

SUI PIANI FORMANTICI ACUSTICI E Uditivi DELLE VOCALI DI UOMO, DONNA E BAMBINO

F.E. Ferrero, E. Magno Caldognetto, P. Così

Centro di Studio per le Ricerche di Fonetica (CNR), Padova

INTRODUZIONE

Le due principali caratteristiche acustiche di derivazione fisiologica delle vocali che, come è noto, sono connesse al *nesso/età* (ossia al *genere*) del parlante, sono la *frequenza fondamentale* (F0), determinata primariamente dalla grandezza della laringe, e le risonanze del condotto vocale, cioè le *frequenze formantiche* (Fn, con n=1-3), determinate essenzialmente dalla lunghezza del condotto vocale stesso.

Questi sono i parametri che tradizionalmente vengono oggi utilizzati per la classificazione acustica dei sistemi vocalici delle varie lingue e che con trasformazioni di scala dette "uditivè", aventi lo scopo di rimuovere la dipendenza dal genere di voce, sono alla base delle strategie di *normalizzazione* degli spazi di distribuzione delle vocali.

I tentativi di normalizzazione sono stati dedicati principalmente alla dicotomia uomo/donna, a cui solo in rari casi si sono aggiunte le voci infantili per il semplice fatto che nel passato sono stati raccolti pochi dati base per questa categoria di voce.

La storica esclusione fino ai tempi più recenti (per es.: Hillenbrand et al., 1995) dagli studi di Fonetica acustica dei bambini, oltre che delle donne, pur con le note eccezioni (per es.: Peterson e Barney, 1952), come si sa, era dovuta essenzialmente a ragioni tecniche, che peraltro condizionarono profondamente l'impostazione originale della teoria acustica di produzione della parola, basata di fatto sulle caratteristiche adulte maschili di disaccoppiamento tra sorgente e filtro (Fant, 1960).

Per molto tempo si ritenne che i dati acustici per le donne e i bambini potessero essere agevolmente estrapolati dai dati raccolti per gli uomini applicando a questi degli opportuni fattori di scala (cfr., per es.: Fant, 1975). Ma, anche se si fossero potuti determinare agevolmente dei fattori ben definiti e normalizzati, sussisterebbe il fatto che, oltre alle differenti dimensioni lineari e volumetriche dell'apparato fonoarticolatorio, donne e bambini presentano ulteriori peculiari caratteristiche di produzione che non possono essere trascurate sia nella teoria che nella pratica analitica (Kent e Read, 1992). È infatti di esperienza quotidiana il fatto che si è in grado di individuare il sesso di una voce adulta anche in presenza di una fonazione *bisbigliata* o

sussurrata, quando cioè le diversità intrinseche di altezza vocale sono “bypassate”; come anche generalmente si distingue agevolmente il sesso di una voce infantile in età prepubere, quando cioè le differenze di grandezza degli organi fonoarticolatori sono trascurabili.

Oggi è abbastanza diffusa la consapevolezza che nel caso di donne e bambini l'ipotesi di disaccoppiamento sorgente-filtro è considerata probabilmente non molto appropriata: spesso si è osservata, seppure informalmente, un'interazione sorgente-filtro abbastanza consistente ed anche un accoppiamento tracheale non trascurabile (Fant, 1985; Klatt e Klatt, 1990).

La caratteristica peculiare delle voci femminili adulte e di quelle infantili prepuberi (Kent e Read, 1992), responsabile dell'accoppiamento biomeccanico ed acustico tra sorgente e condotto vocale da un lato, e tra sorgente e tubo tracheale dall'altro lato, può essere individuata nella qualità “*soffiata*” (“breathy”) di queste voci. Tale tipo di fonazione implica una persistente diastasi delle aritenoidi (circa il terzo posteriore dei processi vocali) in presenza di un regolare movimento vibratorio dei circa due terzi anteriori delle corde vocali (la parte, cosiddetta, “membranosa” delle c.v.).

Dal momento che con tale assetto la rima glottica posteriormente non è mai completamente chiusa, durante ciascun ciclo di vibrazione vi è una perdita d'aria continuativa. La conseguente coesistenza di un treno di impulsi periodici e di un flusso d'aria turbolento continuo, determina la generazione di un segnale sorgente di eccitazione del condotto vocale in cui alla struttura armonica si sovrappongono componenti aperiodiche di rumore. Inoltre, contrariamente a quanto avviene nella fonazione “modale” maschile adulta, le corde vocali, anziché affrontarsi simultaneamente sulla loro lunghezza (il che implica una brusca cessazione del flusso d'aria con conseguente asimmetria del segnale glottico e quindi forte eccitazione delle armoniche di alta frequenza nell'istante di chiusura), dapprima si chiudono anteriormente e successivamente la chiusura si propaga posteriormente con un conseguente “arrotondamento” della forma d'onda di portata d'aria nell'istante di chiusura e aumento del quoziente di apertura delle c.v. (Klatt e Klatt, 1990). In tal modo il segnale glottico assume una forma più di tipo sinusoidale in cui la componente armonica fondamentale risulta relativamente più intensa e con le armoniche di alta frequenza significativamente attenuate in ampiezza e parzialmente e progressivamente sovrastate dalle componenti aperiodiche del rumore turbolento.

A causa di queste peculiarità di fonazione la misurazione delle frequenze formantiche di donne e bambini può risultare alquanto difficoltosa; infatti, specie per fondamentali elevate, la prima armonica relativamente molto intensa può essere confusa con una formante. Inoltre l'accoppiamento con la trachea può, da un lato, causare un aumento della larghezza di banda di F1 e quindi ridurre la “visibilità” e, dall'altro lato, fare comparire extra-formanti nello spettro (spesso a circa 600, 1400 e/o 2200 Hz) facilmente confondibili con le formanti effettive.

L'accoppiamento tracheale, introducendo coppie extra di zeri/poli, pone anche problemi circa il grado di attendibilità delle misure effettuate con un algoritmo LPC, in quanto esso si basa, come è ben noto, su un modello a “tutti poli”.

Per mantenere l'omogeneità con i dati raccolti in precedenza su database di adulti maschi e femmine, questo studio utilizza la tecnica di misura LPC anche per il data base di voci infantili, essendo però stata posta particolare cura nella individuazione dei valori formantici (la stessa peraltro già posta nell'analisi del data base delle donne) onde non incorrere nei possibili errori di attribuzione testé menzionati.

I dati formantici ricavati e rappresentati in piani cartesiani con diversi tipi di coordinate cartesiane, sono stati utilizzati per valutare il grado di *normalizzazione* possibile per sesso/età delle vocali di un Italiano regionale.

MATERIALI E METODI

Sono state raccolte le produzioni “spontanee” delle vocali (pronunciate prolungate per pochi secondi) di alunni delle ultime tre classi di una Scuola Elementare di Padova. Le registrazioni sono avvenute nella sala di lettura della Scuola, utilizzando un registratore DAT portatile, modello *Sony TCD-D7*, e un microfono dinamico *Shure PROLOG* posizionato lateralmente a circa 3 cm dalle labbra.

La “spontaneità” delle produzioni è consistita nell’invitare ciascun soggetto a pronunciare le 7 vocali Italiane nella successione alfabetica e cercando di realizzare per “e” ed “o” la distinzione aperto/chiuso, secondo l’esemplificazione fornita dall’operatore che suggeriva contemporaneamente anche un modello di “prolungamento”, ma cercando comunque di non indurre alcun tipo di “forzatura”.

Tale “spontaneità” ha avuto come conseguenza innanzitutto che non tutti i bambini/e hanno prodotto emissioni accettabili per livello o costanza dell’intensità e/o per stabilità dell’altezza, e pertanto non sono stati presi in considerazione nell’allestimento del data base. In secondo luogo, tra i “non esclusi”, in generale un notevole numero di soggetti ha realizzato le vocali non tutte in modo “accettabile” (secondo i criteri di cui sopra) e, in particolare, la stragrande maggioranza dei soggetti non è riuscita ad evidenziare “spontaneamente” la dicotomia aperto/chiuso, di essa realizzando quasi esclusivamente /e/ e, rispettivamente, /ɔ/.

Alla fine il data base di vocali infantili è risultato costituito da: 63 /i/ (38 bambine + 25 bambini), 59 /e/ (34 bambine + 25 bambini), 12 /ɛ/ (solo bambine), 74 /a/ (48 bambine + 26 bambini), 42 /ɔ/ (42 bambine + 25 bambini), 6 /o/ (solo bambine), e 56 /u/ (33 bambine + 23 bambini).

Il materiale verbale di questo database è stato acquisito col sistema *KayElemetrics CSL-4300 vers. 5.01* con conversione A/D a 10240 Hz e 16 bit. Il procedimento di analisi si è svolto in due fasi, di cui la prima è consistita nella selezione manuale di un segmento stazionario di circa 100 ms da ciascuna vocale pronunciata isolatamente e prolungata.

Per effettuare questa selezione, insieme alla forma d’onda del segnale, si sono visualizzati: lo stesso segnale amplificato e segmentato in periodi fondamentali, la curva di frequenza fondamentale, la curva di intensità, lo spettrogramma a banda larga e i tracciati formantici. La “stazionarietà” è stata valutata come sussistenza di una relativa costanza o uniformità di andamento contemporaneamente di tutti questi parametri.

Il segmento prescelto con tale procedura è stato trasferito alla seconda fase di misura automatica con la quale sono stati calcolati, con un algoritmo LPC-12 “pitch-asincrono” con una finestra temporale di larghezza costante di 20 ms, il valore medio, la deviazione standard e altri valori statistici descrittivi della frequenza fondamentale (F0) e della frequenza delle prime 4 formanti (F1, F2, F3, F4).

In modo analogo erano state analizzate le vocali sostenute di informatori padovani adulti, 10 maschi e 10 femmine (per il dettaglio si veda Ferrero et al., 1995).

Questi dati numerici sono stati elaborati statisticamente con il programma *SYSTAT 5.0 for Windows*.

RISULTATI

La prima verifica effettuata è stata, evidentemente, se vi sia differenza fra le voci infantili di bambini e bambine nei valori di F0 e delle prime 4 formanti, con un T-test che confronta, per ogni vocale, i valori medi dei due gruppi fornendo la probabilità che essi appartengano alla stessa “popolazione”. E’ risultato che per le voci infantili dei due sessi del nostro campione i valori medi di questi 5 parametri acustici sono generalmente, per ogni vocale, significativamente simili, con una probabilità frequentemente molto superiore a circa lo 0,005 (il dettaglio di questo confronto, nonché i valori dei singoli campioni e le loro statistiche, sono riportati in Pacetto e Ferrero, 1994-95).

Pertanto i dati di bambine e bambini si sono considerati congiuntamente come gruppo unificato di voci infantili (**bambini**) da confrontare con le altre due categorie di età/sexo di voce adulta femminile (**donne**) e maschile (**uomini**).

Le coordinate cartesiane dei piani formantici usualmente utilizzate per disegnare le aree di esistenza delle vocali sono o di tipo “acustico”, cioè F1 e F2 in Hz, o di tipo *uditivo*”, cioè (F1-F0) e (F2-F1), oppure (F3-F2), misurate in Bark (la formula di trasformazione Hz/Bark è quella già usata in Ferrero et al., 1995).

Avendo avuto preventivamente modo di osservare (Pacetto e Ferrero, 1994-95) che agli effetti di una *normalizzazione inter-categoria* la coordinata contenente F3 “funziona” meglio di quella contenente F2 per un certo gruppo di vocali e viceversa per un altro gruppo di vocali, si è ritenuto di considerare anche la *seconda formante “effettiva”*, F2’, misurata anch’essa in Bark (e F2’-F1), nella formulazione di Carlson et al. (1975) secondo i quali trattasi di un «*approccio intuitivo avente l’intento di posizionare F2’ circa tra F2 e (F3*F4)^{1/2}. Il limite inferiore F2’=F2 si ha quando F2 è vicina a F1 come nelle vocali posteriori [dunque, per queste vocali F2’ “funziona” come F2]. Il limite superiore, F2’ = (F3*F4)^{1/2} si ha quando (F2-F1) è grande e F3 è molto più vicina a F4 che a F2, come nella vocale /i/. D’altro canto, quando (F3-F2) è molto piccola, F2’ si colloca giusto sopra F2. Dislocazioni intermedie sono realizzate da una appropriata pesatura*».

Dunque è lecito ritenere che F2’ debba conglobare al meglio il peso individuale di F2 e F3 nelle altre varie combinazioni di “differenza lineare” usualmente considerate.

Pertanto come coppie di coordinate cartesiane dei piani formantici consideriamo, oltre a quelle acustiche “grezze” F2|F1 in Hz, le seguenti, tutte misurate in Bark:

$$(F2-F1)|(F1-F0), (F3-F2)|(F1-F0), (F2’-F1)|(F1-F0) \text{ e } F2’|(F1-F0).$$

Il nostro scopo primario essendo quello di verificare quale di questi piani formantici **normalizza** meglio le tre categorie di voce, sulle varie coordinate formantiche di esse e per ogni vocale, si dovrebbe effettuare un’analisi della varianza (ANOVA). Però questa tecnica di analisi statistica è opportuna e pienamente significativa solo nel caso in cui i gradi di libertà, ossia la numerosità, dei vari campioni da confrontare siano simili (il che non è nel nostro caso, data la notevole disomogeneità del numero di casi di ciascuna vocale per categoria) e presentino una dispersione a distribuzione normale (il che nemmeno vale la pena di verificare per i due campioni di voci adulte e per /e/ e /o/ delle voci infantili, dove i casi sono non più di una decina).

Si è quindi effettuata una semplice analisi di “uguaglianza” o meno dei valori medi dei 7 tipi di coordinata per ogni vocale, confrontando a due a due le tre categorie di voce, con un T-test “a varianza diversa”. I risultati sono presentati in **TAB. 1**.

<i>Confr.i</i>	<i>Coord.e</i>	i	e	ε	a	ɔ	o	u
u / d	<u>F1 (Hz)</u>	.088	.000	.000	.000	.000	.009	.139
u / d	<u>F1-F0 (Bark)</u>	.001	.012	.018	.001	.290	.609	.000
d / b	<u>F1 (Hz)</u>	.000	.000	.001	.000	.000	.776	.000
d / b	<u>F1-F0 (Bark)</u>	.036	.000	.018	.001	.000	.186	.004
u / b	<u>F1 (Hz)</u>	.000	.000	.000	.000	.000	.004	.000
u / b	<u>F1-F0 (Bark)</u>	.020	.539	.000	.000	.000	.070	.102
u / d	<u>F2 (Hz)</u>	.000	.000	.000	.000	.022	.012	.168
u / d	<u>F2-F1 (Bark)</u>	.000	.000	.006	.896	.883	.240	.509
u / d	<u>F3-F2 (Bark)</u>	.997	.709	.027	.377	.006	.006	.002
u / d	<u>F2' (Bark)</u>	.000	.000	.000	.478	.451	.829	.656
u / d	<u>F2'-F1 (Bark)</u>	.000	.000	.008	.155	.639	.474	.883
d / b	<u>F2 (Hz)</u>	.000	.001	.006	.000	.000	.000	.000
d / b	<u>F2-F1 (Bark)</u>	.001	.105	.973	.075	.002	.001	.351
d / b	<u>F3-F2 (Bark)</u>	.107	.001	.000	.000	.002	.016	.000
d / b	<u>F2' (Bark)</u>	.002	.000	.000	.000	.000	.003	.000
d / b	<u>F2'-F1 (Bark)</u>	.007	.000	.000	.000	.000	.005	.001
u / b	<u>F2 (Hz)</u>	.000	.000	.000	.000	.000	.000	.000
u / b	<u>F2-F1 (Bark)</u>	.000	.000	.024	.055	.000	.000	.066
u / b	<u>F3-F2 (Bark)</u>	.179	.000	.000	.000	.000	.000	.000
u / b	<u>F2' (Bark)</u>	.000	.000	.000	.000	.000	.006	.000
u / b	<u>F2'-F1 (Bark)</u>	.000	.000	.000	.000	.000	.021	.001

TAB. 1 Confronto a due a due dei campioni di **uomini**, **donne** e **bambini** con un T-test “a varianza diversa”. Per ciascuna vocale, di ogni tipo di coordinata formantica è indicata la probabilità che i due campioni appartengano alla stessa popolazione. Si assume una significatività per $p \geq 0,001$.

In essa si può vedere come la coordinata (F1-F0) in Bark normalizzi efficacemente tutte le vocali, eccetto /u/, di adulti maschi e femmine (confermando i risultati di Ferrero et al., 1995). Viceversa nel confronto donne/bambini la normalizzazione esclude le vocali /e/ e /ɔ/, e nel confronto uomini/bambini non vengono normalizzate invece, ancora /ɔ/, e in più /ε/ e /a/. Quindi su questa coordinata la normalizzazione “età/sex” ha successo sempre per /i/ e /o/, rimane sufficientemente efficace per /e/, /ε/, /a/ e /u/, ma è poco robusta nel caso di /ɔ/.

Osservando la coordinata (F2-F1) in Bark, si rileva la sua generale efficacia nel normalizzare donne/bambini, mentre la normalizzazione uomini/donne fallisce per /i/ e /e/, cui nel confronto uomini/bambini si aggiungono nel fallimento /ɔ/ e /o/. Questa è la seconda coordinata usualmente utilizzata per definire le aree di esistenza delle vocali.

In alternativa ad essa è talvolta utilizzata la coordinata (F3-F2) in Bark che, come si conferma anche con questi dati, sembra più robusta nella normalizzazione uomini/donne (Ferrero et al., 1995). Nei confronti uomini/bambini, invece, essa si dimostra notevolmente meno efficace in assoluto, e anche rispetto alla precedente coordinata, in quanto riesce a normalizzare solamente la vocale /i/. Tale inefficacia, generale o

relativa, è confermata anche per il confronto donne/bambini, ma in modo meno drastico in quanto, come si può vedere, la normalizzazione fallisce solo sulle vocali /ε/, /a/ e /u/.

Considerando ora le due ulteriori coordinate basate sulla *seconda formante effettiva* (che si è ritenuto di “riesumere”), si osserva che, sia F2’ che (F2’-F1) in Bark, non mostrano sostanziali miglioramenti rispetto ai due tipi di ascisse “differenza” (cioè, in cui si presentano differenze tra formanti) testé esaminati. Nel confronto uomini/donne entrambe “funzionano” globalmente circa allo stesso modo di (F2-F1) e quindi peggio di (F3-F2). Nel confronto donne/bambini esse “funzionano” globalmente meno bene di (F3-F2) e quindi molto peggio di (F2-F1). Infine nel confronto uomini/bambini esse si comportano globalmente circa come gli altri due tipi di coordinate e quindi sono largamente inefficaci come “normalizzatore”.

Questi risultati, talvolta controversi, dimostrano che, quantomeno con i campioni di voce utilizzati, nessuna coordinata riesce a normalizzare in modo efficace congiuntamente tutte le 3 categorie di voce su tutte le 7 vocali, anche se alcune parzialmente “funzionano” meglio di altre, quantomeno in raffronti fra due delle tre categorie.

DISCUSSIONE

Le aree di esistenza delle vocali sui vari tipi di piano formantico tradizionalmente sono definite con le ellissi di dispersione equiprobabili, per es. al 75%, dei singoli valori misurati. Ma per gli stessi motivi di disomogenea “numerosità” dei vari campioni di vocale per categoria di voce che hanno sconsigliato di effettuare una ANOVA, non sembra pertinente valutare la dispersione dei singoli valori misurati di ciascun campione, mentre invece appare sensata la discussione della dispersione teorica del valore medio di ciascun campione, che graficamente è rappresentata dal cosiddetto **centroide** di equiprobabilità, per es. al 95%.

Il centroide è formalmente e concettualmente analogo ad una ellisse di dispersione in quanto esso è calcolato sostituendo alla “deviazione standard” della popolazione lo “errore standard della media”, cioè la deviazione standard teorica del valore medio del campione (calcolato come rapporto fra la deviazione standard della popolazione e la radice quadrata della “numerosità” del campione). Dunque il centroide è un’ellisse il cui centro è il valore medio del campione e i cui assi sono dati da ± 1 errore standard (e l’orientamento è determinato dal tasso di covarianza, o di correlazione, delle due coordinate): il valore medio di qualsiasi altro campione estratto dalla stessa popolazione ha il 95% di probabilità di cadere entro il centroide.

Dato che l’errore standard è inversamente proporzionale alla radice quadrata della “numerosità” del campione, esso sarà tanto più piccolo quanti più sono i singoli valori misurati che costituiscono il campione; di conseguenza (a parità di deviazione standard di campioni diversi) tanto più piccola sarà l’area del centroide, che comunque, per la stessa definizione di errore standard, è sempre notevolmente meno estesa dell’area della corrispondente ellisse di dispersione (anche se solo al 75%).

In **FIG.1**, **2** e **3** sono riportati i “centroidi di esistenza” delle vocali secondo i 5 tipi di coppie di coordinate di cui si è finora discusso, sia a categorie unificate (a destra) che distinte (a sinistra): in quest’ultimo caso i simboli fonetici, circa al centro dei “centroidi” di dispersione teorica al 95% del valore medio di ciascun campione, sono connessi da *segmenti di diverso spessore* a seconda della categoria di sesso/età di appartenenza (*grosso* = uomini, *medio* = donne e *sottile* = bambini).

A conferma di quanto detto nel precedente penultimo capoverso, si noti che in tutti i diagrammi delle figure i centroidi di 5 vocali infantili sono quasi sistematicamente più piccoli di quelli delle corrispondenti vocali delle due categorie adulte, proprio perché fra i rispettivi campioni vi è un rapporto di “numerosità” circa di 5 a 1: mentre per /ε/ ed /o/, essendo la “numerosità” ridotta e grossomodo equivalente per tutte tre le categorie, i centroidi sono sensibilmente più grandi ed anche generalmente abbastanza equivalenti intra-categoria ed intra-vocale.

In **FIG.1** i diagrammi con le coordinate acustiche “grezze” F2|F1 in Hz mostrano, a sinistra, una chiara differenziazione inter-categoria di voce delle vocali anteriori, e contemporaneamente una buona reciproca discriminazione in ordinata, il che giustifica la loro rappresentazione “a categorie di voce unificate” nel diagramma di destra.

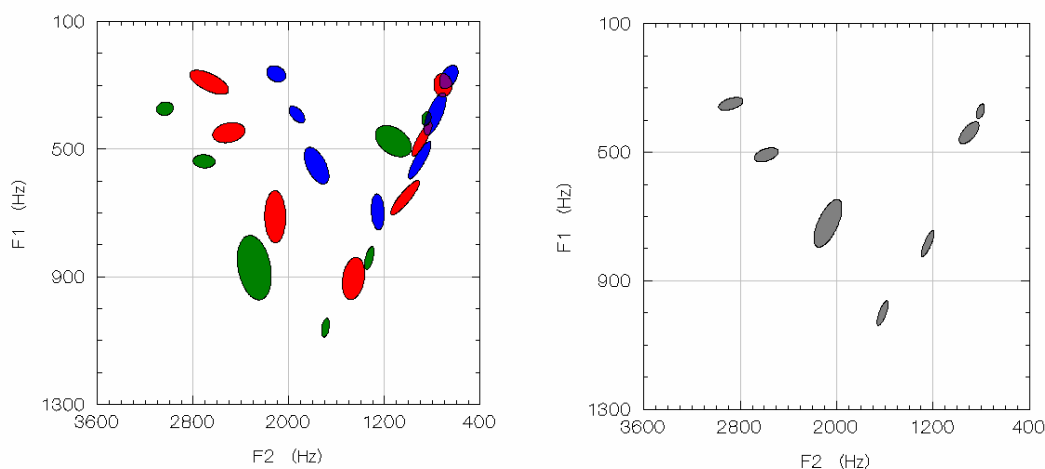


FIG. 1 Piano formantico “acustico” F2|F1 per “*categorie separate*”, uomini, donne e bambini (a sinistra), e “*unificate*” (a destra). [Cfr. testo].

Anche la vocale centrale /a/ è ben differenziata inter-categoria (con una discriminazione rispetto alle altre vocali principalmente dovuta all’ascissa), giustificando dunque la sua collocazione nel diagramma a destra “a categorie di voce unificate”: però fra la sue versioni adulte maschile e femminile si “intromette” il centroide della /ɔ/ infantile. Malgrado ciò quest’ultima vocale appare ancora ben differenziata dalle altre nel caso “a categorie di voce unificate”, anche se in effetti risulta centralizzata e abbassata a causa del peso residuo ancora consistente della “prevaricante numerosità” del campione infantile. Le restanti vocali posteriori appaiono “confusamente ravvicinate e/o sovrapposte”, financo su tipi vocalici diversi per appartenenza a differenti categorie di voce: infatti, “a categorie di voce unificate”, /o/ ed /u/ rimangono molto ravvicinate e quindi, di fatto, non normalizzate per categoria di voce, né discriminate per categoria vocalica. Dunque, il sostanziale “fallimento” delle coordinate acustiche “grezze” nella normalizzazione e nella discriminazione si verifica principalmente sulle vocali posteriori.

In **FIG.2** sono rappresentati i diagrammi con le coordinate uditive in Bark, sopra (F2-F1)|(F1-F0) e sotto (F3-F2)|(F1-F0). Rispetto alle precedenti coordinate acustiche, la prima coppia di coordinate mostra a destra una persistente, ma minore, differenziazione per “categoria di voce” delle vocali anteriori, con una discriminazione inter-vocalica migliorata dal contributo anche dell’ascissa. Anche le vocali centrale e posteriori sono molto meglio disambiguate, essendo da un lato ravvicinate inter-

categoria di voce e nel contempo abbastanza discriminate per categoria vocale, essenzialmente grazie alla ordinata: fa eccezione la /ɔ/ infantile non più “intrusa” nella zona delle /a/ adulte, però ad esse anomalamente ravvicinata, e la /o/ infantile che appare centralizzata sull’ascissa. Tutto questo si ripropone nel diagramma di sinistra dove, a “categorie di voce unificate”, se /o/ e /u/ appaiono ora ben distanziate, di converso /ɛ/ e /ɔ/ appaiono piuttosto abbassate. Inoltre /a/ e /ɔ/ sono eccessivamente posteriorizzate.

La seconda coppia di coordinate, che si differenzia dalla precedente solo sull’ascissa che ora è (F3-F2), migliora proprio sull’ascissa (a sinistra) la discriminazione delle vocali posteriore e centrale, che risultano anche meglio normalizzate sempre grazie anche al contributo dell’ordinata, fatta eccezione per la “intrusione”, questa volta, della /ɛ/ infantile fra le /a/ adulte. Il grafico di destra a “categorie di voce unificate” mostra di fatto solo una anomala centralizzazione di /o/.

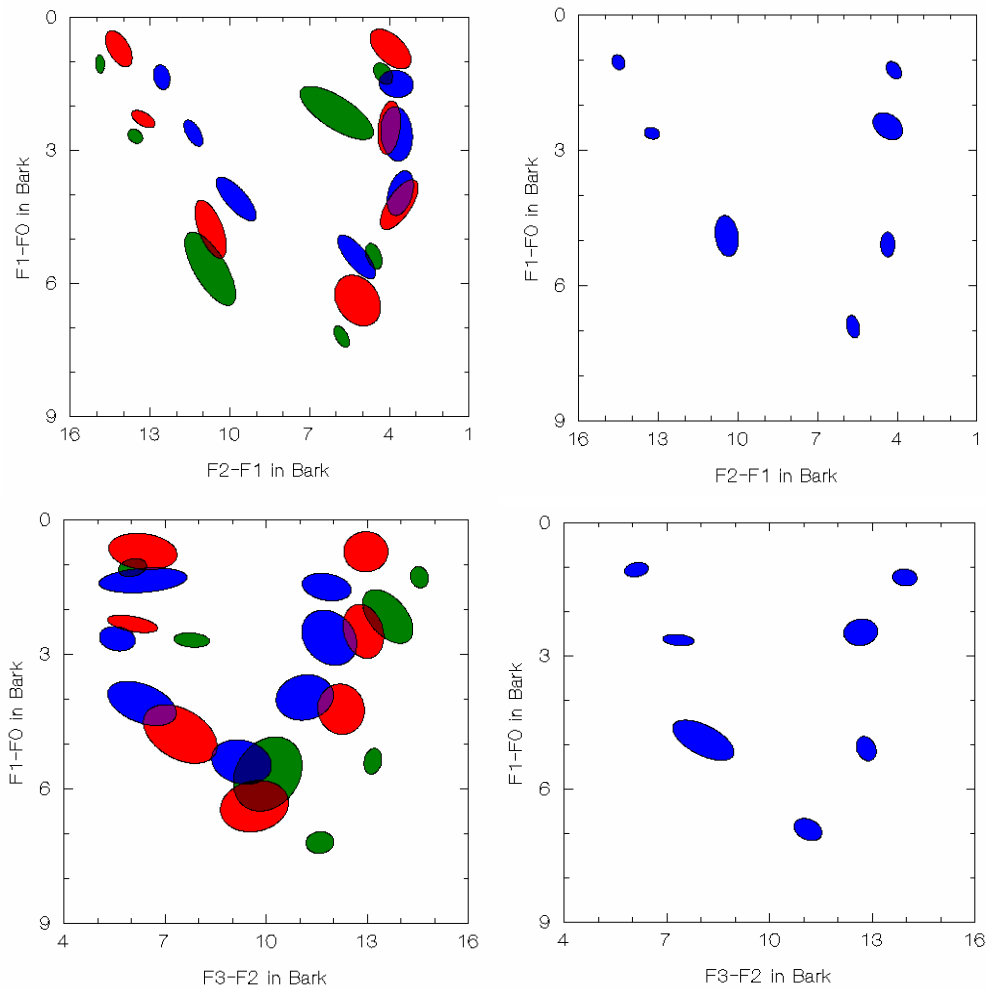


FIG. 2 Piani formantici “uditivi”, sopra (F2-F1)|(F1-F0), sotto (F3-F2)|(F1-F0), per “*categorie separate*”, uomini, donne e bambini (a sinistra), e “*unificate*” (a destra). [Cfr. testo].

In **FIG.3** sono rappresentati i diagrammi con le coordinate uditive in Bark e con l’uso della *seconda formante effettiva*, con la stessa ordinata (F1-F0) e in ascissa rispettivamente, sopra F2’ e sotto (F2’-F1). In effetti le configurazioni diagrammatiche

si modificano abbastanza sensibilmente rispetto a quella delle figure precedenti, sostanzialmente per una apparente generale anteriorizzazione complessiva delle vocali anteriori e centrale. L'ascissa $F2'$ determina una rilevante centralizzazione delle vocali posteriori e una iper-anteriorizzazione delle vocali anteriori infantili rispetto alla due configurazioni di voci adulte; invece l'ascissa $(F2'-F1)$ opera una iper-posteriorizzazione delle vocali posteriori e centrale dei due tipi di voce adulta rispetto a quella infantile.

Malgrado ciò, in entrambi i tipi di grafico, complessivamente la dislocazione di ciascuna vocale delle tre categorie di voce appare ben distinta da tutte le altre, eccezione fatta per la /ɔ/ infantile che cade nella zona della /a/ adulta maschile. Nei diagrammi a destra a “categorie unificate”, questa anomalia si traduce in una centralizzazione e un abbassamento di /ɔ/, anche /ɛ/ risulta relativamente abbassata a causa della preponderante influenza della sua versione infantile.

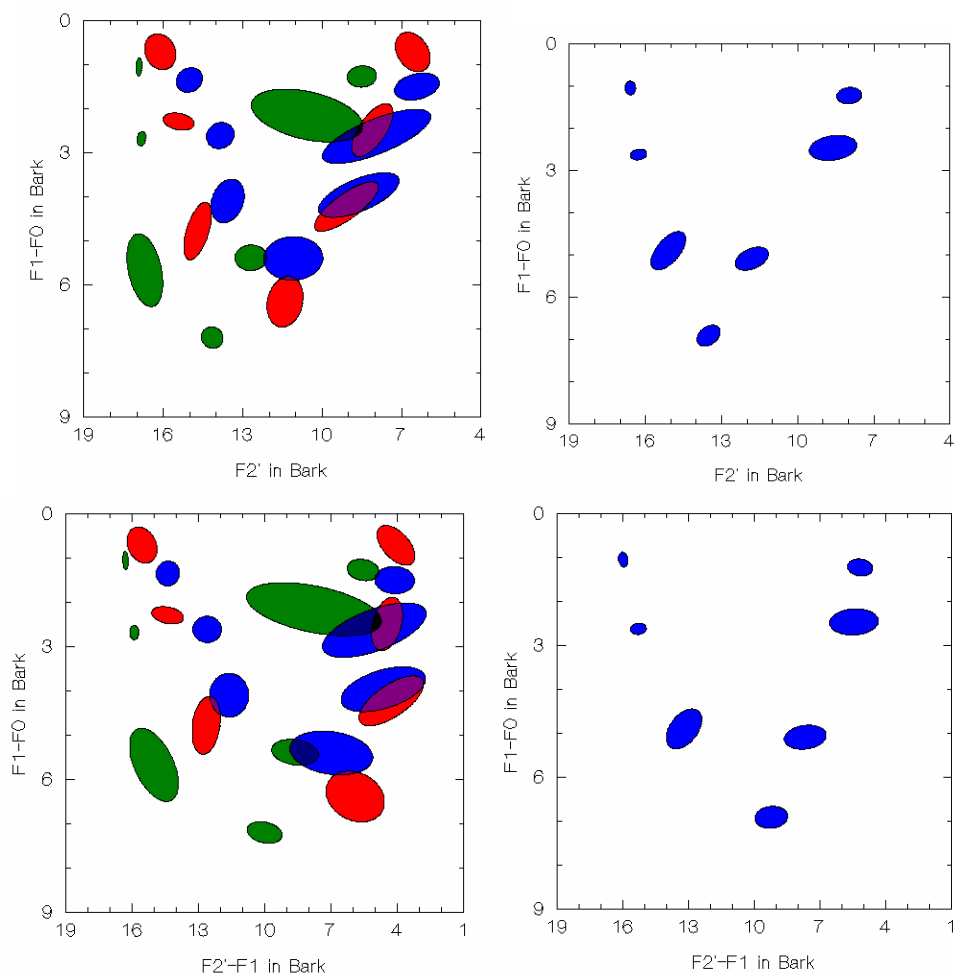


FIG. 3 Piani formantici “uditivi”, sopra $F2'|(F1-F0)$, sotto $(F2'-F1)|(F1-F0)$, per “*categorie separate*”, uomini, donne e bambini (a sinistra), e “*unificate*” (a destra). [Cfr. testo].

CONCLUSIONI

Questi risultati dei confronti di *normalizzazione inter-sesso/età* e di *discriminazione intervocalica* devono essere considerati solamente come “indicatori di possibili tendenze” in quanto questo tipo di studio potrà essere validato solo utilizzando data base sufficientemente omogenei nel numero di soggetti per vocale e per sesso/età.

Comunque a conclusione di questo studio pilota si ritiene pertinente analizzare la morfologia del *triangolo vocalico fonetico* di questo Italiano regionale, così come risulta dai vari piani formantici considerati (i grafici a destra di ciascuna figura).

Innanzitutto, le coordinate acustiche di **FIG. 1** si confermano inadeguate in quanto non collocano le vocali posteriori a distanze reciproche sufficientemente corrette.

Le coordinate uditive di **FIG. 2**, con l'ascissa (F2-F1), pur allontanando /o/ rispetto a /u/, quantomeno posteriorizzano /ɔ/ e /a/ in modo anomalo rispetto alla tradizionale categorizzazione fonologica. Invece, con l'ascissa (F3-F2), /a/ appare meglio centralizzata, ma di contro lo è anche /o/ in contrasto con la categorizzazione fonologica.

Le coordinate uditive basate sulla *seconda formante effettiva* di **FIG. 3** mostrano con entrambi i tipi di ascissa un leggero abbassamento della /ɔ/ e inoltre anche una anteriorizzazione/centralizzazione delle vocali basse quando l'ascissa è F2': questo spostamento non è realizzato dalla ascissa (F2'-F1) che dunque presenta un triangolo vocalico in assoluto più simile a quello fonetico.

La natura "esplorativa" di questo studio, dovuta, come si è detto, alla disomogeneità numerica dei campioni di voce/vocale, è anche imputabile alla scelta contingente di privilegiare la ottimalità dei campioni congiuntamente alla precisione delle misure (effettuate, infatti, in modo *semiautomatico*) rispetto alla quantità di dati utilizzati: Anche da ciò deriva il fatto che il materiale usato non può essere statisticamente rilevante come in altri studi (per es.: Cosi et al., 1995) in cui si è adottata una metodologia completamente *automatica*, con ciò privilegiando la quantità di dati (e quindi la significatività delle statistiche) rispetto alla precisione delle misurazioni. Si ritiene che si potrà giungere ad un ottimale compromesso tra precisione di misura e validità statistica dei risultati, portando la numerosità di ciascuna vocale/voce a circa 50 realizzazioni. Solo allora si potranno confermare o revisionare le linee di tendenza provvisorie messe in luce in questo studio.

BIBLIOGRAFIA

- COSI P., FERRERO F.E. e VAGGES K., "*Rappresentazioni acustiche e uditive delle vocali italiane*", in Atti del 23° Convegno Naz. AIA, Bologna, **1995**, 151-156.
- FANT G., "*Acoustic Theory of Speech Production*", Mouton, The Hague (NL), 1960.
- FANT G., "*Non-uniform Vowel Normalization*", STL-QPSR 2-3, **1975**, 1-19.
- FANT G., "*The Voice Source Theory and Acoustic Modelling*", in "Vocal Fold Physiology, Acoustics and Phonatory Control" edited by I.R. Titze and R.C. Scherer, The Dever Center for the Performing Arts, Denver (CO), **1985**, 453-464.
- FERRERO F.E., MAGNO CALDOGNETTO E. e COSI P., "*Le vocali al femminile*", in: Atti del Convegno Internazionale "Dialectologia al femminile", Sappada (BL), 26-30 giugno 1995, CLUEP Padova, **1995**, 413-436.
- HILLENBRAND J., GETTY L.A., CLARK M.J. e WHEELER K., "*Acoustic Characteristics of American English Vowels*", JASA 97, **1995**, 3099-3111.
- KENT R.D. e READ C., "*The Acoustic Analysis of Speech*", Singular Publishing Group. San Diego (CA), **1992**.
- KLATT D.H. e KLATT L.C., "*Analysis, Synthesis and Perception of Voice Quality Variations among Female and Male Talkers*", JASA 87 (2), **1990**, 820-857.
- PACETTO M.G. e FERRERO F.E. (Relatore), "*Diagrammi di esistenza acustici e uditivi delle vocali infantili italiane*", Tesi di Diploma in Logopedia, Università di Padova, aa. **1994-95**.

PETERSON G.E. e BARNEY H.L., "*Control Methods used in a Study of the Vowels*", JASA 24, **1952**, 175-184.