

Patrizia Soriano

Per una rappresentazione uditiva dei segmenti vocalici: il caso del senese
(versione estesa dell'articolo *Un tentativo di classificazione uditiva delle vocali toniche senesi*, in *Così P., Magno Caldognetto E. & A. Zamboni, Scritti in onore di Franco Ferrero, Padova, Unipress, in stampa*)

1. Premessa

La rappresentazione acustica di un sistema vocalico avviene, tradizionalmente, mediante l'elaborazione di diagrammi cartesiani aventi come coordinate i valori frequenziali medi della prima e della seconda formante di una vocale. Generalmente, gli assi cartesiani dei piani formantici impiegati per visualizzare le aree di esistenza delle vocali sono di tipo acustico, cioè espressi in Hz. Tali configurazioni acustiche presentano però un limite, in quanto non rappresentano fedelmente il modo con il quale l'orecchio umano discrimina le componenti frequenziali di suoni diversi. Come è noto, il sistema uditivo dell'uomo percepisce un suono secondo complessi meccanismi fisiologici e psicologici. La relazione tra il dato acustico e la sensazione di altezza realmente indotta non è lineare, giacché il nostro orecchio percepisce le variazioni di frequenza secondo una scala logaritmica. Stando a quanto enunciato dalla psicoacustica, legge di Weber-Fechner, ciò che è importante ai fini della discriminazione di due suoni è il rapporto esistente tra le frequenze piuttosto che la loro distanza assoluta: la risposta psicologica fra due grandezze equivale infatti ad una differenza costante del loro valore logaritmico¹. Basti pensare che ogni volta che la frequenza di un suono è raddoppiata si percepisce sempre una stessa differenza, cioè un'ottava. Di conseguenza, le variazioni che avvengono a bassa frequenza hanno una maggiore rilevanza percettiva rispetto a quelle che avvengono alle alte frequenze.

Per questo motivo, nel corso degli anni, l'elaborazione di scale percettive per la rappresentazione acustica dei diagrammi vocalici, in sostituzione delle coordinate acustiche in Hz, è stata avvertita dagli studiosi come una forte esigenza teorica. Si passò gradualmente all'adozione di scale in Hz semilogaritmiche², all'impiego della scala in Mel (cfr. Fant 1973) o più recentemente in Bark. Il bark costituisce l'unità di misura della scala uditiva a bande critiche (ingl. *critical bandwidth*), attualmente considerata la più efficace per la classificazione delle vocali. La natura di tale scala uditiva, teorizzata da Zwicker nel 1961 dopo una lunga serie di esperimenti di psicoacustica, è essenzialmente fisiologica, in quanto strettamente dipendente dall'anatomia dell'orecchio umano³. Più precisamente, il sistema uditivo dell'uomo analizza un suono come una successione in parallelo di più filtri passa banda. Questi filtri non hanno una larghezza di banda costante, bensì variabile a seconda della frequenza centrale occupata dalla banda medesima e, quindi, più piccola nelle fasce di frequenza più basse, più ampia alle alte frequenze (cfr. Zwicker *et alii* 1957). La larghezza di banda di ogni filtro corrisponde per l'appunto ad un bark, ovvero ad una banda critica. Il campo uditivo dell'uomo risulta così scomponibile in 24 bark, solo i suoni che

¹ Tale legge è così formulabile: $R = K \log S$, dove R rappresenta la risposta psicologica ad uno stimolo sonoro, S la grandezza oggettiva dello stimolo, mentre K è una costante di proporzionalità (cfr. Fechner 1966). Sulla complessa tematica relativa alla percezione dei suoni si consulti anche l'importante contributo di Stevens (1975).

² Si ricorda a tale proposito la scala proposta da Koenig (1949), lineare fino a 1000 Hz e poi logaritmica.

³ Più nel dettaglio, un bark corrisponde ad uno spostamento sistematico della membrana basilare del canale cocleare di circa 1.5 mm e a 1200 fibre nervose. Cfr. Zwicker (1961), Zwicker e Terhardt (1980).

ricadono in bande critiche diverse avranno salienza percettiva, i suoni che rientrano in una stessa banda critica sono per converso percepiti come identici (cfr. Schroeder *et alii* 1979).

Diversi sono stati i vantaggi derivati dall'impiego di questa scala uditiva. Uno fra tutti, rivelatosi fondamentale ai fini dell'analisi vocalica tanto da aver assunto ormai grande importanza interpretativa, è stato certamente quello di aver messo in evidenza l'esistenza di alcune relazioni costanti, somme o differenze, che i valori formantici delle vocali instaurano reciprocamente nel sistema. Numerosi studi condotti su lingue diverse hanno provato infatti come la discriminazione percettiva delle vocali migliorasse sostanzialmente sostituendo i parametri formantici F1 e F2 con la distanza tra coppie di formanti, per l'esattezza F1-Fo, F2-F1, F3-F2. Tra i primi studiosi ad aver verificato sperimentalmente l'utilità di sostituire F1 con F1-Fo ricordiamo Traunmüller (1981; 1985) il quale, avvalendosi di stimoli sintetici, trovò che la differenza in bark tra F1 e Fo di una stessa vocale costituisse il criterio fondamentale per una corretta percezione dell'altezza vocalica (cfr. anche Syrdal 1985, Syrdal e Gopal 1986).

Più recentemente, Di Benedetto (1991; 1994), in una ricerca svolta sulle vocali dell'anglo-americano, pronunciate da tre locutori e successivamente sintetizzate, osservava come il valore di Fo agisse da fattore di normalizzazione uditiva, attenuando le differenze individuali riscontrate, specie quelle delle vocali basse.

Analogamente, le variazioni che avvengono lungo la dimensione anteriore/posteriore risultano meglio discriminate, a livello percettivo, dalla distanza esistente tra le prime due formanti (F2-F1) piuttosto che da F2 (cfr. Ladefoged 1975).

In un modello generale di rappresentazione uditiva delle vocali, l'efficacia del rapporto tra coppie di formanti quale strumento di classificazione è, d'altra parte, una diretta conseguenza dell'effetto del *Centro di Gravità Spettrale*, o soglia critica, identificato da Chistovich e collaboratrici nel 1979 e pari esattamente a 3.5 bark. Questi, in sintesi, i risultati conseguiti dalle autrici in una ricerca compiuta sulla percezione di vocali russe sintetiche: le prime formanti di una vocale sono percettivamente riconosciute solo quando tra i due valori vi è una distanza superiore a 3.5 bark, una distanza critica quest'ultima capace di annullare gli effetti del *Centro di Gravità Spettrale*. Qualora la distanza tra le due formanti sia invece inferiore a 3.5 bark, l'orecchio umano, attuando una sorta di media aritmetica, tende a identificare una sola zona di risonanza dal valore pressoché intermedio rispetto ai due originari. Tale fenomeno di risposta percettiva dà luogo ad una integrazione spettrale, o *Centro di Gravità*, tra valori frequenziali non sufficientemente distanziati. Da qui il peso teorico della nozione di soglia critica che si prefigura quindi come un forte confine divisorio al di sopra e al di sotto del quale gli effetti del *Centro di Gravità Spettrale* sarebbero rispettivamente nulli o presenti.

Il concetto teorico della distanza critica si rivela così fondamentale anche ai fini della descrizione e della classificazione di un sistema vocalico. Studi effettuati da Syrdal (1985) e da Syrdal e Gopal (1986) sull'anglo-americano dimostrano che una distanza F1-Fo inferiore a 3.5 bark caratterizza le vocali alte, una distanza superiore le vocali basse. Relativamente al luogo di articolazione, differenze F2-F1 inferiori a 3.5 bark sono tipicamente presenti nei timbri posteriori, mentre differenze superiori sono presenti nelle vocali anteriori. Anche la somma F1+F2 rappresenta una valida dimensione formantica capace di distinguere, posto un medesimo grado di altezza, vocali arrotondate da vocali non arrotondate (cfr. Fant 1983 e per l'italiano Uguzzoni 1988; 1994).

2. *Gli studi sull'italiano.*

Limitatamente alla lingua italiana, gli studi acustici sul vocalismo che hanno applicato tali criteri di classificazione uditiva non sono molti. La finalità di questi lavori non è sempre direttamente comparabile, così come non lo sono i protocolli sperimentali impiegati. Ciò nonostante, i risultati sperimentali ai quali si perviene procedono, significativamente, verso un'unica direzione, ad ulteriore conferma della validità metodologica di tale approccio uditivo.

Un obiettivo propriamente linguistico è perseguito da Uguzzoni (1988; 1994) in due studi incentrati sul sistema vocalico del dialetto frignanese di Crocette, una frazione di Pavullo in provincia di Modena. L'autrice, creando delle matrici classificatorie in bark dei valori formantici, individua gli indici qualitativi delle nove vocali lunghe del dialetto a partire dai risultati acustici del campione⁴. La ricerca dimostra come la differenza tra coppie di formanti costituisca un parametro adeguato per la distinzione della qualità vocalica. In special modo, la coordinata F2-F1 riesce a collocare, nella giusta posizione articolatoria, il timbro posteriore centralizzato /u:/, una vocale non correttamente distinta a livello acustico dalla serie posteriore. Particolarmente valida risulta inoltre la somma delle prime due formanti, F1+F2, giacché permette la separazione, a parità di altezza vocalica, tra i timbri arrotondati e quelli non arrotondati presenti nel sistema frignanese. Secondo l'autrice: "i parametri F2-F1, F1+F2, F1 si rivelano dunque efficaci strumenti di categorizzazione e di sottocategorizzazione e di essi è consigliabile avvalersi soprattutto per la descrizione di sistemi vocalici complessi" (Uguzzoni 1988: 104).

L'impiego della scala percettiva in bark e soprattutto l'elaborazione grafica di diagrammi cartesiani con coordinate F1-Fo, F2-F1, F3-F2 è utile anche nella normalizzazione della variabilità formantica dipendente dal sesso e dall'età, come documentano gli studi svolti da Così *et alii* (1995), Ferrero *et alii* (1995), Ferrero *et alii* (1996)⁵. Queste ricerche, il cui obiettivo essenziale è costituito dalla descrizione acustica e uditiva dei sistemi vocalici appartenenti a uomini, donne e bambini, evidenziano, da un lato, l'inadeguatezza delle coordinate acustiche in Hz F1/F2 nella normalizzazione delle voci e, dall'altro, la capacità detenuta dalla scala uditiva di ridurre, in modo rilevante, le covariazioni imputabili al genere. A tali conclusioni si giunge solo dopo aver comparato statisticamente il potere descrittivo e differenziante di più rappresentazioni vocaliche, elaborate tanto su scala frequenziale lineare che logaritmica. Ferrero *et alii* (1995) osservano, ad esempio, che il miglior grado di normalizzazione sia quello derivato dall'impiego degli assi cartesiani F1-Fo/F3-F2 in bark. Particolarmente forte, ai fini della normalizzazione, si rivela anche il parametro della seconda formante effettiva (F2') nella coordinata F2'-F1, in quanto, come testato da Ferrero *et alii* (1996), riesce ad abbassare il grado di variabilità dipendente dal sesso, senza per questo alterare la disposizione fonetica dei timbri all'interno del triangolo vocalico⁶.

⁴ Il sistema vocalico tonico di questo dialetto, come esplicitato da Uguzzoni (1988) è costituito da 13 timbri, di cui 9 lunghi e 4 brevi.

⁵ Nel panorama degli studi italiani svolti in prospettiva uditiva, non può essere ignorata l'importante rassegna di Ferrero (1994) tesa alla valutazione critica dei risultati ottenuti, per le vocali dell'italiano, dall'autore medesimo, ben 25 anni prima.

⁶ Per il concetto teorico relativo alla seconda formante effettiva (F2'), intesa come una combinazione dei valori di F2, F3 e F4, si legga quanto riportato da Carlson *et alii* (1975).

3. *Materiali e metodi*

Lo scopo di questa ricerca è quello di classificare in termini uditivi il sistema vocalico, tonico e atono, dell'italiano senese. Il materiale linguistico preso in considerazione, tre testi preconfezionati, di cui due dialoghi, è stato realizzato da tre locutori maschi nati e vissuti a Siena, aventi un'istruzione superiore (studenti universitari o laureati) e un'età variabile tra 25 e 30 anni. Agli informatori è stata preliminarmente concessa la facoltà di visionare i due dialoghi strutturati e di apportare, se necessario, modifiche di natura sintattica e lessicale, allo scopo di far aderire il testo da somministrare alle loro abitudini linguistiche. In seguito, a ciascun locutore è stato chiesto di interpretare i dialoghi nel modo più naturale possibile, avendo come riferimento il testo scritto opportunamente modificato⁷. La produzione del testo non è avvenuta quindi tramite una vera e propria lettura. La metodologia adottata ha consentito una riduzione del grado di controllo, inevitabilmente presente nella lettura, inducendo nel contempo una maggiore naturalezza enunciativa.

Dalle registrazioni così ottenute è stato estratto un campione di 1700 ricorrenze, delle quali 1000 vocali toniche e 700 vocali atone. Le vocali analizzate, circa 45/50 per timbro, ricorrono sia in sillaba aperta che chiusa e in diversi contesti consonantici. Dall'analisi è stata sistematicamente esclusa la ricorrenza delle vocali realizzate alla fine di una unità tonale, di natura sia continuativa che conclusiva.

L'analisi acustica è stata svolta per mezzo del *software Kay Multispeech 3700* versione 2.2, previo campionamento digitale dei *corpora* di parlato a 22050 Hz e conversione a 16 bit. Per ogni vocale è stato rilevato il valore frequenziale di FO, F1, F2, F3 e la durata temporale dell'intero segmento⁸. L'estrazione dei valori formantici è stata effettuata sull'involuppo di uno spettro medio *FFT (Long Term Average Power Spectrum)* con filtro a banda larga in corrispondenza della parte stazionaria del segmento. Tutti i valori formantici del campione sono stati successivamente convertiti in Bark applicando la seguente formula proposta da Traunmüller (1990):

$$1) \quad \text{Bark} = [26.81 F / (1960 + F)] - 0.53$$

Per la rappresentazione grafica del sistema vocalico sono stati elaborati dei diagrammi cartesiani F1/F2 con coordinate uditive in bark. Per una migliore lettura dei grafici riportati di seguito nel testo, è utile specificare che ciascun area di esistenza vocalica è un'ellisse equiprobabile al 68%. Il centro dell'area corrisponde infatti al valore frequenziale medio delle prime due formanti, mentre la lunghezza dei semiassi è data dal valore di una deviazione standard.

4. *Le vocali toniche*

Il piano formantico uditivo del vocalismo tonico senese (cfr. Fig. 1) appare di poco più ristretto rispetto a quello acustico già riportato in Soriano (in stampa). L'impiego della scala uditiva in bark determina infatti una maggiore compattezza delle coordinate F1 e F2 e, di conseguenza, una riduzione dello spazio articolatorio occupato dall'intero sistema. Pur tuttavia, non si rilevano sostanziali variazioni nella disposizione delle ellissi dei 7 timbri accentati. La caratteristica più evidente del triangolo vocalico, oltre ad un buon grado di simmetria, è data dall'assenza di ampie zone di sovrapposizione. Le aree di dispersione delle vocali, ad eccezione di /a/, presentano comunque diversi punti di contatto tanto sul

⁷ Per questo motivo, nei tre locutori esaminati, la realizzazione dei testi non è coincidente.

⁸ Per quel che riguarda Fo, è stato considerato sia il valore iniziale, finale e medio della medesima porzione selezionata per l'estrazione dello spettro, sia il valore di *onset* e di *offset* dell'intera vocale

versante anteriore che su quello posteriore. I dati relativi alle vocali sono riportati, rispettivamente in Hz e in Bark, nelle Tabb. 1 e 2. Da questi dati empirici si osserva come i valori in bark di F1 siano direttamente proporzionali al grado di apertura della vocale. Il fonema basso /a/ detiene infatti il valore di F1 più alto (5.9), per contro le vocali alte, aventi il minimo grado di apertura, sono connotate da valori nettamente inferiori, 3.2 per /i/ e 3.4 per /u/. Per le vocali di media altezza si registrano valori intermedi. Anche per quanto concerne F2, si nota una chiara relazione tra l'indice uditivo delle vocali e il loro luogo di articolazione. Relativamente a questa dimensione formantica, i timbri anteriori si collocano ai vertici della scala con i valori più alti, in netta opposizione a quelli posteriori che si collocano nella fascia più bassa del parametro F2.

Per far sì che la rappresentazione uditiva rispecchiasse, il più possibile, il modo in cui l'orecchio umano discrimina le variazioni frequenziali presenti lungo il parametro dell'altezza vocalica e lungo l'asse anteriore/posteriore, si è proceduto con il calcolo della distanza tra coppie di formanti (cfr. Tab. 2). I risultati di questa procedura sono stati impiegati per l'elaborazione grafica di un nuovo diagramma (cfr. Fig. 2).

Quest'ulteriore normalizzazione ha come effetto diretto uno spostamento generale di tutte le aree di esistenza vocalica verso valori più bassi della scala in bark, per quanto riguarda sia F1 che F2. Tutti i timbri vocalici mostrano una diversa disposizione rispetto al diagramma F1/F2 della Fig. 1. Sul versante anteriore, si osserva una graduale, ma significativa, convergenza verso il centro delle tre vocali, un movimento che conferisce al sistema la tipica configurazione triangolare e che coinvolge, in misura progressivamente maggiore, i timbri più bassi. Sul versante posteriore, invece, si nota una sorta di allineamento verticale delle vocali che interessa anche /a/, motivo per cui tali fonemi risultano differenziati quasi esclusivamente per altezza, ma non per luogo di articolazione. Ciò determina un restringimento a destra dello spazio vocalico che raggiunge il suo punto minimo tra /ε/ e /ɔ/. Pur tuttavia, non si rilevano zone di significativa sovrapposizione tra le ellissi, ogni timbro sembra quindi conservare una propria identità fonetica anche dopo la normalizzazione uditiva effettuata.

Per valutare in modo più accurato il rapporto che ciascuna vocale instaura all'interno del sistema, i nuovi valori formantici ottenuti dalla distanza F1-Fo e da F2-F1, posti in ordine crescente, sono stati valutati anche rispetto al valore della soglia critica, che ricordiamo essere equivalente a 3.5 bark (cfr. § 1). Esaminando più analiticamente la dimensione F1-Fo (cfr. Tab. 3), si evince che il valore di Fo procede in senso diametralmente opposto a quello della prima formante, essendo intrinsecamente alto nelle vocali chiuse, notoriamente connotate da una F1 alta, ma basso nelle vocali aperte, caratterizzate queste ultime da valori di F1 tendenzialmente alti. Di conseguenza, il calcolo della distanza tra F1 e Fo ha come esito un indice progressivamente crescente con l'aumentare del grado di apertura della vocale. In linea con le nostre attese, in tutte le vocali, ad eccezione di /a/, il valore medio F1-Fo è inferiore a quello della soglia critica. Più complessa, la situazione che si delinea analizzando il rapporto tra F2-F1 (cfr. Tab. 4). Secondo quanto teorizzato da Syrdal (1985) e Syrdal e Gopal (1986), la distanza tra F2-F1 dovrebbe essere superiore alla soglia differenziante nelle vocali anteriori, ma inferiore nelle vocali posteriori. Ebbene, nell'italiano di Siena tale rapporto formantico è sempre superiore a 3.5 bark in tutte le vocali, anche se decresce gradualmente dal timbro anteriore più periferico /i/ fino a /ɔ/, la vocale che, seguendo questa interpretazione, risulta caratterizzata dal maggior grado di posteriorità.

E' ragionevole supporre che le vocali posteriori del senese siano connotate da un certo grado di avanzamento articolatorio e che pertanto siano, per così dire, poco posteriori, un aspetto quest'ultimo già emerso nella rappresentazione acustica del sistema vocalico e che significativamente permane anche dopo la conversione dei dati su scala uditiva⁹. Particolarmente eloquente si dimostra a tal fine la configurazione grafica visualizzata nella Fig. 2, in cui si ravvisa una disposizione verticale, e pressoché centrale, delle ellissi dei timbri posteriori i quali sembrano perdere la propria posizione periferica mossi da una spinta centripeta (cfr. *ultra*). Prima di trarre conclusioni, si ritiene utile passare in rassegna anche i dati inerenti al sistema vocalico non accentato.

5. Le vocali atone

Dal diagramma uditivo F1/F2 relativo al vocalismo atono (cfr. Fig. 3), emerge una generale riduzione dello spazio vocalico utilizzato dal sistema. La tendenza alla centralizzazione colpisce ambedue le dimensioni, F1 e F2, causando un restringimento dello spazio vocalico e il conseguente ravvicinamento di tutte le aree di dispersione verso il centro del triangolo. Complessivamente, il vocalismo atono utilizza uno spazio più compresso; la distanza tra i vertici del triangolo, definito dai valori formantici delle vocali cardinali /i/ /u/ e /a/ è minore rispetto a quello del sistema tonico. Ad esempio, la distanza tra F1 di /i/ e di /u/ rispetto a F1 della /a/ è di 326 Hz (bark 3.2) nel vocalismo tonico, ma di soli 203 Hz (bark 1.9) in quello atono. La distanza invece tra il valore di F2 di /i/ e quello di /u/ è di 1076 Hz (bark 8.9) nel sistema tonico, ma di 817 (bark 7.3) in quello atono. Le vocali anteriori subiscono un minore spostamento attestandosi su spazi più periferici. Sovrapponendo il diagramma della Fig. 3 con quello delle vocali toniche (cfr. Fig. 1) si osserva che il timbro che converge di più verso il centro è /a/; /i/ ed /e/, per converso, sembrano mantenere la loro posizione originaria, risultando più forti agli effetti di centralizzazione. Per un confronto diretto dei valori formantici delle vocali non accentate si riportano nella Tab. 6 le frequenze medie in bark delle prime due formanti. Nella Tab. 5 sono invece presenti, come utile riferimento, i valori medi in Hz di Fo, F1 e F2.

Nonostante l'azione di questa chiara tendenza di tipo centripeto, i valori in bark delle vocali poste in contesto atono delineano una situazione del tutto simile a quella riscontrata per le toniche. Nuovamente, i valori di F1 decrescono progressivamente dalla vocale bassa /a/ alle vocali alte /i/ ed /u/. Relativamente a F2, i timbri anteriori presentano l'indice bark più alto, mentre le posteriori il valore più basso; /a/ assume una posizione pressoché intermedia tra le due serie articolatorie, per quanto concerne sia F1 che F2, anche se si nota un lieve spostamento verso l'asse posteriore. Al fine di valutare la posizione articolatoria delle vocali atone, si è proceduto con il calcolo della distanza tra coppie di formanti. Consideriamo innanzitutto la distanza esistente tra F1 e Fo. In accordo con le nostre attese, solo il valore in bark della vocale bassa /a/ (3.9) è superiore, anche se di poco, all'indice di soglia critica; per le vocali alte e per quelle di media altezza si registrano valori inferiori alla soglia (cfr. Tab. 7).

Meno chiara la situazione che emerge dal computo della distanza calcolata tra F2 e F1: tutti i timbri atoni mostrano difatti un valore medio in bark superiore a quello della distanza critica (cfr. Tab. 8). Tale comportamento potrebbe essere imputato al processo di riduzione

⁹ In merito a questo aspetto, va osservato che Di Benedetto (1991) fa uso di un diagramma con assi cartesiani diversamente specificati, ovvero F1-Fo in bark, ma F2 in Hz, ritenendo quest'ultimo parametro già capace di assegnare una giusta collocazione ai timbri vocalici lungo l'asse antero/posteriore.

timbrica che investe il sistema atono. L'effetto di centralizzazione sarebbe però asimmetrico, più forte cioè lungo la dimensione anteriore/posteriore (F2) che non lungo la dimensione dell'altezza vocalica (F1), in quanto determina una convergenza più vistosa e marcata dell'asse posteriore verso il centro del triangolo. Quanto detto risulta ancora più palese dalla rappresentazione grafica del diagramma vocalico con coordinate in bark F1-Fo e F2-F1 (cfr. Fig. 4), in cui il movimento verso il centro di tutti i timbri in atonia appare significativo. Il sistema atono occupa uno spazio ancora più compresso, la riduzione in questo caso coinvolge in modo sostanziale la dimensione dell'altezza (F1), come è evidente confrontando la posizione detenuta dall'ellisse di /a/ rispetto al medesimo timbro tonico. Nell'insieme, il vocalismo non accentato sembra riprodurre su scala ridotta le stesse tendenze già rinvenute per le vocali accentate.

6. Osservazioni.

La rappresentazione uditiva del vocalismo senese, sia tonico che atono, sembra evidenziare una diversa caratterizzazione dei due assi articolatori: le vocali anteriori risultano infatti maggiormente differenziate lungo la dimensione F2 di quanto non lo siano le posteriori. Le aree di dispersione dei timbri posteriori si dispongono, all'interno del triangolo vocalico, in un'area piuttosto centrale. La distanza tra coppie di formanti calcolata per queste vocali è, quale diretta conseguenza, sempre superiore a quella della soglia critica. La pronuncia avanzata delle vocali posteriori potrebbe, ragionevolmente, essere assunta quale tratto tipico del vocalismo senese. Tuttavia, nell'ampia letteratura sul parlato toscano non vi sono riferimenti diretti sulla pronuncia delle vocali senesi che possano confortare il nostro dato. Tale ipotesi appare però suffragata, sul piano acustico, dai risultati dell'esame comparativo effettuato in Soriano (in stampa) tra i dati senesi e quelli riportati da Ferrero *et alii* (1978) e da Calamai (1999-2000), rispettivamente per Firenze e Livorno¹⁰. Da questa valutazione contrastiva emerge che le vocali posteriori senesi hanno, nel sistema, una disposizione più avanzata di quella detenuta dai medesimi timbri del fiorentino e del livornese¹¹.

La situazione del senese rimane pressoché immutata anche ad un confronto uditivo con le varietà toscane sopraindicate. Tale esame, dal carattere tendenziale e puramente indicativo, si è reso necessario per disporre di una preliminare base contrastiva con la quale comparare i nostri dati e verificare la validità di alcune ipotesi interpretative. Ciò ha implicato la conversione in bark dei valori formantici medi ottenuti da Ferrero *et alii* (1978) e da Calamai (1999-2000) e il successivo computo della distanza F2-F1¹². Nella Fig. 5 abbiamo proiettato, sul sistema vocalico tonico di Siena, le medie delle vocali disponibili per queste varietà di parlato toscano. Da questo raffronto intraregionale emerge che tutte le vocali senesi toniche si discostano da quelle fiorentine per quanto riguarda il valore di F2, i timbri senesi anteriori sono infatti lievemente più posteriori, mentre quelli posteriori

¹⁰ I dati sperimentali che abbiamo considerato per Livorno sono quelli relativi al campione maschile e alla pronuncia di parole bersaglio poste in posizione interna di frase cornice, il contesto enunciativo tra quelli esaminati da Calamai (1999-2000) più vicino al nostro materiale linguistico.

¹¹ La posizione dei fonemi anteriori senesi manifesta, al contrario, un buon grado di approssimazione con quella delle varietà toscane assunte come termini di confronto.

¹² Poiché tali valori sono espressi in Hz, è stato necessario, in via preliminare, convertire i dati fiorentini e livornesi in Bark. Occorre sottolineare che la trasformazione ha riguardato solo i valori medi, mentre per Siena le aree di esistenza vocalica in bark sono state ottenute mediante la conversione in bark di tutti i valori frequenziali assoluti e non solo di quelli medi.

appaiono decisamente più anteriori. La comparabilità è invece maggiore per il parametro F1, si nota tuttavia la maggiore apertura della /i/ senese e l'articolazione più arretrata della /a/. Nel complesso, tra Siena e Firenze c'è una buona coincidenza degli spazi acustici occupati dalle vocali anteriori, ma non di quelle posteriori, approssimabili a quelle fiorentine per altezza, ma non per luogo di articolazione. Tra Siena e Livorno la convergenza concerne invece l'articolazione di /i/ e /e/ e il carattere velare del fonema /a/.

Gli esiti sperimentali sembrano quindi collocare l'italiano di Siena in una posizione intermedia tra il toscano occidentale livornese e il fiorentino¹³. Del resto, il pisano-livornese costituisce una varietà marcata per ciò che concerne proprio il vocalismo. Tale varietà risulta infatti connotata, rispetto all'italiano di Siena e di Firenze, da vocali più aperte e rilassate, specie quelle medio-basse, oltre che più lunghe (cfr. Giannelli 1976: 50)¹⁴.

Nel quadro generale che si è andato delineando, un aspetto, a nostro avviso, merita di essere approfondito. L'esame comparativo dei dati uditivi ha contribuito ad evidenziare come nelle vocali posteriori del senese, la distanza F2-F1 superi sistematicamente il valore della soglia critica. Non può, tuttavia, passare inosservato il fatto che tale rapporto (F2-F1) superi i 3.5 bark in buona parte delle vocali posteriori fiorentine e livornesi. Numerose sono le riflessioni che si impongono a questo punto. La prima ipotesi interpretativa, puramente linguistica, è che l'italiano toscano sia caratterizzato da una pronuncia postero-centrale delle vocali posteriori (cfr. anche Canepari 1986: 57), un tratto articolatorio questo comunque più forte nel parlato senese, come documentano gli esiti del confronto. Ciò nonostante, la diversa tipologia dei materiali linguistici impiegati in questi studi non garantisce che il minore grado di posteriorità delle vocali senesi non debba ascriversi proprio a tale variabile. Generalmente, gli studi uditivi sul vocalismo si basano su campioni di parole isolate o, in alternativa, su vocali sintetiche¹⁵. Il presente studio è il solo, tra quelli compiuti sulla lingua italiana, ad applicare i criteri di classificazione uditiva in bark a vocali estratte da brani di parlato connesso, due dei quali realizzati mediante una 'pseudolettura' (cfr. § 3). Nel nostro campione si ravvisano per questo fenomeno di coarticolazione, allungamenti vocalici, ritmi elocutivi variabili e, per quanto concerne Fo, condizionamenti reciproci tra fattori di micro e di macro prosodia.

L'analisi di un campione di parlato continuo potrebbe aver oscurato parte delle relazioni formantiche, nella fattispecie il rapporto F2-F1, rinvenute in altre ricerche. Sul versante anteriore, la dimensione F1-Fo, interpretando correttamente la posizione articolatoria dei timbri senesi, non sembra subire il condizionamento del materiale linguistico considerato. L'inadeguatezza metodologica, è quindi da riconoscere al solo parametro F2-F1, poiché non differenzia i due assi articolatori del triangolo vocalico. D'altra parte, già Syrdal (1985) e Syrdal e Gopal (1986) osservarono come nell'anglo-americano, a differenza dello svedese (cfr. Fant 1983), la distanza F2-F1 non riuscisse a distinguere fonemi anteriori e fonemi posteriori. Tale evidenza empirica portò gli autori a sostituire la coordinata F2-F1 con il parametro F3-F2, risultato più adeguato allo scopo, non senza aggiungere la seguente

¹³ Il raffronto effettuato, in special modo quello con il fiorentino, è da assumersi con una certa cautela a causa della diversità dei protocolli sperimentali impiegati.

¹⁴ Si veda anche quanto rilevato sull'argomento da Calamai (1999-2000) nella esaustiva ricerca fonetico-acustica incentrata sul vocalismo tonico di Livorno e di Pisa.

¹⁵ Tutti gli studi italiani realizzati in questa direzione, fanno uso di parole isolate (cfr. Uguzzoni 1988, 1994, Così *et alii* 1995, Ferrero *et alii* 1995, Ferrero *et alii* 1996). L'uso di parole isolate e/o di stimoli sintetici si riscontra anche nelle ricerche condotte su altre lingue, ad esempio Chistovich *et alii* (1979) per il russo, Traunmüller (1981) per il bavarese, Fant (1983) per lo svedese, Di Benedetto (1991) per l'anglo-americano.

raccomandazione¹⁶: “While the auditory Bark-difference dimensions are useful in relating acoustic and phonetic features, one must exercise caution in claiming universality between the auditory dimensions and specific phonetic categories” (Syrdal e Gopal 1986: 128)¹⁷.

L'applicazione di questo ulteriore parametro di normalizzazione uditiva (F3-F2), insieme allo svolgimento di accurati test di percezione, costituiscono due obiettivi fondamentali che intendiamo perseguire nel prossimo futuro, nel tentativo di riuscire ad assegnare, nello studio delle vocali senesi, il giusto peso interpretativo alla variabile diatopica e a quella metodologica.

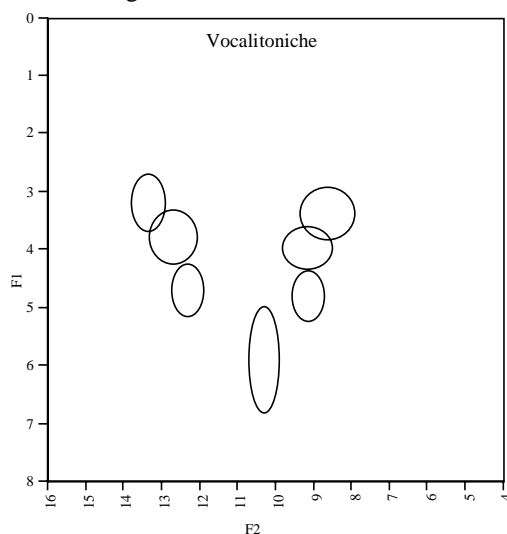


Fig. 1: diagramma F1/F2 in bark delle aree di esistenza delle vocali toniche senesi.

Timbri vocalici	Fo Hz	F1 Hz	F2 Hz	ms
/i/	140 (20)	326 (39)	2089 (165)	80 (21)
/e/	130 (7)	384 (40)	1910 (159)	94 (23)
/ɛ/	130 (17)	479 (52)	1794 (122)	95 (25)
/a/	121 (13)	652 (87)	1323 (81)	115 (32)
/ɔ/	134 (20)	488 (48)	1086 (94)	97 (26)
/o/	133 (20)	420 (57)	1097 (116)	97 (27)
/u/	133 (22)	341 (35)	1013 (142)	77 (29)

Tab. 1: valori medi e deviazioni standard in Hz relativi a Fo, F1, F2 e durata temporale in ms delle vocali toniche senesi.

¹⁶ Dall'applicazione del rapporto formantico F3-F2 si evince che le vocali anteriori sono caratterizzate da valori inferiori alla soglia critica, mentre le vocali posteriori da valori superiori. Nel tentativo di classificare le vocali dell'anglo-americano, Syrdal e Gopal (1986) valutarono anche la bontà delle distanze F4-F3 e F4-F2.

¹⁷ E' utile ricordare che anche nello studio di Ferrero *et alii* (1995) la coordinata F3-F2 mostrò una maggiore capacità differenziante rispetto a F2-F1.

Vocali	F1 bark	F2 bark	F1-F0 Bark	F2-F1 Bark
/i/	3.2 (0.5)	13.3 (0.5)	1.8 (0.4)	10.1 (0.9)
/e/	3.8 (0.5)	12.7 (0.7)	2.6 (0.4)	8.9 (0.9)
/ɛ/	4.7 (0.5)	12.3 (0.5)	3.5 (0.5)	7.6 (1)
/a/	5.9 (1)	10.3 (0.5)	5.2 (0.8)	4.4 (0.7)
/ɔ/	4.8 (0.5)	9.1 (0.5)	3.5 (0.5)	4.3 (0.8)
/o/	4.2 (0.6)	9.1 (0.7)	2.9 (0.5)	4.9 (1)
/u/	3.4 (0.5)	8.6 (0.8)	2.0 (0.5)	5.2 (1.2)

Tab. 2: valori medi e deviazione standard in bark relativi a F1, F2 e alla distanza F1-Fo e F2-F1 delle vocali toniche sinesi.

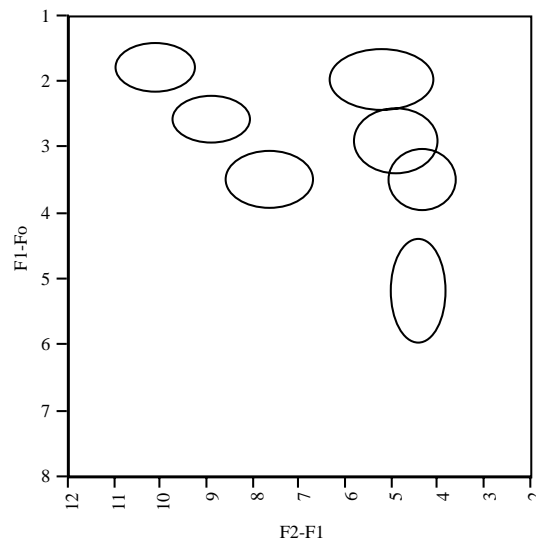


Fig. 2: Diaframma F1-Fo e F2- F1 in bark delle aree di esistenza delle vocali toniche.

	/i/	/u/	/e/	/o/	/ɛ/	/ɔ/	/a/
F1-Fo	1.8	2.0	2.6	2.9	3.5	3.5	5.2
+ 3.5	-	-	-	-	-	+	+
- 3.5	+	+	+	+	+	+	-

Tab. 3: valori in bark F1-Fo delle vocali toniche posti in ordine crescente e relativa specificazione positiva (+) o negativa (-) rispetto al valore della soglia critica.

	/ɔ/	/a/	/o/	/u/	/ɛ/	/e/	/i/
F2-F1	4.3	4.4	4.9	5.2	7.6	8.9	10.1
+ 3.5	+	+	+	+	+	+	+
- 3.5	-	-	-	-	-	-	-

Tab 4: valori in bark F2-F1 delle vocali toniche posti in ordine crescente e relativa specificazione positiva (+) o negativa (-) rispetto al valore della soglia critica.

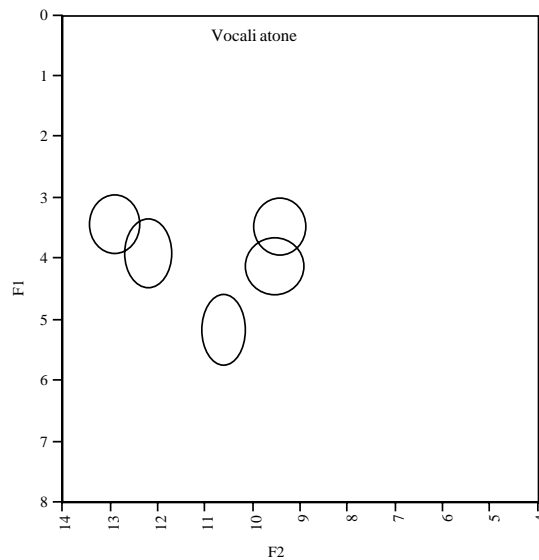


Fig. 3: diagramma F1/F2 in bark delle aree di esistenza delle vocali atone senesi

Timbri vocalici	Fo Hz	F1 Hz	F2 Hz	Durata ms
/i/	126 (25)	339 (39)	1947 (39)	51 (15)
/e/	128 (29)	395 (49)	1809 (147)	55 (46)
/a/	125 (22)	542 (69)	1407 (96)	64 (21)
/o/	124 (25)	422 (46)	1172 (144)	56 (18)
/u/	125 (30)	351 (40)	1130(124)	45 (17)

Tab. 5: valori medi e deviazioni standard in Hz relativi a F0, F1, F2 e valori di durata temporale in ms delle vocali atone.

Vocali	F1 bark	F2 bark	F1-Fo Bark	F2-F1 Bark
/i/	3.3 (0.5)	12.9 (0.7)	2.3 (0.4)	9.6 (1)
/e/	3.9 (0.6)	12.2 (0.6)	3.0 (0.7)	8.3 (1.1)
/a/	5.2 (0.6)	10.6 (0.5)	3.9 (0.4)	5.4 (0.8)
/o/	4.1 (0.5)	9.5 (0.7)	3.3 (0.5)	5.4 (1.2)
/u/	3.5 (0.5)	9.4 (0.6)	2.4 (0.5)	5.9 (1)

Tab. 6: valori medi e deviazioni standard in bark relativi ai valori di F1, F2, e alla distanza F1-Fo e F2-F1 delle vocali atone.

	/i/	/u/	/e/	/o/	/a/
F1-Fo	2.3	2.4	3.0	3.3	3.9
+ 3.5	-	-	-	-	+
- 3.5	+	+	+	+	-

Tab. 7: valori in bark F1-Fo delle vocali atone posti in ordine crescente e relativa specificazione positiva (+) o negativa (-) rispetto al valore della soglia critica.

	/a/	/o/	/u/	/e/	/i/
F2-F1	5.4	5.4	5.9	8.3	9.6
+ 3.5	+	+	+	+	+
- 3.5	-	-	-	-	-

Tab. 8: valori in bark F2-F1 delle vocali atone posti in ordine crescente e relativa specificazione positiva (+) o negativa (-) rispetto al valore della soglia critica.

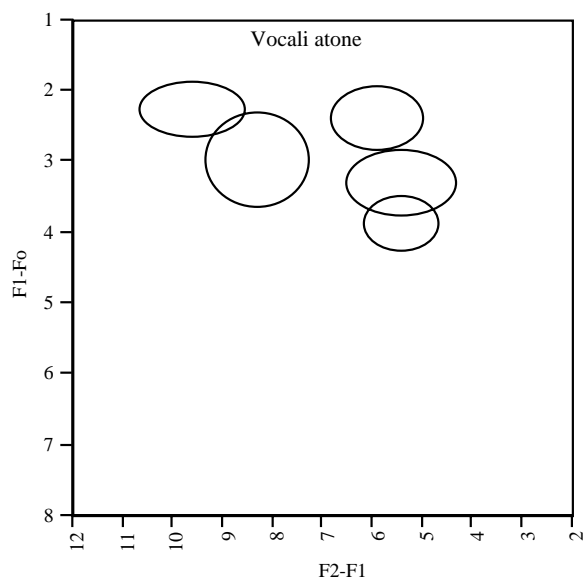


Fig. 4: diagramma F1-Fo e F2-F1 in bark delle aree di esistenza delle vocali atone sinesi.

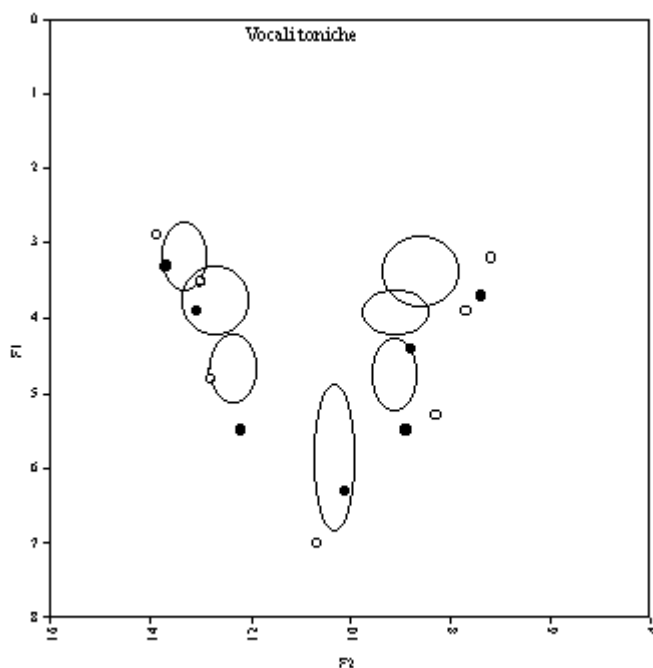


Fig. 5: proiezione dei valori medi in bark di F1 e di F2 delle vocali toniche del fiorentino (indicatori bianchi) e del livornese (indicatori neri).

Bibliografia:

- Calamai S., *Il vocalismo tonico dell'area pisana e livornese. Aspetti storici, percettivi, acustici*, Tesi di Dottorato, Voll. 2, 1999-2000, Università degli Studi di Perugia.
- Canepari L., *Italiano standard e pronunce regionali*, 3° ediz., Padova, 1986, Cleup.
- Carlson R., Fant G. & Granstrom, Two-formant models, pitch and vowel perception, in Fant G. & M.A. Tatham (eds), *Auditory analysis and perception of speech*, New York and London 1975, Academic Press, pp. 55-82.
- Chistovich L. A., Sheikin R.L. & V.V. Lublinskaya, 'Centers of gravity' and spectral peaks as determinants of vowel quality, in Lindblom B. & S. Ohman (eds), *Frontiers of speech communication research*, London 1979, Academic Press, pp. 143-157.
- Cosi P., Ferrero F. & K. Vaggies, Rappresentazioni acustiche e uditive delle vocali italiane, *Atti del XXIII Convegno Nazionale dell'Associazione Italiana di Acustica*, Bologna, 1995, pp. 151-156.
- Di Benedetto M. G., Complex relation between F1 and Fo in determining vowel height: acoustic and perceptual evidence, *Studi Italiani di Linguistica Teorica e Applicata* 1991, pp 579-603.

- Di Benedetto M.G., Acoustic and perceptual evidence between F1 and F0 in determining vowel height, *Journal of Phonetics* 22, 1994, pp. 205-224.
- Fant G., *Speech sounds and features*, Cambridge (Ma), 1973, Mit Press.
- Fant G., Feature analysis of Swedish vowels. A revisit, *S.T.L.-Q.P.S.R.*, 2/3, 1983, pp. 1-19.
- Fechner G., *Elements of psychophysics*, New York, 1966.
- Ferrero F., Riflessione sui “diagrammi di esistenza delle vocali italiane” dopo 25 anni: evoluzione delle ricerche e prospettive, *Atti delle 3° Giornate del Gruppo di Fonetica Sperimentale Padova* 1992, Unipress 1994, pp. 9-25.
- Ferrero F. E., Magno Caldognetto E., Vagges K. E C. Lavagnoli, Some acoustic characteristics of the Italian vowels, *Journal of Italian Linguistics* 3, 1978, pp. 87-96.
- Ferrero F., Magno Caldognetto E. & P. Cosi, Le vocali al femminile, *Atti del Convegno internazionale di Studi Dialettologia al femminile*, Sappada (BL), Padova 1995, Cleup, pp. 413-436.
- Ferrero F., Magno Caldognetto E. & P. Cosi, Sui piani formantici acustici e uditivi delle vocali di uomo, donna e bambino, *Atti del XXIV Convegno Nazionale dell’A.I.A.*, Trento, 1996, pp. 169-178.
- Koenig W., A new frequency scale for acoustic measurements, *Bell Telephone Lab. Record* 27, 1949, pp. 299-301.
- Ladefoged P., *A course in phonetics*, New York 1975, Harcourt Brace Jovanovich.
- Schroeder M.R., Atal B.S. & J.L. Hall, Objective measure of certain speech signal degradations based on masking properties of human auditory perception, in Lindblom B. & S. Ohman (eds), *Frontiers of speech communication research*, London 1979, Academic Press, pp. 217-229.
- Stevens S. S., *Psychophysics*, J. Wiley & Son, London, 1975.
- Sorianello P., Il vocalismo dell’italiano senese: un’indagine sperimentale, *Atti delle XII° Giornate di Studio del G.F.S., La fonetica acustica come strumento di analisi della variazione linguistica in Italia*, Università degli Studi di Macerata, Macerata 14-17 dicembre 2001, in stampa.
- Sorianello P., Un tentativo di classificazione uditiva delle vocali senesi, in Cosi P. Magno Caldognetto E. & A. Zamboni (acd), *Scritti in onore di Franco Ferrero*, Padova, Unipress., in stampa.
- Syrdal A. K., Aspects of a model of the auditory representation of American English, *Speech Communication* 4, 1985, pp. 121-135.
- Syrdal A. K. & H. S. Gopal, A perceptual model of vowel recognition based on the auditory representation of American English vowels, *Journal of Acoustic Society of America*, 79, 1986, pp. 1086-1100.
- Trautmüller H., Perceptual dimension of openness in vowels, *Journal of Acoustic Society of America* 69, 1981, pp. 1086-1100
- Trautmüller H., Analytical expression for the tonotopic sensory scale, *Journal of Acoustic Society of America* 88, 1990, pp. 97-100.

- Uguzzoni A., Verso un'analisi parametrica del vocalismo di un dialetto frignanese, *Italia Dialettale*, 25, 1988, pp. 86-110
- Uguzzoni A., Rappresentazioni e classificazioni delle vocali in termini uditivi, *Atti delle 3° Giornate del Gruppo di Fonetica Sperimentale Padova* 1992, Unipress 1994, pp. 27-35.
- Zwicker E., Subdivision of the audible frequency range into critical bands, *Journal of Acoustic Society of America*, 33, 1961, pp. 248.
- Zwicker E., Flottrop G. & S.S. Stevens, Critical bandwidth in loudness summation, *Journal of Acoustic Society of America* 29, 1957, pp. 458-557.
- Zwicker E. & E. Terhardt, Analytical espressions for critical band rate and critical bandwidth as a function of frequency, *Journal of Acoustic Society of America* 68, 1980, pp. 1523-1525.