

# I VALORI DI H1-A2 E H1-A3 COME CORRELATI DELLA INTENSITÀ “RIVISITATA”. ASPETTI E PROBLEMI.

Arianna Uguzzoni  
Alma Mater Studiorum - Università di Bologna  
[arianna.uguzzoni@unibo.it](mailto:arianna.uguzzoni@unibo.it)

## 1. SOMMARIO

Negli ultimi anni gli studiosi hanno guardato con occhi nuovi alla intensità sia come grandezza fisica sia come dimensione percettiva. Questa “rivisitazione” è di indubbio interesse, da un lato per le implicazioni di carattere generale, dall’altro per le ricadute positive su molteplici settori di indagine.

Il *new look* deriva dallo spostamento dell’attenzione dall’intensità dell’intero spettro alla distribuzione dell’intensità in differenti parti dello spettro. Ricerche condotte su lingue diverse e su fatti di varia natura hanno trovato una caratteristica comune. Le differenze di intensità tra coppie di fenomeni contrapposti si manifestano prevalentemente nelle parti dello spettro che sono al di sopra di 500 Hz. Ne consegue una variazione nel rapporto tra la zona bassa e le zone superiori della scala delle frequenze, e a ciò viene dato il nome di *spectral balance*. I fenomeni di volta in volta sotto esame sono pertanto assegnati a due categorie distinte che vengono caratterizzate o in termini di *mid-to-high-frequency emphasis* o in termini di *spectral balance*.

Uno dei modi per quantificare le differenze di intensità che si riscontrano nelle bande medie e alte consiste nel sottrarre ai valori dell’intensità della prima armonica (H1), da una parte, i valori dell’intensità della seconda formante (A2), dall’altra, quelli dell’intensità della terza formante (A3): si ottengono così i parametri H1-A2 e H1-A3. Risulta per esempio che le vocali accentate presentano valori di H1-A2 e H1-A3 che sono più piccoli rispetto a quelli delle non accentate in olandese, in inglese americano, in tedesco. Ciò equivale a dire che le prime, confrontate con le seconde, sono caratterizzate dalla presenza di maggiore energia alle medie e alte frequenze.

Oggi si sostiene che questa enfasi spettrale (*mid-to-high-frequency emphasis*) è un affidabile correlato acustico e percettivo dell’accento lessicale (*stress*) nelle suddette lingue: un correlato che è più robusto dell’intensità globale (*overall intensity*), che usualmente si misurava e si studiava prima della svolta metodologica di cui stiamo parlando. La diversa entità dei parametri H1-A2 e H1-A3, è una manifestazione acustica riscontrata anche in altri fenomeni e svolge un ruolo nel segnalare distinzioni come le seguenti: sillabe focali *vs.* non focali, vocali rilassate *vs.* tese, vocali brevi *vs.* lunghe.

Nel presente articolo si fa il tentativo di mettere in connessione i dati acustici con i meccanismi pneumo-fono-articolatori che vi sono sottesi. Valori relativamente bassi di H1-A2 e H1-A3 caratterizzano le vocali accentate, le sillabe focali, le vocali rilassate, le vocali brevi, rispetto alle loro controparti. I fatti fisiologici responsabili di tali differenze fisiche non sono gli stessi in tutti i casi: pertanto si sente il bisogno di ulteriori studi che mettano meglio in luce il contributo di stati e attività a tre livelli: sottoglottidale, glottidale, sopraglottidale.

## 2. PREMESSA

Questo articolo è sostanzialmente una rassegna critica di studi e ricerche che ruotano intorno al tema della intensità. Sono affrontati aspetti e problemi riguardanti l’accento in

olandese, l'accento in inglese americano, l'accento in tedesco, il *focus* in svedese, la distinzione teso/rilassato in tedesco, la distinzione *abrupt cut/smooth cut* in tedesco, la distinzione breve/ lungo in una varietà italo-romanza. Per quest'ultima si presentano sia ricerche precedenti sia ricerche in corso sulla "quantità vocalica" in area emiliana, che è un fenomeno plurifacciale e pluriparametrico. Una faccia è costituita dalle differenze di intensità che sembrano contribuire alla distinzione tra vocali accentate brevi e vocali accentate lunghe.

### 3. PANORAMA DELLE RICERCHE

#### 3.1 Sull'accento in olandese

Sluijter & van Heuven (1996) nelle indagini sperimentali sull'accento linguistico in olandese hanno tenuti separati gli effetti prodotti sull'intensità dall'accento lessicale (*stress*) da quelli prodotti dall'accento intonativo collegato al *focus* (*accent*), essendo convinti che si tratta di due dimensioni distinte (v. Uguzzoni, 2003; Tamburini, 2005). La dimensione *stress* è una "structural, linguistic property of a word that specifies which syllable in the word is, in some sense, stronger than any of the others" e le sillabe *stressed* hanno "an accent-lending pitch movement associated with them when they occur within a single word in a narrow focus" (Sluijter & van Heuven, 1996: 2471). Con questa impostazione i due studiosi hanno trovato che la differenza di intensità globale tra le vocali accentate e le vocali non accentate è lieve in condizione non focale, mentre è consistente in condizione focale. L'intensità globale non è dunque un correlato acustico affidabile del fenomeno *stress*, ma del fenomeno *accent*. D'altra parte le misure dell'intensità in quattro bande di frequenza contigue (0-500 Hz, 500-1000 Hz, 1000-2000 Hz, 2000-4000 Hz) hanno mostrato che l'effetto esercitato sull'intensità dall'accento lessicale è trascurabile nella banda bassa (0-500 Hz), mentre è considerevole nelle bande superiori (500-4000 Hz). Comparando vocali accentate e vocali non accentate si vede chiaramente che nella regione bassa dello spettro non c'è differenza di intensità, mentre nelle regioni superiori dello spettro le accentate presentano un aumento di intensità di 5-10 dB rispetto alle loro controparti non accentate.

Da un punto di vista generale è di grande interesse cercare di connettere in modo plausibile i dati acustici con i meccanismi pneumo-fono-articolatori della produzione. Nel caso di questa "difference in spectral balance between stressed and unstressed vowels, stressed vowels having more high-frequency emphasis than unstressed vowels" (Sluijter & van Heuven, 1996: 2483) è necessario domandarsi quali ne possano essere le cause. In sintonia con Glave & Rietveld (1975), Sluijter & van Heuven (1996) sostengono che il ruolo principale è svolto da fattori che si collocano a livello sottoglottidale e a livello glottidale. L'incremento dell'intensità del segnale che si constata nelle regioni aventi frequenze superiori a 500 Hz viene fatta dipendere dal più grande sforzo fisiologico (*vocal effort*) con cui sono prodotte le vocali accentate rispetto alle vocali non accentate. A ciò si collegano importanti caratteristiche riguardanti la configurazione della glottide e delle corde vocali durante la fonazione.

Avendo in mente un destinatario che, come me, non si è mai occupato esplicitamente di questa tematica, ritengo utile riassumere qui alcuni elementi di base. Un aspetto cruciale è la modalità di chiusura delle corde vocali, che influisce fortemente sulla forma d'onda e sullo spettro della sorgente sonora. Al variare della velocità con cui avviene l'evento di chiusura varia l'inclinazione dell'involuppo dello spettro. Se la chiusura è meno veloce l'inclinazione spettrale è ripida, se invece essa è più veloce l'inclinazione spettrale è

graduale. Con “inclinazione spettrale” intendo riferirmi al *negative spectral slope* e al *negative spectral tilt* della bibliografia originaria. Ciò che risulta modificato comparando i due casi è la distribuzione dell’energia lungo la scala delle frequenze: le armoniche situate nelle zone superiori dello spettro hanno un’intensità minore nel primo caso e maggiore nel secondo caso. Riassumendo, più veloce è la chiusura delle corde vocali, più graduale è l’inclinazione spettrale, più grande è l’ammontare dell’energia alle medie e alte frequenze. Aumenta così la forza e la efficacia della sorgente sonora che fa da *input* all’attività del condotto sovraglottidale (v. Stevens, 1977; Fant, 1985; Holmberg *et al.*, 1988; Gauffin & Sundberg, 1989; Childers & Lee, 1991; Holmberg *et al.*, 1995; Stevens, 1998).

Le conclusioni della ricerca sull’olandese sono due. Da una parte, le differenze d’intensità rilevate nell’analisi degli spettri acustici di vocali accentate *vs.* non accentate (v. Sluijter & van Heuven, 1996: 2481, fig. 5), sono da attribuire principalmente all’influenza di differenti caratteristiche della sorgente sonora, che a loro volta dipendono dal diverso grado di sforzo fisiologico implicato nella produzione di vocali con accento lessicale *vs.* vocali senza accento lessicale. D’altra parte, sembra invece trascurabile l’influsso esercitato sulla *spectral balance* da differenti valori delle frequenze formantiche connessi con la presenza *vs.* assenza di tale accento.

Sia nell’articolo di Sluijter & van Heuven (1996) sia in quello di Sluijter *et al.* (1997) vengono riabilitate le posizioni più tradizionali che classificavano come “dinamico” o “espiratorio” l’accento di alcune lingue germaniche e indicavano come suo correlato fonetico primario la maggiore intensità percepita (*loudness*). Il ruolo dell’intensità nella realizzazione e nella percezione dell’accento lessicale riacquista l’importanza che per un po’ di tempo era stata messa in dubbio (v. Uguzzoni, 2003). E ciò è stato reso possibile da quella “rivisitazione” che consiste essenzialmente nel porre in primo piano non tanto l’intensità dell’intero spettro quanto l’intensità relativa delle parti superiori dello spettro: “*a stressed syllable might be perceived as louder, and therefore more prominent, than unstressed one due to the increased intensity levels in the higher part of the spectrum*” (Sluijter & van Heuven, 1996: 2483).

### 3.2 Sull’accento in inglese americano

Una ricerca di Sluijter *et al.* (1995) dimostra che una maggiore enfasi alle medie e alte frequenze è un robusto correlato acustico e percettivo dell’accento lessicale (*stress*) in inglese americano. Sono esaminate vocali *stressed* e *unstressed* sia in *focus* (cioè con un *pitch accent*) sia fuori *focus* (cioè senza un *pitch accent*). Risulta che anche in condizione non focale le vocali accentate si distinguono dalle loro controparti non accentate soprattutto per una diversa distribuzione dell’energia nelle parti superiori dello spettro (quantificata come H1-A2 e H1-A3). Secondo l’interpretazione proposta tale caratteristica è dovuta a differenze nella forma dell’impulso glottidale piuttosto che a differenze nella configurazione delle cavità sovraglottidali. Il fenomeno *accent*, ma non il fenomeno *stress*, è addizionalmente caratterizzato da un lieve aumento del quoziente di apertura glottidale (stimato in base alla differenza di intensità tra le prime due armoniche: H1-H2) e da un aumento dell’energia dell’impulso glottidale (stimato in base alla misura dell’intensità della prima armonica: H1).

Questo lavoro, esposto in modo stringato nello spazio delle canoniche quattro pagine congressuali, solleva a mio parere una domanda di carattere generale. Con quale metodo sono enucleati i fattori fisiologici coinvolti nel processo della produzione? L’impostazione è esplicitamente espressa in Stevens & Hanson (1995), Hanson (1997), Hanson & Chuang (1999), Hanson *et al.* (2001), che hanno sviluppato una metodologia specifica per inferire le caratteristiche glottidali da misurazione acustiche fatte sul suono radiato.

Tornando all'articolo sull'inglese americano, è molto interessante il quadro riassuntivo delle corrispondenze tra dimensioni fisiologiche e parametri acustici: “*an overview of the physiological dimensions in which glottal pulses of stressed and unstressed syllables can differ, and the acoustic parameters from which these differences can be derived*” (Sluijter *et al.*, 1995: 632). Per fare un esempio, l'informazione sulla velocità della chiusura delle corde vocali e sulla asimmetria dell'impulso glottidale viene ricavata da H1-A2 e da H1-A3, cioè sottraendo ai valori dell'intensità della prima armonica (H1) i valori dell'intensità della seconda formante (A2) e della terza formante (A3). L'entità sia di H1-A2 sia di H1-A3 risulta minore nelle vocali accentate rispetto alle vocali non accentate corrispondenti (v. Sluijter *et al.*, 1995: 633, fig. 1).

Da questo dato acustico del segnale emesso si inferisce che le suddette caratteristiche glottidali, *rate of closure* e *skewness of glottal pulse*, differiscono a seconda che le vocali prodotte siano accentate o non accentate. Si ipotizza che il movimento di chiusura della glottide nel caso delle vocali accentate sia più veloce e influisca direttamente sulla asimmetria dell'impulso glottidale. In conseguenza di tale assetto si ha nel campo spettrale un rafforzamento delle armoniche situate nelle zone medie e alte.

Bisogna precisare che nella tabella citata i valori di H1, H2, A2, A3 sono contrassegnati da un asterisco in esponente. Le misure asteriscate H1\*, H2\*, A2\*, A3\* derivano da correzioni fatte allo scopo di eliminare gli influssi di F1 su H1 e H2, di F1 su A2, di F1 e F2 su A3. Gli aspetti teorici e tecnici di queste normalizzazioni sono trattati in Stevens & Hanson (1995), Hanson (1997), Jessen & Marasek (1997), Jessen (2002), Jessen & Roux (2002). Si arriva così alle misure asteriscate H1\*-A2\* e H1\*-A3\* che costituiscono il mezzo per quantificare l'enfasi spettrale alle medie e alte frequenze, considerata oggi come un affidabile correlato dell'accento lessicale in olandese, in inglese, in tedesco.

### 3.3. Sull'accento in tedesco

Studiosi dell'Università di Stuttgart hanno lavorato molto sulla lingua tedesca. Anche in questo caso è emersa la necessità di passare dalla misurazione dell'energia globale (*overall signal amplitude*) alla comparazione dei valori dell'energia presente in differenti zone della scala delle frequenze (*frequency-sensitive amplitude measurements*). Il primo tipo di analisi non consente infatti di rilevare che l'impatto dell'accento lessicale sull'intensità è molto più forte nelle zone superiori dello spettro che nella sua zona bassa. La procedura adottata consiste nel mettere in relazione l'intensità della seconda (A2) e della terza (A3) formante con l'intensità della prima armonica (H1). Si sottraggono ai valori di H1 da una parte i valori di A2 dall'altra i valori di A3 e si ottengono i parametri acustici H1-A2 e H1-A3 (*H1-based parameters*).

Questi sono spesso sostituiti da H1\*-A2\* e H1\*-A3\*, dopo che sono state apportate alcune correzioni ai singoli valori dell'intensità. Dato che le misurazioni vengono fatte sullo spettro del segnale emesso, si mira a raggiungere una migliore stima delle caratteristiche della sorgente sonora eliminando gli influssi delle frequenze formantiche. Con le parole di Hanson *et al.* (2001), queste correzioni “*are based on the acoustic theory of speech production*” e sono “*necessary to allow comparison of the measures across vowels and speakers*” (Hanson *et al.*, 2001: 459).

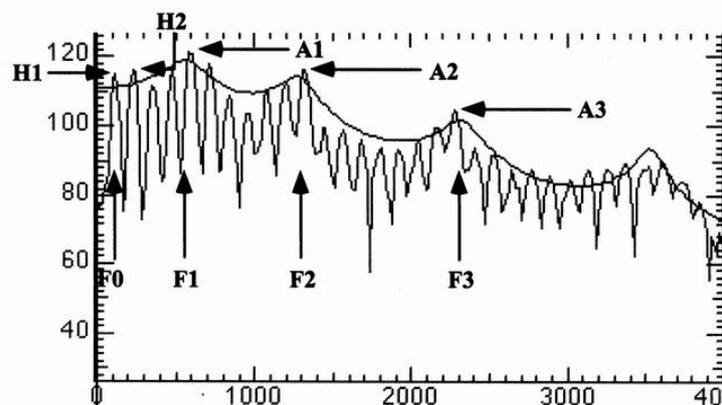


Figura 1: Misure acustiche (da Jessen & Marasek, 1997: 128).

Gli studi di carattere acustico sull'accento lessicale in tedesco (*Wortbetonung*) sono esposti in modo specifico in Jessen & Marasek (1997), Claßen *et al.* (1998), Schneider (2001). Per ogni vocale del corpus sono creati uno spettro DFT e uno spettro LPC. Nella fig. 1, dove gli spettri sono sovrapposti l'uno all'altro, le frecce verticali e le frecce orizzontali fanno vedere, rispettivamente, le misurazioni delle frequenze (in Hz) e le misurazioni dell'intensità (in dB).

Come si è già accennato, con H1, H2, A1, A2, A3 si indicano le intensità delle prime due armoniche e delle prime tre formanti. A questo riguardo è utile ricordare che si opera sullo spettro DFT misurando, da un lato (per H1, H2), l'ampiezza della prima e della seconda armonica, dall'altro (per A1, A2, A3), l'ampiezza dell'armonica più forte che si trova ad essere più vicina in frequenza al rispettivo picco formantico (F1, F2, F3).

I risultati delle ricerche sul tedesco sono analoghi a quelli raggiunti per l'olandese e per l'inglese americano. Le vocali accentate presentano valori di  $H1^*-A2^*$  e  $H1^*-A3^*$  che sono più piccoli rispetto a quelli delle non accentate, come si vede per esempio nella figura in cui sono riportati gli spettri di una [a] accentata e di una [a] non accentata (Claßen *et al.*, 1998: 222, fig. 4). In modo sintetico si può dire che le vocali in "*betonten Silben mit einem geringer Spectral Tilt, daß heißt mit einem grösseren Energieanteil im Bereich der mittleren und hohen Frequenzen, realisiert werden als Vokale in unbetonten Silben*" (Claßen *et al.*, 1998: 227). Per rendere più chiaro il confronto tra gli studi sull'accento lessicale nelle tre lingue germaniche mi sembra conveniente passare sopra alle divergenze di terminologia (*spectral balance*, *Spectral Tilt*) e mettere invece in primo piano il loro comune denominatore: la *mid-to-high-frequency emphasis* e gli aspetti della qualità della voce (*voice quality*) di cui la suddetta enfasi spettrale è una conseguenza acustica.

Attraverso quali mezzi di indagine gli studiosi tedeschi cercano di scoprire gli eventi che si verificano sul piano sottoglottidale e sul piano glottidale quando il parlante produce vocali accentate e vocali non accentate? A grandi linee: da una parte viene adottato il metodo che consiste nell'inferire tali eventi dalle proprietà acustiche rilevate nell'analisi del segnale verbale (in sintonia con Sluijter, 1995; Stevens & Hanson, 1995; Sluijter *et al.*, 1995; Sluijter & van Heuven, 1996). Dall'altra si fa ricorso alla tecnica elettroglottografica che fornisce informazioni attendibili sulla qualità del ciclo vibratorio delle corde vocali.

I valori di  $H1^*-A2^*$  e di  $H1^*-A3^*$ , che in Sluijter (1995) e in Sluijter *et al.* (1995) sono presentati e discussi come un gruppo senza un esplicito ordine interno, in Claßen *et al.*

(1998) sono invece separati. Precisamente, il valore di H1\*-A2\* è collegato con la asimmetria dell'impulso glottidale (*skewness of glottal pulse* (SK)), mentre il valore di H1\*-A3\* è collegato con la velocità della chiusura glottidale (*rate of closure* (RC)) (Claßen *et al.*, 1998: 210, tabella 1). Questa connessione più diretta tra manifestazioni acustiche colte nel segnale verbale e dimensioni della fisiologia della vibrazione delle corde vocali si appoggia sui risultati di studi elettroglottografici (v. Marasek, 1995; Marasek, 1996; Marasek, 1997; Jessen & Marasek, 1997; Jessen, 2002).

Dello stesso materiale sono stati registrati simultaneamente il segnale acustico e il segnale EGG. Le misurazioni fatte sull'uno e sull'altro sembrano corroborarsi a vicenda. In questa sede accenno solo al fatto che ai parametri acustici H1\*-A2\* (SK) e H1\*-A3\* (RC) corrispondono questi parametri EGG: la pendenza della chiusura (*closing slope*) e la durata della fase di chiusura (*duration of closing*) (v. Jessen & Marasek, 1977: 127 e Jessen, 2002: 167 per le variabili EGG). Nelle vocali accentate, comparate con le non accentate, i valori più piccoli di H1\*-A2\* e di H1\*-A3\* si associano, rispettivamente, ad una maggiore ripidità della pendenza ed a una minore durata della fase di chiusura del ciclo EGG.

Anche per il tedesco si sostiene che le differenze glottidali emerse dipendono dal grado di sforzo fisiologico (*vocal effort*) e quindi dalla entità della pressione sottoglottidale (*subglottal pressure*), che sono più grandi quando si producono vocali accentate. Le conclusioni delle ricerche sull'accento tedesco si allineano con quelle raggiunte nei casi dell'olandese e dell'inglese americano: "*betonte Silben mit einem erhöhten subglottal druck produziert werden, der sich akustisch in einer Erhöhung des Energieanteils in den mittleren und hohen Frequenzen (d.h. in verminderten SK- und RC-Werten) äußert*" (Claßen *et al.*, 1998: 230). E ciò consente di vedere in nuova luce il concetto tradizionale di *expiratorische Verstärkung* che era considerato un correlato fisiologico dell'accento in alcune lingue germaniche. Secondo la nuova prospettiva che si è delineata, la manifestazione acustica di un maggior sforzo fisiologico non è un aumento dell'intensità globale ma un aumento dell'intensità che si concentra solo nelle parti medie e alte dello spettro.

#### 3.4 *Sul focus in svedese*

Fant e i suoi colleghi hanno fatto numerosi studi sulla prominente nella lingua svedese (Fant & Kruckenberg, 1989, 1994, 1995; Fant, 1997; Fant *et al.* 1999, 2000a, 2000b), esaminando non solo la frequenza fondamentale, la durata, l'intensità, ma anche parametri connessi con il sistema respiratorio, la sorgente glottidale, la composizione segmentale, cioè l'intero processo della produzione (Fant *et al.*, 2000a: 55).

Dicotomizzando alla maniera tradizionale si parla solitamente di *stressed vs unstressed*, *accented vs unaccented*, *focal vs non-focal*. La proposta innovativa consiste nell'introdurre nell'analisi una scala continua per la misurazione della prominente percepita: la scala chiamata Rs. Una corrispondenza approssimativa tra la scala soggettiva Rs e le categorie prosodiche consuete è presentata in una tabella che specifica cinque livelli di prominente, senza una separazione acustica netta e con possibilità di sovrapposizione di valori. Nelle regioni più basse si situano le sillabe *unstressed* (Rs < 12) e le sillabe che sono *stressed* ma *unaccented* (Rs = 12-15). Nelle regioni più alte si trovano le sillabe *accented* con *non-focal accentuation* (Rs = 15-20) e le sillabe *accented* con *focal accentuation* (Rs = 20-25); il termine accentuazione implica la presenza di una significativa modulazione di f<sub>0</sub>. Se i gradi sono ancora maggiori (Rs > 25) si tratta di una *extra high prominence* (Fant *et al.* 2000a: 80; Fant *et al.* 2000b: 115).

Uno dei compiti affrontati dagli studiosi svedesi è quello di stabilire relazioni quantitative tra valori Rs e parametri fisici: la frequenza fondamentale, la durata, la

pressione sottoglottidale, l'intensità. Per quel che riguarda l'intensità sono usate due misure. Una è il livello di pressione sonora: SPL; l'altra, chiamata SPLH, è caratterizzata dalla introduzione di una pre-enfasi che cresce al crescere della frequenza: l'aumento è di 3dB a 200 Hz, 14 dB a 1000 Hz, 25 dB a 5000 Hz. La misura SPLH è più sensibile della misura SPL alle variazioni nelle zone della seconda e della terza formante (F2 e F3) (Fant & Kruckenberg, 1995; Fant 1997; Fant *et al.* 2000a, 2000b). La differenza tra SPLH e SPL mette in evidenza il livello spettrale delle formanti superiori a F1, che è associato sia a caratteristiche della sorgente sia a caratteristiche del filtro (cioè alla particolare configurazione formantica): *“at constant articulation, variations in the SPLH-SPL measure accordingly bring out variations in the high-frequency contents of the source which in turn is related to the concept of spectral tilt”* (Fant *et al.*, 2000b: 117).

Si osserverà che il parametro SPLH-SPL è analogo ai parametri H1-A2 e H1-A3 (non normalizzati e normalizzati) di cui si è parlato nelle pagine precedenti: un mezzo per quantificare le differenze di intensità che si manifestano nelle parti dello spettro corrispondenti alla collocazione della seconda e della terza formante (cioè la *mid-to-high-frequency emphasis*). Ma leggendo gli articoli di Fant e colleghi non si può non notare il frequente e forte richiamo a tenere sempre conto delle interazioni complesse che esistono tra la sorgente e il filtro. Riguardo all'aumento del valore di SPLH-SPL, per esempio, si sottolinea che esso in parte è il riflesso acustico di una sorgente sonora più efficiente (per effetto di una maggiore velocità della chiusura glottidale), in parte è dovuto all'influenza di cambiamenti nella configurazione formantica delle vocali. Ritengo di un certo interesse ricordare qui che le dicotomie sorgente/filtro, fonazione/articolazione sono simili ma non identiche, come mostra il fatto che un evento articolatorio può influire sia sul filtro sia sulla sorgente (Fant *et al.* 1985; Fant & Li, 1987; Fant, 1997; Stevens, 1998; Fant *et al.*, 2000a, 2000b). L'influsso di un gesto articolatorio sullo spettro della sorgente non è limitato alle consonanti, ma si constata anche nel caso di vocali: una maggiore costrizione del condotto sovraglottidale porta ad una *“reduction of voice source strength and of its high-frequency content”* (Fant *et al.* 2000b: 125; cf. Bickley & Stevens, 1986; Stevens, 1998: 94 e 297).

I dati di una indagine sulla accentuazione focale in svedese mostrano una chiara covarianza tra i seguenti parametri:  $f_0$ , durata, pressione sottoglottidale, SPL, SPLH-SPL. Mancano elementi per stabilire una graduatoria fra tali parametri e per decretare quale di essi contribuisce di più a percepire una sillaba come focale, anche se sembra che in svedese *“the dominant correlate of focal accentuation is the size of the local  $f_0$  modulation”* (Fant *et al.*, 2000b: 120).

Dallo studio delle interrelazioni tra i vari fattori coinvolti si possono trarre alcune conclusioni degne di nota. La pressione sottoglottidale e la  $f_0$  hanno un ruolo di prim'ordine nel determinare l'intensità globale (SPL). La adduzione delle corde vocali contribuisce a rendere la sorgente sonora più efficiente e quindi a fare crescere sia SPL sia SPLH-SPL, a parità di pressione sottoglottidale. L'accentuazione focale sembra richiedere un gesto attivo che aumenti la forza polmonare. Sull'entità dei valori di SPL e di SPLH-SPL agiscono anche variazioni nell'articolazione delle vocali (cioè nella configurazione formantica). Se SPL aumenta di 1 dB SPLH-SPL di solito aumenta di 0.5 dB. Il parametro SPLH-SPL, che per comodità espositiva si può considerare il pendant della *mid-to-high-frequency emphasis*, porta un indubbio contributo alla segnalazione della differenza tra sillabe focali e sillabe non focali. Se si prende come riferimento la scala Rs, a cui si è accennato, si osserva che passando da  $R_s = 15$  (il grado più debole di accentuazione) a  $R_s = 25$  (un livello di evidente accentuazione focale) si ha un aumento di: 4-8 semitoni in  $f_0$ , 125 ms di durata, 6 dB in SPL, 9 dB in SPLH e quindi 3 dB in SPLH-SPL.

In una recente indagine Heldner (2003) si è occupato espressamente di intensità globale e di enfasi spettrale come correlati acustici del fenomeno chiamato *focal accent*. Secondo la consuetudine degli studi sulla prosodia svedese con il termine *focal accent* si intende “*an accent signalling that a word (or some other constituent within a phrase which may be smaller or larger) is ‘focused’ or ‘in focus’*” (Heldner, 2003: 39). Risulta che le parole focali differiscono dalle parole non focali per un aumento della intensità globale e della enfasi spettrale che è rispettivamente, di circa 3 dB e di circa 2 dB. Gli aumenti in entrambi i parametri si riscontrano in tutte le parole, in tutte le posizioni nella frase, in tutti i soggetti.

In base a questa differenza statisticamente significativa Heldner sostiene che intensità globale e enfasi spettrale sono affidabili correlati acustici del *focal accent* in svedese (cosa che non equivale a dire che siano affidabili correlati percettivi). Quale dei due aspetti dell’intensità è il correlato acustico più affidabile? I dati della ricerca di Heldner sullo svedese centrale inducono ad ipotizzare che l’enfasi spettrale alle medie e alte frequenze sia un correlato migliore rispetto alla intensità globale, giacché è meno influenzata da fattori intraindividuali, come la posizione nella frase e la composizione segmentale delle parole esaminate, e sembra anche poco influenzata da differenze interindividuali.

### 3.5 Sulla distinzione teso/rilassato in tedesco

Gli studi sulla lingua tedesca visti in 3.3 si collocano nel quadro complessivo di una esplorazione dei molteplici correlati acustici dell’accento lessicale, da una parte, e della distinzione teso/rilassato, dall’altra. Mentre alcuni parametri sono comuni, altri “*are largely specific to the expression of either stress or tenseness*” (Jessen *et al.*, 1995: 428). Nell’ambito della presente panoramica è naturale domandarsi se le vocali tese (per esempio [i:] in *Miete*, [o:] in *Schote*) si differenziano dalle vocali rilassate (per esempio [ɪ] in *Mitte*, [ɔ] in *Schotte*) anche in base a proprietà acustiche che si collegano alla qualità della voce (*voice quality, Stimmqualität*).

I risultati per teso *vs.* rilassato si trovano, accanto a quelli per accentato *vs.* non accentato, nelle tabelle di Jessen & Marasek (1997) e di Claßen *et al.* (1998). Le misure acustiche mostrano valori di H1\*-A2\* (SK) e di H1\*-A3\* (RC) più grandi nelle vocali tese rispetto alle vocali rilassate. D’altro canto dalle misure elettroglottografiche risulta che le tese, comparate con le rilassate, sono prodotte con una minore ripidità della pendenza ed una maggiore durata della fase di chiusura del ciclo EGG. L’analisi statistica indica che il parametro H1\*-A3\* (RC) è più significativo del parametro H1\*-A2\* (SK); e ciò è in consonanza con i dati del segnale EGG, dove “*the effect is more consistent and more often significant for duration of closing, than closing slope*” (Jessen & Marasek, 1997: 130). Si può concludere che la distinzione teso/rilassato in tedesco ha tra i suoi correlati anche una differenza nella qualità della voce che si manifesta acusticamente nei valori di H1\*-A2\* e di H1\*-A3\*. Essi sono più elevati nelle vocali tese che nelle corrispondenti vocali rilassate. Questo equivale a dire che le tese rispetto alle rilassate presentano minore energia nel campo delle medie e alte frequenze.

Dal punto di vista percettivo le vocali tese “suonano” meno intense/forti (*loud*) delle vocali rilassate? La risposta è affermativa se si guarda ai risultati emersi dalle indagini che hanno adottato l’impostazione secondo la quale ciò che conta è la diversa distribuzione della energia lungo la scala delle frequenze. La suddetta risposta può meravigliare chi invece si aspetterebbe che le tese siano più intense/forti delle rilassate in base all’idea che la produzione di un suono teso implichi uno sforzo più grande. A questo proposito Jessen sottolinea giustamente che il tratto teso/rilassato non deve essere usato come una semplice metafora fonetica, ma deve essere riempito di contenuto fonetico ricavato da ricerche

sperimentali (Jessen, 2002:157, nota 5).

La distinzione tra vocali tese e vocali rilassate ha molteplici correlati acustici: differenze di durata, differenze formantiche, differenze glottidali; e a seconda delle lingue i meccanismi fisiologici sottostanti a tale distinzione sono diversi. Senza entrare nei particolari di un tema così complesso, mi sembra pertinente ricordare qui ciò che dice Stevens riferendosi a vocali non basse dell'inglese: “*a more breathy or spread glottal configuration for tense vowels would reduce the spectrum amplitudes at high frequencies, whereas a pressed or constricted glottal configuration for lax vowels would enhance the high-frequency spectrum amplitude*” (Stevens, 1998: 297; cfr. Halle & Stevens, 1969). Si ipotizza che modificazioni della sorgente glottidale contribuiscano a esaltare quelle differenze di intensità nelle zone superiori dello spettro che sono indotte dal diverso valore della frequenza della prima formante (: esso è più piccolo nelle tese che nelle rilassate). Vari studiosi del fenomeno, presente in più lingue, concordano nel ritenere che accanto alla influenza di F1 sia in gioco un diverso tipo di fonazione.

Per tornare agli studi sulla lingua tedesca, è evidente dunque che i valori di H1\*-A2\* e di H1\*-A3\* sono fatti acustici che differenziano sia le vocali accentate vs. le non accentate sia le vocali tese vs. le rilassate; ma i fatti fisiologici soggiacenti a tali manifestazioni acustiche non sono gli stessi nei due casi. Mentre nel primo caso si ritiene che il fattore più importante sia l'entità della pressione sottoglottidale, nel secondo caso si invoca come causa o concausa una differente modalità di fonazione. In breve, sembra che le tese e le rilassate siano prodotte, rispettivamente con *breathy or slack voice* e con *modal voice* (Jessen, 2002: 157-159; Ladefoged & Maddieson, 1996; Jessen & Roux, 2002).

### 3.6 Sulla distinzione abrupt cut/smooth cut in tedesco

Non è esagerato dire che la distinzione teso/rilassato nelle vocali tedesche è la più dibattuta nella bibliografia sia fonetica sia fonologica. Si nota, per esempio, che in posizione accentata le vocali tese sono lunghe e le vocali rilassate sono brevi: /i:, y:, e:, ø:, u:, o:, a:/ vs. /ɪ, ʏ, ε, œ, ʊ, ɔ, a/. In un'ottica segmentale i fonologi assegnano il primato ora alla quantità, ora alla quantità oppure sostengono che entrambe le caratteristiche collaborano a distinguere le due classi di vocali riportate sopra, eccetto il caso di /a:/ vs. /a/ dove opera solo la differenza di quantità. Ma la situazione del tedesco standard può essere guardata anche in un'altra angolatura. La questione si riassume nella domanda: si tratta di un fatto segmentale circoscritto alla vocale accentata o di un fatto soprasegmentale che ha il suo campo nella sequenza “vocale accentata + consonante”?

La seconda interpretazione ha le sue radici in Sievers (fine dell'Ottocento), ma sono stati trovati prodromi anche più antichi (per esempio Ickelsamer nel Cinquecento) (v. Uguzzoni 2002). Nel quadro della trattazione dei fenomeni prosodici e delle loro proprietà Trubetzkoy (1939) attribuisce alla lingua tedesca una correlazione di taglio sillabico (*Silbenschnittkorrelation*), caratterizzata da una differenza che riguarda il modo di concatenare la vocale accentata e la consonante successiva: un legame stretto (*fester Anschluss*) vs. un legame lento (*loser Anschluss*). La brevità e la lunghezza associate, rispettivamente, ad una vocale con legame stretto e ad una vocale con legame lento sono considerate un fenomeno fonetico collaterale. Secondo Vennemann (1991 e 2000), che riprende il concetto di taglio sillabico (*syllable cut*), il tedesco deve essere classificato non come una *quantity language*, ma come una *syllable-cut language*. Questa lingua oppone *abruptly cut syllables* a *smoothly cut syllables*: le prime sono caratterizzate da “*an energy contour with a sharp drop at the very end of the syllable nucleus*”, mentre le seconde presentano “*a slow drop of the energy contour extending over a large portion of the*

*syllable nucleus*” (Vennemann, 2000: 252). Le principali manifestazioni fonetiche delle due prosodie di taglio sillabico sono le seguenti: le vocali nel primo caso sono centralizzate, rilassate, brevi; nel secondo caso sono invece periferiche, tese, lunghe. La teoria del taglio sillabico (*Silbenschnitt, syllable cut*) è ancorata ad alcune argomentazioni che per il tedesco si possono riassumere così: distinzioni di quantità vocalica si trovano solo in sede accentata; vocali accentate brevi non sono ammesse in sillabe aperte e richiedono la presenza di una consonante successiva nell’ambito della stessa parola; una consonante intervocalica quando è preceduta da una vocale accentata breve è sempre ambisillabica (v. Uguzzoni *et al.*, 2003).

Non è questa la sede per parlare delle indagini sperimentali che sono state svolte, nel versante acustico e nel versante articolatorio, alla ricerca di correlati fisici e fisiologici del fenomeno percettivo per cui si “sente”, per esempio che la parola *Mitte* è diversa da *Miete* e la parola *Wonne* è diversa da *wohne* (v. Fischer-Jørgensen & Jørgensen, 1969; Mooshammer, 1998; Spiekermann, 2000 e 2002; Hoole & Mooshammer, 2002). Mi soffermerò invece sul contributo portato al convegno di Friburgo su *Silbenschnitt und Tonakzent* da Jessen (2002), perché si collega strettamente al tema della “intensità rivisitata”. Nell’ambito della ricerca di un correlato acustico specifico del *syllable cut* altri studiosi hanno concentrato l’attenzione sull’andamento dell’energia lungo l’asse del tempo e si sono basati su misure dell’energia globale (*overall signal amplitude*). Jessen si è proposto invece di verificare se il nuovo modo di intendere e di misurare l’intensità (*frequency-sensitive amplitude measurements*) può essere utile anche nello studio della distinzione *abrupt cut/smooth cut* in tedesco.

Mentre per l’aspetto elettroglottografico Jessen riprende i dati di studi precedenti che ho menzionato in 3.3 e 3.5 (Marasek, 1997; Jessen & Marasek, 1997), per l’aspetto acustico si serve dei dati di una indagine più recente condotta su nove soggetti maschili e nove soggetti femminili (cfr. Schneider, 2001). Sono analizzate parole in cui compaiono /ɪ, ε, u, ɔ, a/ denominate *abruptly cut vowels* o *abrupt vowels* e /i:, e:, u:, o:, a:/ denominate *smoothly cut vowels* o *smooth vowels*: Le misurazioni di  $f_0$ , F1, F2, F3, H1, H2, A1, A2, A3 sono fatte con il metodo che ho già esposto in 3.3 con l’appoggio della fig. 1. Ai valori di H1, H2, A2, A3 vengono applicate procedure di normalizzazione (a cui si è accennato in 3.2 e 3.3) allo scopo di ottenere “*a better estimate of pure source characteristics, i.e., of laryngeal information without influence from supralaryngeal differences*” (Jessen, 2002: 162).

I risultati riguardanti le vocali in posizione accentata consentono di affermare che esse (eccetto il caso delle vocali basse) si distinguono tra di loro in maniera statisticamente significativa per valori differenti dei parametri H1-A2 e H1-A3: questi sono più piccoli nelle vocali con taglio brusco (*abrupt cut*) che nelle corrispondenti vocali con taglio piano (*smooth cut*). Come ho sottolineato più volte in queste pagine, valori più piccoli di H1-A2 e H1-A3 indicano che una classe di vocali presenta nel campo delle medie e alte frequenze un grado di energia maggiore rispetto alla classe di vocali contrapposta.

La differenza tra *abruptly cut vowels* e *smoothly cut vowels* è illustrata nella parte superiore e nella parte inferiore della fig. 2. In ciascuna parte sono sovrapposti uno spettro DFT e uno spettro LPC: la linea disegnata da Jessen congiunge l’ampiezza della prima armonica (H1) e la ampiezza della armonica più forte che si trova nell’area della terza formante (A3). Il valore di H1-A3 è chiaramente più piccolo nella vocale *abrupt* /ɔ/ (sopra) confrontata con la vocale *smooth* /o:/ (sotto).

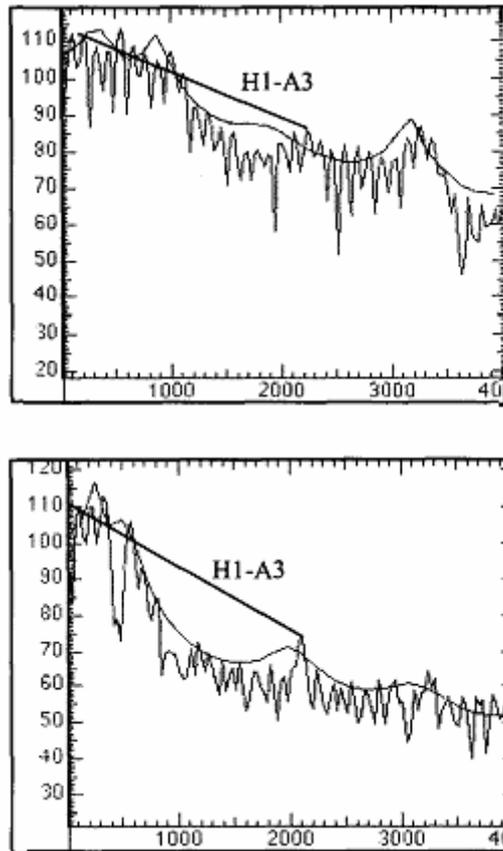


Figura 2: Differenti valori di H1 - A3 in vocali accentate tedesche; sopra la vocale /ɔ/ *abruptly cut*, sotto la vocale /o:/ *smoothly cut* (adattamento da Jessen, 2002: 166).

Mi pare che l'esempio addotto si inquadri nella tendenza generale rappresentata dalla sezione della tabella in cui sono inseriti, per tutte le vocali analizzate, i valori medi di H1-A2 e H1-A3 (senza e con normalizzazioni) (Jessen, 2002: 163, tabella 2). Il valore medio di H1-A3 per le vocali /ɔ/ e /o:/ dell'intero *corpus* è, rispettivamente, 29.4 dB e 42.2 dB. Continuando a scuriosare dentro alla tabella trovo che per le suddette vocali il valore medio di H1-A2 è da un lato 20.8 dB, dall'altro 26.6 dB: l'entità della differenza nel parametro H1-A2 è dunque minore di quella osservata nel parametro H1-A3, ma anch'essa risulta statisticamente significativa.

In linea generale i dati sperimentali inducono a rispondere affermativamente alla domanda: tra vocali con taglio secco e vocali con taglio piano in tedesco ci sono differenze di intensità in termini di *spectral balance* o di *mid-to-high-frequency emphasis*? E in consonanza con questa conclusione si può anche ipotizzare che le vocali *abrupt* siano percettivamente più forti (*loud*) delle loro controparti *smooth*.

Più difficile è dare una risposta alla domanda: le differenze di intensità nell'area della seconda e della terza formante, espresse da H1-A2 e H1-A3, sono il riflesso acustico di

differenze di carattere glottidale o dipendono invece da differenze nelle frequenze formantiche? Per comprendere bene l'ampia discussione di Jessen è di aiuto una figura ad albero in cui sono schematizzate le possibili spiegazioni delle *spectral balance differences* (Jessen, 2002: 169, fig. 4). Sulla base di vari elementi sembra ragionevole concludere che tali differenze non sono il prodotto di gesti glottidali controllati attivamente dal parlante, ma sono un effetto secondario di gesti sopraglottidali “*more directly involved in the implementation of the smooth/abrupt vowel distinction*” (Jessen, 2002: 170).

Riguardo alla possibilità di associare da una parte vocali *abrupt* e fonazione modale, dall'altra vocali *smooth* e fonazione aspirata, Jessen sottolinea che in realtà in tedesco “*breathiness levels can only be weak at best and the most that can be found is some slight degree of slack voice*” (Jessen, 2002: 169).

### 3.7 Sulla distinzione breve/lungo in una varietà italo-romanza: ricerche precedenti

Ai problemi connessi con la distinzione vocale breve/vocale lunga nella Italo-romania settentrionale sono stati dedicati vari studi, in parte presentati anche a Congressi internazionali di scienze fonetiche (Uguzzoni & Busà, 1995a; Uguzzoni *et al.*, 1999, Uguzzoni *et al.*, 2003) e a Giornate del Gruppo di Fonetica Sperimentale (Uguzzoni, 2000a; Zmarich *et al.*, 2003).

Costituisce un problema decidere con sicurezza se le varietà italo-romanze settentrionali in oggetto appartengono tipologicamente alle lingue con opposizioni di quantità (*quantity languages*) o alle lingue con opposizioni di taglio sillabico (*syllable-cut languages*). E se invece di giurare su una netta divisione fra il tratto della quantità vocalica e il tratto del taglio sillabico si provasse ad avanzare e puntellare l'ipotesi che i due fenomeni non si escludono a vicenda ma possono essere compresenti in una stessa lingua?

Porto come esempio soltanto due criteri di classificazione abbastanza noti. Uno è la indipendenza *vs.* la dipendenza dall'accento lessicale: in base a questo fattore lingue come, da un lato, il finnico e il ceco, dall'altro, il tedesco e l'olandese, formano due gruppi distinti, giacché le opposizioni breve/lungo nel primo gruppo ricorrono sia in sillaba accentata sia in sillaba non accentata, mentre nel secondo ricorrono solo in sillaba accentata. L'altro è la proprietà distribuzionale che riguarda le vocali in sillabe accentate aperte: anche in base a questo fattore i suddetti gruppi si differenziano tra di loro, giacché in tale contesto il primo gruppo ammette sia vocali brevi sia vocali lunghe, invece il secondo ammette vocali lunghe ed esclude vocali brevi.

Le varietà romanze dell'Italia settentrionale di cui mi sono occupata, insieme a più colleghi, presentano una *crux* (v. Uguzzoni *et al.*, 2003). Alla luce del primo criterio esse risultano simili al tedesco e all'olandese, dato che le opposizioni breve/lungo sono vincolate dalla presenza dell'accento; invece alla luce del secondo criterio risultano simili al finnico e al ceco, dato che in sillabe accentate aperte sono ammesse sia vocali brevi sia vocali lunghe. Se si guarda ai dati dei dialetti del Medio Frignano (Appennino modenese) la dicotomia *quantity languages vs. syllable-cut languages* non soddisfa, visto che essi condividono proprietà dell'uno e dell'altro tipo di lingue.

Anche se può sembrare trasgressiva, non è priva di fondamento la proposta di superare la suddetta biforcazione tipologica, ammettendo la possibilità che in una stessa lingua convivano proprietà solitamente attribuite al primo e al secondo tipo. Riguardo alla distinzione vocale breve/vocale lunga, la situazione dei dialetti frignanesi esaminati si riassume in questi termini: da una parte essa è soggetta allo stesso vincolo del tedesco in quanto è operante solo se le sillabe sono accentate; dall'altra essa è svincolata dalla restrizione per cui in tedesco vocali brevi sono escluse in sillaba aperta. Infatti in questa specifica area italo-romanza vocali brevi si oppongono a vocali lunghe in tre tipi di parole:

monosillabi terminanti in vocale (/CV(:)/, es. /pe/ vs. /pe:/) , monosillabi terminanti in consonante (/CV(:)C/, es. /pel/ vs. /pe:l/), bisillabi accentati nella prima sillaba (/CV(:)CV/, es. /'pela/ vs. /'pe:la/ (v. Uguzzoni *et al.*, 1999; Uguzzoni, 2000a; Zmarich *et al.*, 2003).

Se si sposta l'attenzione al piano della realizzazione nascono alcuni problemi che sollevano dubbi sulla validità di una netta separazione tra lingue germaniche da un lato e lingue romanze e slave dall'altro (v. Uguzzoni *et al.*, 1999; Uguzzoni, 2002; Uguzzoni *et al.*, 2003). Faccio ora riferimento solo a due proprietà che per tradizione sono considerate tipiche delle lingue germaniche o almeno di alcune lingue germaniche. Una è l'accento dinamico o espiratorio (cfr. 3.1); l'altro è la distinzione tra due modi di concatenare una vocale accentata e la consonante successiva: legame stretto vs. lento o taglio brusco vs. piano (*fester/loser Anschluss, close/open contact, scharfer/sanfter Schnitt, abrupt/smooth cut*; cfr. 3.6). Secondo alcuni c'è una connessione tra i due fatti, nel senso che il legame stretto (o taglio brusco) si troverebbe solo in lingue con forte accentuazione dinamica o espiratorio. In questa prospettiva a lingue germaniche come il tedesco settentrionale e l'olandese vengono contrapposte le lingue romanze e slave, che non possiedono tale distinzione tra due modi di legame (o di taglio), ma hanno sempre legame lento (o taglio piano).

Se le dicotomie classificatorie viste sopra fossero rigide e invalicabili, per gli emiliani si creerebbe un caso di crisi di identità. Le varietà italo-romanze parlate in Emilia presentano caratteristiche fonologiche e fonetiche che agli studiosi potrebbero sembrare strane in tale area linguistica e invece ovvie in area linguistica germanica. Propendo per l'ipotesi che nei dialetti del Medio Frignano coesistano due fenomeni: da una parte una opposizione vocale accentata breve vs. vocale accentata lunga in parole /CV(:)/, /CV(:)C/, /CV(:)CV/; dall'altra una distinzione legame stretto vs. legame lento tra la vocale accentata e la consonante successiva in parole /CV(:)C/, /CV(:)CV/.

Tutto ciò è limitato a un sottoinsieme del ricco vocalismo frignanese che in posizione accentata è costituito di tredici unità: mentre le vocali /i:, y:, u:, ε:, o:/ sono sempre lunghe, alle vocali lunghe /e:, ø:, ɔ:, a:/ si oppongono le brevi /e, ø, ɔ, a/. La trascrizione usata si fonda sulla constatazione che l'entità delle differenze qualitative tra le quattro brevi e le quattro corrispondenti lunghe, pur essendo sistematiche non sono tali da giustificare l'uso di simboli diversi. Per questo aspetto rinvio a Uguzzoni & Busà (1995a e 1995b) che riportano in tabelle e diagrammi i valori di F1 e di F2 (in Hz e in mel) delle tredici vocali frignanesi e forniscono una visione d'insieme delle loro collocazioni nello spazio acustico. Le differenze formantiche che si accompagnano alle differenze di durata variano a seconda del tipo vocalico e a seconda dei soggetti, ma mostrano tendenze generali simili a quelle trovate in altre lingue. Segnalo per esempio questa tendenza sistematica: confrontata con la /e:/, la /e/ presenta un valore maggiore di F1 e un valore minore di F2; confrontata con la /ɔ:/, la /ɔ/ presenta un valore maggiore di F1 e un valore maggiore di F2. Un altro esempio: dai dati dei tre soggetti relativi alla coppia /a/ vs. /a:/ emerge sempre un valore F1 minore per la vocale breve: si manifesta qui lo stesso *trend* osservato in lingue come tedesco, islandese, ceco, serbo-croato, hausa; in altre lingue invece, come svedese e danese, il valore di F1 nella vocale breve /a/ è maggiore.

Le differenze più considerevoli tra /e, ø, ɔ, a/ e /e:, ø:, ɔ:, a:/ compaiono nel campo della durata. Considerando insieme parole /CV(:)C/ e /CV(:)CV/ e mediando le realizzazioni dei soggetti SG e GB risulta che la durata media delle vocali brevi è un po' meno della metà di quella delle vocali lunghe (v. Uguzzoni & Busà, 1995b: 26): il rapporto medio V/V: è pari a 49.28 e si allinea con rapporti riscontrati in lingue come l'islandese (V/V: = 51.48) e il

danese (V/V: = 50.50).

In questo *excursus* non posso passare sotto silenzio un dato frignanese che suscitò una certa sorpresa già nelle prime analisi sperimentali. La misurazione della durata della consonante successiva alla vocale accentata ha sempre portato a risultati che sono diversi a seconda che la consonante si trovi alla fine della parola o all'interno della parola. Solamente in posizione finale, cioè in parole monosillabiche (/CV(:)C/), si riscontra una differenza di durata tra la C collocata dopo vocale lunga e la C collocata dopo vocale breve. Nel primo contesto la consonante ha durata minore che nel secondo: in media il rapporto (C dopo V:)/(C dopo V) è pari a 81.29. Il comportamento temporale della consonante postonica è chiaramente diverso in posizione interna, cioè in parole bisillabiche (/CV(:)CV/), dove non si manifestano differenze fisiche degne di nota tra la durata della C che si trova dopo vocale lunga e quella della C che si trova dopo vocale breve: in media il rapporto (C dopo V:)/(C dopo V) è pari a 98.57. Una situazione analoga è stata segnalata ad esempio per l'olandese e per il bolognese.

La sorpresa derivava e deriva non dalle misure strumentali delle durate delle consonanti finali in parole /CV(:)C/, bensì da quelle delle durate delle consonanti interne in parole /CV(:)CV/. La maggiore durata di una consonante finale nella condizione /CVC/ (vs. /CV:C) non è inaspettata, né in termini di comparazione con altre lingue (penso per esempio a islandese, olandese, lecchese, bolognese), né in termini di fonetica generale: può trattarsi di un fenomeno collegabile a fatti fisiologici inerenti al processo della produzione di una sequenza “vocale accentata breve+consonante finale”. L'interrogativo più interessante mi sembrava e mi sembra quello concernente gli esiti nella condizione /CVCV/: quali le possibili ragioni della divergenza tra valori oggettivi misurati a livello acustico e valori soggettivi colti a livello uditivo? Strumentalmente nella sequenza “vocale accentata breve+consonante interna” l'estensione temporale della consonante non risulta maggiore. Uditivamente invece si ha l'impressione che la C interna di /CVCV/ sia un po' più lunga e un po' più forte della C interna di /CV:CV/ (v. Uguzzoni & Busà, 1995b; Uguzzoni, 2000a e 2000b). Si può ipotizzare che l'impressione uditiva di una maggiore salienza della consonante interna nella condizione /CVCV/ sia un aspetto della “sindrome” del tipo di legame (*Anschlussart*) e del tipo di taglio sillabico (*Silbenschnitt, syllable cut*). Parole frignanesi come, ad esempio, /'pepa/, /'føsa/, /'bøta/, /'pana/ all'orecchio di nativi e non nativi danno la sensazione che la consonante interna sia connessa in modo “stretto” con la precedente vocale breve (*fester Anschluss, close contact, scharfer Schnitt, abrupt cut*).

### 3.8 Sulla distinzione breve/lungo in una varietà italo-romanza: ricerche in corso

Nella rapida rassegna appena conclusa sono state sottolineate molteplici condizioni e proprietà che disegnano un quadro complesso del fenomeno della “quantità vocalica” in area emiliana. Usando ora le virgolette intendo dire che, a mio parere, il fenomeno è plurifacciale e pluriparametrico. È evidente che esso non coinvolge soltanto il segmento vocalico e la sua durata. Sulle quattro coppie frignanesi in opposizione /e, ø, ɔ, a/ vs. /e:, ø:, ɔ:, a:/ bisogna continuare a indagare e a discutere, possibilmente senza trascurare nulla di ciò che sta attorno alla vocale accentata, prima e dopo (v. Uguzzoni, 2000a e 2000b, anche per le figure che esemplificano parole /CV(:)CV/ analizzate con il *Computerized Speech Lab*, 4300b). So bene che è un sogno difficile da realizzare, ma come in ogni ricerca degna di questo nome lo scopo ultimo dovrebbe essere quello di integrare “*all relevant knowlwdge concerning production, acoustic pattern, and perception within a linguistic-phonetic frame*” (Fant *et al.*, 2000a: 57).

In siffatta cornice ideale si inquadrano gli ultimi studi intrapresi e ancora in corso. Per

esempio l'analisi cinematica è stata motivata inizialmente dall'esigenza di trovare su questo piano indizi strumentali che non venivano forniti dall'analisi acustica al fine di fare un po' di luce sulla *vexata quaestio* della interazione tra vocale accentata breve vs. lunga e consonante successiva in area emiliana. I primi risultati della ricerca condotta a Padova con il sistema optoelettronico ELITE sono stati presentati e discussi in Zmarich *et al.* (2003), a cui rinvio per gli specifici aspetti teorici e metodologici.

In questa nuova prospettiva le vocali frignanese acusticamente brevi, confrontate con le lunghe corrispondenti, risultano caratterizzate dalle seguenti proprietà: il gesto di apertura mostra una durata minore, una ampiezza (*displacement*) minore, una velocità maggiore, una *stiffness* maggiore, una fase di accelerazione più lunga rispetto alla fase di decelerazione. Un'altra proprietà significativa concerne la percentuale del ciclo articolatorio completo che è compresa tra il picco di velocità del gesto di apertura e il picco di velocità del gesto di chiusura: tali percentuali sono 34.43 in /CVC/ vs. 54.98 in /CV:C/ e 35.15 in /CVCV/ vs. 51.74 in /CV:CV/: quindi il ciclo articolatorio nel caso di *target* breve, rispetto al caso di *target* lungo, è caratterizzato da una collocazione più ravvicinata dei picchi di velocità dei due gesti. Infine si osserva una differenza nel numero dei picchi di accelerazione che si trovano tra i picchi di velocità: il *target* breve presenta di solito un unico picco di accelerazione, mentre il *target* lungo presenta due o più picchi di accelerazione.

I risultati dell'analisi cinematica dei gesti labiali consentono di avanzare l'ipotesi che per esempio in parole del tipo /pV(:)pa/ la distinzione V breve/V lunga in posizione accentata sia da attribuire a una diversa coordinazione temporale (*phasing*) tra il gesto di apertura (: dalla consonante /p/ alla vocale accentata) e il gesto di chiusura (: dalla vocale accentata alla consonante /p/). Anticipare l'inizio del gesto di chiusura e con ciò stesso troncato il gesto di apertura: questi potrebbero essere i meccanismi intergestuali che soggiacciono sia alla brevità del segmento vocalico misurata acusticamente sia alla impressione uditiva di un legame stretto tra il segmento vocalico breve e il segmento consonantico successivo. Rinvio a Hoole & Mooshammer (2002) per una importante discussione sui risultati delle indagini cinematiche svolte in questi anni in Germania per mezzo di EMMA (*electrographic midsagittal articulography*). Da un lato sono stati trovati correlati articolatori che fanno da sostrato fonetico alla distinzione teso/rilassato (cfr. paragrafo 3.5), dall'altro sono stati individuati alcuni parametri che illuminano sulla distinzione *abrupt cut/smooth cut* (cfr. paragrafo 3.6). La secolare *syllable-cut intuition* (v. Uguzzoni, 2002) comincia così ad assumere contorni più precisi e fondati strumentalmente.

Con il sistema ELITE sono stati registrati il segnale articolatorio e il segnale acustico; inoltre il segnale acustico è stato registrato anche con un dispositivo indipendente, cioè un DAT. Sullo stesso materiale, prodotto da un soggetto femminile, sono state eseguite sia analisi cinematiche sia analisi acustiche. Di queste ultime, realizzate con *MultiSpeech*, finora sono stati pubblicati i risultati concernenti la durata della prima /p/, della vocale accentata, della seconda /p/ nei *target* brevi e nei *target* lunghi di monosillabi /CV(:)C/ e di bisillabi /CV(:)CV/ (v. Zmarich *et al.*, 2003: 301). In questa sede presento ulteriori misurazioni che riguardano vari aspetti delle vocali accentate brevi vs. lunghe, limitandomi al settore delle parole del tipo /pV(:)pa/. I valori medi di alcuni parametri sotto esame sono visualizzati in altrettante figure che parlano da sé. Ma, seguendo la consuetudine, farò qualche commento su durata, F1, F2, f<sub>0</sub>, H1, A2, A3, H1-A3.

Le durate assolute delle vocali analizzate mostrano una correlazione con il grado di altezza, in consonanza con il fenomeno generale noto col nome di durata vocalica intrinseca (cfr. Uguzzoni & Busà, 1995a). La differenza temporale tra l'elemento breve e l'elemento lungo di ciascuna delle quattro coppie è considerevole, come si vede bene nella fig. 3. Il

calcolo della V/V: ratio fatto sulle misure delle durate dei singoli tipi vocalici porta a questi esiti: I 40.85 per a/a:, II 41.62 per e/e:, III 42.11 per ø/ø:, IV 46.55 per o/o:. In altri termini, la distinzione breve/lungo si manifesta in modo un po' diverso a seconda della qualità della vocale: l'entità della differenza tra V e V: diminuisce progressivamente da I a II, a III, a IV. In generale i dati di questa ricerca confermano che le vocali brevi frignanesi hanno una durata che è inferiore alla metà della durata delle loro controparti lunghe (cfr. 3.7).

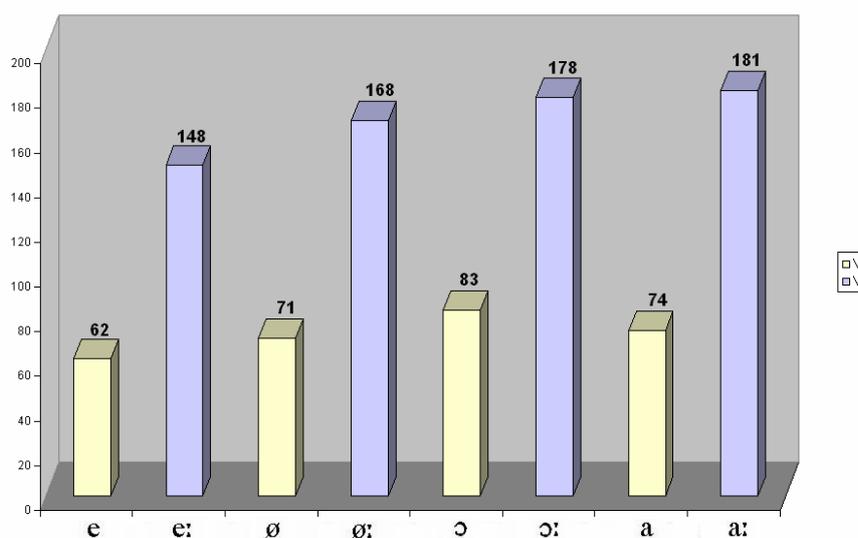


Figura 3: Valori medi delle durate (ms) in parole bisillabiche.

La misurazione della frequenza delle formanti è stata laboriosa e talora problematica soprattutto a causa del fatto che il soggetto è femminile (v. Klatt & Klatt, 1990; Hanson, 1997; Traunmüller & Eriksson, 2000). Sono state utilizzate tutte le metodiche fornite dal programma *MultiSpeech* e sono state seguite le preziose indicazioni di Ferrero (1997a e 1997b). Per lo spettrogramma, a banda larga, si è scelto il numero di punti equivalenti a una larghezza di banda di analisi che superasse di un quarto il valore della  $f_0$  dei segnali verbali esaminati.

Per la tecnica LPC è stato scelto un "ordine" del predittore (da 12 a 18) che fosse soddisfacente in relazione alla frequenza di campionamento del segnale (11.025 Hz e 16.000 Hz). Nell'effettuare la sovrapposizione dei tracciati formatici allo spettrogramma si è usata l'opzione che mostra nel *display* la larghezza di banda delle formanti, cosa che è di aiuto immediato per vedere nella "storia formantica" i momenti di banda allargata o di banda ristretta.

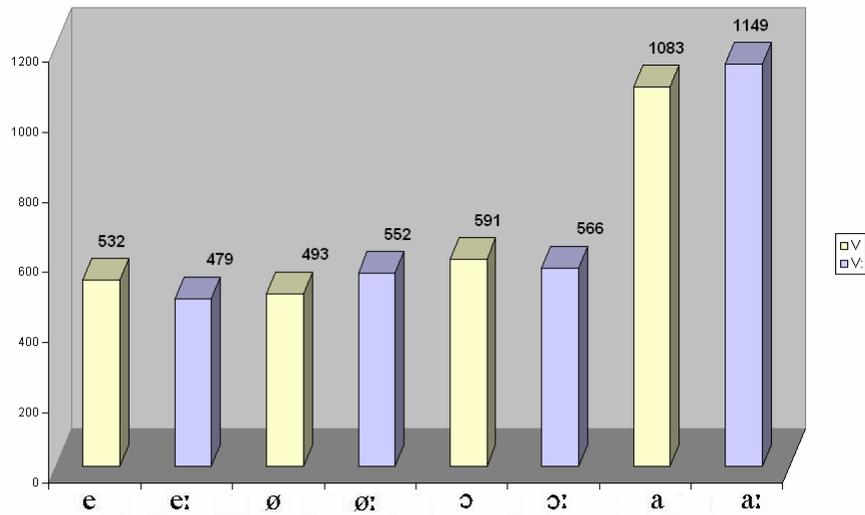


Figura 4: Valori medi di F1 (Hz) in parole bisillabiche.

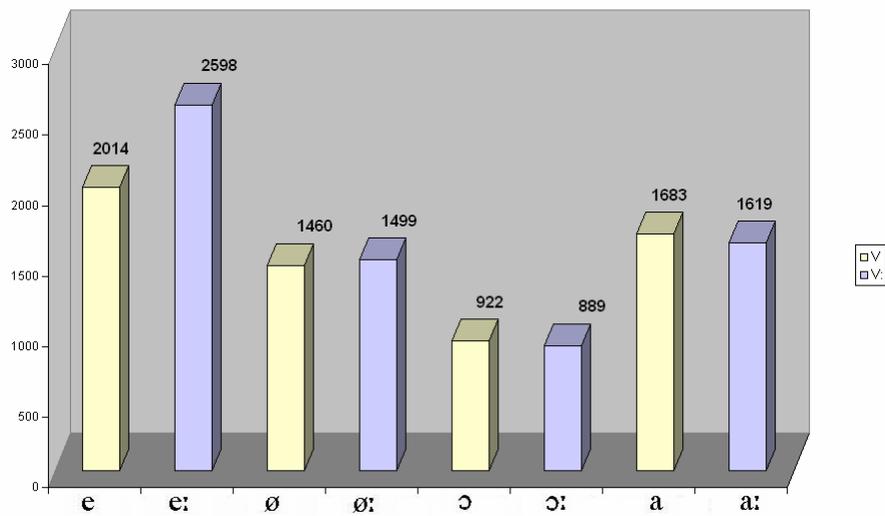


Figura 5: Valori medi di F2 (Hz) in parole bisillabiche.

I risultati ottenuti per F1 e per F2 compaiono nella fig.4 e nella fig.5. Le tendenze generali concordano in linea di massima con quelle emerse da precedenti ricerche su materiali frignanesi prodotti da soggetti maschili (v. Uguzzoni & Busà, 1995a e 1995b). È noto che differenze qualitative tra vocali brevi e vocali lunghe si riscontrano in varie lingue, come tedesco, svedese, islandese, danese, ceco, serbo-croato, thailandese, hausa, navajo. Come indicano i dati della presente analisi, in frignanese le differenze formantiche che sono associate

alla distinzione breve/lungo sono modeste, ma non inferiori alla soglia della discriminazione uditiva che si suppone intorno ai 14 Hz per F1 e ai 30-35 Hz per F2. Le figure mostrano che l'entità delle differenze formantiche varia dall'uno all'altro dei quattro tipi vocalici. Riguardo ai valori di F1 lo scarto più piccolo (25 Hz) si ha nella coppia /ɔ/ vs. /ɔ:/ e quello più grande (66Hz) si ha nella coppia /a/ vs. /a:/. Riguardo ai valori di F2 nella coppia /ɔ/ vs. /ɔ:/ si osserva la differenziazione minima (33 Hz), mentre la coppia /e/ vs. /e:/ presenta la differenziazione massima (584 Hz).

A proposito dell'orientamento delle differenze formantiche si conferma ciò che ho segnalato nel paragrafo 3.7. In modo sistematico la /e/ ha F1 maggiore e F2 minore rispetto alla /e:/; d'altra parte la /ɔ/ ha F1 maggiore e F2 maggiore rispetto alla /ɔ:/. Pertanto nello spazio vocalico le due brevi si collocano in aree più vicine al centro, mentre le lunghe corrispondenti occupano posizioni più periferiche. Se a ciò si aggiunge il fatto che la /a/ presenta una F1 minore rispetto alla /a:/ si può dire che nell'insieme lo spazio acustico delle vocali brevi è comparativamente meno periferico.

Per la quantificazione dei parametri  $f_0$ , H1, A2, A3 ho seguito la metodologia illustrata nel paragrafo 3.3 con la fig.1 (Jessen & Marasek, 1997) e nel paragrafo 3.6 con la fig.2 (Jessen, 2002), tenendo conto anche delle indicazioni di Jessen & Roux (2002). Per ogni vocale del *corpus* è stato creato uno spettro FFT a 512 punti. Uno spettro istantaneo FFT di questo tipo rappresenta, come è noto, il contenuto armonico spettrale di una determinata porzione della forma d'onda del segnale. Lo spettro FFT consente una buona stima delle ampiezze: su esso ho misurato sia l'ampiezza della prima armonica (H1) sia le ampiezze della seconda e della terza formante (A2, A3). Si considerano ampiezze delle formanti "the peak amplitudes of those harmonics that occur closest in frequency to the respective formants" (Jessen & Roux, 2002: 13): con A2 e A3 si fa dunque riferimento all'armonica più forte che nella scala delle frequenze è più vicina, rispettivamente, al secondo e al terzo picco formantico (cioè F2 e F3 del corrispondente spettro LPC che viene sovrapposto allo spettro FFT). Per quel che riguarda la determinazione della  $f_0$ , si è preso il valore della frequenza della prima armonica nello spettro FFT, ma si sono fatti anche controlli tramite un altro metodo che stima la  $f_0$  misurando la durata del periodo intorno al quale è stato creato lo spettro FFT e calcolando il reciproco di tale durata.

I valori di  $f_0$  riportati nella fig.6 da una parte sono conformi alle aspettative, dall'altra portano informazioni nuove. Ciò dipende dal punto di vista adottato per guardare il grafico. Se guardiamo alle brevi e alle lunghe separatamente, troviamo nell'ambito di ciascuna classe differenti valori di  $f_0$  secondo queste graduatorie: /ø > e > ɔ > a/ e /ø: > e: > ɔ: > a:/. Siamo di fronte al fenomeno denominato *intrinsic f<sub>0</sub>*: in numerose lingue si è trovato che la frequenza fondamentale è in correlazione negativa con l'altezza della vocale (v. Fischer-Jørgensen, 1990; Whalen & Levitt, 1995). I dati frignanesi riguardanti  $f_0$  devono essere messi in collegamento con i valori di F1 delle brevi e delle lunghe, riportati nella fig.4, dove si hanno graduatorie inverse rispetto a quelle di  $f_0$ : precisamente /ø < e < ɔ < a/ e /ø: < e: < ɔ: < a:/.

Se usiamo un altro criterio per leggere la fig. 6 emerge una regolarità di tipo diverso che è degna di rilievo. Confrontando i valori medi delle singole vocali brevi e lunghe osserviamo che la  $f_0$  sistematicamente ha un valore più grande in ciascuna breve rispetto alla sua controparte lunga. Poiché sono vocali prodotte da un soggetto femminile è opportuno esprimere queste differenze in percentuale. Comparate con le vocali /a:/, /e:/, /ɔ:/, /ø:/, le vocali /a/, /e/, /ɔ/, /ø/ mostrano, nell'ordine, i seguenti incrementi nel valore di  $f_0$ : 18.69%, 18.32%, 15.68%, 7.88%.

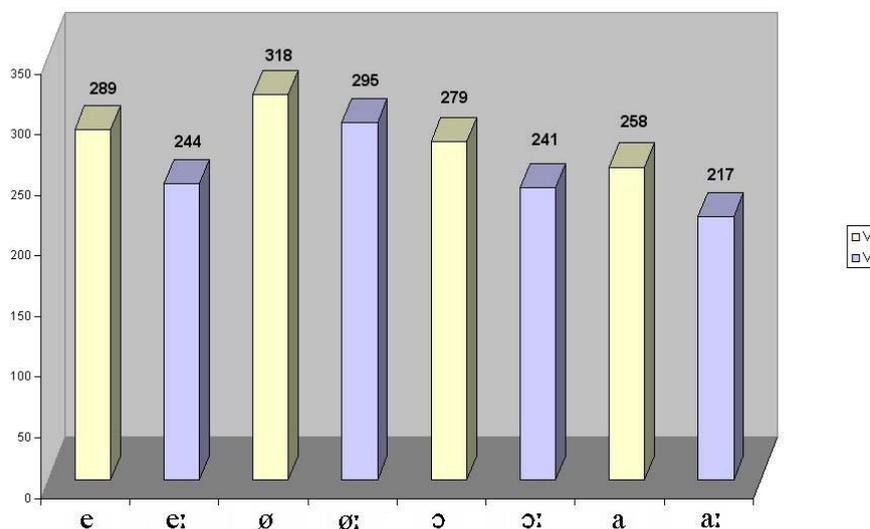


Figura 6: Valori medi di  $f_0$  (Hz) in parole bisillabiche.

Ritengo che l'informazione fornita da questi dati frignanesi sia interessante, alla luce del confronto con analisi di vocali danesi e tedesche. Da una ricerca sulle brevi /i, e, ε, a/ vs. le corrispondenti lunghe /i:, e:, ε:, æ:/ del danese è risultato che le prime hanno un valore di  $f_0$  maggiore rispetto alle seconde. Accanto a ciò è utile sottolineare che le differenze qualitative tra le vocali brevi e le vocali lunghe anteriori sono minime: per esempio “*the difference in F1 between short and long /i/ and between short and long /e/ was on the average 10 Hz, and the difference in F2 16 Hz*” (Fischer-Jørgensen, 1990: 132). In frignanese e in danese c'è dunque una connessione tra durata breve e valore maggiore di  $f_0$ . I problemi che sorgono nelle due lingue sono gli stessi e concernono il perché e il come: “*why short vowels should be produced on a higher  $f_0$ , and how it is brought about*” (Fischer-Jørgensen, 1990: 133). Le risposte possibili sono più di una, ma nessuna è del tutto convincente: di qui l'esigenza di approfondire lo studio delle relazioni fra tre parametri: durata,  $f_0$ , intensità.

In tedesco la situazione è più complicata e controversa (cfr. 3.5 e 3.6). Jessen e colleghi, in base ai loro dati sulla differenziazione tra vocali rilassate e brevi da una parte e vocali tese e lunghe dall'altra, hanno concluso che lo *status* distintivo è da attribuire alla qualità nel caso delle coppie di vocali non basse e alla durata nel caso della coppia di vocali basse (cioè /a/ vs. /a:/). Riguardo a  $f_0$  la tabella non presenta differenze sistematiche e significative tra i membri di ciascuna delle cinque coppie analizzate, ma gli studiosi tedeschi auspicano che ulteriori ricerche chiariscano “*why  $f_0$ std for male speakers depends significantly on tenseness in most conditions involving stressed vowels*” (Jessen *et al.*, 1995: 431). In Jessen (2002) si vede che i valori di  $f_0$  non differiscono in maniera significativa comparando le singole vocali *abrupt* e brevi con le corrispondenti vocali *smooth* e lunghe.

In un accuratissimo e stimolante studio su /I, ε, a/ e /i:, e:, a:/ del tedesco settentrionale Fischer-Jørgensen (1990) mostra quanto sia difficile render conto del fatto che nelle vocali rilassate e brevi il valore di  $f_0$  risulta più grande di quello che ci si aspetterebbe in base alla

teoria della *intrinsic f<sub>0</sub>*, che si fonda sulla correlazione negativa tra  $f_0$  e altezza della vocale (nel senso di *tongue height*). Sono discusse le relazioni tra l'incremento di  $f_0$ , trovato nelle vocali del primo gruppo rispetto a quelle del secondo gruppo, e i seguenti parametri: *jaw opening, airflow, duration, intensity*. Nessuna di queste connessioni riesce a spiegare in modo soddisfacente il fenomeno e la studiosa ritiene che si debbano svolgere studi approfonditi per aumentare le nostre conoscenze, comprese quelle sulla pressione sottoglottidale in vocali brevi vs vocali lunghe. Hoole & Mooshammer (2002) confermano che la situazione tedesca è ingarbugliata e progettano analisi EMG allo scopo di verificare se si può invocare una partecipazione attiva della muscolatura laringale per spiegare la proprietà, da essi denominata *raised f<sub>0</sub>*, che caratterizza tutta la serie delle vocali rilassate (e brevi) del tedesco. Penso che sarà molto utile a livello generale conoscere al più presto i risultati delle indagini sperimentali auspicate e programmate su un tema che affascina per i segreti che ancora nasconde.

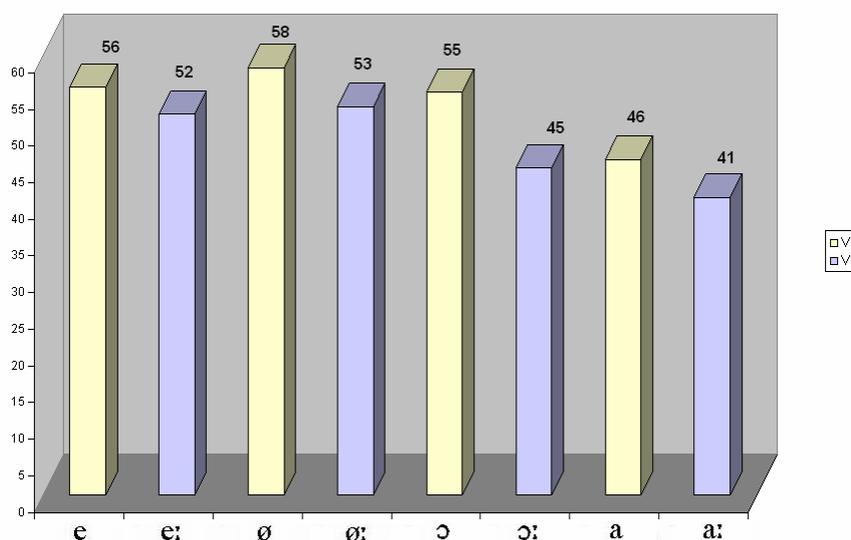


Figura 7: Valori medi di H1 (Hz) in parole bisillabiche.

Le misure di H1 costituiscono la base per il procedimento di sottrazione da cui vengono ricavati i parametri H1-A2 e H1-A3, che secondo vari studiosi sono i correlati della intensità "rivisitata" (cfr i paragrafi 3.2, 3.3, 3.5, 3.6). Ma esse possono essere considerate anche in sè e per sè: quelle frignanesi, riportate nella fig. 7, mostrano che l'ampiezza della prima armonica nelle vocali brevi è maggiore che nelle vocali lunghe. Non dispongo di elementi sufficienti per abbozzare un'ipotesi esplicativa.

Posso richiamare il fatto che un aumento di H1 è ritenuto uno degli effetti acustici di una più forte pressione sottoglottidale, come fanno notare Sluijter *et al.* (1995) a proposito della *amplitude of voicing* stimata in base alla misura di H1. Dalla bibliografia si sa che nelle donne i valori dell'ampiezza della prima armonica superano quelle dei maschi (Klatt & Klatt, 1990; Holmberg *et al.*, 1995; Hanson, 1997). E ciò da un lato viene collegato con un più largo quoziente glottidale, dall'altro viene additato come uno dei correlati acustici di voce non modale, specificamente di *breathy voice* (Gordon & Ladefoged, 2001). Ma queste

osservazioni non sono di aiuto nel nostro caso, dal momento che il materiale del presente studio è stato prodotto da un soggetto femminile. Resta quindi aperto il problema delle regolari differenze in H1 che si riscontrano tra vocali brevi e vocali lunghe pronunciate dallo stesso soggetto. Il grafico fa vedere che nelle quattro coppie tali differenze sono pari a 4dB, 5dB, 10dB, 5dB.

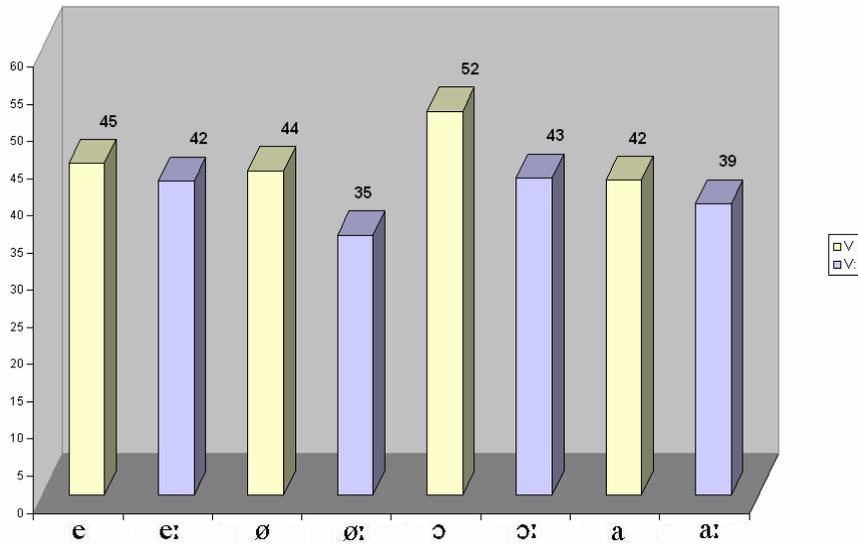


Figura 8: Valori medi di A2 (dB) in parole bisillabiche.

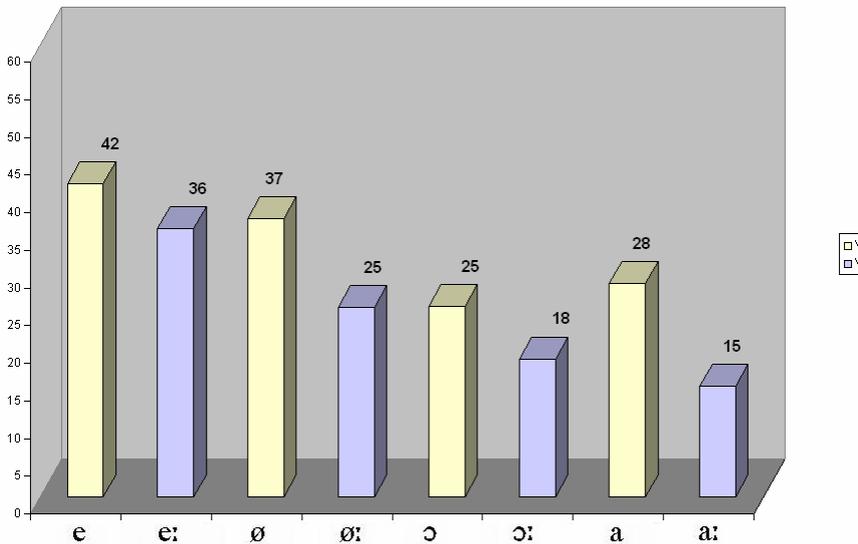


Figura 9: Valori medi di A3 (dB) in parole bisillabiche

I risultati delle misurazioni di A2 e A3 sono rappresentati nella fig. 8 e nella fig. 9. Il confronto tra ciascuna unità breve e la sua controparte lunga consente di enucleare una regolarità sistematica riguardo alle ampiezze della seconda e della terza formante. Leggendo i grafici da sinistra a destra la differenza di ampiezza tra la breve e la lunga risulta pari a 3dB, 9dB, 9dB, 3db nel caso di A2 e 6dB, 12dB, 7dB, 13dB nel caso di A3. L'entità della differenza è quindi più rilevante in A3 che in A2, eccetto che nella coppia /ɔ/ vs. /ɔ:/. I dati frignanesi complessivamente sembrano indicare che, rispetto alle vocali lunghe, le vocali brevi presentano un grado maggiore di energia nello spettro acustico corrispondente a F2 e F3.

Si è proceduto a sottrarre dai valori di H1, da una parte i valori di A2, dall'altra i valori di A3, sulla scia degli studi esposti nelle pagine precedenti. Qui vengono riportati solo gli esiti concernenti H1-A3, dato che nel caso di H1-A2 non si manifestano tendenze sistematiche. Come si vede nella fig. 10, i valori di H1-A3 sono differenti nelle vocali brevi vs. le vocali lunghe: essi sono più piccoli nell'elemento breve delle coppie /e/ vs. /e:/, /ø/ vs. /ø:/, /a/ vs. /a:/, mentre nella coppia /ɔ/ vs. /ɔ:/ si ha il contrario. Siccome valori piccoli di H1-A3 segnalano un grado più elevato di energia nel campo delle alte frequenze, si può affermare che le vocali brevi, confrontate con le lunghe, sono caratterizzate da una maggiore intensità nella parte alta dello spettro. Non è escluso che l'eccezione della coppia /ɔ/ vs. /ɔ:/ dipenda da errori di misurazione dei valori di H1 e di A3 nell'elemento breve che potrebbero essere, in media, rispettivamente, inferiori a 55dB e superiori a 25dB, portando, di conseguenza, a un risultato della sottrazione minore di quello della fig. 10 (che è pari a 30dB). I valori attuali potrebbero però non essere errati e riflettere una aporia effettiva inerente alla produzione di tale coppia da parte del soggetto analizzato.

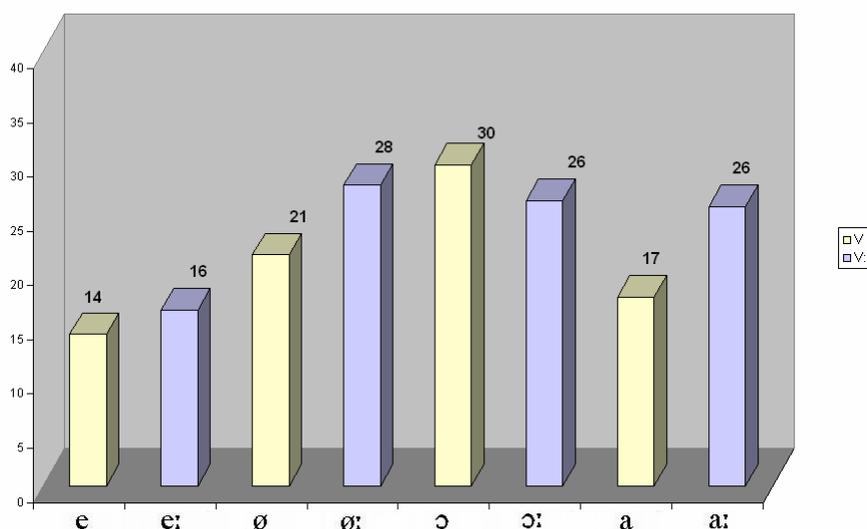


Figura 10: Valori medi di H1-A3 (dB) in parole bisillabiche

Quando si studierà il comportamento di altri soggetti si avranno conferme o smentite di

ciò che è saltato fuori da questa ricerca pilota. E si potrà quindi fare una discussione approfondita intorno alle differenze di intensità che sembrano essere associate alla distinzione tra vocali accentate brevi e vocali accentate lunghe in frignanese. Le brevi sono più intense delle lunghe? Perché lo sono? Quali meccanismi sottostanno a tale effetto acustico? La differenza misurata strumentalmente viene percepita dai parlanti-ascoltatori? Questa proprietà come si combina con le altre proprietà che caratterizzano le brevi rispetto alle lunghe corrispondenti?

A parte l'articolo di Gordeeva *et al.* (2003), che mette in relazione durata vocalica estrinseca e *spectral balance* in scozzese, per ora ho come unici punti di riferimento le importanti indagini di studiosi tedeschi sulle quali mi sono soffermata nei paragrafi 3.5 e 3.6 e che mi hanno fatto da guida. Le vocali del tedesco settentrionale sono state interpretate ora in termini segmentali (rilassate *vs.* tese, brevi *vs.* lunghe) ora in termini soprasegmentali (*abruptly cut vowels vs. smoothly cut vowels*).

Prescindendo dalle controversie sullo *status* linguistico della distinzione che intercorre tra vocali come quella di *Mitte* e vocali come quella di *Miete*, in questa sede l'attenzione è rivolta soprattutto agli aspetti strettamente fonetici che si combinano nella implementazione di vocali accentate. Eccetto il caso delle basse (cioè /a/ *vs.* /a:/), si può dire che vocali "rilassate-brevi-*abruptly cut*" (A) si contrappongono a vocali "tese-lunghe-*smoothly cut*" (B) e si manifestano come diverse, le une rispetto alle altre, attraverso una pluralità di attributi che sono stati oggetto di ricerche sperimentali.

Da analisi acustiche eseguite in questi ultimi anni è scaturito che le suddette classi di vocali sono differenziate "anche" dal punto di vista della intensità "rivisitata". Una proprietà che caratterizza la classe (A) rispetto alla classe (B) è la *mid-to-high-frequency emphasis* quantificata da valori più piccoli di H1-A2 e H1-A3. La nuova impostazione, secondo cui si considera non l'intensità globale ma la distribuzione dell'intensità in differenti parti dello spettro, ha portato dunque a individuare una differenza significativa tra le vocali (A) e le vocali (B): le prime, confrontate con le seconde, presentano nella zona delle medie e alte frequenze un grado più elevato di energia.

Sul risultato acustico in sé non si discute. Il dibattito invece si accende quando si affronta il problema dei fattori che possono essere causa o concausa di tale differenza acustica. Rinvio alla specifica e articolata trattazione di Jessen (2002) e mi limito a sottolineare che anche in questo caso ogni tentativo di spiegazione deve essere accompagnato dalla consapevolezza delle complicate interazioni che si attivano fra tre ordini di fatti: il sottoglottidale, il glottidale, il sopraglottidale (cfr. 3.4, 3.5., 3.6).

#### **RINGRAZIAMENTI**

Grazie a Michael Jessen, Cristina Villari, Gabriele Azzaro, Matteo Balasso per il loro aiuto prezioso e paziente.

#### **4. BIBLIOGRAFIA**

Bickley, C.; Stevens, K. N., 1986. Effects of a vocal tract constriction on the glottal source: experimental and modelling studies. *Journal of Phonetics*, 14, 373-382.

Childers, D. G.; Lee, C. K., 1991. Vocal quality factors: analysis, synthesis, and perception. *Journal of the Acoustical Society of America*, 90, 2394-2401.

Claßen, K.; Dogil, G.; Jessen, M.; Marasek, K.; Wokurek, W., 1998. Stimmqualität und Wortbetonung im Deutschen. *Linguistische Berichte*, 174, 202-245.

- Fant, G., 1985. The voice source: theory and acoustic modelling. In I. R. Titze e R. C. Scherer (a c. d.) *Vocal fold physiology: biomechanics, acoustics and phonatory control*, Denver, CO: The Denver Center for the Performing Arts, 453-464.
- Fant, G., 1997. The voice source in connected speech. *Speech Communication*, 22, 125-139.
- Fant, G.; Kruckenberg, A., 1989. Preliminaries to the study of Swedish prose reading and reading style. *Speech Transmission Laboratory Quarterly Progress and Status Report*, Royal Institute of Technology, Stockholm, 2, 1-83.
- Fant, G.; Kruckenberg, A., 1994. Notes on stress and word accent in Swedish. *Speech Transmission Laboratory Quarterly Progress and Status Report*. Royal Institute of Technology, Stockholm, 2-3, 125-144.
- Fant, G.; Kruckenberg, A., 1995. The voice source in prosody. In *Proceedings of the XIII International Congress of Phonetic Sciences*, Stockholm, 2, 622-625.
- Fant, G.; Kruckenberg, A.; Liljencrants, J., 1999. Prominence correlates in Swedish prosody. In *Proceedings of the XIV International Congress of Phonetic Sciences*, San Francisco, 1749-1752.
- Fant, G.; Kruckenberg, A.; Liljencrants, J., 2000a. Acoustic-phonetic analysis of prominence in Swedish. In A. Botinis (a c. d.) *Intonation analysis, modelling and technology*. Dordrecht: Kluwer Academic Publishers, 55-85.
- Fant, G.; Kruckenberg, A.; Liljencrants, J., 2000b. The source-filter frame of prominence. *Phonetica*, 57, 113-127.
- Fant, G.; Liljencrants, J.; Lin, Q., 1985. A four-parameter model of glottal flow. *Speech Transmission Laboratory Quarterly Progress and Status Report*, Royal Institute of Technology, Stockholm, 4, 1-13.
- Fant, G.; Lin, Q. G., 1987. Glottal source-vocal tract acoustic interaction. *Speech Transmission Laboratory Quarterly Progress and Status Report*, Stockholm, Royal Institute of Technology, 1, 13-27.
- Ferrero, F. E., 1997a. Problemi spettroacustici di classificazione e di misurazione delle vocali: un contributo. In *Atti delle VII Giornate di Studio del Gruppo di Fonetica Sperimentale*, Napoli: Esagrafica, 235-264.
- Ferrero, F. E., 1997b. Semiologia avanzata della funzione vocale (laringe e condotto vocale). *Quaderni del Centro di Studio per le Ricerche di Fonetica*, 16, 192-210.
- Fischer-Jørgensen, E., 1990. Intrinsic F0 in tense and lax vowels with special reference to German. *Phonetica*, 47, 99-140.
- Fischer-Jørgensen, E.; Jørgensen, H. P., 1969. Close and loose contact (Anschluss) with special reference to North German. *Annual Report of the Institute of Phonetics of the University of Copenhagen*, 4, 43-80.
- Glave, R.; Rietveld, A. C. M., 1975. Is the effort dependence of speech loudness explicable on the basis of acoustical cues?. *Journal of the Acoustical Society of America*, 58, 875-879.
- Gauffin, J.; Sundberg, J., 1989. Spectral correlates of glottal voice source waveform

- characteristics. *Journal of Speech and Hearing Research*, 32, 556-565.
- Gordeeva, O.; Mennen, I.; Scobbie, J. M., 2003. Vowel duration and spectral balance in Scottish English and Russian. In *Proceedings of the XV International Congress of Phonetic Sciences*, Barcelona, 3, 3193-3196.
- Gordon, M.; Ladefoged, P., 2001. Phonation types: a cross-linguistic overview. *Journal of Phonetics*, 29, 383-406.
- Halle, M.; Stevens, K. N., 1969. On the feature 'advanced tongue root'. *MIT Research Laboratory of Electronics Quarterly Progress Report*, 94, 209-215.
- Hanson, H. M., 1997. Glottal characteristics of female speakers: acoustic correlates. *Journal of the Acoustical Society of America*, 101, 466-481.
- Hanson, H. M.; Chuang, E. S., 1999. Glottal characteristics of male speakers: acoustic correlates. *Journal of the Acoustical Society of America*, 106, 1064-1077.
- Hanson, H. M.; Stevens, K. N.; Kuo, H.-K. J.; Chen, M. Y.; Slifka, J., 2001. Toward models of phonation. *Journal of Phonetics*, 29, 451-480.
- Heldner, M., 2003. On the reliability of overall intensity and spectral emphasis as acoustic correlates of focal accents in Swedish. *Journal of Phonetics*, 31, 39-62.
- Holmberg, E. B.; Hillman, R. E.; Perkell, J. S., 1988. Glottal airflow and transglottal air pressure measurements for male and female speakers in soft, normal and loud voice. *Journal of the Acoustical Society of America*, 84, 511-529.
- Holmberg, E. B.; Hillman, R. E.; Perkell, J. S.; Guiod, P.; Goldman, S. L., 1995. Comparisons among aerodynamic, electroglottographic, and acoustic spectral measures of female voice. *Journal of Speech and Hearing Research*, 38, 1212-1223.
- Hoole, P.; Mooshammer, C., 2002. Articulatory analysis of the German vowel system. In P. Auer *et al.* (a c. d.) *Silbenschnitt und Tonakzente*, Tübingen: Niemeyer, 129-152.
- Jessen, M., 2002. Spectral balance and its relevance for syllable cut theory. In P. Auer *et al.* (a c. d.) *Silbenschnitt und Tonakzente*, Tübingen: Niemeyer, 153-179.
- Jessen, M.; Marasek, K.; Schneider, K.; Claßen, K.; 1995. Acoustic correlates of word stress and the tense/lax opposition in the vowel system of German. In *Proceedings of the XIII International Congress of Phonetic Sciences*, Stockholm, 4, 428-431.
- Jessen, M.; Marasek, K., 1997. Voice quality correlates of word stress and tense versus lax vowels in German. In *Proceedings of LARYNX 97*, Marseille, June 16-18, 127-130.
- Jessen, M.; Roux, J. C., 2002. Voice quality differences associated with stops and clicks in Xhosa. *Journal of Phonetics*, 30, 1-52.
- Klatt, D.; Klatt, L., 1990. Analysis, synthesis, and perception of voice quality variation among female and male talkers. *Journal of the Acoustical Society of America*, 87, 820-857.
- Ladefoged, P.; Maddieson, I., 1996. *The sounds of the world's languages*. Oxford: Blackwell.
- Marasek, K., 1995. An attempt to classify Lx signals. In *Proceedings of Eurospeech*, Madrid, 18-21 September, 1995, 1729-1732.

- Marasek, K., 1996. Glottal correlates of the word stress and the tense/lax opposition in German. In *Proceedings of the Fourth International Conference in Spoken Language Processing*, Philadelphia, October 3-6, 1996, 1573-1576.
- Marasek, K., 1997. *Electroglottographic description of voice quality*, Habilitationsschrift, Universität Stuttgart.
- Mooshammer, C., 1998. Experimentalphonetische Untersuchungen zur artikulatorischen Modellierung des Gespanntheitsopposition im Deutschen. *Forschungsberichte des Instituts für Phonetik und sprachliche Kommunikation der Universität München*, 36, 3-192.
- Schneider, K., 2001. *Eine akustische Studie der Frequenz-und Amplitudenparameter deutscher Vokale*, Diplomarbeit, Universität Stuttgart.
- Sluijter, A. M. C., 1995. *Phonetic correlates of stress and accent*. The Hague: Holland Academic Graphics.
- Sluijter, A. M. C.; Shattuck-Hufnagel, S.; Stevens, K. N.; van Heuven, V. J., 1995. Supralaryngeal resonance and glottal pulse shape as correlates of stress and accent in English. In *Proceedings of the XIII International Congress of Phonetic Sciences*, Stockholm, 2, 630-633.
- Sluijter, A. M. C.; van Heuven, V. J., 1996. Spectral balance as an acoustic correlate of linguistic stress. *Journal of the Acoustical Society of America*, 100, 2471-2485.
- Sluijter, A. M. C.; van Heuven, V. J.; Pacilly, J. J. A., 1997. Spectral balance as a cue in the perception of linguistic stress. *Journal of the Acoustical Society of America*, 101, 503-513.
- Spiekermann, H. 2000. *Silbenschnitt in Deutschen Dialekten*. Tübingen: Niemeyer.
- Spiekermann, H. 2002. Ein akustisches Korrelat des Silbenschnitt: Formen des Intensitätsverlaufs in Silbenschnitt- und Tonakzentsprachen. In P. Auer *et al.* (a c. d.) *Silbenschnitt und Tonakzente*, Tübingen: Niemeyer, 181-199.
- Stevens, K. N., 1977. Physics of larynx behavior and larynx modes. *Phonetica*, 34, 264-279.
- Stevens, K. N., 1998. *Acoustic Phonetics*. Cambridge, MA: MIT Press.
- Stevens, K. N., Hanson, H. M., 1995. Classification of glottal fold vibration from acoustic measurements. In O. Fujimura *et al.*, (a c. d.) *Vocal fold physiology. Voice quality control*, San Diego: Singular, 147-170.
- Tamburini, F., 2005. *Fenomeni prosodici e prominenza: un approccio acustico*. Bologna: BUP.
- Trautmüller, H.; Eriksson, A., 2000. Acoustic effects of variation in vocal effort by men, women, and children. *Journal of the Acoustical Society of America*, 107, 3438-3451.
- Trubetzkoy, N. S., 1939. Grundzüge der Phonologie. *Travaux du Cercle Linguistique de Prague*, 7, 1-268.
- Uguzzoni, A., 2000a. Aspetti -emici e aspetti -etici del fenomeno della “quantità vocalica” nei dialetti dell’Italia settentrionale. In *Atti delle X Giornate di Studio del Gruppo di Fonetica Sperimentale*, Napoli: Istituto Universitario Orientale, 227-235.

- Uguzzoni, A., 2000b. Fonologia e fonetica della quantità vocalica in area italo-romanza. Il caso dei dialetti del Medio Frignano (provincia di Modena). *Studi Orientali e Linguistici*, 7, 339-349.
- Uguzzoni, A., 2002. Fester vs. loser Anschluss. Appunti per una storia di un concetto secolare. *Lingue e Linguaggio*, 1, 327-340.
- Uguzzoni, A., 2003. In margine ad una rivisitazione della intensità. In P. Cosi *et al.* (a c. d.) *Voce, canto, parlato. Studi in onore di Franco Ferrero*, Padova: Unipress, 299-302.
- Uguzzoni, A.; Azzaro, G.; Schmid, S., 2003. Short vs long and/or abruptly vs smoothly cut vowels. New perspectives on a debated question. In *Proceedings of the XV International Congress of Phonetic Sciences*, Barcelona, 3, 2717-2220.
- Uguzzoni, A.; Busà, M. G., 1995a. Acoustic correlates of vowel quantity contrasts in an Italian dialect. In *Proceedings of the XIII International Congress of Phonetic Sciences*, Stockholm, 3, 390-393.
- Uguzzoni, A.; Busà, M. G.; 1995b. Correlati acustici della opposizione di quantità vocalica in area emiliana. *Rivista Italiana di Dialettologia*, 19, 7-39.
- Uguzzoni, A.; Pettorino, M.; Filipponio, L., 1999. On stressed vowel durations, vowel-consonant contact types and syllable shapes in the Italo-Romance area. In *Proceedings of the XIV International Congress of Phonetic Sciences*, San Francisco, 3, 2209-2210.
- Vennemann, T., 1991. Syllable structure and syllable cut prosodies in modern standard German. In P. M. Bertinetto *et al.* (a c. d.) *Certamen Phonologicum II. Papers from the 1990 Cortona Phonology Meeting*, Torino: Rosenberg & Sellier, 211-243.
- Vennemann, T., 2000. From quantity to syllable cuts. On so-called lengthening in the Germanic languages. *Rivista di Linguistica*, 12, 251-282.
- Whalen, D. H.; Levitt, A. G., 1995. The universality of intrinsic F<sub>0</sub> of vowels. *Journal of Phonetics*, 23, 349-366.
- Zmarich, C.; Uguzzoni, A.; Ferrari, V., 2003. Controllo articolatorio della opposizione di quantità vocalica in area emiliana: analisi cinematica dei gesti labiali. In *Atti delle XIII Giornate di Studio del Gruppo di Fonetica Sperimentale*, Pisa: Edizioni ETS, 295-306.