrici le variazioni percentuali che interessano VSA. I risultati sono congruenti con quanto già descritto, dato che lo scarto più cospicuo concerne CI e GM, tanto nella lettura ($\pm 41\%$) quanto nello spontaneo ($\pm 33,7\%$). In quest'ultimo stile, lo scenario è poco sistematico, in CI e RV la dimensione di VSA è approssimabile ($\pm 11\%$), contraddicendo il *trend* riscontrato per la lettura.

C'è inoltre un condizionamento dipendente dal registro stilistico (Tabella 2), sebbene gli effetti prodotti siano eterogenei. Prevedibilmente, la dimensione di VSA è più ampia nella lettura che nello spontaneo; la distanza tra le due forme di parlato è rilevante, in special modo per CI e GM. Nel dettaglio, in RV, VSA subisce una riduzione del 16,7%, nelle locutrici CI e GM la variazione di VSA dal passaggio tra letto e spontaneo è maggiore.

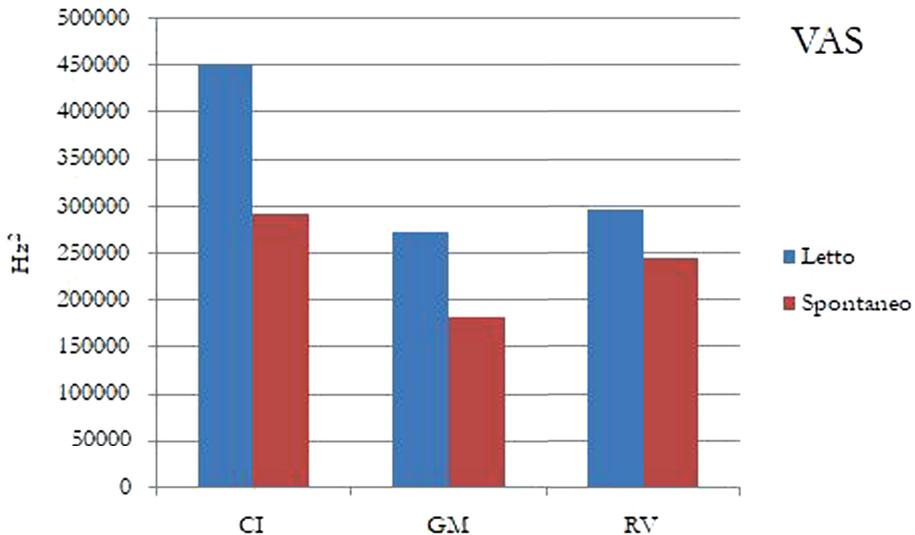
Figura 5 - Valori medi di VSA (in Hz²) per le tre locutrici nel parlato letto (L) e spontaneo (Sp)

Tabella 1 - Differenze percentuali tra la dimensione della VSA nei due stili

Soggetti	Letto vs Spontaneo
CI	± 47,8%
GM	+ 36,7%
RV	+ 16,7%

Tabella 2 - Differenze percentuali tra la dimensione della VSA tra parlanti

Soggetti	Letto	Spontaneo
CI/RV	+ 39,8%	+ 11%
CI/GM	+ 42%	+ 33,7%
GM/RV	+ 11,2%	+ 22%

6.3 Intensità e velocità elocutiva

L'intensità ha un comportamento nel complesso significativo, [Lettura, $F(2,33) = 14,595$, $p = .000$; Spontaneo $F(2,33) = 10,586$, $p = .000$]. I confronti multipli dimostrano che nella lettura lo scarto tra CI e GM si pone sulla soglia di accettabilità ($p = .05$); tutte le altre comparazioni sono però rilevanti. Nel parlato spontaneo, solo il confronto tra GM e RV non è significativo ($p = .201$). Non si evincono apprezzabili differenze dipendenti dallo stile: nel complesso, CI e GM mostrano valori di poco più elevati nel parlato spontaneo, per cui lo scarto tra i due stili per le due locutrici non è significativo, (rispettivamente, $p = .40$ e $p = .12$), mentre in RV la distanza tra letto e spontaneo è trascurabile ($p = .42$), si veda in merito la Figura 6.

La velocità elocutiva è più sostenuta nella lettura, un risultato atteso (Figura 7), considerato il numero maggiore di pause silenziose che caratterizza il parlato spontaneo. Tutte le differenze sono significative sia a un confronto tra parlanti, sia a un confronto tra stili [Lettura, $F(2,21) = 7,988$, $p = .003$; Spontaneo $F(2,21) = 31,790$, $p = .000$]. Il test post-hoc di Tukey raggiunge sempre la soglia di significatività, ad eccezione del confronto tra CI e RV nel letto ($p = .824$).

Figura 6 - Valori medi dell'Intensità (in dB) per le tre locutrici nel parlato letto (L) e spontaneo (Sp)

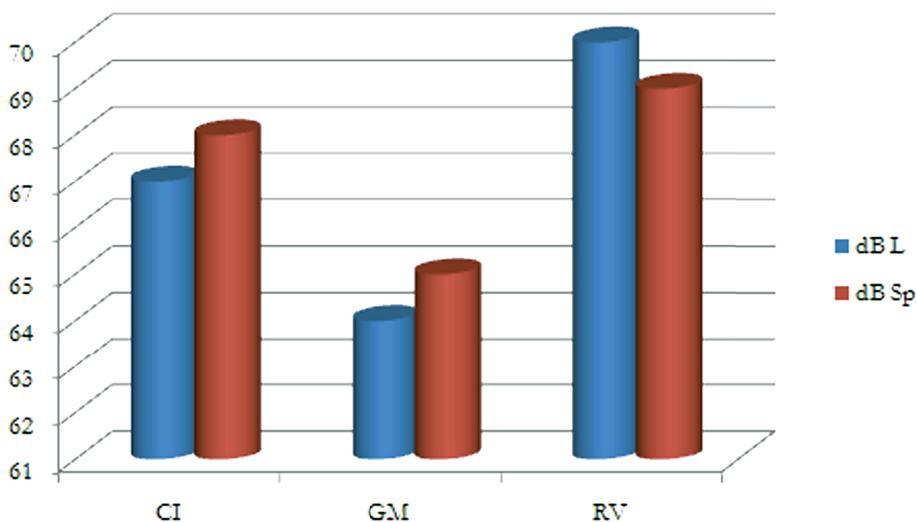
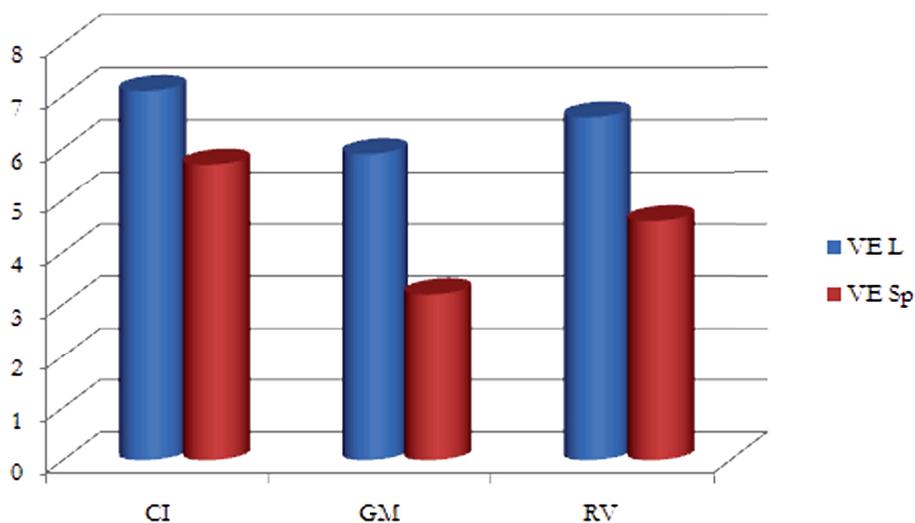


Figura 7 - Valori medi della Velocità di Eloquio (in sillabe/secondo) per le tre locutrici nel parlato letto (L) e spontaneo (Sp)

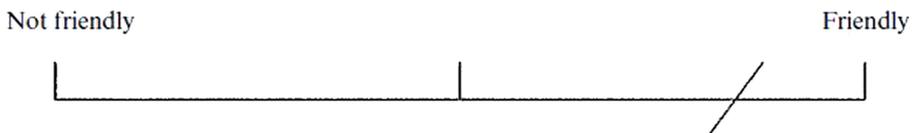


7. *Il test percettivo*

Il test ha avuto lo scopo di verificare l'esistenza di una relazione tra la costituenza acustica delle voci, in primo luogo il *pitch*, e alcuni marcatori della personalità. Il test è stato somministrato a un gruppo di 50 studenti universitari (36 F e 14 M) provenienti dal territorio barese e peri-barese, aventi un'età media di 23.8 anni. Agli uditori è stata inizialmente fornita una scheda cartacea contenente cinque coppie di tratti semantici rappresentati mediante una *Visual Analogue Scale* (VAS)⁵, (Chen *et al.*, 2004). Per la percezione dei tratti della personalità, abbiamo scelto una scala concepita in modo continuo, la VAS per l'appunto, che consiste di un asse graduato utilizzato prevalentemente in ambito clinico per misurare quegli stati variabili come l'umore o il dolore. Nel dettaglio, la VAS è una scala di 10 cm ai cui estremi, a sinistra 0 e a destra 10, sono posizionati in modo antitetico gli elementi lessicali di una coppia semantica associata a un tratto della personalità o della fisicità. Un esempio di VAS è riportato nella riproduzione grafica della Figura 8. Diversamente da quanto effettuato da Chen *et al.* (2004), all'estremità sinistra della VAS abbiamo posto i termini denotanti la sfera della sottomissione, mentre a destra quelli relativi alla sfera della dominanza. Gli aggettivi selezionati per il test sono i seguenti: *Piccolo/Grande, Insicuro/Sicuro di sé, Sottomesso/Dominante, Vulnerabile/Aggressivo, Desideroso di benevolenza/Deciso*.

Agli uditori è stato chiesto di ascoltare un sottocampione di stimoli audio delle tre voci, ordinati in modo casuale, e di formulare un giudizio sulla parlante, assumendo quale riferimento le coppie di aggettivi rappresentate. Durante l'ascolto, ogni giudice è stato invitato a marcare il punto della VAS in cui a suo avviso si collocava, nella voce proposta, il tratto semantico considerato. I punteggi sono stati ottenuti sommando i millimetri intercorrenti tra l'estremità sinistra, ossia il polo contrassegnato con il valore 0, e il punto marcato dall'ascoltatore.

Figura 8 - *Rappresentazione grafica della Visual Analogue Scale*
(fonte: Chen *et al.*, 2004: 320)



L'ipotesi sperimentale è che le voci, essendo differenziate per *pitch*, suscitino negli uditori giudizi diversi. Precisamente, ci aspettiamo che la voce grave (RV) sia associata più frequentemente, nella valutazione percettiva, a quei tratti della personalità che trasmettono dominanza e sicurezza e che pertanto superi sulla VAS il punteggio

⁵ Diversamente, la scala Likert, a 5 o a 7 punti, originariamente impiegata in campo sociologico per la misurazione delle opinioni, si serve di una batteria di affermazioni rispetto alle quali i soggetti devono dichiarare il loro grado di accordo (Marradi, Gasperoni, 2005). Avendo carattere discreto, la scala Likert ci è sembrata quindi meno adatta rispetto alla VAS, almeno in questa fase esplorativa della ricerca.

semanticamente neutro, cioè cinque. Diversamente, per la voce acuta (CI), ci attendiamo che i giudizi, inferiori a cinque, si collochino in modo rilevante nella sezione associata ai tratti di benevolenza e sottomissione.

7.1 I giudizi percettivi

I giudizi raccolti non si prestano a una lettura univoca. Com'è evidente dai dati riportati nella Tabella 3, tanto nel parlato letto che nello spontaneo, non c'è una decisa polarizzazione degli esiti. In ambedue gli stili, sebbene in modo più evidente nello spontaneo, si nota una concentrazione consistente dei giudizi intorno all'area centrale della VAS. I risultati medi computati per ogni locutrice e per ogni coppia semantica non sono sempre significativamente differenziati. Nel parlato letto, la situazione è nel complesso più definita, i giudizi sono statisticamente significativi [$F(2,736) = 102,629, p = .000$], come pure tutte le comparazioni multiple. Limitandoci al confronto tra CI e RV, osserviamo che nello stile lettura le tendenze attese sono per buona parte rispettate, anche se le differenze sono spesso di modesta entità. Una parte dei giudizi espressi per la voce più grave RV si colloca, sulla VAS, nella sezione dei valori superiori a 6 (media 6,4, dev.st. 2); la valutazione propende, anche se non in modo netto, verso il riconoscimento in questa voce dei tratti di sicurezza e dominanza. Al contrario, CI detiene i giudizi meno elevati, in media 5,6 (dev.st. 1,7). Ciò nonostante, nelle coppie *insicuro/sicuro di sé* e *desideroso di benevolenza/deciso* gli *scores* percettivi, a un confronto tra CI e RV, sono statisticamente irrilevanti (t-test), rispettivamente: *insicuro/sicuro di sé* Lettura: $p = .89$, Spontaneo: $p = .30$; *desideroso di benevolenza/deciso*, Lettura: $p = .17$, Spontaneo: $p = .72$.

Tabella 3 - Valori medi e deviazione standard (tra parentesi) dei giudizi percettivi ottenuti per ogni coppia di tratti semantici, nel parlato letto (L) e spontaneo (Sp)

	Piccolo/ Grande		Insicuro/ Sicuro di sé		Sottomesso/ Dominante		Vulnerabile/ Aggressivo		Desideroso Benevolenza/ Deciso		Media	
	L	Sp	L	Sp	L	Sp	L	Sp	L	Sp	L	Sp
CI	5 (1,7)	4,6 (1,9)	6,6 (1,8)	5,5 (2,7)	5,4 (1,5)	6 (1,9)	4,5 (1,3)	5,1 (2,1)	6,5 (2)	5,6 (1,7)	5,6 (1,7)	5,4 (2,1)
GM	5,3 (2,5)	5,9 (2,6)	3,2 (2)	3,4 (2,5)	3,3 (1,6)	4 (2,3)	2,8 (1,7)	3,7 (2,5)	3,1 (1,7)	5,4 (2,7)	3,5 (1,9)	4,1 (2,4)
RV	6,1 (2)	5,8 (1,9)	6,7 (1,9)	6 (2,3)	6,8 (1,7)	5,5 (2)	5,8 (1,7)	5 (1,9)	6,7 (2,3)	5,8 (2,4)	6,4 (2)	5,6 (2,1)

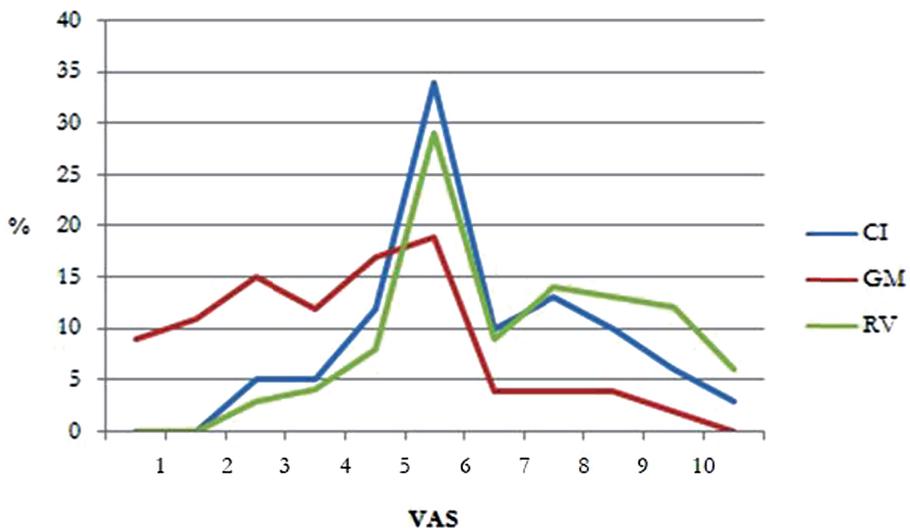
Nel parlato spontaneo, la situazione è nell'insieme significativa [$F(2,737) = 28,64, p = .000$]; dai confronti multipli emerge però che le differenze tra CI e RV sono trascurabili ($p = .591$). Se si esclude la dicotomia *piccolo/grande*, i giudizi mostrano una sostanziale omogeneità, il loro confronto non è infatti statisticamente significativo. Mediamente, CI ottiene sulla VAS giudizi percettivi pari a 5,4 (dev.st. 2,1), un indice di poco inferiore rispetto a quello attribuito durante lo stile lettura ($p = 0.41$),

mentre RV ha un valore pari a 5,6 (dev.st. 2,1); in quest'ultimo caso, la differenza tra letto e spontaneo è significativa ($p = 0.01$). C'è un aspetto che comunque non passa inosservato: i dati dello spontaneo sono complessivamente più vicini al valore mediano della VAS, ossia 5. Ciò sembra denotare la presenza di una sorta di incertezza da parte degli uditori nel valutare quanto richiesto, che pertanto scelgono, molto spesso, di apporre un valore pressoché neutro, per questo anche poco indicativo.

In decisa controtendenza si pone la voce dell'informatrice GM che, pur non essendo la più acuta, ottiene, nei due stili di parlato, i valori più bassi, rispettivamente 3,5 (dev.st. 1,9) nella lettura e 4,1 (dev.st. 2,4) nello spontaneo ($p = 0,01$), indice dell'attribuzione in questa voce di quei tratti della personalità che convogliano indecisione, sottomissione e vulnerabilità. Lo scarto è significativo, sia rispetto allo stile (letto vs. spontaneo: $p = .011$) sia rispetto alle altre voci ($p = .002$). È plausibile pensare che i giudizi degli ascoltatori siano stati in qualche modo deviati, in questa parlante, dalla concomitante presenza di una più bassa intensità sonora e di una velocità elocutiva rallentata (cfr. Figura 6 e 7), elementi che hanno finito per conferire al suo eloquio un certo alone di insicurezza, un risultato che conferma quanto già emerso in Sorianello, Gurrado (in stampa).

Con l'intento di accertare la distribuzione dei giudizi rispetto ai punteggi della VAS, abbiamo computato gli indici di frequenza e li abbiamo rappresentati graficamente in termini percentuali (v. Figura 9 e 10).

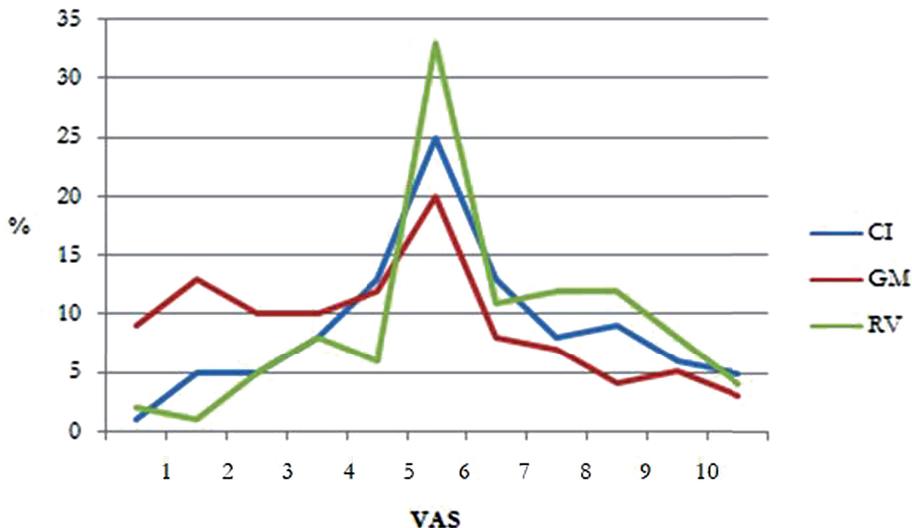
Figura 9 - *Parlato Letto: distribuzione percentuale dei giudizi percettivi assegnati alle tre voci*



Da ciò risulta che nel parlato letto (Figura 9), la distribuzione dei giudizi tra CI e RV mostra una certa diversificazione solo nella sezione sinistra della VAS; precisamente il 34% (CI) e il 29% (RV) di tutti i giudizi ottenuti per queste due locutrici coincide con il valore mediano di 5. Tuttavia, il 13% dei giudizi formulati per RV si colloca

tra i punteggi 7 e 9, una percentuale che, per i medesimi indici numerici, scende al 10% in CI. Una simile tendenza si ravvisa anche nello spontaneo (Figura 10).

Figura 10 - *Parlato Spontaneo: distribuzione percentuale dei giudizi percettivi assegnati alle tre voci*



Nuovamente, i giudizi pari a 5 sono maggiori in RV (33%) che non in CI (25%), laddove quelli superiori a 5 sono pari al 12% in RV, ma all'8% in CI, limitatamente al punteggio 8. La rappresentazione grafica mette in luce anche la posizione particolare dell'informatrice GM. Tanto nella lettura che nello spontaneo, i giudizi riguardanti GM che ricadono nella parte inferiore della VAS sono decisamente maggiori, in termini percentuali, se confrontati a quelli assegnati dagli uditori alle altre due parlanti: precisamente, nella lettura i punteggi 2 e 3 sono rappresentati con percentuali pari al 12 e 15%, nelle altre locutrici tali punteggi, invece, raggiungono appena il 5%, nello spontaneo. Va inoltre osservato che i giudizi percettivi formulati per GM si mantengono intorno al 10%, con una punta del 13% per il punteggio 1, diversamente da CI e RV che mostrano percentuali più basse. In aggiunta, si noti che gli indici ottenuti da GM rientranti nella sezione mediana sono percentualmente minori, 19% nella lettura, 20% nello spontaneo.

8. Riflessioni

La domanda che ha ispirato l'intera ricerca è se voci femminili diverse siano percettivamente valutabili, in modo altrettanto diverso, con riferimento alla sfera della dominanza fisica e sociale della parlante. La caratterizzazione acustica delle voci ha costituito un tassello preliminare, al fine di poter determinare quanto le voci fossero effettivamente diverse e quindi procedere con un'interpretazione più consapevole dei giudizi percettivi. Un primo aspetto degno di rilievo conferma che l' f_0 media

dei soggetti considerati è realmente differente. Ciò è sufficiente a conferire alle voci una propria identità percettiva. Tuttavia, i dati sulle frequenze formantiche e sulla dimensione complessiva della VSA tendono solo in parte nella medesima direzione. La parlante avente la frequenza fondamentale più alta, CI, mostra anche la maggiore superficie vocalica, nettamente al di sopra rispetto all'area rinvenuta per RV, la locutrice dalla f_0 media più bassa. Due aspetti confondono però il quadro interpretativo: nel parlato spontaneo, le differenze timbriche a volte si attenuano (CI vs. RV), a volte si rinforzano (GM vs. RV), probabilmente per via di quella elasticità espressiva, poco controllata, che distingue il parlato spontaneo. Inoltre, in GM si registra l'area vocalica più piccola, a prescindere dallo stile, sebbene tale locutrice abbia una $f_0\bar{x}$ perfettamente in linea con la media prevista per una femmina adulta. Nel complesso, il *pitch* e il timbro vocalico sembrano prefigurarsi dei solidi *discrimen* acustici, ma non sono gli unici. A intricare il quadro, va aggiunto che altri parametri concorrono alla diversità vocale delle informatrici, come il volume e la velocità elocutiva. Il peso svolto da tutti questi indici non sembra risentire in modo incisivo dell'effetto stile, giacché tutte le tendenze emerse nel parlato letto si ripresentano, seppure in misura diversa, anche nello spontaneo. Né d'altra parte possiamo escludere che altri fattori, da noi non esplorati, condizionino ulteriormente la situazione descritta; è il caso dei contorni intonativi oppure della qualità della voce⁶.

L'analisi acustica ha dimostrato che le voci del nostro campione sono diverse e che tale 'diversità' sia riconducibile a più parametri. Ciò nonostante, durante il test percettivo, i giudizi degli ascoltatori tesi a saggiare il carattere più o meno dominante di una voce non si distribuiscono in modo uniforme. L'associazione di un *pitch* grave con una personalità dominante, e viceversa, di un *pitch* acuto con una personalità non dominante, assume, infatti, solo un carattere tendenziale. I punteggi si sfrangano, distribuendosi su tutta la lunghezza della VAS, sebbene alcune sezioni siano emblematicamente più rappresentate. A un confronto con la precedente ricerca (Soriano, Gurrado, in stampa), limitata solo al parlato letto, si osserva che la riduzione del campione delle informatrici, da cinque a tre, non ottimizza i risultati. Gli ascoltatori sembrano spesso indecisi, non propendendo in modo netto verso l'uno o l'altro tratto della personalità proposto attraverso le cinque scale semantiche. Questa ipotesi sembra avvalorata anche dall'alta densità di giudizi che ricadono nell'area mediana della VAS, rappresentata dal valore 5, un punteggio che, nei casi implicati, sembra neutralizzare il giudizio formulato dagli ascoltatori che pertanto lo sospendono, assegnando alla personalità della parlante un significato né dominante, né sottomeso. L'ipotesi che vi sia una correlazione tra il *pitch* della voce e il suo significato in termini di 'dominanza/sottomissione' trova dunque solo una debole conferma.

La questione affrontata in questa seconda parte della ricerca è per alcuni versi insidiosa. Dire che le voci delle tre locutrici analizzate sono acusticamente diverse non

⁶ In merito a quest'ultimo aspetto, abbiamo già avviato una prima verifica, analizzando, per un campione di vocali pronunciate in modo prolungato, i parametri di *shimmer*, *jitter* e *HNR*. Le locutrici non sembrano tuttavia manifestare differenze rilevanti, ad eccezione di GM in cui il valore di *jitter* è più alto di quello attestato negli altri soggetti.

presume automaticamente che ognuna di esse richiami nel giudizio di un ascoltatore gradazioni della personalità altrettanto differenti. Eppure sappiamo che proprio attraverso la voce, strumento semiotico molto potente, ‘generatrice di sensi’ come direbbe Albano Leoni (2009), viaggiano informazioni di varia natura, affettive e sociali, oltre che squisitamente linguistiche. Si tratta di fondamentali indizi sociofonetici relativi, ad esempio, alla provenienza geografica del parlante, alla sua appartenenza sociale. Accanto a questi, troviamo pure indizi di natura paralinguistica, spesso involontari e incontrollabili, si pensi agli stati d’animo, e finanche informazioni extralinguistiche, come quelle inerenti al sesso, all’età o allo stato di salute. Ora, da questa complessa mescolanza di elementi, stabilire quanta parte l’ascoltatore decida di trattenere, assegnandovi un precipuo significato, è un’operazione tutt’altro che semplice. Pur restringendo il campo di osservazione alle componenti della personalità, e nello specifico alla dimensione della dominanza, l’aspetto che ci riguarda più da vicino, resta da capire come avvenga, nella mente dell’ascoltatore, l’associazione tra una categoria sonora, la voce, e la scelta di una o più qualità della personalità dello stesso locutore. C’è da chiedersi, dunque, cosa induca un ascoltatore a classificare una voce come più o meno dominante, in assenza di contatto visivo o di conoscenza diretta del parlante. È d’altra parte plausibile ritenere che i significati affettivi e sociali che affiorano dall’ascolto di una semplice voce siano per buona parte guidati da stereotipi culturali e sociali. In questo ambito di indagine, che per molti aspetti sconfinava anche verso una dimensione più prettamente psicologica, non sarà fuori luogo menzionare la ben nota tecnica del *matched-guise*, una procedura che coinvolge gruppi di ascoltatori ignari per giudicare più versioni di una stessa voce ‘travestita’, al fine di esplorare la dinamica degli stereotipi sociali. In questo caso, giudizi diversi, ma in realtà espressi per uno stesso individuo, tradiscono involontariamente la presenza dello stereotipo. Nella nostra ricerca, l’assenza di un confronto diretto tra voci femminili e voci maschili potrebbe aver in parte disorientato l’uditorio che ha cercato, non sempre riuscendoci, di estrapolare dal campione femminile una gamma di significati relativi al tratto di ‘dominanza/sottomissione’ posizionandoli su un asse continuo. C’è tuttavia un aspetto che potrebbe gettare altra luce su questa complessa tematica, ovvero lo studio degli effetti ‘percettivi’ prodotti per una stessa voce dalla manipolazione, tramite sintesi, di alcuni parametri acustici. A tale argomento intendiamo dedicarci nel prossimo futuro.

Riferimenti bibliografici

- ALBANO LEONI, F. (2009). *Dei suoni e dei sensi. Il volto fonico delle parole*. Bologna: Il Mulino.
- APPLE, W., STREETER, L.A. & KRAUSS, R.M. (1979). Effects of pitch and speech rate on personality attributions. In *Journal of Personality Social Psychology*, 37, 715-727.
- ARONOVITCH, C.D. (1976). The voice of personality: stereotyped judgements and their relation to voice quality and sex of speakers. In *Journal of Social Psychology*, 99, 201-220.
- BOERSMA, P., WEENINK, D. (2013). *Praat: doing phonetics by computer*, <http://www.praat.org>.

- CHEN, A., GUSSENHOVEN, C. & RIETVELDT, T. (2004). Language specificity in perception of paralinguistic intonational meaning. In *Language and Speech*, 47(4), 311-349.
- COLLINS, S.A. (2000). Men's voices and women's choices. In *Animal Behaviour*, 60(6), 773-780.
- COLLINS, S.A., MISSING, C. (2003). Vocal and visual attractiveness are related in women. In *Animal Behaviour*, 65(5), 997-1004.
- DARWIN, C.R. (1872). *The expressions of the emotions in man and animals*. London: John Murray (trad. it. *L'espressione delle emozioni nell'uomo e negli animali*, Torino: Bollati Boringhieri, 1992).
- D'ERRICO, F., SIGNORELLO, R. & POGGI, I. (2013). The perception of charisma from voice. A cross-cultural study. In *Proceedings of the IEEE Conference on Affective Computing and Intelligent Interaction*. Geneva, Switzerland, 552-557.
- FEINBERG, D.R., JONES, B.C., LITTLE, A.C., BURT, D.M. & PERRET, D.I. (2005). Manipulation of fundamental and formant frequencies influence the attractiveness of human male voices. In *Animal Behaviour*, 69, 561-568.
- FERRERO, F.E., MAGNO CALDOGNETTO, E. & COSI, P. (1995). Le vocali al femminile. In MARCATO, G. (Ed.), *Donna e linguaggio*. Padova: CLUEP, 413-436.
- FERRERO, F., MAGNO CALDOGNETTO, E. & COSI, P. (1996). Sui piani formantici acustici e uditivi delle vocali di uomo, donna e bambino. In PERETTI, A., SIMONETTI, P. (Eds.), *XXIV Convegno Nazionale dell'Associazione Italiana di Acustica*. Trento: Arti Grafiche Padovane, 169-178.
- FLIPSEN, P., LEE, S. (2012). Reference data for the American English acoustic vowel space. In *Clinical Linguistics and Phonetics*, 26, 926-933.
- GUSSENHOVEN, C. (2002). Intonation and interpretation: phonetics and phonology. In BEL, B., MARLIEN, I. (Eds.), *Proceedings of the International Conference on Speech Prosody 2002*. Aix-en-Provence, Université de Provence, 47-57.
- HENTON, C. (1989). Fact and fiction in the description of female and male pitch. In *Language and Communication*, 9, 299-311.
- HIGGINS, C., HODGE, M. (2002). Vowel area and intelligibility in children with and without dysarthria. In *Journal of Medical Speech and Language Pathology*, 10, 271-277.
- JACEWICZ, E., FAX, R.A. & SALMONS, J. (2007). Vowel space areas across dialects and gender. In TROUVAIN, J., BARRY, W.J. (Eds.), *Proceedings of the 16th International Congress of Phonetic Sciences*. Saarbrücken, Germany, 1465-1468.
- JIANG, H. (2011). Gender differences in English intonation. In WAI-SUM, L., ZEE, E. (Eds.), *Proceedings of the 17th International Congress of Phonetic Sciences*. Hong Kong, China, 974-977.
- MARRADI, A., GASPERONI, G. (2005). *Costruire il dato, 3: Le scale Likert*. Milano: Franco Angeli.
- MORTON, E.S. (1977). On the occurrence and significance of motivation-structural rules in some bird and mammal sounds. In *The American Naturalist*, 111, 855-869.
- OHALA, J.J. (1983). Cross-language use of pitch: an ethological view. In *Phonetica*, 40, 1-18.
- OHALA, J.J. (1984). An ethological perspective on common cross-language utilization of F0 of voice. In *Phonetica*, 41, 1-16.

- OHALA, J.J. (1994). The frequency code underlies the sound symbolic use of voice pitch. In HINTON, L., NICHOLS, J. & OHALA, J.J. (Eds.), *Sound symbolism*. Cambridge: Cambridge University Press, 325-347.
- PUTS, D.A., GAULIN, S.J.C. & VERDOLINI, K. (2006). Dominance and the evolution of sexual dimorphism in human voice pitch. In *Evolution and Human Behaviour*, 27, 283-296.
- ROSENBERG, A., HIRSCHBERG, J. (2009). Charisma perception from text and speech. In *Speech Communication*, 51(7), 640-655.
- SCHERER, K.R. (1972). Judging personality from voice: a cross cultural approach to an old issue in interpersonal perception. In *Journal of Personality*, 40, 191-210.
- SCHERER, K.R. (1979). Personality markers in speech. In SCHERER, K.R., GILES, H. (Eds.), *Social markers in speech*. Cambridge: Cambridge University Press, 147-201.
- SIGNORELLO, R., D'ERRICO, F., POGGI, I. & DEMOLIN, D. (2012). How charisma is perceived from speech. A multidimensional approach. In *ASE/IEEE International Conference on Social Computing*. Amsterdam, The Netherlands, 435-440.
- SIGNORELLO, R., DEMOLIN, D. (2013). The physiological use of the charismatic voice in political speech. In *Proceedings of Interspeech 2013*. Lyon, France, 987-991.
- SIMPSON, A.P. (2009). Phonetic differences between male and female speech. In *Language and Linguistics Compass*, 3, 621-640.
- SORIANELLO, P., GURRADO, G. (in stampa). Il Codice della Frequenza: una valutazione acustica e percettiva della voce femminile in rapporto ad alcuni tratti della personalità. In DE MEO, A., DOVETTO, F. (Eds.), *La Comunicazione Parlata*, Napoli, 2016.
- TITZE, I.R. (1989). Physiological acoustic differences between male and female voices. In *Journal of the Acoustical Society of America*, 85, 1699-1707.
- TUSING, K.J., DILLARD, J.P. (2000). The sounds of dominance: vocal precursors of perceived dominance during interpersonal influence. In *Human Communication Research*, 26, 148-171.
- VORPERIAN, H.K., KENT, R.D. (2007). Vowel acoustic space development in children: a synthesis of acoustic and anatomic data. In *Journal of Speech, Language, and Hearing Research*, 50(6), 1510-1545.