

Statistiche di f_0 per 200 parlanti di italiano

Antonio Romano & Valentina De Iacovo
LFSAG, Dip. di Lingue e L.S. e C.M. – Università di Torino

Abstract

This paper presents statistical data for the fundamental frequency of 200 Italian speakers producing the same 10 sentences. Since, in general literature, mean values for this variable are usually referred to data not supported by an extensive statistical measurement program, we propose here a sample of more than 1,000,000 values which allow to start accounting for the pitch variation in Italian speech, at least within the modal voice register and reading tasks. The material comes from the *CALL-UniTO* database (see Romano *et alii*, in this volume), for which sentences with different intonation functions were selected. The analysis of the distribution of f_0 in a large sample of speakers is intended to fill a knowledge gap for the variation of mean pitch and pitch span for the entire Italian population. In order to do this, we propose here an assessment based on a *probability density function* (PDF) of short-term f_0 values (Jassem 1971, Romano 2000, Hudson *et alii* 2007).

Introduzione¹

Non sono poche le pubblicazioni di colleghi stranieri che negli ultimi anni hanno indagato l'altezza media e l'estensione tonale della voce di parlanti di lingue diverse e in diverse condizioni di produzione². La manualistica italiana, che riporta dati orientativi per caratterizzare voci italiane maschili e femminili (generalmente senza altre distinzioni), resta tuttavia ancora poco

1. L'articolo rappresenta la versione trascritta e corredata di grafici e riferimenti bibliografici della comunicazione presentata in occasione della Giornata Mondiale della Voce 2021 (16 aprile) svoltasi presso il Laboratorio di Fonetica Sperimentale «Arturo Genre» (https://www.lfsag.unito.it/locandine/programma_wvd2021.pdf). L'intervento è presente in formato video all'indirizzo: <https://youtu.be/j03Lz0nPmkM> e il testo qui pubblicato è il risultato di un sommario adattamento a opera dei due autori (AR per il §1 e VD per il §2; entrambi per le parti comuni).

2. Come si vedrà, in quest'articolo si tiene conto della variabile acustica relativa alla frequenza della prima armonica (frequenza fondamentale, f_0) responsabile della percezione di altezza di una voce. Vale la pena ricordare che questa variabile può essere misurata con algoritmi diversi, più o meno adatti a seconda della qualità della registrazione e delle caratteristiche della voce (fisiologica, in un registro di fonazione modale). Una certa accortezza è necessaria per monitorare i salti d'ottava (rilevazione della frequenza della seconda armonica).

supportata da dati concreti. Inoltre, gli stessi descrittori della variabilità d'uso della voce nell'ambito del parlato si presentano incostanti, con limiti di rappresentazione e con approssimazioni numeriche talvolta dubbie che rendono difficilmente comparabili anche i dati prodotti sulle stesse lingue (cfr. Patterson 2000, Hudson *et alii* 2007), al punto da incoraggiare studi come quello di Vogel *et alii* (2009)³.

Le fonti più facilmente accessibili si riferiscono, infatti, a campioni di parlanti di altre lingue. La variabile più frequentemente (e spesso quasi esclusivamente) considerata è la media della frequenza fondamentale (*mean pitch*), misurata su campioni di parlato variabili da studio a studio e con scarse distinzioni intra-campione, in base agli usi e ai gruppi sociali (ma v. dopo, Hess 1983). Un altro parametro importante (sin da Jassem 1971) è l'estensione vocale (*pitch span* o *pitch range*), con valori riportati su diverse scale, e in molti casi soggetti a varie forme di normalizzazione (Shriberg *et alii* 1996, Patterson 2000)⁴.

La mancanza di dati e di rilevamenti affidabili si riflette nelle poche, ma in-

dicative, informazioni che si hanno riguardo alle voci italiane, anche in uno dei più importanti manuali di prosodia degli ultimi anni. In Soriano (2006, p. 42) abbiamo ad esempio la definizione di estensione tonale come *pitch range* riferito ai valori minimi e massimi «di un contorno intonativo» (in riferimento a uno spazio melodico ripartito «tradizionalmente» in tre fasce: bassa, media e alta)⁵.

5. Date le finalità dello studio e il profilo più generale dei suoi destinatari, quello che qui non si dice riguarda il fatto che proprio questi valori (minimi e massimi, spesso a cavallo tra *registri* o *meccanismi* di fonazione diversi, v. dopo) sono particolarmente soggetti a errori di misura. Tuttavia, altre informazioni utili sono esplicitate: la variabilità di questo parametro di caratterizzazione vocale è infatti legata ad alcuni fattori che dipendono dalle caratteristiche fisiche, dal sesso, e dall'età del parlante. A parità di contesto enunciativo si ha un'escursione maggiore nei parlanti femminili che esplorano l'«intero campo tonale» verso le regioni più alte del *range* (Soriano 2006: 42), ancora più alti nei bambini (intorno a «300 Hz»). Si precisa inoltre, utilmente, che un *pitch range* più alto corrisponde a stati d'animo di rabbia, sorpresa etc., mentre un *pitch range* compresso appare in un parlato caratterizzato da stati emotivi di tristezza o noia. È inoltre interessante la considerazione che riferisce il raggiungimento del massimo «più alto» all'enfasi e/o a circostanze emotive, visto che l'escursione tonale varia quasi esclusivamente verso l'alto (p. 43). Come si vedrà, risulta anche determinante che la delimitazione della regione inferiore sia vincolata dalle potenzialità della voce nel restare nel registro modale (la realizzazione di suoni laringalizzati, con valori spesso incostanti di f_0 , dipende infatti dal passaggio a un altro meccanismo fonatorio).

3. Sempre per l'inglese si veda anche, più recentemente, Goy *et alii* (2013).

4. Seppur valutate su scale musicali, gli studi sulla voce cantata vantano classificazioni delle estensioni vocali che risalgono anche alla seconda metà dell'Ottocento (Gutzmann 1909; cfr. Garde 1965). Per una riflessione sulle applicazioni artistiche e gli stili v. Uberti (2005).

Anche se queste definizioni meriterebbero ulteriori approfondimenti ed esemplificazioni, ci limitiamo qui a citare il dato essenziale di Sorianello (2006) che distingue il *pitch range* di locutori italiani con valori di riferimento per voci maschili, con estensione 80-160 Hz, e femminili, 160-320 Hz⁶.

D'altra parte, in virtù delle ragioni che mostreremo (e già incluse nelle considerazioni riportate nella letteratura pionieristica sull'argomento, riassunta in Romano 2001), è molto limitante continuare a riportare solo indicativamente un valore medio di f_0 per distinguere voci maschili, femminili o infantili (riferendosi ad es., come fanno alcuni manuali, a valori medi fittizi di tipo 100, 200 e 300 Hz). Aggiungere valori di estensione aiuta senz'altro, ma non consente di stabilire le reali condizioni di distribuzione dei valori, la cui concentrazione in certe regioni può lasciare emergere la sovrapposizione di statistiche diverse.

1. Qualche riferimento

Come anticipavamo, non è facile confrontare i dati di f_0 media ed estensione vocale offerti dai vari autori, perché i

6. Sebbene sia chiaro lo scopo orientativo, di massima, di questi riferimenti, soprattutto il limite inferiore di questo secondo range pare contrastare con l'affermazione finale della n. prec. e coi dati dell'esperienza di laboratorio sul parlato di locutrici, la cui frequenza fondamentale assume spessissimo valori che sono ben al di sotto di questo minimo indicativo.

valori dipendono prima di tutto dal tipo di materiale usato e dal metodo di estrazione (a quest'argomento dedicheremo infatti il §2). Un'altra fonte di variabilità, come mostra bene la rassegna di Patterson (2000), è nella scelta di rappresentare i valori considerati su scala assoluta (Hz) o musicale (semitoni, in riferimento a basi diverse), per normalizzare differenze nelle medie e nelle fasce tonali esplorate (già da voci maschili vs. femminili ad es.)⁷. Ancor più variabile è poi la scelta di esprimere il *range* di variazione come estensione totale (da un minimo a un massimo, rappresentati da valori ai limiti del registro fisiologico di produzione della voce modale e che quindi sono soggetti a errori di misurazione), oppure come deviazione standard ($\pm\sigma$; o multipli, ad es. $\pm 2\sigma$), oppure col ricorso a vari indicatori statistici elementari, legati alla posizione relativa di media, mediana e moda (v. tra gli altri, Hudson *et alii* 2007), oppure infine, più opportunamente, con altri indicatori della dispersione statistica dei valori rilevati (densità di probabilità, *ddp* o *PDF*) che includono l'asimmetria (*skewness*) e la curtosi (*kurtosis*). In questi ultimi due casi non si dà per scontato che la curva sia gaussiana (cioè che la dispersione dei valori sia

7. Le ricadute su una valutazione fonologica del riferimento a modalità diverse di considerare l'estensione sono discusse anche da Shriberg *et alii* (1996).

Speech	Fairbanks (1940)	65 - 450 Hz
	Risberg (1961b)	50 - 310 Hz
	Hadding-Koch (1961)	50 - 500 Hz
	Shaffer (1964)	110 - 500 Hz
	Hollien (1972)	80 - 300 Hz
	Rabiner et al. (1976)	50 - 500 Hz
	Monsen and Engebretson (1977)	110 - 250 Hz
Phonation	Catford (1964)	28 - 1100 Hz
	Mörner et al. (1964)	33 - 3100 Hz
	Hollien (1972)	27 - 1200 Hz
	Keating and Buhr (1978)	30 - 2500 Hz
	Monsen and Engebretson (1977)	35 - 1200 Hz
Singing	Classical music (Mozart)	50 - 1800 Hz
	Mörner et al. (1964)	65 - 1350 Hz

Fig. 1. Estensioni tonali delle voci studiate da vari autori, distinguendo in base all'attività (ma ignorando distinzioni di genere) (da Hess 1983: 64).

Tab. I. Estensioni tonali di campioni di voci di diversi parlanti (distinti per genere)⁹.

	Hollien <i>et alii</i> (1971)	Baken & Orlikoff (2000)
Voci maschili	78-698 Hz _z	85-800 Hz _z
Voci femminili	139-1108 Hz _z	160-1300 Hz _z

normale, centrata sulla media e simmetrica, cfr. Boë *et alii* 1975). La cautela è necessaria soprattutto nel caso in cui non si sia sicuri che la voce sia «fisiologica» e/o che non siano coinvolti usi e meccanismi fonatori diversi a seconda dei compiti vocali richiesti ai soggetti osservati. Ad ogni modo, fermo restando il potere descrittivo di questi parametri, l'osservazione diretta della *ddp* ha l'indubbio vantaggio di fare emergere regolarità e irregolarità nella distribuzione o preferenze per fasce frequenziali diverse a seconda dell'ambito d'uso della voce e del profilo sociofonetico del parlante (simili

valutazioni, espresse in altri termini, sono anticipate in Romano 2000)⁸.

8. Con queste premesse appaiono alcuni limiti nella rappresentatività dei dati che offriamo in questo studio (v. §2).

9. Riportiamo a parte i valori prodotti da Boë *et alii* (1975) per la sola statistica di f_0 media in voci maschili, 118 ± 18 Hz, e femminili, 207 ± 20 Hz. Anche i dati di Vogel *et alii* (2009) sono offerti per un campione limitato di voci maschili, con f_0 media di 115 ± 26 Hz e femminili con f_0 media di 170 ± 21 Hz. Nonostante lo scopo di queste ricerche sia di illustrare possibilità di normalizzazione, i bassi valori di deviazione standard – della sola f_0 media – sono indicativi di dati (in molti casi di campioni poco popolosi) che risultano relativi a condizioni di produzione scarsamente variate. Questo vale anche

Per dare anche solo alcuni riferimenti essenziali, riproduciamo qui i dati proposti da alcuni autori (v. Fig. 1 e *Tab. I*).

Dal confronto tra i dati in tabella, espressi come intervallo di valori rilevati nel parlato ed estensioni distinte per attività di produzione (general, di fonazione e/o canto) in Fig. 1, si osserva una discrepanza per quanto concerne l'estensione considerata, soprattutto – ancora una volta – riguardo il limite inferiori di voci femminili. Si noti inoltre il distinto riferimento a valori grezzi o arrotondati (alla decina, o al centinaio di Hz, nel caso del limite massimo).

2. Materiali e metodo

Riservandoci di sviluppare in altra occasione l'esemplificazione delle modalità di verifica delle misurazioni effettuate, ci limitiamo qui a precisare che l'estrazione dei valori è avvenuta con uno *script* per *PRAAT* (Boersma & Weenink 2005) realizzato presso il *LIMSI/Orsay* e messi a disposizione da Philippe Boula de Mareüil (che qui approfittiamo per ringraziare). Lo *script* sfrutta l'algoritmo di estrazione di f_0 del software applicandolo con passo di 10 ms al segnale di ogni file caricato e producendo i risultati in una lista (formattata con l'indicazione della

posizione temporale in cui è avvenuta la rilevazione) affissa in una finestra di testo. Da questa lista i valori sono poi esportati in un foglio elettronico come quello usato in Romano (2000) che li classifica con passo di 2 Hz da 70 a 600 Hz definendo una funzione distribuzione di probabilità (cumulativa) da cui, per sottrazione progressiva, di ottiene poi l'istogramma di densità di probabilità (*ddp*).

Come anticipato sopra, questo grafico, valutato individualmente per ciascun parlante, presenta una dispersione di valori indicativa dell'uso della voce nel corso dell'attività richiesta, ma può risentire di errori di misurazione soprattutto ai limiti dell'intervallo d'uso o in corrispondenza di passaggi da un meccanismo fonatorio all'altro (Henrich 2021)¹⁰.

Con le premesse del §1 risulta limitante, ai fini della rappresentatività dei risultati, il fatto che i dati siano stati raccolti in un'unica modalità di pro-

10. Non ci soffermiamo qui nella valutazione di quanto l'uso della voce in quest'intervallo corrisponda a un'esplorazione della fascia di maggiore «comodità» (che nel canto viene indicata come *tessitura*, v. Romano *et alii* 2011) o si estenda in regioni in cui il controllo dell'attività vocale divenga «malfermo» (come tipicamente si direbbe nel caso di voci «non educate»). I nostri parlanti sono infatti persone normali (non professionisti della dizione) e occasionalmente producono variazioni melodiche che «deborzano» in registri di fonazione non modali (per i quali rimandiamo a Gordon & Ladefoged 2001 e Gerratt & Kreiman 2001).

per i dati di Traunmüller & Eriksson (1995) che offrono una rassegna su lingue diverse ma includono valutazioni in diverse condizioni di attivazione di campioni di parlanti svedesi.

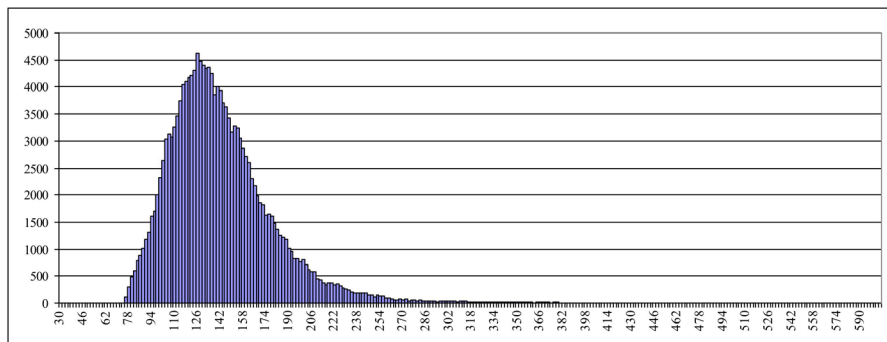


Fig. 2. Curva di ddp relativa al campione M01 costituito da 29 voci maschili (classificate come giovanili). No. valori rilevati: 165718; tempo di parlato: 27'52"; unica distribuzione apparente con ddp completa in un registro medio: f_0 media 141 Hz ($\sigma = 44$ Hz; $\mu \pm 2\sigma = 53$ -229 Hz); moda 126 Hz; pitch-span 76-318 Hz (ultimo valore con più di 16 occorrenze, pari allo 0.01%).

duzione (*task* di lettura, più o meno interpretata). Un'altra limitazione deriva dal fatto che i dati riguardano un campione di voci di una rete sociale che, per quanto raggiunta con metodi di passa-parola, si estende a parlanti su livelli medio-alti di istruzione¹¹ e su fasce d'età che vedono poco rappresentati i giovanissimi (al di sotto dell'età in cui si produce eventualmente la cosiddetta muta vocale)¹².

Il campione è costituito prevalentemente da dati vocali, e valutati frase

per frase (18900 enunciati ascoltati due volte, da entrambi gli autori) e testati a campione con *ddp* parziali in un centinaio di casi. I valori rilevati sono stati rappresentati su grafici cumulativi dell'intera produzione del parlante, senza escludere i valori ai limiti dell'estensione individuale. Una serie di sei *ddp* cumulative (v. Figg. 2-7) ottenute distinguendo due gruppi (voci maschili, §2.1, vs. femminili, §2.2) e sottogruppi per fasce d'età (giovanili, adulte, anziane) si caratterizza poi per diversi valori di f_0 media e moda e per un *pitch-span* delimitato da un minimo non necessariamente corrispondente al limite dell'intervallo (v. dopo) e da un massimo individuato per mezzo di una soglia percentuale di affidabilità¹³.

11. Tutti i parlanti erano in grado di usare lo *smartphone* per mezzo del quale hanno inviato i loro dati vocali (attraverso sistemi di raccolta e digitalizzazione del dato acustico variabili in base alle caratteristiche tecniche di questo). La qualità delle registrazioni conferite si è mostrata in tutti casi più che soddisfacente per un'estrazione affidabile dei valori di f_0 (v. Romano *et alii*, in questo volume).

12. V. diversi contributi in Schindler (2009). Cfr. Vernero & Guiot in Romano *et alii* (2012).

13. La dispersione dei valori monitorata sul grafico *ddp*, dà garanzie sul fatto che una diminuzione progressiva dei valori misurati conduca a un certo punto sotto una soglia che li

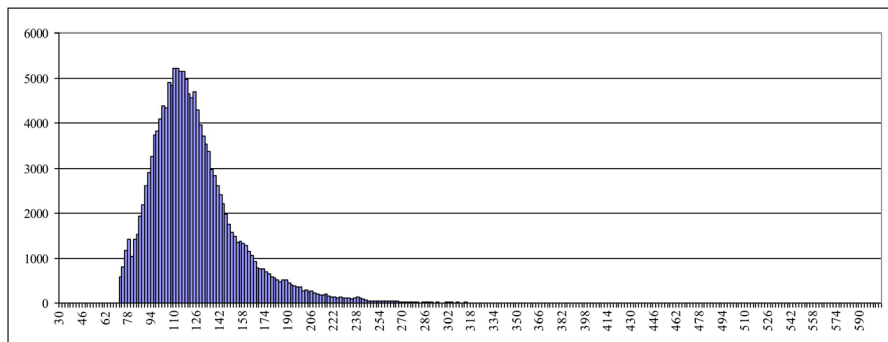


Fig. 3. Curva di *ddp* relativa al campione M02 di 29 voci maschili (class. come adulte). No. valori rilevati: 152090; tempo di parlato: 25'35"; unica distribuzione apparente (limitata a 70 Hz) nel registro medio: f_0 media 126 Hz ($\sigma = 44$ Hz; $\mu \pm 2\sigma = 37$ -215 Hz); moda 110 Hz; pitch-span (66)-308 Hz (ultimo valore con più di 15 occorrenze, pari allo 0.01%).

Ci limitiamo qui alla discussione dei dati rilevati per questi distinti gruppi nella certezza che nessun altro studio finora pubblicato in Italia abbia mai conseguito la descrizione dei principali indicatori statistici per un numero così elevato di voci. Nel complesso sono stati infatti valutati i dati di 200 parlanti e, per quanto la loro numerosità sia variabile da gruppo a gruppo, nel caso di voci femminili giovanili (78 parlanti) le considerazioni riportate si riferiscono a più di 400 mila campioni misurati¹⁴.

rende rari e, potenzialmente, inaffidabili a rappresentare un limite (soprattutto nel caso di quello superiore).

14. Sin dal 1999, nell'esperienza dell'autore AR, sono state diverse le occasioni d'incontro, nel corso di convegni o visite in centri di ricerca (di enti pubblici o aziende private) anche in settori disciplinari diversi, in cui singoli ricercatori hanno riferito la disponibilità di ingenti quantitativi di dati utili per simili finalità descrittive. A tutt'oggi però, a nostra conoscenza, nessuno studio ave-

2.1. Le voci maschili di CALL-UniTO

I tre distinti gruppi analizzati (M01 di voci classificate come giovanili, M02 voci adulte e M03 voci anziane, qui considerate per simmetria, senza ambizioni di rappresentatività) hanno presentato le *ddp* raffigurate nelle Figg. 2-4.

Il campione la cui *ddp* è riportata in Fig. 2 è quello maschile più popoloso e si caratterizza per una f_0 media di 141 Hz ($\sigma = 44$ Hz; $\mu \pm 2\sigma = 53$ -229 Hz) e un'estensione vocale inclusiva di massimi più attendibili nell'intervallo 76-318 Hz. La concentrazione più rilevante si arresta a circa 270 Hz e, se la scala delle ascisse fosse logaritmica, apparirebbe grosso modo gaussiana. Resta tuttavia una *coda* di dispersione nelle alte frequenze decisamente molto prolungata e si riferisce a valori

va fatto emergere pubblicamente dati di questa portata (cfr. anche Cataldo *et alii* 2017, il cui studio deve rifarsi a dati relativi a lingue straniere).

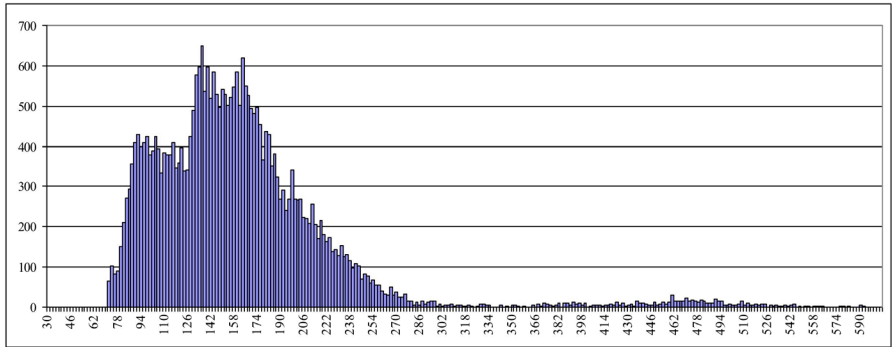


Fig. 4. Curva di ddp relativa al campione M03 di sole 4 voci maschili (classificate come anziane). No. valori rilevati: 31826; tempo di parlato: 5'30"; appaiono distinte distribuzioni sovrapposte (limitate a 70 Hz) nel registro medio: f_0 media 161 Hz ($\sigma = 65$ Hz; $\mu \pm 2\sigma = 30$ -291 Hz); moda 136 Hz; pitch-span (64)-296 Hz (ultimo valore verificato manualmente).

realmente presenti (verificati manualmente a campione).

La *ddp* delle voci maschili adulte (Fig. 3), basata su un campione ugualmente numeroso, presenta caratteristiche simili ma si rivela molto più concentrata fino a circa 250 Hz. La f_0 media è più bassa, 126 Hz ($\sigma = 44$ Hz; $\mu \pm 2\sigma = 37$ -215 Hz) con *pitch-span* più incerto nel limite inferiore, (66)-308 Hz. Come mostra l'abbassamento che interessa anche la moda (126 > 110 Hz), le voci maschili adulte si presentano quindi più basse di un tono ($Si_1 \rightarrow La_1$). Sebbene siano meno numerosi e non diano sufficienti garanzie statistiche per trarre conclusioni, più di 30000 valori rilevati per le voci anziane confermano invece che l'invecchiamento della voce si rifletta in un generale allargamento dell'estensione verso le alte frequenze, con un innalzamento significativo della media (161 Hz, in un campione di soli 4 parlanti)¹⁵.

2.2. Le voci femminili di CALL-UniTO

Le voci femminili analizzate sono nettamente più numerose e i dati anche più interessanti per illustrare le proprietà osservate nelle curve di *ddp* ottenute e l'attendibilità di misure che tengano conto genericamente di minimi e massimi.

Già all'osservazione del grafico di Fig. 5 emergono alcuni dati interessanti: la dispersione di valori del registro modale pare infatti limitata in alto da una coda che si stempera gradualmente fino a circa 380 Hz. Alle basse frequenze invece, sebbene siano presenti migliaia di valori misurati tra i 70 e i 120 Hz, la *ddp* della voce modale sem-

15. Le diverse centinaia di valori disperse nell'intervallo 350-550 Hz non sembrano relative a misure aberranti (*outliers*), ma descrivono occasionali perdite di controllo della voce che si innalza temporaneamente in queste fasce di frequenza (cfr. Ramig 1986 e Caruso *et alii* 1995).

bra esaurire la sua coda inferiore intorno a 126 Hz. Questo spiega il limite inferiore delle voci femminili indicato da diversi studi (v. §1).

Queste voci, che potremmo descrivere quindi con f_0 media a 209 Hz e *pitch-span* di 126-384 Hz, raggiungono tuttavia con una certa sistematicità anche frequenze più basse con una distribuzione che sembra avere un suo massimo intorno ai 100 Hz a indicare la frequente esplorazione di un altro meccanismo fonatorio di tipo *M0* di voce laringalizzata / cricchiata (Henrich 2021). Troviamo infatti ormai descritto in diversi studi un ricorrente modello d'uso della voce che, soprattutto nelle aree urbane del mondo occidentale (soprattutto negli Stati Uniti d'America), spinge le giovani donne a sfoggiare spesso voci di tipo *creaky* (Henton & Bladon 1985, 1988, Yuasa 2010)¹⁶. I dati mostrano come a

16. Le preferenze di un campione di valutatori canadesi sono state osservate da O'Connor & Barclay (2017) nei confronti di queste voci (in campo economico, affettivo o sociale in genere), in esperimenti di ascolto dai quali è risultata un'indifferenza da parte dei soggetti femminili in alcuni dei campi testati e una sfiducia generale nei confronti delle voci più gravi. Da diversi anni infatti è ormai opinione comune che, come dimostra anche il caso dell'imprenditrice statunitense Elizabeth Holmes (tornato in auge proprio mentre andiamo in stampa), che – per sembrare più convincenti – molte giovani americane abbassano volontariamente la loro voce fino a sconfinare nelle regioni frequenziali in cui si innesca un meccanismo fonatorio qui indicato come *M0*.

questa tendenza non sfuggano alcune delle voci italiane analizzate. Nonostante una maggiore irregolarità e una minore salienza nella curva, valori in questa fascia di frequenze sono però presenti anche negli altri due gruppi di parlanti (Figg. 6 e 7). Per le voci femminili classificate come adulte si ha infatti una *ddp* molto regolare con media a 202 Hz (discosta in maniera non significativa da quella del gruppo precedente) e *pitch-span* 94-380, quindi essenzialmente caratterizzata proprio da un abbassamento del limite inferiore che oscura il discrimine tra le due potenziali distribuzioni relative ai due meccanismi *M0* e *M1*.

Sono infine le voci anziane, seppur meno numerose, quelle che presentano un abbassamento sensibile della media (194 Hz) e dell'intero intervallo di estensione della *ddp* con limite inferiore a 92 e superiore a 360 Hz¹⁷. Per queste voci, alla presenza di flessi e gobbe nella *ddp*, con massimi localizzati in diverse posizioni, si associa comunque una coda molto regolare alle alte frequenze che, contrariamente a quanto osservato per le voci maschili (§2.1), corrisponde bene a quella rilevata per gli altri due gruppi, mostrando un'area con una certa permanenza di caratteristiche al variare dell'età.

17. Se l'algoritmo di misurazione non fosse stato limitato a 78 Hz (per ragioni di stabilità nelle stime di f_0) sarebbero probabilmente apparsi con una certa regolarità anche valori più bassi.

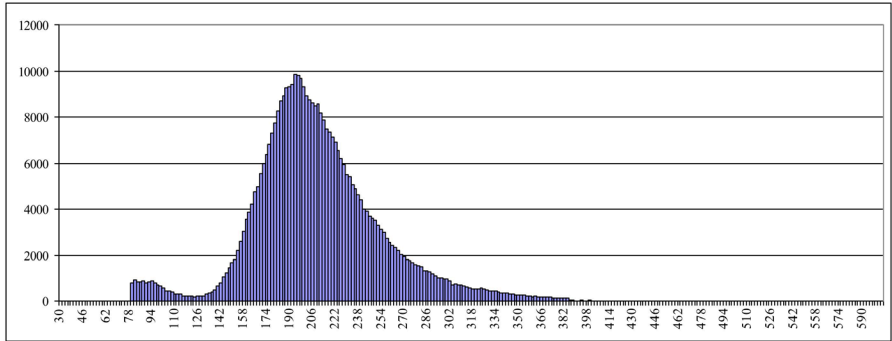


Fig. 5. Curva di ddp relativa al campione F01 di 78 voci femminili (classificate come giovanili). No. valori rilevati: 404500; tempo di parlato: 67'48". La curva appare come il risultato della sovrapposizione di due distribuzioni: una distribuzione di valori emerge irregolarmente alle basse frequenze e si aggiunge a una ddp perfettamente delimitata nel registro medio con f_0 media 209 $H\zeta$ ($\sigma = 49 H\zeta$; $\mu \pm 2\sigma = 111-308 H\zeta$); moda 194 $H\zeta$ e pitch-span 126-384 $H\zeta$.

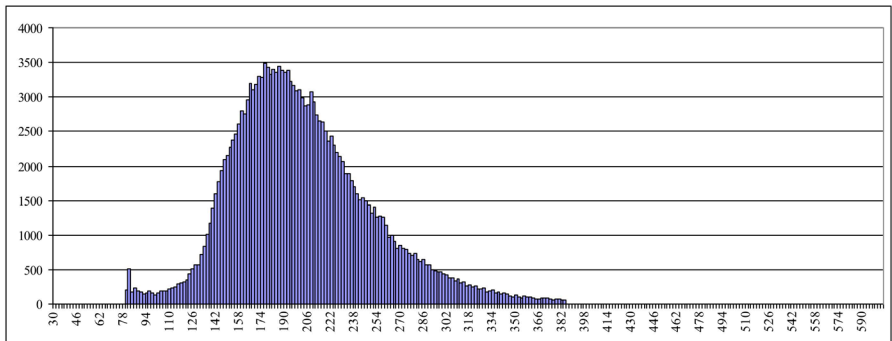


Fig. 6. Curva di ddp relativa al campione F02 di 35 voci femminili (classificate come adulte). No. valori rilevati: 182689; tempo di parlato: 30'45". In questo caso si osserva una distribuzione completa con valori più irregolari in corrispondenza della soglia inferiore di rilevazione. La ddp si presenta con f_0 media 202 $H\zeta$ ($\sigma = 52 H\zeta$; $\mu \pm 2\sigma = 98-307 H\zeta$); moda 176 $H\zeta$ e pitch-span 94-380 $H\zeta$.

Conclusioni

In questo studio preliminare abbiamo offerto una stima sommaria dei valori di *mean pitch* e *pitch span* rilevati in tre gruppi di parlanti italiani distinti per genere. Escludendo i risultati ottenuti nelle produzioni dei parlanti anziani, ancora relativamen-

te poco rappresentati nel corpus di dati, il gruppo cumulativo di 58 voci maschili analizzate presenta una f_0 media (μ) attestata su valori intorno ai 130 Hz (cfr. Romano 2000) e con estensione indicativa ($\mu \pm 2\sigma$) tra i 45 e i 220 Hz, sebbene in molti casi i parlanti più giovani raggiungano

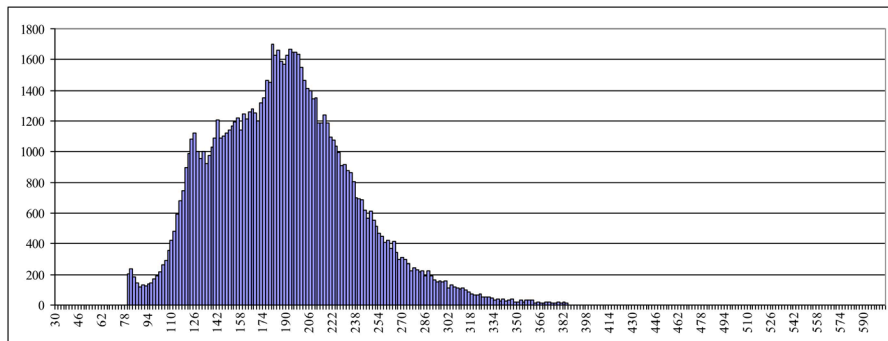


Fig. 7. Curva di ddp relativa al campione F03 di 14 voci femminili (classificate come anziane). No. valori rilevati: 92518; tempo di parlato: 15'45". Sembra trattarsi anche in questo caso di più distribuzioni, con una concentrazione di valori (non completamente visualizzata) nel registro di cricchiato. La curva complessiva si caratterizza per f_0 media 194 Hz ($\sigma = 42$ Hz; $\mu \pm 2\sigma = 109-278$ Hz); moda 180 Hz e pitch-span 92-360 Hz.

spesso, con una coda statistica non trascurabile, valori fino ai 320 Hz.

Lo studio della ddp per le voci femminili ha evidenziato la presenza di diverse distribuzioni statistiche che includono un frequente ricorso a voci nel registro di voce cricchiata. Nel registro di voce modale il gruppo cumulativo delle voci analizzate (113) ha presentato una f_0 media a circa 205 Hz, con *pitch-span* indicativo che si estende in modo più uniforme nell'intervallo 110-380 Hz.

Al di là delle differenze riscontrate tra i distinti gruppi (che necessitano ancora di una più accurata analisi statistica), l'articolo propone quindi alcuni dati rappresentativi del parlato italiano per fornire un primo contributo di conoscenze in questo campo da affiancare ai descrittori valutati nella letteratura internazionale sul tema.

Riferimenti bibliografici

- Baken R.J. & Orlikoff R.F. (2000). *Clinical measurement of speech and voice*. San Diego: Singular Thomson Learning (2nd ed., 1st ed. 1993).
- Boë L.J., Contini M. & Rakotofringa H. (1975). «Étude statistique de la fréquence laryngienne. Application à l'analyse et à la synthèse des faits prosodiques du français». *Phonetica*, 32, 1-23.
- Boersma P. & Weenink D. (2005). *Praat: Doing Phonetics by Computer* [www.praat.org].
- Caruso A., Mueller P. & Shadden B.B. (1995). «Effects of aging on speech and voice». *Physical and Occupational Therapy in Geriatrics*, 13, 63-80.
- Cataldo V., Orrico R. & Savy R. (2017). «Phonetic variations of f_0 range in L1 and L2: a comparison between Italian, English, and Spanish native and nonnative speakers». In C. Bertini *et alia* (a cura di), *Fattori Sociali*

e *Biologici nella Variazione Fonetica / Social and Biological Factors in Speech Variation*, Milano: Officinaventuno, 227-256.

Garde É. (1965). *La voix*. Paris: PUF (*Que sais-je ?* 3^a ed. 1965).

Gerratt B. & Kreiman J. (2001). «Toward a taxonomy of nonmodal phonation», *Journal of Phonetics*, 29, 365-381.

Gordon M. & Ladefoged P. (2001). «Phonation types: a cross-linguistic overview». *Journal of Phonetics*, 29, 383-406.

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.

Gutzmann H. (1909). *Physiologie der Stimme und Sprache*. Braunschweig: Friedrich Vieweg und Sohn.

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]

Henton C.G. & Bladon R.A.W. (1985). «Breathiness in normal female speech: inefficiency versus desirability». *Language and Communication*, 5, 221-227.

Henton C.G. & Bladon R.A.W. (1988). «Creak as a sociophonetic marker». In: L.M. Hytam & C.N. Li (eds.). *Language, speech, and mind: studies in honour of Victoria A. Fromkin*, London: Routledge, 3-29.

Hess W. (1983). *Pitch Determination of Speech Signals*. Berlin: Springer, 1983.

Hollien H., Dew D. & Philips P. (1971). «Phonational frequency ranges of adults». *Journal of Speech and Hearing Research*, 14(4), 755-760.

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 August 2007), 1809-1812 [www.icphs2007.de ID 1570].

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

O'Connor J.M. & Barclay P. (2017). «The influence of voice pitch on perceptions of trustworthiness across social contexts». *Evolution and Human Behavior*, 38, 506-512.

Patterson D. (2000). «A Linguistic Approach to Pitch Range Modeling». *PhD Thesis*, University of Edinburgh [era.ed.ac.uk/bitstream/handle/1842/6746/492996.pdf].

Ramig L.A. (1986). «Aging speech: Physiological and sociological aspects». *Language and Communication*, 6, 25-34.

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., Cesari U., Mignano M., Schindler O. & Vernero 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.).

Romano *et alii*, in this volume.

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

Shriberg E., Ladd D.R., Tenken J. & Stolcke A. (1996). «Modelling pitch range variation within and across speakers: predicting f_0 targets when “speaking up”». *Proc. Addendum of the 4th International Conference on Spoken Language Processing* (ICSLP, Philadelphia, 3-6 October 1996), 1-4.

Sorianello P. (2006). *Prosodia: modelli e ricerca empirica*. Roma: Carocci.

Traunmüller H. & Eriksson A. (1995). «The perceptual evaluation

of F_0 -excursions in speech as evidenced in liveliness estimations». *Journal of the Acoustical Society of America*, 97, 1905-1915 (v. Idd., «The frequency range of the voice fundamental in the speech of male and female adults» [www.researchgate.net/publication/240312210]).

Uberti M. (2005). «Acustica della voce». In S. Cingolani, R. Spagnolo (a cura di), *Acustica musicale e architettonica*, Torino: UTET, 495-526.

Vogel A.P., Maruff P., Snyder P.J. & Mundt J.C. (2009). «Standardization of pitch range settings in voice acoustic analysis». *Behavioral Research Methods*, 41(2), 318-324 [www.ncbi.nlm.nih.gov/pmc/articles/PMC2669687/].

Yuasa I.P. (2010). «Creaky Voice: A New Feminine Voice Quality for Young Urban-oriented Upwardly Mobile American Women?». *American Speech*, 85, 315-337.