

Lunghezza o frequenza: quale parametro per la prominenz?

Giovanna Marotta[°], Anna Molino[°], Chiara Bertini[^]

[°]Università di Pisa, [^]Scuola Normale Superiore di Pisa *

gmarotta@ling.unipi.it, annamolino@hotmail.it, c.bertini@sns.it

1. ABSTRACT

Vengono qui presi in considerazione alcuni campioni di parlato spontaneo relativi a tre varietà di italiano regionale (pisano, crotonese e vastese) per verificare la rilevanza dei parametri “durata” e “frequenza” nell’espressione e nella percezione della prominenz. L’analisi acustica conferma una differenza statisticamente significativa tra le vocali prominenti e quelle non prominenti per quanto riguarda la durata, mentre per la frequenza i risultati sono meno sistematici.

D’altro lato i risultati del test percettivo effettuato mostrano una maggiore sensibilità alle variazioni di frequenza. Inoltre, l’“educazione musicale” degli ascoltatori si dimostra parametro significativo. I dati raccolti costituiscono un primo punto di partenza per una ricerca futura a più ampio spettro.

2. INTRODUZIONE

Il fulcro di questo studio è la nozione di prominenz, intesa come grado di salienza percettiva assegnato ad alcune parole o sillabe all’interno dell’enunciato. Di norma, il segmento prominente presenta una modifica significativa dei tre parametri acustici: durata, intensità e frequenza. Con particolare riferimento alla lingua italiana, “la varia mescolanza dei tre parametri acustici fondamentali determina comunque quell’effetto uditivo di maggior volume, che viene associato alla percezione della prominenz” (Bertinetto 1981).

L’obiettivo del nostro lavoro può dirsi duplice: in primo luogo, intendiamo esaminare i parametri acustici che veicolano la prominenz. A tale scopo, sono state oggetto di misurazione acustica la “durata” e la “frequenza fondamentale”, poiché studi precedenti hanno sottolineato la particolare salienza di questi due parametri nell’esprimere la prominenz (cfr. ad es. Rietveld e Gussenhoven 1985, Kohler 2008). In secondo luogo, vorremmo estendere la nostra indagine al versante percettivo, allo scopo di studiare l’interazione tra i due parametri sopra menzionati nella percezione della prominenz.¹ La domanda cruciale cui vorremmo provare a dare una prima plausibile risposta è pertanto: qual è il parametro acustico basilare per la percezione della prominenz? Nella nostra indagine sono stati infatti presi in considerazione campioni vocalici tratti dal parlato spontaneo di tre varietà di italia-

* Questo lavoro è stato discusso dalle tre Autrici in maniera congiunta. Ai soli fini accademici, si fa presente che A. Molino ha svolto l’analisi acustica e percettiva, mentre C. Bertini ha condotto l’analisi statistica dei dati, oltre ad aver elaborato lo *script* di *Presentation* per la gestione degli stimoli percettivi. Nella redazione del contributo, A. Molino andrà ritenuta direttamente responsabile per le sezioni da 3 a 7, G. Marotta per le sezioni 2 e 8.

¹ Qualche spunto di riflessione è reperibile in Jones & Munhall (2000; 2005).

no regionale - pisano, crotonese e vastese – per verificare l’influenza della variabile “*provenienza del parlante*” nella produzione e nella percezione della prominente.

A quanto ci risulta, il nostro lavoro rappresenta il primo studio a carattere sperimentale su dati relativi alla lingua italiana volto a tentar di comprendere come vengono percepiti i correlati acustici della prominente.²

3. ALCUNE DEFINIZIONI DI PROMINENZA

E’ opportuno soffermarsi brevemente sul concetto di prominente, dal momento che si tratta di un evento prosodico particolarmente complesso e di difficile definizione. Liberman e Pierrehumbert (1984) identificano la prominente come a “local degree of stress or emphasis”. In questo studio, il termine entra ufficialmente nel novero dei fattori che regolano l’intonazione in inglese, insieme al contorno intonativo (*tune*), la declinazione finale (*declination line*) e l’escursione in frequenza (*pitch range*). La prominente prosodica non ha solo carattere quantitativo, ma anche relativo, poiché, come ascoltatori, percepiamo un elemento come più rilevante solo se viene messo a confronto con gli altri presenti nell’enunciato. Tra le più note definizioni di prominente, si trova infatti quella di Terken (1991), secondo cui “the prominence is the property by which linguistic units are perceived as standing out from their environment”. La prominente è dunque un fenomeno percettivo in grado di consentire a determinate unità linguistiche di “emergere” in un dato contesto.

La prominente è certamente un fenomeno complesso, dal momento che alla sua realizzazione concorrono in maniera integrata la sintassi, la semantica e la pragmatica. L’emittente, infatti, assegna una particolare salienza prosodica a quelle parti del messaggio intorno alle quali intende concentrare il proprio interesse comunicativo (Soriano, 2006; De Dominicis, 2010). L’innalzamento del tono, ad esempio, identifica un punto della stringa sonora su cui il parlante intende trasmettere informazione aggiuntiva a quella veicolata dal lessico; in prospettiva percettiva, si dirà che i punti della sequenza fonica su cui si realizzano i bersagli alti di F0 sono quelli verso cui si concentra l’attenzione dell’ascoltatore, in quanto in contrasto con il profilo melodico discendente non marcato, fisiologicamente determinato (Marotta, 2003).

A tal proposito, un altro aspetto interessante riguarda la relazione tra prominente e piano articolatorio. In un recente studio, Avesani *et al.* (2009) hanno notato che diverse proprietà cinematiche distinguono in modo sistematico le sillabe atone da quelle toniche e da quelle focalizzate. I gesti di apertura e di chiusura delle sillabe prominenti sono progressivamente più lunghi in durata, più ampi in estensione e più veloci rispetto ai movimenti relativi alle sillabe atone. Come era logico attendersi, la prominente consiste pertanto anche in un aumento globale dello sforzo articolatorio.

Il termine prominente può assumere, dunque, diverse accezioni: espediente pragmatico utile a segnalare un elemento ritenuto rilevante dal parlante (Liberman e Pierrehumbert, 1984; Marotta e Sardelli, 2009), porzione percettivamente rilevante (Terken, 1991), fino a meccanismo intorno al quale ruota l’intera struttura ritmica dell’enunciato (Niebuhr, 2009). In questa indagine non intendiamo soffermarci sugli aspetti pragmatici che la prominente può assumere, quanto piuttosto sulle sue caratteristiche acustiche e percettive.

² Gli unici studi in chiave percettiva che conosciamo sono quelli di Calamai e Ricci (2005a; 2005b), con stimoli relativi ad alcune varietà toscane.

4. MATERIALI E METODI

Il corpus utilizzato nell'indagine è stato ricavato da alcuni dialoghi di parlato spontaneo prodotti da tre parlanti, uno per ognuna delle varietà indagate in questa sede. In particolare:

- varietà di Pisa: locutrice di sesso femminile di cinquantacinque anni, proveniente da famiglia pisana, nata e sempre vissuta nella medesima area;
- varietà di Vasto (CH): locutrice di sesso femminile, di sessantadue anni, residente da sempre nel territorio e di famiglia proveniente dalla medesima zona;
- varietà di Crotona: locutore di sesso maschile, studente universitario di venticinque anni, residente a Pisa, con caratteristiche meno marcate in diatopia rispetto a quelle delle parlanti analizzate per le altre due località.

Per quanto riguarda le ragioni della selezione di queste tre varietà, va in primo luogo ricordato che la varietà pisana è stata già oggetto di indagini sperimentali sia nel campo del vocalismo che della prosodia. Com'è ormai noto, il vocalismo tonico della Toscana occidentale presenta alcune caratteristiche peculiari: la vocale /ε/ è spesso resa come [æ] (Calamai, 2004) e le vocali accentate hanno in genere una lunghezza ed una modulazione di F0 maggiori rispetto a quelle delle altre varietà toscane (cfr. Marotta *et al.* 2004). Sulla prosodia del vastese, invece, non ci risultano studi a carattere sperimentale; la varietà di Vasto è stata inserita in questo lavoro fondamentalmente perché si tratta del luogo di provenienza della seconda autrice di questo contributo. La varietà crotonese, infine, è stata considerata per il particolare allungamento che si verifica sulle vocali toniche, come rilevato in studi precedenti che hanno coinvolto anche varietà calabresi limitrofe (cfr. Romito e Trumper, 1989; Mendicino e Romito, 1991; Romito, 1993; Marotta e Sardelli, 2009).

Le vocali analizzate sono state suddivise in tre classi:

- vocali toniche prominenti (Classe A);
- vocali toniche non prominenti (Classe B);
- vocali atone (Classe C).

Le vocali e i loro sintagmi di appartenenza sono stati oggetto di segmentazione ed etichettatura; è stata misurata la loro durata in millisecondi e i valori di F0 in tre punti della porzione stabile della vocale, oltre a quelli massimi e minimi. I dati ottenuti sono stati sottoposti a procedura di normalizzazione, per la "durata" secondo la formula seguente in cui il denominatore è un indice indiretto della velocità d'eloquio:

$$(1) \frac{\text{Durata Vocale}}{(\text{Durata Sintagma}/n.\text{sillabe})}$$

Per la frequenza fondamentale, invece, sono stati calcolati i valori medi tra i tre punti. L'operazione di normalizzazione è stata condotta prendendo in considerazione i valori frequenziali massimi e minimi sia della vocale che del sintagma, secondo la formula:

$$(2) \frac{(\text{max} - \text{min}) \text{Vocale}}{(\text{max} - \text{min}) \text{Sintagma}}$$

5. ANALISI ACUSTICA – RISULTATI

5.1. Durata

Per quanto riguarda il parametro "durata", l'analisi acustica ha confermato l'esistenza di una gerarchia secondo cui le vocali toniche prominenti sono più lunghe delle toniche non

prominenti, che, a loro volta, hanno una durata maggiore delle atone. Questa tendenza si manifesta in tutti i timbri vocalici e in tutte le varietà considerate. Il trattamento statistico dei risultati conferma che il rapporto tra “durata” e “classe vocalica” è significativo nella maggior parte dei casi. Nelle tabelle seguenti riportiamo i risultati dell’applicazione del test *t di Student* ai nostri dati; per una più facile lettura, diamo di seguito la legenda delle sigle impiegate nelle Tabelle: *Gr. Lib.* = gradi di libertà ($N - 1$); *NA*, *NB* = occorrenze vocaliche relative alla classe A e alla classe B; *Media A*, *Media B* = media delle durate di classe A e classe B; *DS A*, *DS B* = Deviazione Standard di classe A e classe B; *t* = valore di *t di Student*; *P* = indice di significatività (*p value*); i dati sottolineati sono quelli significativi; *E.S.*: errore standard sulla differenza tra i valori medi.

	<i>Gr. Lib.</i>	<i>NA</i>	<i>NB</i>	<i>Media A</i>	<i>DS A</i>	<i>Media B</i>	<i>DS B</i>	<i>t</i>	<i>P</i>	<i>E.S.</i>
Durata /a/	58	51	9	100	24	78	29	2.4039	<u>0.0194</u>	9.062
Durata /e/	34	29	7	106	25	61	14	4.4794	<u>>0.0001</u>	9.970
Durata /e/	14	8	8	93	25	54	13	3.9739	<u>0.0014</u>	9.908
Durata /i/	17	13	6	77	19	52	17	2.6209	<u>0.0179</u>	9.235
Durata /o/	21	17	6	76	24	55	18	1.9490	0.0648	10.689
Durata /o/	16	8	10	92	25	54	14	4.2532	<u>0.0006</u>	9.117
Durata /u/	10	5	7	94	10	54	11	6.4244	<u>>0.0001</u>	6.324

Tabella 1: Risultati test t-Student sulle durate normalizzate del parlato spontaneo (Pisa).

	<i>Gr. Lib.</i>	<i>NA</i>	<i>NB</i>	<i>Media A</i>	<i>DS A</i>	<i>Media B</i>	<i>DS B</i>	<i>t</i>	<i>P</i>	<i>E.S.</i>
Durata /a/	34	24	12	85	21	58	9	4.1532	<u>0.0002</u>	6.330
Durata /e/	20	13	9	82	20	54	11	3.7717	<u>0.0012</u>	7.351
Durata /e/	23	16	9	76	23	55	12	2.5194	<u>0.0192</u>	8.305
Durata /i/	26	18	10	102	76	54	10	1.9506	0.0620	24.881
Durata /o/	16	8	10	68	19	52	16	1.8421	0.0841	8.333
Durata /o/	18	11	9	82	17	55	14	3.6595	<u>0.0018</u>	7.293
Durata /u/	10	7	5	65	15	56	12	1.1812	0.2649	7.958

Tabella 2: Risultati test t-Student sulle durate normalizzate nel parlato spontaneo (Vasto).

	<i>Gr. Lib.</i>	<i>NA</i>	<i>NB</i>	<i>Media A</i>	<i>DS A</i>	<i>Media B</i>	<i>DS B</i>	<i>t</i>	<i>P</i>	<i>E.S.</i>
Durata /a/	13	10	5	88	26	49	10	3.2143	<u>0.0068</u>	12.227
Durata /e/	19	11	10	69	20	53	27	1.5062	0.1485	10.176
Durata /i/	12	9	5	69	19	43	9	2.8055	<u>0.0159</u>	9.275
Durata /o/	14	9	7	86	21	58	27	2.4187	<u>0.0298</u>	11.846
Durata /u/	6	2	6	58	4	42	14	1.5420	<u>0.1740</u>	10.700

Tabella 3: Risultati test t-Student sulle durate normalizzate nel parlato spontaneo (Crotone).

Risulta degna di attenzione la distribuzione degli indici di significatività. Nel caso della varietà pisana il *p value* più significativo è quello relativo a /e/, foneticamente equivalente a

Lunghezza o frequenza: quale parametro per la prominenzza?

[æ], in linea con quanto già documentato in letteratura (Calamai, 2004; Marotta et al., 2004). Il *p value* più alto si riscontra, invece, per [a] sia nel vastese che nel crotonese, confermando che quanto più la vocale è bassa, tanto più è lunga. Il dato era del resto atteso, in linea con le nozioni basilari della fonetica tradizionale (cfr. ad es. Grammont, 1933).

5.2. Frequenza

Per quanto riguarda il parametro “frequenza”, le vocali toniche prominenti sono di norma realizzate a frequenze più alte e risultano mediamente più modulate rispetto alle toniche non prominenti e alle atone; tuttavia, il rapporto tra “frequenza” e “classe vocalica” è statisticamente significativo solo in pochi casi; cfr. le tabelle seguenti:

	<i>Gr. Lib.</i>	<i>NA</i>	<i>NB</i>	<i>Media A</i>	<i>DS A</i>	<i>Media B</i>	<i>DS B</i>	<i>t</i>	<i>P</i>	<i>E.S.</i>
F0 /a/	58	51	9	55	49	30	26	1.5033	0.1382	16.716
F0 /e/	34	29	7	51	25	21	17	3.0924	<u>0.0040</u>	9.905
F0 /e/	14	8	8	51	31	27	19	1.7769	0.0973	12.944
F0 /i/	17	13	6	52	23	28	14	2.3375	<u>0.0319</u>	10.218
F0 /o/	21	17	6	50	24	33	19	1.5489	0.1363	10.969
F0 /o/	16	8	10	39	22	30	28	0.7319	0.4748	12.263
F0 /u/	10	5	7	49	27	30	16	1.5347	0.1559	12.306

Tabella 4: Risultati test t-Student su F0 normalizzata del parlato spontaneo (Pisa).

	<i>Gr. Lib.</i>	<i>NA</i>	<i>NB</i>	<i>Media A</i>	<i>DS A</i>	<i>Media B</i>	<i>DS B</i>	<i>t</i>	<i>P</i>	<i>E.S.</i>
F0 /a/	34	24	12	37	23	25	14	1.6947	0.0993	7.229
F0 /e/	20	13	9	42	24	19	19	2.4228	<u>0.0250</u>	9.652
F0 /e/	23	16	9	42	25	24	13	1.8946	0.0708	9.068
F0 /i/	26	18	10	41	33	28	18	1.1049	0.2793	11.293
F0 /o/	16	8	10	35	23	30	17	0.4640	0.6489	9.590
F0 /o/	18	11	9	38	23	26	10	1.4696	0.1589	8.227
F0 /u/	10	5	7	49	27	30	16	1.5347	0.1559	12.306

Tabella 5: Risultati test t-Student su F0 normalizzata del parlato spontaneo (Vasto).

	<i>Gr. Lib.</i>	<i>NA</i>	<i>NB</i>	<i>Media A</i>	<i>DS A</i>	<i>Media B</i>	<i>DS B</i>	<i>t</i>	<i>P</i>	<i>E.S.</i>
F0 /a/	13	10	5	27	21	22	18	0.5622	0.5823	9.605
F0 /e/	19	11	10	35	24	18	15	1.9802	0.0624	8.787
F0 /i/	12	9	5	30	27	16	13	1.1014	0.2923	12.811
F0 /o/	14	9	7	40	42	34	35	0.2928	0.7740	19.790
F0 /u/	6	2	6	58	4	42	14	1.5420	0.1740	10.700

Tabella 6: Risultati test t-Student su F0 normalizzata del parlato spontaneo (Crotona).

Il primo dato degno di nota è di nuovo relativo alla vocale /e/ del pisano, che mostra un indice statisticamente significativo per la modulazione di frequenza, collimando con il *p value* significativo già riscontrato nella durata. Questo risultato concorda con quanto rilevato

nella letteratura precedente sulla varietà pisana (cfr. Marotta, Calamai e Sardelli, 2004): la vocale /ɛ/, prodotta come [æ] dalla nostra locutrice, risulta ancora una volta particolarmente lunga e modulata.

Va altresì sottolineata la mancanza di un indice statisticamente significativo in rapporto alla variazione di F0 per la varietà crotonese. Tale risultato può essere dovuto alle caratteristiche vocali del soggetto, il cui *pitch range* si attesta su livelli mediamente bassi, anche in considerazione del suo genere maschile.

Sulla base dei dati acustici rilevati, sembrerebbe dunque che l'indice primario di prominenza per tutte e tre le varietà sia la "durata", coerentemente con quanto riportato in letteratura per l'italiano (Bertinetto, 1981).

6. ANALISI PERCETTIVA – MATERIALI E METODI

6.1. *Materiali*

Poiché il fenomeno della prominenza non può prescindere dalla dimensione uditiva, abbiamo deciso di testare sul piano percettivo i risultati ottenuti nell'analisi acustica. A tale proposito, abbiamo messo a punto un test sperimentale, con lo scopo di verificare quale dei due indici presi in considerazione ("durata" e "frequenza") avesse maggiore rilevanza per gli ascoltatori italiani. I dati di questo esperimento possono inoltre essere utili per verificare quale sia la soglia minima di discriminazione della durata vocalica, dal momento che in questo ambito ancora oggi si fa riferimento ai lavori ormai datati di Lehiste (1970) e Rossi (1972), relativi alle cosiddette *Just Noticeable Differences*, secondo cui la soglia minima di percezione si colloca tra i 10 e i 40 ms.³

Sono state pertanto selezionate una serie di vocali toniche prominenti in sillaba aperta, in cui sono stati modificati i parametri "durata" e "frequenza" (per i dettagli, cfr. Marotta et al., in corso di stampa). La durata è stata ridotta in due passaggi successivi di 30 ms ciascuno (per cui la prima risulta più breve di 30 ms, la seconda di 60 ms), grazie all'utilizzo di uno *script* del software *Praat*. Dopo la manipolazione, i segmenti sono stati reinseriti nel contesto originario. In questo modo, sono state proposte agli ascoltatori le parole originali, con la sola eventuale modifica della durata vocalica, anziché logotomi o altri succedanei del parlato, per ridurre al minimo il rischio che il soggetto percepisse la parola *target* come innaturale (cfr. Van Dommelen, 1993; Lehnert - Le Houillier, 2007). La frequenza fondamentale è stata manipolata livellando la curva intonativa sui valori, rispettivamente, massimo e minimo, ed invertendo l'andamento di F0 rispetto all'originale (per cui, ad es., se l'andamento originale era ascendente, lo stimolo modificato aveva un andamento discendente). Anche in questo caso, ci siamo avvalsi del software *Praat*, selezionando l'opzione *Stylize Pitch* (2 ST). Abbiamo quindi creato tre blocchi distinti di stimoli, ciascuno corrispondente ad una varietà.

I partecipanti al test hanno ascoltato 161 coppie di parole, uguali oppure diverse per durata o frequenza. Il soggetto doveva indicare se i due stimoli ascoltati fossero uguali o diversi. Tra uno stimolo e l'altro intercorreva un secondo, mentre il tempo di risposta era di due secondi, seguiti da una pausa di silenzio di altri due secondi. È stato consentito ascoltare gli stimoli una sola volta. Al termine di ogni blocco, il soggetto poteva fare una pausa. Il

³ Alcuni studi più recenti sono dedicati all'interazione tra *JND* e velocità di elocuzione; cfr. ad es., Drake e Botte (1993), Janse (2004), Quené (2007).

Lunghezza o frequenza: quale parametro per la prominenza?

test è stato gestito tramite il software *Presentation*® (www.neurobs.com, versione 14.5), che ha consentito la presentazione delle coppie e dei tre blocchi in ordine casuale e la registrazione delle risposte. I risultati sono stati quindi trasferiti in un foglio *Excel*, mediante il quale è stato possibile procedere all'analisi dettagliata dei dati.

6.2. I partecipanti

I partecipanti hanno svolto il test presso il Laboratorio di Fonetica del Dipartimento di Linguistica dell'Università di Pisa. L'esperimento è stato condotto su un campione composto da 35 ascoltatori, tutti normoudenti, di età compresa tra i 20 e i 40 anni, provenienti da diverse parti d'Italia ma tutti residenti in Toscana, per la maggior parte a Pisa. In particolare, 7 sono originari di Pisa e provincia, 6 provengono da altre zone della Toscana e 22 da altre zone d'Italia.

Ognuno di essi è stato invitato a leggere un foglio di istruzioni e a compilare una scheda sociolinguistica per il reperimento di informazioni utili all'indagine (es. provenienza, grado d'istruzione, varietà materna d'italiano ed eventuali altre varietà conosciute). In particolare, abbiamo concentrato la nostra attenzione sulla variabile "educazione musicale", che abbiamo inserito anche nell'analisi dei risultati. Vari studi (cfr. Rauscher e Hinton, 2003; Schon et al., 2004; Pape, 2005, 2008) hanno infatti dimostrato che i musicisti e, in generale, i soggetti che hanno conoscenze in campo musicale mostrano una maggiore abilità nel discriminare differenze anche minime relative all'escursione tonale. Tra i partecipanti, 17 hanno dichiarato di aver ricevuto un'educazione musicale.

In questa sede per limiti di spazio, non sarà considerata la variabile "provenienza degli ascoltatori", per la quale si rinvia a Marotta et al. (in corso di stampa).

7. ANALISI PERCETTIVA – RISULTATI

7.1. Durata

I risultati ottenuti mettono in evidenza in primo luogo che la maggior parte degli errori è stata commessa su coppie di stimoli in cui A è diverso da B. In particolare, i partecipanti hanno incontrato maggiori difficoltà nel riconoscimento delle differenze del parametro "durata", se si fa riferimento alle due varietà di Pisa e Vasto; la situazione si ribalta nel caso della varietà crotonese, come dimostrano le figure seguenti:

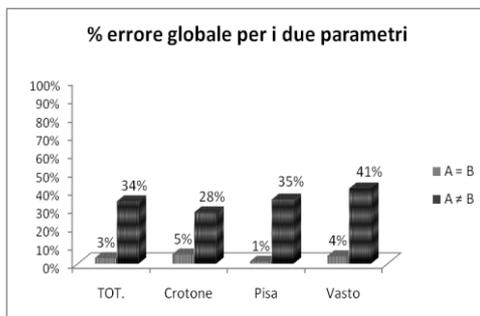


Figura 1: % errore "durata" vs. "frequenza"

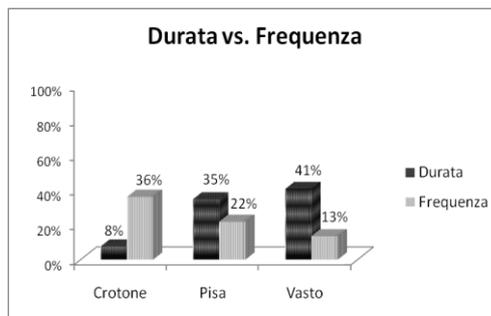


Figura 2: % errore "varietà"

Andrà in primo luogo osservato che, sulla base dei risultati ottenuti nell'analisi acustica, percentuali d'errore così elevate per la durata erano del tutto inattese. Se si entra nel detta-

glio dei risultati, si può notare come, comprensibilmente, lo scarto di 60 ms sia stato riconosciuto con minore difficoltà rispetto a quello di 30 ms. Va tuttavia osservato che la percentuale di errore aumenta sensibilmente (76%) soprattutto quando vengono messi a confronto due stimoli già manipolati (cfr. Figura 3). A tal proposito, abbiamo ritenuto opportuno eliminare i risultati provenienti da quest'ultimo tipo di combinazione, per evitare di falsare l'esito generale dell'esperimento. La proporzione resta stabile anche facendo riferimento alle singole varietà, come evidenziato nella Figura 4, in cui emerge di nuovo la percentuale d'errore notevolmente più bassa nel caso del crotonese, rispetto alle altre due varietà considerate.

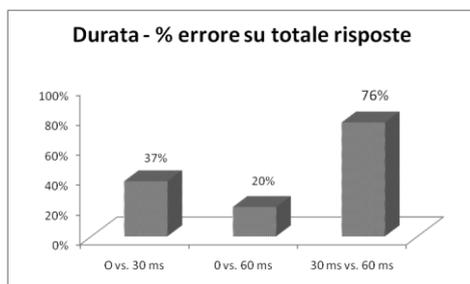


Figura 3: % errore "durata"

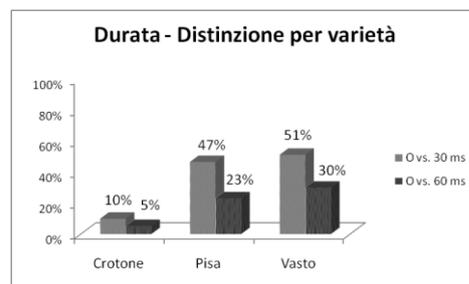


Figura 4: % errore "varietà"

7.2. Frequenza

Per quanto riguarda i risultati ottenuti sul parametro "frequenza", si osserva in primo luogo una percentuale d'errore più bassa nel caso in cui gli ascoltatori hanno dovuto riconoscere gli andamenti melodici in cui il *pitch* era stato invertito. Come si può rilevare dalla Figura 5 (in cui O = stimolo originale, PA = *pitch* innalzato, PB = *pitch* abbassato, PI = *pitch* invertito; cfr. § 6.1), la percentuale di errore nel compito di discriminazione è invece notevolmente più alta se si tratta di coppie in cui l'originale è confrontato con il *pitch* livellato sul valore massimo o minimo.

In secondo luogo, abbiamo riscontrato una percentuale di errore particolarmente elevata nell'ascolto e conseguente valutazione degli stimoli crotonesi (cfr. Figura 6), diversamente da quanto rilevato per la durata (cfr. § 7.1.).

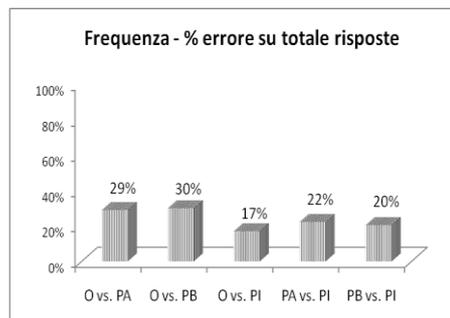


Figura 5: % d'errore "frequenza"

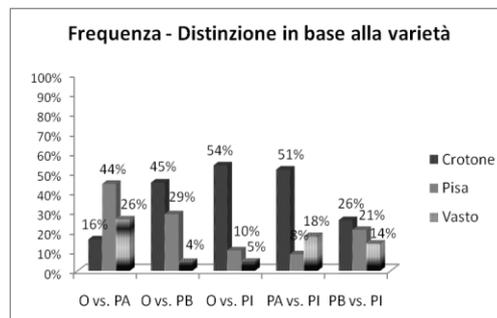


Figura 6: % d'errore "varietà".

7.3. L'educazione musicale

I risultati ottenuti sulla variabile “*educazione musicale*” hanno in buona parte confermato l'ipotesi di partenza: questo fattore incide sulla percezione degli stimoli linguistici da parte dei partecipanti, come ben illustra la Figura 7.

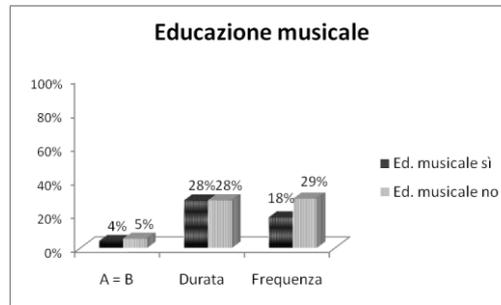


Figura 7: % errore “*educazione musicale*”

I soggetti che hanno ricevuto un'educazione musicale sono più sensibili alle variazioni tra i due stimoli ascoltati per quanto riguarda la “*frequenza*”, come evidenziato nella Figura 7. Le percentuali d'errore relative al parametro “*durata*” sono invece equivalenti. A livello statistico, l'esito del test per il calcolo del χ^2 conferma che non c'è dipendenza tra le variabili “*durata*” e “*educazione musicale*” ($\chi^2 = 0.563$, $p = 0.453$), mentre il rapporto diventa significativo con la “*frequenza*” ($\chi^2 = 31.527$, $p = 0.000$).

8. CONCLUSIONI

L'analisi acustica condotta sui materiali relativi alle tre diverse varietà di italiano considerate (pisano, vastese e crotonese) ha indicato la “*durata*” vocalica quale indice primario di prominenzza. D'altro canto, l'esperimento percettivo condotto indica piuttosto la rilevanza del parametro frequenziale: gli ascoltatori si sono infatti rivelati più sensibili alle variazioni di F0 rispetto a quelle di durata, contrariamente alle previsioni che si potevano formulare sulla base dei risultati ottenuti dall'analisi acustica. Ciò sembra valere in particolare per le varietà di Pisa e Vasto, mentre la tendenza si ribalta se si fa riferimento ai dati relativi al crotonese. In merito a questa discrepanza, pare opportuno fare riferimento alla variabile “*genere del parlante*”: come si ricorderà, per le varietà di Pisa e Vasto, si tratta di parlanti di genere femminile, mentre il parlante di Crotona è di genere maschile, dunque presenta *per se* una frequenza fondamentale media più bassa ed un'escursione tonale assai inferiore rispetto a quella delle altre due locutrici.

I nostri risultati relativi agli stimoli del pisano e del vastese sono del tutto compatibili con quelli recentemente ottenuti da Kohler (2008) in un test percettivo in cui il bisillabo logotomico “*baba*” è stato modificato in durata, F0 e intensità: la “*frequenza*” risulta l'indice più potente sul piano percettivo. Sulla base dei nostri risultati, sembra inoltre di poter evincere che in un compito di discriminazione qual è quello qui adottato, l'inversione dell'andamento melodico viene più facilmente riconosciuto dai soggetti, mentre il livellamento sui valori massimi o minimi del *pitch* pone maggiori problemi (cfr. § 7.2.).

Le risultanze empiriche raccolte sono del resto coerenti con quanto riportato tradizionalmente in letteratura per l'italiano (cfr. ad es. Bertinetto 1981); vale inoltre la pena di ricordare che un recente esperimento percettivo condotto da Vanrell *et al.* (in corso di stampa) ha dimostrato che la durata è il parametro fondamentale anche nella discriminazione tra due diversi accenti tonali, dal momento che gli ascoltatori basano il loro giudizio sulle informazioni di durata + allineamento più che su quelle del solo allineamento o di allineamento + *scaling*.

Un risultato originale, che ci preme pertanto sottolineare, concerne la variabile “*educazione musicale*”. I nostri risultati sono infatti in linea con quanto sostenuto recentemente da Rauscher e Hinton (2003), da Schon *et al.* (2004) e Pape (2008): i soggetti che hanno studiato musica mostrano capacità percettive molto fini in ambito prosodico. Infatti, una precoce esposizione alla musica è in grado di influenzare la soglia di discriminazione uditiva soprattutto nei compiti di discriminazione nelle variazioni tonali. In generale, riteniamo che negli esperimenti percettivi siano da tener presenti anche le variabili dipendenti dal soggetto, in primo luogo quelle culturali come nel nostro caso.

In conclusione, possiamo affermare che sul piano acustico la prominenzza sembra essere veicolata *in primis* dalla durata, anche se affiancata di solito dalla modulazione di frequenza. In particolare, le vocali toniche prominenti sono le più lunghe in assoluto, seguite dalle vocali toniche non prominenti (di durata, in genere, compresa tra i 100 e i 120 ms.), mentre le atone sono le più brevi (inferiori in media ai 100 ms.). Anche per F0 è stato rilevato uno scarto notevole tra i valori registrati per le vocali toniche prominenti rispetto alle vocali non prominenti e alle atone, così come ci attendevamo. Le prime, infatti, presentano un andamento più modulato rispetto alle altre, oltre a frequenze medie più elevate. Tuttavia, mentre la relazione tra prominenzza e durata è significativa nella maggior parte dei casi, il rapporto tra prominenzza e F0 non fa registrare risultati altrettanto rilevanti a livello statistico.

Sul piano percettivo il quadro si presenta più complesso: stando a quanto è emerso da questo studio, che per molti aspetti può considerarsi quale studio-pilota, la rilevanza di un parametro rispetto all'altro (lunghezza *versus* F0) sembra essere relativa più alla natura dello stimolo che alla varietà ascoltata. Si conferma così ancora una volta il noto carattere relativo delle grandezze prosodiche. Ad esempio, uno scarto come quello di 30 ms è percepito con notevole difficoltà sulle vocali iperlunghe della parlante pisana, ma con relativa facilità sui dati del parlante crotonese, per il quale gli stimoli originali sono più brevi rispetto a quelli delle altre due varietà. La situazione risulta speculare per le variazioni di frequenza: l'inversione dell'andamento non viene riconosciuta alle frequenze basse in cui sono prodotti gli stimoli del parlante crotonese, contrariamente a quanto avviene per gli stimoli delle locutrici pisana e vastese, che presentano escursioni del *pitch* più marcate.

Inoltre, è necessario sottolineare che lunghezza e frequenza sono grandezze intrinsecamente diverse e diversamente manipolabili (cfr. Jones e Munhall 2000; 2005). In particolare, nei test che intendiamo svolgere in futuro, riteniamo idoneo apportare variazioni di durata in termini percentuali più che assoluti.

Per concludere, alla domanda che ci eravamo poste in apertura, cioè “Qual è il parametro basilare per la percezione della prominenzza?”, riteniamo che i nostri risultati non consentano di rispondere in modo chiaro ed esaustivo. I dati raccolti possono esser considerati come un punto di partenza per altri studi, di maggiore ampiezza e più mirati. Anche per la determinazione di una soglia minima di discriminazione delle variazioni di durata, si rendono necessarie ulteriori analisi percettive, che dovranno a nostro parere prendere in esame

Lunghezza o frequenza: quale parametro per la prominenzza?

in particolare le vocali toniche non prominenti, nelle quali non si verifica quel fenomeno di iperallungamento tipico dei segmenti prominenti, che può mascherare effetti non secondari nella percezione della lunghezza segmentale.

BIBLIOGRAFIA

- Avesani, C., Vayra, M., Zmarich, C. (2009), Coordinazione vocale – consonante e prominenza accentuale, in Ferrari, G., Mosca M. (eds.), *Linguistica e modelli tecnologici di ricerca*, Roma, Bulzoni, pp. 353-386.
- Bertinetto, P. M. (1981), *Strutture prosodiche dell'italiano. Accento, quantità, sillaba, giuntura, fondamenti metrici*, Firenze, Accademia della Crusca.
- Calamai, S. (2004), *Il vocalismo tonico dell'area pisana e livornese. Aspetti storici, percettivi e acustici*, Alessandria, Edizioni dell'Orso.
- Calamai, S., Ricci, I. (2005a), Sulla percezione dei confini vocalici in Toscana: primi risultati, in Cosi, P. (ed.), *Atti del I Convegno AISV*, Torriana (RN), EDK Editore.
- Calamai, S., Ricci, I. (2005b), Un esperimento di matched-guise in Toscana, *Studi Linguistici e Filologici Online* 3.1, pp. 63-105 (www.humnet.unipi.it/slifo.html).
- De Dominicis, A. (2010), *Intonazione*, Roma, Carocci.
- Drake, C., Botte, M. C. (1993), Tempo sensitivity in auditory sequences: evidence for a multiple-look model, *Perception and Psychophysics*, 54/3, pp. 277 - 286.
- Grammont, M (1933), *Traité de phonétique*, Paris, Delagrave.
- Janse, E. (2004), Word perception in fast speech, *Speech Commun*, 42, pp. 155 – 173.
- Jones, A. J., Munhall, K. G. (2000), Perceptual calibration of F0 production: evidence from feedback perturbation, *JASA* 108, pp. 1246-1251.
- Jones, A. J., Munhall, K. G. (2005), Remapping Auditory-Motor Representations in Voice Production, *Current Biology*, 15, pp. 1768-1772.
- Kohler, K.J. (2008), The Perception of Prominence Patterns, *Phonetica* 65, pp. 257-269.
- Lehiste, I. (1970), *Suprasegmentals*, Cambridge, Massachusetts, Cambridge, MIT Press.
- Lehnert – Le Houillier, H. (2007), The influence of dynamic F0 on the perception of vowel duration: cross linguistic evidence, in *Proceedings of the 16th International Congress of Phonetic Sciences*, Saarbrücken, Germany, pp. 757 – 760.
- Lieberman, M., Pierrehumbert, J. (1984), Intonational Invariance under changes in pitch range and length, in Aronoff, M., Oherle, R. T. (eds.), *Language Sound Structure*, Cambridge, MIT Press, pp. 157 – 233.
- Marotta, G. (2003), “L’illusione prosodica”, in “Studi e saggi linguistici”. Atti del Convegno di Studi in memoria di Tristano Bolelli, Pisa, ETS, pp. 237 – 258.
- Marotta, G., Calamai, S., Sardelli, E. (2004), Non di sola lunghezza. La modulazione di F0 come indice socio-fonetico, in De Dominicis, A., Mori, L. e Stefani, M. (eds.), *Costituzione, gestione e restauro di corpora vocali*, Roma, Esagrafica, pp. 210-215.
- Marotta, G., Sardelli, E. (2009), Prosodiatopia: parametri prosodici per un modello di riconoscimento diatopico, in Ferrari, G., Mosca M. (eds.), *Linguistica e modelli tecnologici di ricerca*, Roma, Bulzoni, pp. 411 – 436.

Lunghezza o frequenza: quale parametro per la prominenzza?

Marotta, G., Molino, A., Bertini, C. (in corso di stampa), Sulla percezione dei parametri prosodici in italiano, *Italia Dialettale*, vol. LXXI.

Mendicino, A., Romito, L. (1991), Isocronia e base di articolazione: uno studio su alcune varietà meridionali, *Quaderni del Dipartimento di Linguistica, Università della Calabria*, Serie Linguistica 3, pp. 49 – 67.

Niebuhr, O. (2009), F0 – based rhythm effects on the perception of local syllable prominence, *Phonetica*, n. 66, pp. 95 – 112.

Pape, D. (2005), Is pitch perception and discrimination of vowels language-dependent and influenced by the vowels spectral properties?, in *Proceedings ICAD*, Limerick, Ireland.

Pape, D. (2008), The native language influence on perceptual Intrinsic Pitch: Crosslinguistic data from German, Italian, Portuguese, and Spanish, in *Proceedings of the 4th Conference on Speech Prosody*, Campinas, SP, Brazil.

Quené, H. (2007), On the Just Noticeable Difference for tempo in speech, *Journal of Phonetics*, 35, pp. 353 – 362.

Rauscher, F.H., Hinton, S.C. (2003), Type of music training selectively influences perceptual processing, in *Proceedings of the European Society for the Cognitive Sciences of Music*, Hannover, Germany, 2003.

Rietveld, A. C. M., Gussenhoven, C. (1985), On the relation between speech excursion size and pitch prominence, *Journal of Phonetics*, 13, pp. 299 – 308.

Romito, L. (1993), Cenni sui correlati elettroacustici dell'accento in alcune varietà di italiano, in *Atti delle IV Giornate di Studio del GFS*, pp. 107 – 119.

Romito, L., Trumper, J. (1989), Un problema della coarticolazione: l'isocronia rivisitata, in *Atti del XVII Convegno dell'Associazione Italiana di Acustica*, pp. 449 – 455.

Rossi, M. (1972), Le seuil differential de durée, in Valdman, A. (ed.), *Papers in Linguistics and Phonetics in memory of Pierre Delattre*, Paris, The Hague, pp. 431-50.

Schon, D., Magne, M., Besson, M. (2004), The music of speech: Music training facilitates pitch processing in both music and language, *Psychophysiology*, vol. 41, n. 3, pp. 341-349.

Sorianello, P. (2006), *Prosodia. Modelli e ricerca empirica*, Roma, Carocci.

Terken, J. (1991), Fundamental Frequency and Perceived Prominence, *Journal of the Acoustical Society of America*, 89, pp. 1768 – 1776.

Van Dommelen, W. (1993), Does dynamic F0 increase perceived duration? New light on an old issue, *Journal of Phonetics*, 21, pp. 367 – 386.

Vanrell M.M., Stella, A., Gili Fivela, B., Prieto P. (in corso di stampa), Universal and language-specific strategies in the prosodic marking of contrastive focus, *Language and Speech*.