SCALING E ALLINEAMENTO DEI BERSAGLI TONALI: L'IDENTIFICAZIONE DI DUE ACCENTI DISCENDENTI

Barbara Gili Fivela
Università degli Studi di Lecce
gili@sns.it

1. SOMMARIO

Gli esperimenti descritti in questo lavoro sono volti a verificare l'esistenza di un salto categoriale nella percezione di configurazioni che risultano distinte dal punto di vista delle caratteristiche e della funzione ad essi associata in fase di produzione, ma che effettivamente, in alcuni contesti, sembra possano differenziarsi anche solo per la forza illocutoria che esprimono. Nello studio sono considerati due accenti che presentano toni diversi associati ed allineati alla sillaba – rispettivamente un tono alto ed un tono basso. Considerando un approccio autosegmentale-metrico, gli accenti sono analizzati come bitonali e categorizzati come H*+L e H+L*.

I risultati mostrano che nel caso di un compito di identificazione relativo a questi due accenti discendenti, che in fase di produzione sono normalmente realizzati in contesti diversi ed esprimono funzioni diverse, i soggetti non mostrano di percepire categorie distinte. Nel contributo sono discusse diverse possibilità in relazione all'interpretazione di questi risultati, tra le quali si considerano sia quelle comunque coerenti con l'esistenza di due categorie fonologiche distinte, che quelle in linea con la considerazione dei due accenti come varianti allofoniche di una stessa categoria. In ogni caso, la variazione di altezza tonale svolge un ruolo nella percezione dei parlanti. Questo ruolo può corrispondere a quello di un correlato di per sé rilevante nell'identificazione delle configurazioni, ma anche a quello di un parametro legato alla variazione paralinguistica – in questo caso alla forza illocutoria – che incide solo indirettamente sulla percezione delle configurazioni accentuali.

2. INTRODUZIONE

Gli esperimenti descritti in questo contributo rientrano in un più ampio progetto di ricerca volto ad indagare le relazioni tra variazione tonale ed interpretazione funzionale, indipendentemente da considerazioni relative alla percezione della prominenza degli accenti, legata sicuramente anche a parametri come la durata e l'intensità. Una prima indagine è descritta in Gili Fivela (2005) in relazione a due accenti caratterizzati dalla presenza di un bersaglio tonale alto associato alla sillaba accentata. In quella sede, così come in questo contributo, gli esperimenti descritti consistevano in test di identificazione a risposta forzata, effettuati al fine di verificare l'esistenza di categorie e di individuare i loro confini percettivi. I test di identificazione permettono di verificare che un '*continuum*' di stimoli venga effettivamente categorizzato dai soggetti mediante il riconoscimento di due insiemi discreti ai quali sono riportati tutti gli esempi, anche quelli dalle caratteristiche intermedie. Come già osservato nel contributo menzionato in precedenza, tuttavia, data la natura paralinguistica di molte variazioni intonative e, comunque, la corrispondenza non biunivoca tra modulazione melodica e funzione linguistica, gli esperimenti percettivi di identificazione risultano particolarmente problematici. La difficoltà principale consistite solitamente nel trovare il giusto quesito da porre ai soggetti per fare in modo che riescano a concentrare l'attenzione sulla variazione delle caratteristiche funzionali a seconda della

modificazione delle caratteristiche formali degli accenti. Tuttavia, come risulterà chiaro dalla discussione dei risultati, non si tratta dell'unica difficoltà. Nella presente indagine, infatti, sono considerati accenti che in produzione risultano formalmente e funzionalmente distinti, ma questo, in unione ad un quesito apparentemente adatto a far emergere la differenza tra le due configurazioni, non è sufficiente a indurre una percezione di tipo categoriale. Il problema può essere ovviamente dovuto a limiti del disegno sperimentale o al fatto che i due accenti non rappresentino categorie fonologiche diverse, ma non si può escludere che sia invece legato ad una intrinseca inadeguatezza dei test di identificazione a risposta forzata per lo studio dell'intonazione.

3. CATEGORIZZAZIONE LINGUISTICA E VARIAZIONI DI ALLINEAMENTO E 'SCALING' DEGLI EVENTI TONALI

L'importanza della temporizzazione tra eventi tonali ed unità prosodiche nella categorizzazione degli eventi intonativi è stata ampiamente dimostrata, anche in lavori sviluppati all'interno di quadri teorici diversi (Pierrehumbert, 1980; Cruttenden, 1986; 't Hart *et al*., 1990). Nelle analisi di tipo autosegmentale-metrico (Bruce, 1977; Pierrehumbert, 1980), quadro teorico di riferimento del presente contributo, l'allineamento di eventi tonali ed unità prosodiche è fondamentale al fine di stabilire quale evento sia associato all'unità metrica rilevante. Anche se diversi lavori in letteratura evidenziano la presenza di casi nei quali è impossibile individuare l'esatta sincronizzazione di un evento tonale con l'unità prosodica di riferimento, ad esempio la sillaba accentata, e sottolineano la necessità di non considerare l'allineamento degli eventi tonali come un indice acustico dell'associazione fonologica – a titolo esemplificativo si veda il lavoro di Arvaniti e colleghi (Arvaniti *et al*., 2000) – resta il fatto che l'allineamento degli eventi tonali rispetto all'unità prosodica permette di distinguere gli accenti ed influenza in modo rilevante la scelta delle etichette con le quali indicare le categorie accentuali stesse. A differenza di ciò che avviene in ambito segmentale, infatti, il sistema di trascrizione intonativo di tipo autosegmentale, per quanto di impianto fonologico, codifica in modo esplicito alcune caratteristiche fonetiche, ad esempio, il fatto che il tono sia alto, basso, ascendente, ecc. (per una discussione più ampia si veda Gili Fivela, in stampa).

Alla variazione di altezza dei bersagli tonali è invece spesso attribuito un ruolo secondario rispetto a quello dell'allineamento. Sicuramente l'altezza tonale è il termine di riferimento per differenziare gli accenti nei termini di livelli tonali, ad esempio L* ed H* (Gussenhoven & Rietveld, 2000), ma normalmente non è considerata come rilevante per differenziare categorie accentuali in cui i toni associati all'unità di riferimento siano realizzati sullo stesso livello, ad esempio H* e H*+L. Fanno eccezione alcuni lavori relativi alla differenziazione, in inglese, di accenti come H* e L+H* che, peraltro, molti studiosi non considerano come categorie fonologiche differenti (si consideri Bartels & Kingston (1994) *vs.* Ladd & Schepman (2003)).

In letteratura, è stato anche proposto che la variazione di altezza tonale possa essere interpretata sulla base di veri e propri codici biologici. Gussenhoven (2002) propone l'esistenza di un codice relativo alla frequenza (*frequency code*, per il quale l'autore prende spunto da una prima osservazione di Ohala (1983, 1994)), di un codice legato allo sforzo articolatorio (*effort code*), e di uno dovuto alla dissipazione di energia durante la fonazione (*production code*). Il codice di frequenza è fondamentalmente correlato a differenze anatomiche, in particolare per quanto riguarda le dimensioni dell'apparato fonatorio. In base a questo codice biologico, valori bassi di frequenza fondamentale (f_0) sono associati ad individui dominanti, sono interpretati come indicativi di sicurezza e, quindi,

linguisticamente associati all'espressione di certezza. La grammaticalizzazione di questi aspetti porta ad associare valori bassi di frequenza ad enunciati dichiarativi e valori alti ad enunciati interrogativi. Il codice dello sforzo articolatorio, invece, si origina a partire dalle differenze dovute allo sforzo legato alla produzione verbale. In base a questo codice, l'utilizzo di un ampio intervallo di frequenze è associato ad un evidente sforzo articolatorio ed è interpretato come volto, ad esempio, ad evidenziare elementi particolarmente importanti del messaggio. Per via di un processo di grammaticalizzazione, quindi, l'uso di un ampio spettro di frequenze è associato ai processi di focalizzazione. Il codice di produzione, invece, è legato alla dissipazione di energia durante la fonazione. La diminuzione di energia nella fonazione è associata alla prossimità della fine della produzione verbale, ed interpretata, grazia alla grammaticalizzazione, come espressione di finalità, anche se non si esclude la continuazione. Al contrario, l'alta energia fonatoria è associata alla posizione iniziale all'interno delle parti del discorso e univocamente interpretata come indice della volontà di continuazione.

Nel caso dell'italiano, e del pisano in particolare, è già stata evidenziata l'importanza delle informazioni di *scaling* nell'identificazione di due accenti che appartengono all'inventario fonologico della varietà (Gili Fivela, 2005), ma non si può escludere che nella distinzione delle configurazioni accentuali considerate in quello studio qualche contributo fosse dovuto ad una differenza di tipo pragmatico, cioè al fatto che uno dei due accenti è frequentemente realizzato nella parte iniziale dell'enunciato. Il lavoro suddetto, infatti, prendeva in considerazione due accenti distinti dal punto di vista formale e generalmente prodotti in contesti funzionali differenti. Formalmente, i due accenti, analizzati come H* e H*+L – foneticamente trascritti come [L+]H* e [L+]H*+L, si veda Gili Fivela (in stampa) – differiscono in modo significativo relativamente alle caratteristiche di allineamento e *scaling* dei bersagli tonali, ma sono entrambi caratterizzati dalla presenza di un picco alto, raggiunto grazie ad un innalzamento dei valori di f_0. Inoltre, nel contesto scelto per lo studio (in prossimità di un confine prosodico di sintagma intermedio), gli accenti sono entrambi seguiti da una fase discendente. Dal punto di vista funzionale, H*+L è utilizzato per la focalizzazione contrastiva, mentre H*, nelle dichiarative, può essere realizzato in contorni di continuazione, in posizione prenucleare o per indicare focalizzazione. In quest'ultimo caso, la focalizzazione può anche essere frutto di interpretazione contrastiva, ma solo in contesti di minore forza illocutiva e/o in presenza di un contrasto sintagmatico piuttosto che paradigmatico (Gili Fivela, 1999). L'esperimento percettivo descritto in Gili Fivela (2005) è stato effettuato manipolando i correlati acustici di un enunciato e creando un *corpus* di stimoli che rappresentassero stadi diversi del passaggio graduale dalle caratteristiche di un accento a quelle dell'altro. I soggetti udivano ogni stimolo per due volte e dovevano classificarlo come adatto a rappresentare la correzione perentoria e conclusiva di un enunciato precedente (esempio di interpretazione e realizzazione contrastiva della frase), oppure come enunciato neutro (esempio di interpretazione e realizzazione neutra). In quest'ultimo caso, se anche l'elemento fosse stato interpretabile come utile ad indicare una sorta di correzione, non sarebbe stato percepito come perentorio e conclusivo, perché avrebbe potuto essere seguito e completato con materiale strettamente connesso dal punto di vista prosodico e semantico. I risultati mostrano che l'allineamento del picco permette di identificare le due categorie accentuali – differenziando le caratteristiche dell'enunciato – e che lo *scaling* dei bersagli tonali interagisce con l'allineamento, modificando i valori di riferimento per le due categorie.

Lo studio ha quindi dimostrato che la variazione di altezza tonale incide sulla categorialità della percezione dei soggetti. Tuttavia questo può significare sia che lo *scaling*

rappresenta un correlato rilevante per l'identificazione dei due accenti – nel caso specifico, di accenti 'simili', caratterizzati da un tono alto associato alla sillaba – sia che esso incide solo nel senso di rendere esplicita la variazione paralinguistica, amplificando le differenze tra i due accenti. In questo senso, le caratteristiche di *scaling*, almeno in relazione all'altezza del picco, potrebbero evidenziare l'interpretazione di una delle due configurazioni come indicativa di un atteggiamento maggiormente assertivo, perentorio e conclusivo. Questo sarebbe coerente con le predizioni possibili in base al riferimento ai codici biologici di Gussenhoven (2002). Infatti, in Gili Fivela (2005), la maggior parte dei giudizi a favore dell'interpretazione contrastiva, conclusiva e perentoria era relativa alla configurazione con minori valori di *scaling*, in linea, rispettivamente, con il codice di frequenza e con quello di produzione. Ne consegue che l'altezza tonale possa anche essere messa in relazione con la forza illocutoria.

Molti studi hanno evidenziato che l'andamento intonativo esplicita le caratteristiche dell'atto illocutivo, il modo in cui la frase deve essere interpretata (Austin, 1962 – per gli aspetti prosodici, si vedano i lavori di Cresti e colleghi, ad esempio, Cresti, 2000 e Firenzuoli, 2000). Tuttavia, è possibile anche considerare come rilevante la relazione tra variazioni intonative, in questo caso relative all'altezza tonale, e gradi di forza illocutoria. Secondo Searle & Vanderveken (1985), la forza illocutoria si compone di un *illocutionary point*, il proposito del parlante, e dell'insieme di *presupposition and attitude* che lo accompagna, e che include, ad esempio, la forza dello stesso *illocutionary point*. Infatti, secondo Searle, gli atti illocutivi possono differire in relazione a varie dimensioni, tra cui lo scopo o ragione d'essere dell'atto, e, rilevante per la nostra discussione, la differenza di energia o intensità con cui si presenta lo scopo illocutivo. L'altezza tonale potrebbe quindi essere un parametro rilevante per esprimere l'energia o l'intensità dell'atto illocutivo.

In questo lavoro, si assumerà che la variazione di allineamento sia fondamentale per l'identificazione degli accenti. Alle variazioni di allineamento si accosteranno però quelle di *scaling* dei bersagli, più spesso associate ad informazioni di tipo pragmatico e paralinguistico. Gli obiettivi sono molteplici. Innanzitutto il lavoro è volto a verificare l'esistenza di un salto categoriale nella percezione di configurazioni che risultano distinte dal punto di vista delle caratteristiche e della funzione ad essi associata in fase di produzione, ma che, almeno in alcuni contesti, sembra possano differenziarsi anche solo per la forza illocutoria che esprimono. In secondo luogo, si intende analizzare puntualmente il ruolo dell'allineamento e dello *scaling* nella fase di identificazione degli accenti.

4. CARATTERISTICHE FUNZIONALI ED ACUSTICHE DEGLI ACCENTI CONSIDERATI

Gli accenti presi in esame in questo esperimento sono categorizzati come H*+L ed H+L* grazie all'analisi di produzioni di parlato letto e semispontaneo – dialoghi *Map-Task*[1] (si veda Gili Fivela (2004), lavoro al quale si rimanda per una descrizione completa delle caratteristiche e dei contesti d'uso degli accenti).

La configurazione H*+L è caratterizzata da un innalzamento dei valori di f_0 e da una fase discendente che termina in prossimità della fine della sillaba. Questo porta alla realizzazione di un picco allineato all'inizio della vocale – nel caso si sillabe CV, a circa

[1] I dialoghi *Map-Task* sono registrati chiedendo a due parlanti, un *instruction-giver* ed un *instruction-follower*, di interagire verbalmente per ricostruire un tracciato su una mappa. Soltanto il *giver* possiede le informazioni complete rispetto al tracciato, mentre il *follower* è solo a conoscenza della posizione di alcune icone tra le quali esso si sviluppa.

metà della sillaba accentata – figura 1, in basso. L'accento si compone quindi di due bersagli tonali bassi, ma mentre il secondo è considerato rilevante nell'analisi fonologica della configurazione, il primo non svolge un ruolo determinante all'interno del sistema (Gili Fivela, in stampa).

L'accento H+L* è caratterizzato da un abbassamento dei valori di frequenza fondamentale nella sillaba accentata, a partire da un bersaglio alto allineato alla sillaba precedente – figura 1, in alto. Questo bersaglio può rappresentare la fine di un *plateau*, il valore massimo nel sintagma prosodico, o può non essere particolarmente visibile sul tracciato di frequenza fondamentale per via di precedenti eventi tonali alti. Si tratta, quindi, di un bersaglio la cui altezza tonale è piuttosto variabile ed è già stata messa in relazione con la forza illocutoria del atto linguistico e, in particolare del proposito illocutivo del parlante (Gili Fivela, 2004)[2]. Anche questa osservazione, come quella fatta nel paragrafo 2 in relazione al confronto tra H* ed H*+L, è coerente con quanto proposto da Gussenhoven (2002). Infatti, una più ampia modulazione di frequenza – causata dall'innalzamento del bersaglio che rappresenta l'inizio della fase discendente – è associata ad un aumento dello sforzo articolatorio ed è quindi interpretata come dovuta ad enfasi. In ogni caso, dai dati raccolti per questo studio, risulta che il valore di frequenza fondamentale del bersaglio alto è in media significativamente minore rispetto a quello riscontrato nella configurazione accentuale descritta in precedenza, cioè H*+L (si veda 5.1).

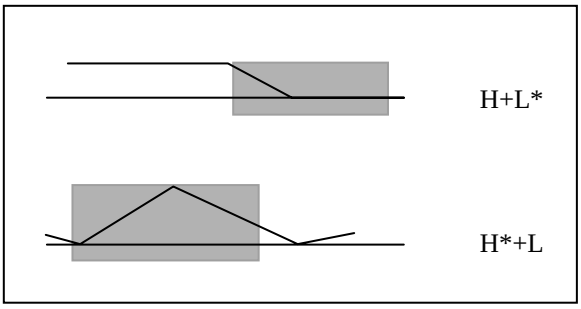

Figura 1: Rappresentazione schematica dell'accento H+L* (in alto) e dell'accento H*+L (in basso) – l'area grigia rappresenta la sillaba nucleare.

Dal punto di vista funzionale, l'accento H*+L esprime focalizzazione contrastiva, e come tale può anche essere realizzato per indicare richieste di conferma nel caso sia utilizzato da parlanti che, come un *instruction follower* nel *Map-Task*, per definizione non possono conoscere con certezza l'informazione relativamente alla quale si esprimono in termini contrastivi. L'utilizzo della configurazione contrastiva da parte di questi parlanti porta ad interpretare l'asserzione come una richiesta di conferma (Gili Fivela, 2004)[3].

[2] Queste osservazioni possono trovare conferma anche nei risultati delle misurazioni compiute da Firenzuoli (2000) sui profili intonativi di ordini ed istruzioni tratti dal *corpus* LABLITA. L'autrice riporta, soprattutto nel caso di locutrici, valori di frequenza media e di escursione di frequenza nettamente superiori negli ordini piuttosto che nelle istruzioni.

[3] Il caso descritto corrisponde quindi all'interpretazione di un tipico enunciato dichiarativo contrastivo per ottenere conferma di una informazione. L'enunciato in questione è caratterizzato dall'accento H*+L e da toni di confine bassi. Nel caso l'accento contrastivo

L'accento H+L* è prodotto in numerosi contesti nella varietà di italiano parlata a Pisa, in unione a diverse combinazioni di toni di confine. Nelle dichiarative, rappresenta l'accento nucleare e può essere seguito sia da toni bassi di confine che da toni alti (o da una sequenza di toni basso-alto), nel caso in cui esprima esplicitamente continuazione. Come già accennato, il bersaglio alto dell'accento è caratterizzato da valori di frequenza fondamentale molto variabili.

Per quanto in fase di produzione verbale i due accenti siano normalmente realizzati in contesti diversi ed esprimano funzioni differenti, in alcuni casi possono anche corrispondere a funzioni parzialmente sovrapponibili. In contesti nei quali si vuole correggere l'interlocutore, ad esempio, è plausibile sia possibile utilizzarli entrambi: H+L*, infatti, come accento nucleare in enunciati dichiarativi neutri, dona prominenza e potrebbe quindi essere interpretato come una correzione 'neutra' o 'poco evidenziata', caratterizzata da una minore forza illocutoria. Peraltro, la potenziale ambiguità tra accenti nucleari *broad* e accenti di focalizzazione – anche se non contrastiva – in posizione finale è già stata evidenziata da Ladd (1996) per la lingua inglese. Il caso quindi non è implausibile, soprattutto qualora H+L* appartenga ad un contesto di elevata forza illocutoria, in cui le affermazioni risultino particolarmente categoriche. Peraltro, i test percettivi descritti nei prossimi paragrafi hanno permesso di verificare che, in un compito di identificazione, i soggetti non mostrano di percepire gli accenti come appartenenti a due categorie distinte.

5. ESPERIMENTI PERCETTIVI

L'obiettivo primario degli esperimenti percettivi descritti nei prossimi paragrafi è di verificare l'effettiva esistenza di un salto categoriale nella percezione di configurazioni che risultano chiaramente distinte dal punto di vista delle caratteristiche e della funzione ad essi associata in fase di produzione verbale. Il secondo intento è di analizzare l'apporto delle caratteristiche di allineamento e di *scaling* nella percezione dei due accenti. Gli esperimenti consistono in test a risposta forzata, effettuati chiedendo a 9 soggetti pisani di giudicare una serie di stimoli dei quali erano state manipolate alcune caratteristiche acustiche. Un primo esperimento è stato effettuato manipolando le caratteristiche di un accento contrastivo, trascritto foneticamente come [L+]H*+L, innalzando gradualmente il livello di f_0 del primo *target* basso ed abbassando gradualmente l'altezza del bersaglio alto al fine di verificare il loro ruolo nell'identificazione dell'accento [L+]H*+L o H+L* [4]. Il secondo esperimento, invece, è stato eseguito manipolando le caratteristiche di un enunciato prodotto con accento H+L*, modificando l'allineamento e lo *scaling* del bersaglio alto ed analizzando l'importanza di questi due fattori nell'identificazione degli accenti H*+L e H+L*. In ogni caso, ai 9 soggetti è stato chiesto di ascoltare ogni stimolo e di stabilire se poteva essere interpretato come un enunciato volto a correggere un potenziale interlocutore in caso di

---

sia realizzato da un parlante che si suppone possa conoscere l'informazione – come un *instruction giver* nel *Map-Task* – l'accento è seguito da toni di confine diversi.

[4] L'unica configurazione che nella varietà pisana di italiano permette di verificare la rilevanza percettiva del bersaglio basso in una sequenza di bersagli basso-alto è l'accento H+L*, unico accento in cui ad un bersaglio alto non corrisponde un precedente bersaglio basso allineato nei pressi della sillaba accentata. Tuttavia le due configurazioni non sono differenti solo per la presenza/assenza di questo bersaglio basso. Differiscono in modo evidente anche per l'allineamento del bersaglio alto, del seguente bersaglio basso e per lo *scaling* del picco.

malinteso legato all'interpretazione della parola bersaglio, quella in corrispondenza della quale erano state eseguite le manipolazioni.

5.1 Materiali

L'enunciato utilizzato come base per le manipolazioni utili alla creazione degli stimoli è *mangia il melone*[5], prodotto da una parlante pisana che non ha poi partecipato all'esperimento percettivo. La parlante ha prodotto almeno tre ripetizioni della frase, con le seguenti interpretazioni:

- Affermazione con focalizzazione neutra: la frase rappresentava la risposta alla domanda *cosa succede?* – si veda un esempio nella figura 2, a sinistra.
- Correzione: la frase rappresentava la risposta alla richiesta di conferma *mangia lamponi?* per la quale era prevista una risposta negativa e l'indicazione dell'oggetto corretto – si veda un esempio nella figura 2, a destra.

Tre ripetizioni di ogni tipo di enunciato sono state etichettate relativamente alla posizione del picco – sia prenucleare, sia nucleare – del bersaglio basso che precede il picco nell'accento contrastivo, e del bersaglio basso che segue il picco in entrambe le configurazioni. Sono stati misurati sia i valori di f_0 dei bersagli tonali che la loro posizione rispetto ai confini sillabici. I valori medi ricavati – riportati nella tabella 1 – sono poi stati considerati per calcolare l'intervallo, in termini di tempo (ms) e di modulazione tonale (Hz), da coprire con la manipolazione delle caratteristiche acustiche al fine di ottenere stimoli che rappresentassero un passaggio graduale dall'accento H+L* all'accento H*+L, e viceversa.

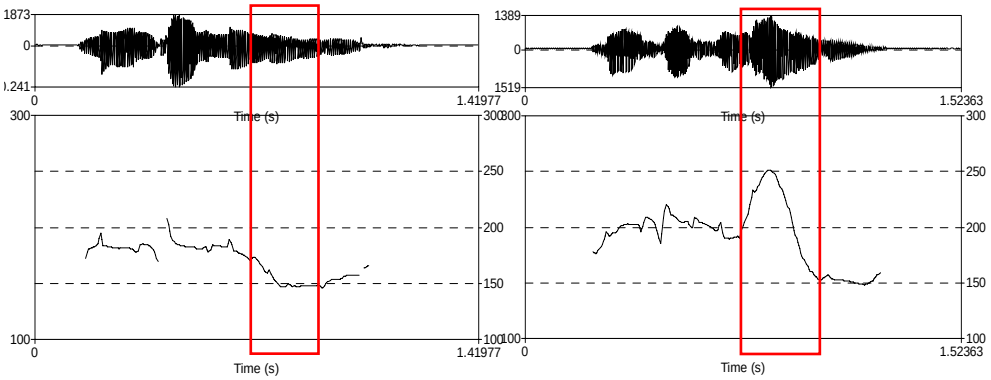

Figura 2: Produzione della frase *Mangia il melone* con interpretazione neutra e realizzazione dell'accento H+L* sulla parola bersaglio (a sinistra); con interpretazione contrastiva e realizzazione dell'accento H*+L sulla parola bersaglio (a destra) – l'area evidenziata in rosso indica la sillaba nucleare.

[5] È stato scelto il termine *melone* piuttosto che *popone*, molto frequente nella varietà di Pisa, per la presenza di soli segmenti sonori che avrebbero reso più attendibile il tracciato di f_0.

	H Prenucleare	L+	H	+L
Neutra	191Hz	-	184 Hz	145 Hz
			Allineamento: fine sillaba prenucleare	
Contrastiva	223 Hz	186 Hz	242 Hz	148 Hz
		Interpolato: 230 Hz	Allineamento: 110ms dopo onset sill.nucl.	

Tabella 1: Valore medio delle misurazioni effettuate in corrispondenza dei bersagli tonali negli enunciati prodotti dalla locutrice pisana.

5.2 Primo esperimento:dall'accento contrastivo all'accento neutro

Un primo esperimento è stato condotto al fine di verificare se l'identificazione degli accenti poteva essere indotta dalle caratteristiche di *scaling* del primo *target* basso in [L+]H*+L, e – alla fine della manipolazione – dal suo allineamento. Innalzando il livello di frequenza del bersaglio basso, infatti, nelle ultime manipolazioni un bersaglio basso, anche se presente, è molto anticipato rispetto alla posizione iniziale. L'ipotesi era che la percezione di due accenti che, foneticamente, sono analizzati come [L+]H*+L e H+L* potesse essere influenzata dalla presenza/assenza di un bersaglio basso allineato a breve distanza dal picco[6]. Inoltre, è stata effettuata un'altra manipolazione al fine di verificare se fosse necessario realizzare una vera e propria fase ascendente per giungere al picco – ovvero che bastasse un *plateau* – e/o che l'altezza del bersaglio alto potesse influire sulla percezione dei parlanti. L'ipotesi era che un accento in cui un bersaglio alto corrispondeva alla fine di un *plateau* ed era realizzato al livello tonale adatto potesse essere interpretato come H+L*, a parte il diverso allineamento del picco.

A partire dall'interpretazione contrastiva della frase *mangia il melone*, in cui un accento [L+]H*+L era allineato sulla parola bersaglio *melone*, sono state quindi effettuate le seguenti manipolazioni:

1) l'altezza del bersaglio basso che precede il picco è stata innalzata in 6 fasi di 9 Hz l'una, al fine di ottenere un'interpolazione diretta tra il picco prenucleare ed il picco realizzato nella sillaba nucleare – si veda la figura 3, schema di sinistra. I 6 stimoli sono stati risintetizzati (sintesi PSOLA) così come lo stimolo di base – per un totale di 7 stimoli che saranno indicati come 0, L1, L2,.., L6[7].

[6] Gili Fivela (in stampa), peraltro, dimostra che il bersaglio basso che precede il picco in un accento [L+]H* realizzato in posizione iniziale assoluta di frase ha caratteristiche di allineamento altamente variabili, e può arrivare a coincidere con il valore di frequenza iniziale dell'enunciato anche nel caso in cui siano disponibili due sillabe atone preaccentuali. Sembra quindi plausibile che le caratteristiche di allineamento di quel bersaglio non siano di fondamentale importanza, e non siano quelle che permettono l'identificazione di due accenti come [L+]H*+L e H+L*. Il ruolo dell'allineamento del picco è studiato nel secondo esperimento – si veda 5.3.

[7] Lo stimolo di base presentava un picco dal valore medio di f_0 più alto del valore medio presentato nella tabella 1. Era realizzato a 250 Hz piuttosto che 242 Hz. Questo ha reso possibile una manipolazione in 6 *step* di 9 Hz l'uno, piuttosto che in 5 *step*.

2) Sono state modificate le caratteristiche dello stimolo corrispondente all'ultima fase di manipolazione della serie precedente – lo stimolo con il valore di f_0 più alto per il primo bersaglio basso (stimolo L6). Al fine di raggiungere il valore medio dell'altezza del bersaglio alto nella configurazione H+L*, è stato abbassato il valore di f_0 del picco in due fasi di 30 Hz ciascuna – si veda la figura 3, schema di destra. I 2 stimoli sono stati risintetizzati (sintesi PSOLA) e saranno indicati come C-pr1 e C-pr2. In C-pr1 il picco ha un valore di frequenza intermedio tra quello di partenza e quello di arrivo, mentre in C-pr2 è realizzato un *plateau* piuttosto che un picco di f_0.

I nove stimoli sintetici sono stati proposti ai 9 soggetti pisani per tre volte in ordine casuale, per un totale di 27 stimoli. I soggetti ascoltavano per due volte ogni stimolo e stabilivano se gli enunciati potessero rappresentare la correzione di un enunciato precedente oppure corrispondessero a enunciati dichiarativi neutri.

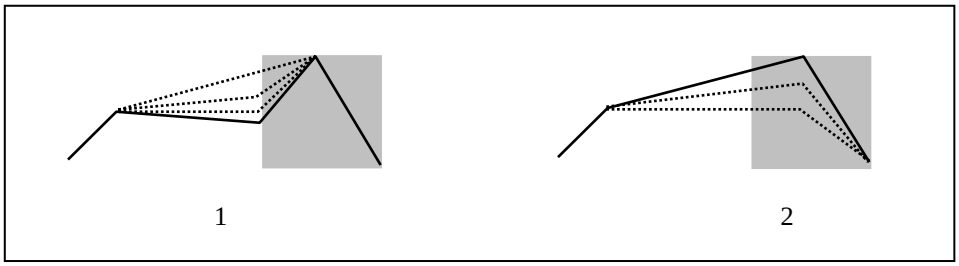

Figura 3: Rappresentazione schematica delle manipolazioni acustiche effettuate – l'area grigia indica la sillaba nucleare.

Nella figura 4 sono riportati due grafici nei quali è indicata la media di tutte le risposte fornite dai soggetti a favore dell'interpretazione contrastiva. Nei grafici, le risposte dei soggetti in relazione ad ogni stimolo sono state espresse sotto forma di un valore compreso tra 0 ed 1, ottenuto dividendo, per ogni soggetto, la somma delle risposte fornite in relazione allo stimolo per il numero totale delle ripetizioni, e sommando i risultati: valori vicino ad '1' indicano che l'enunciato è stato riconosciuto come contrastivo.

Il grafico a sinistra nella figura 4 riporta i dati ricavati per gli stimoli nei quali è stata effettuata la manipolazione descritta in 1. Come si può osservare dal fatto che i valori medi non si attestano mai al di sotto del valore 0,85, aumentare il valore di f_0 del primo bersaglio basso nell'accento [L+]H*+L (manipolazioni in 0, L1…L6) non porta i soggetti a percepire una categoria accentuale diversa da quella di partenza. Lo *scaling* del bersaglio basso, quindi, non influisce sulle risposte dei soggetti.

In modo analogo, il grafico di destra in figura 4 mostra che anche l'abbassamento del valore di altezza del picco – manipolazione in due fasi, descritta al punto 2 – dai valori medi della configurazione contrastiva[8] a quelli medi della configurazione neutra H+L* non porta alla percezione di due categorie diverse. Quindi anche l'altezza del picco non è sufficiente, almeno nel caso in cui rappresenti l'unico correlato, ad influire sui giudizi dei

[8] In questo caso, tuttavia, il picco corrisponde alla parte terminale di una interpolazione lineare con il picco prenucleare – stimolo 'L6' descritto in 1 – e non alla fine di una fase ascendente iniziata nella sillaba pretonica, come accade tipicamente nell'accento [L+]H*+L.

soggetti, che in media si attestano tra lo 0,85 e lo 0,73. Tuttavia, il fatto che lo *scaling* del picco sia simile a quello tipico della configurazione H+L* causa una leggera diminuzione della media di identificazioni a favore dell'interpretazione contrastiva. In particolare, considerando separatamente le risposte date da ogni soggetto, emerge che due pisani su nove non identificano mai stimoli corrispondenti ad una minore altezza tonale del picco come adatti ad una interpretazione contrastiva – si noti, oltre al valore medio, la maggiore deviazione standard nella colonna C-pr2 in figura 4.

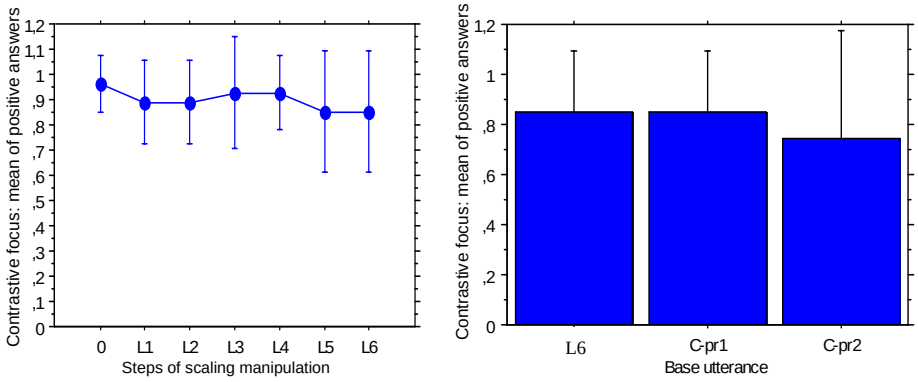

Figura 4: Media e deviazione standard delle risposte dei soggetti a favore dell'interpretazione contrastiva in relazione a stimoli nei quali era stato incrementato il valore di frequenza del bersaglio basso [L+] – step 0, L1…L6 riportati nel grafico di sinistra – e decrementato quello del picco alto – step C-pr1 e C-pr-2 nel grafico di destra.

Le risposte dei parlanti, quindi, non cambiano a seconda che l'altezza del primo bersaglio basso nell'accento [L+]H*+L sia incrementata in termini di frequenza, nemmeno quando si ottenga la diretta interpolazione dal precedente picco di frequenza – un andamento di frequenza che, *scaling* e allineamento del picco a parte, è simile quello riscontrato nella configurazione H+L*. Per quanto riguarda gli stimoli in cui lo *scaling* del picco è stato manipolato in modo da arrivare a corrispondere alle caratteristiche della configurazione neutra H+L*, i valori medi delle risposte dei soggetti indicano soltanto una lieve diminuzione delle interpretazioni a favore dell'intepretazione contrastiva, anche se la deviazione standard dei valori ricavati è maggiore, a dimostrare il fatto che due soggetti su nove attribuiscono a questi stimoli sempre l'interpretazione neutra e mai quella contrastiva. Lo *scaling* del picco sembra quindi giocare un qualche ruolo – o, almeno, sembra essere più importante rispetto all'innalzamento del valore di frequenza del bersaglio basso – anche se sicuramente non permette di ottenere nulla che si avvicini alla percezione categoriale di entità fonologiche differenti.

Il ruolo dell'allineamento del picco è indagato in un secondo esperimento, descritto nel prossimo paragrafo, ma i risultati fin qui riportati evidenziano come l'allineamento del primo bersaglio basso in [L+]H*+L possa non essere di fondamentale importanza nell'identificazione dell'accento. Tuttavia questi risultati potrebbero essere dovuti al fatto che nel sistema fonologico della varietà presa in considerazione mancano accenti che differiscano solo ed esclusivamente per la presenza /assenza del *target* basso [L+]. Infatti, poiché la configurazione di arrivo in questo esperimento era H+L*, la mancanza di identificazione di quest'accento potrebbe essere verosimilmente dovuta al diverso

allineamento del picco in H*+L e H+L*, o a possibili differenze nell'inclinazione della fase discendente dei due accenti.

5.3 Secondo esperimento: dall'accento neutro all'accento contrastivo

Il secondo esperimento è stato effettuato al fine di evidenziare l'eventuale percezione categoriale e di studiare dettagliatamente l'interazione tra lo *scaling* e l'allineamento dei bersagli tonali. In particolare, l'intento era quello di verificare se la presenza di informazioni coerenti sia di allineamento che di *scaling* potesse favorire l'identificazione di un accento piuttosto che dell'altro. Inoltre, l'obiettivo era quello di verificare quali valori di allineamento e di *scaling* corrispondessero alla scelta categoriale. L'ipotesi era che l'allineamento rappresentasse l'informazione più rilevante per l'identificazione degli accenti, non solo per via dello *status* privilegiato attribuitogli in letteratura, ma anche per l'osservazione circa lo stretto legame, nel caso di H+L*, tra lo *scaling* del tono alto e la variazione di forza illocutoria. Coerentemente con quanto discusso nel paragrafo 3, infatti, l'incremento di altezza del bersaglio H nell'accento H+L* potrebbe essere messa in relazione con l'aumento della forza illocutoria dell'enunciato e, in particolare, del proposito illocutivo del parlante. Secondo quanto osservato da Gussenhoven (2002), infatti, una più ampia modulazione di frequenza – causata in questo caso dall'innalzamento del bersaglio che rappresenta l'inizio della fase discendente – può essere associata ad un aumento dello sforzo articolatorio e interpretata come dovuta ad enfasi. L'ipotesi era quindi che l'allineamento rappresentasse il correlato più stabile del cambiamento categoriale, ma che l'informazione relativa allo *scaling* potesse modificare i valori di riferimento per il passaggio da una categoria all'altra, coerentemente con quanto già osservato nel confronto tra due configurazioni come H* e H*+L – si veda il paragrafo 2.

Per il secondo esperimento, gli stimoli manipolati sono stati creati a partire dalla configurazione neutra. Una volta dimostrato, infatti, che l'eliminazione del bersaglio basso di H*+L non porta a modifiche rilevanti nell'identificazione dell'accento, prendere come stimolo di base un enunciato neutro avrebbe avuto il vantaggio di poter escludere un effetto dovuto all'enunciato di base per la manipolazione. Per questo motivo, gli stimoli manipolati sono stati realizzati a partire da una delle ripetizioni della frase *mangia il melone* realizzata con interpretazione neutra, nella quale un accento H+L* era stato prodotto in corrispondenza della parola bersaglio *melone* – si veda 4.1.

Una manipolazione è stata effettuata con l'obiettivo di modificare l'altezza del bersaglio alto e un'altra con l'intento di ritardare gradualmente l'allineamento del picco (e, contemporaneamente, del bersaglio basso seguente) realizzato nell'accento H+L*, in modo da arrivare all'allineamento medio del picco (e del bersaglio basso seguente) rilevato nell'accento H*+L – si veda lo schema in figura 5.

A partire dall'interpretazione neutra della frase *mangia il melone*, in cui un accento H+L* era allineato sulla parola bersaglio *melone*, sono quindi state eseguite le seguenti manipolazioni:

1. l'altezza del bersaglio alto allineato alla sillaba prenucleare è stata innalzata in 9 fasi di 7,2 Hz l'una[9] – figura 5, a sinistra. La risintesi dello stimolo di base e dei 9 manipolati (sintesi PSOLA) sarà indicata come 0, H1, H2…H9;
2. l'allineamento del picco (e del bersaglio basso seguente) è stato ritardato in 5 fasi di 22ms l'una – figura 5, a destra. Questa manipolazione è stata effettuata per lo

[9] Il picco dell'enunciato era realizzato a 177Hz, mentre il suo valore medio nelle tre ripetizioni era di 184Hz – si veda la tabella 1. L'intervallo rispetto al valore medio del picco per la configurazione di arrivo è stato quindi coperto in 9 fasi.

stimolo con *scaling* naturale del picco in H+L* e per ogni stimolo che corrispondesse a due *step* di variazione dello *scaling* del picco, ossia per gli stimoli H2, H4, H6 e H8. I 30 enunciati risintetizzati (sintesi PSOLA) saranno indicati come 0,0-al1,..,0-al5, H2-0, H2-al1,..,H2-al5, H8-0,..,H8-al5.

I quaranta stimoli sintetici più la risintesi dell'enunciato corrispondente all'interpretazione contrastiva – indicato come 'C' – sono stati proposti ai 9 soggetti pisani per tre volte in ordine casuale, per un totale di 123 stimoli. I soggetti ascoltavano per due volte ogni stimolo e stabilivano se gli enunciati potessero rappresentare la correzione di un enunciato precedente oppure corrispondessero a enunciati dichiarativi neutri.

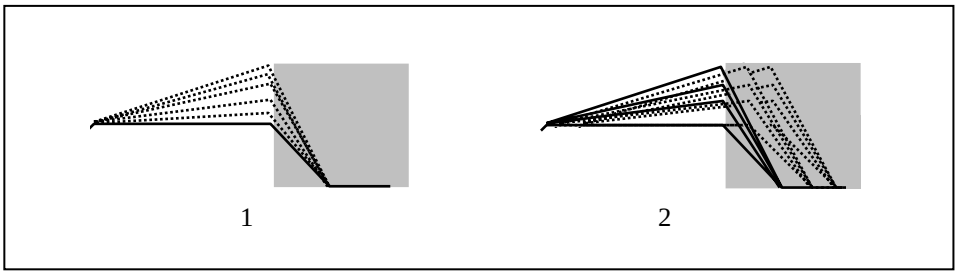

Figura 5: Rappresentazione schematica delle manipolazioni acustiche effettuate – l'area grigia indica la sillaba nucleare.

Nella figura 6 sono riportati due grafici nei quali è indicata la media di tutte le risposte fornite dai soggetti a favore dell'interpretazione contrastiva. Come per l'esperimento precedente, le risposte dei soggetti in relazione ad ogni stimolo sono state espresse sotto forma di un valore compreso tra 0 ed 1, ottenuto dividendo, per ogni soggetto, la somma delle risposte fornite in relazione allo stimolo per il numero totale delle ripetizioni, e sommando i risultati: valori vicino ad '1' indicano che l'enunciato è stato riconosciuto come contrastivo.

Il grafico di sinistra nella figura 6 riporta i dati relativi agli stimoli nei quali è stato incrementato il valore di frequenza fondamentale del bersaglio alto – si veda il punto 1 – e, per confronto, anche quelli relativi allo stimolo corrispondente alla risintesi della naturale interpretazione contrastiva. Come risulta evidente dal grafico e come emerge dall'analisi ANOVA dei dati, l'incremento dell'altezza del picco ha un influenza significativa sulle risposte dei soggetti [$F(88,10)=9.741;p<0.0001$]. Il *post-hoc* di Fisher ha infatti evidenziato una differenza significativa nel numero di risposte a favore dell'interpretazione contrastiva fornito in relazione agli stimoli 'C' e 'H9' e quello dato per tutti gli altri stimoli. Non esiste quindi differenza significativa tra il numero di risposte a favore dell'interpretazione contrastiva fornito per la risintesi della configurazione contrastiva H*+L (stimolo C) e per lo stimolo manipolato in modo da avere il picco allineato come l'H in H+L*, ma realizzato col valore di f_0 medio dell'H in H*+L (stimolo H9); al contrario, le risposte fornite in relazione a 'C' e 'H9' sono significativamente diverse da quelle date per gli altri stimoli. Tuttavia, l'indice della media delle risposte favorevoli in relazione a 'H9' è circa di 0,67, cioè poco superiore allo 0,5 che potremmo far corrispondere alla casualità.

Quindi, aumentare il valore di frequenza del picco in una configurazione come H+L* influenza in modo significativo le risposte dei soggetti, coerentemente con quanto predetto dall'ipotesi di una relazione tra l'altezza del picco nell'accento H+L* e l'aumento della forza illocutoria – si veda il paragrafo 3. Tuttavia, lo *scaling* del picco non è sufficiente a

causare il salto categoriale nella percezione degli stimoli – il numero massimo di risposte a favore dell'identificazione dell'interpretazione contrastiva è 0,67.

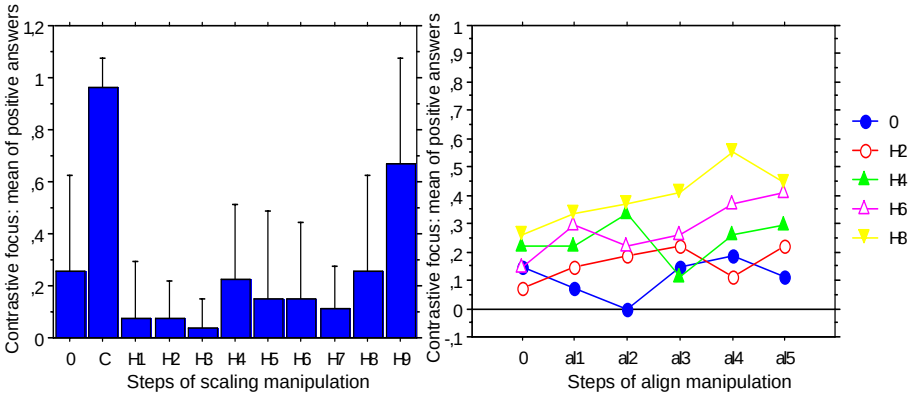

Figura 6: Media e deviazione standard delle risposte dei soggetti a favore dell'interpretazione contrastiva in relazione a stimoli nei quali era stato incrementato il valore di frequenza del bersaglio alto – step 0,H1, …H9 - e allo stimolo contrastvo 'C' per confronto – grafico di sinistra; nel grafico di destra sono riportati i dati relativi agli stimoli nei quali è stato ritardato l'allineamento del picco e del seguente bersaglio basso – step al1, ..., al5, per le diverse variazioni di altezza H2…H8.

Il grafico di destra nella figura 6 riporta i dati relativi agli stimoli nei quali è stato posticipato l'allineamento del picco (e del seguente bersaglio basso) in corrispondenza di cinque diversi livelli di altezza tonale dello stesso – si veda il punto 2. Il test ANOVA, effettuato considerando sia l'allineamento che lo *scaling* come fattori, ha evidenziato che mentre l'allineamento del bersaglio alto non è significativo in relazione alle risposte dei soggetti [$F(240,5)=0.894;p>0.05$] il suo *scaling* lo è [$F(240,4)=6.200;p<0.0001$][10]. In particolare, stimoli nei quali il picco è realizzato ad un livello alto di f_0 sono riconosciuti come contrastivi in un maggior numero di casi[11]. Comunque, anche in caso di allineamento molto ritardato – simile o corrispondente, quindi, a quello medio della configurazione contrastiva – la quantità di risposte a favore dell'accento contrastivo si attesta solo attorno allo 0,5.

Anche nel caso dell'esperimento percettivo volto a confrontare H* ed H*+L (Gili Fivela, 2005) era stato osservato che i risultati non assumevano il tipico andamento a forma di 's' normalmente associato ai contesti nei quali esiste la categorialità della percezione. Come discusso in Gili Fivela (in stampa), questo poteva essere dovuto al fatto che i dati riportati nel grafico erano relativi alle risposte date per stimoli in cui era stato manipolato un solo parametro alla volta, mentre gli altri parametri, ad esempio, lo *scaling*, erano mantenuti costanti. Questo metodo, che permette ovviamente di creare un *continuum* di

[10] Tra i due fattori, inoltre, non emerge alcuna interazione [$F(240,20)=0.637;p>0,05$].

[11] Risultati paragonabili sono stati ottenuti incrementando contemporaneamente il valore di f_0 sia del picco prenucleare sia di quello nucleare, ottenendo quindi un plateau per ogni livello di f_0 considerato. I risultati di questa manipolazione non sono discussi in questa sede.

variazione su una sola dimensione, di fatto porta alla creazione di stimoli in cui sono sistematicamente presenti (altri) correlati con valori ambigui. Tuttavia, mentre i dati relativi al confronto tra H* ed H*+L permettevano comunque di verificare la categorialità delle risposte, osservando i risultati in relazione agli stimoli con le caratteristiche più coerenti (gli estremi del *continuum* di variazione per entrambi i correlati considerati), nel caso dei risultati qui discussi questo non si verifica. Infatti, anche la media delle risposte date in corrispondenza di stimoli nei quali sia l'allineamento che lo *scaling* sono coerenti con l'interpretazione contrastiva si attesta attorno allo 0,5 (si vedano i valori per lo *step* di manipolazione di allineamento 'al5', o 'al4', nella condizione 'H8' – grafico di destra della figura 6). Al contrario, la media delle risposte date in corrispondenza di stimoli nei quali sia l'allineamento che lo *scaling* sono coerenti con l'interpretazione neutra si attesta sul valore 0,1, indicando quindi il riconoscimento dell'accento neutro (si vedano, nel grafico in figura 6, i valori per lo *step* di manipolazione di allineamento '0' nella condizione 'H2' o anche nella '0').

I soggetti, tuttavia, sembrano differire in relazione ai correlati ai quali prestano maggiore attenzione, come già osservato in occasione dell'esperimento percettivo volto a confrontare H* ed H*+L (Gili Fivela, 2005). Anche i dati descritti in questa sede permettono di osservare che alcuni soggetti sembrano prestare maggiore attenzione alle variazioni di allineamento del bersaglio alto (e del basso seguente) – si veda il grafico di sinistra nella figura 7 – mentre altri sembrano focalizzare l'attenzione sullo *scaling* del bersaglio tonale alto – si veda il grafico di destra nella figura 7. Nei grafici, infatti, sono riportate le risposte fornite dal soggetto VA e dal soggetto MA a favore dell'interpretazione contrastiva in relazione a stimoli realizzati manipolando l'allineamento del picco (e bersaglio basso seguente), per ogni manipolazione di altezza del picco considerata.

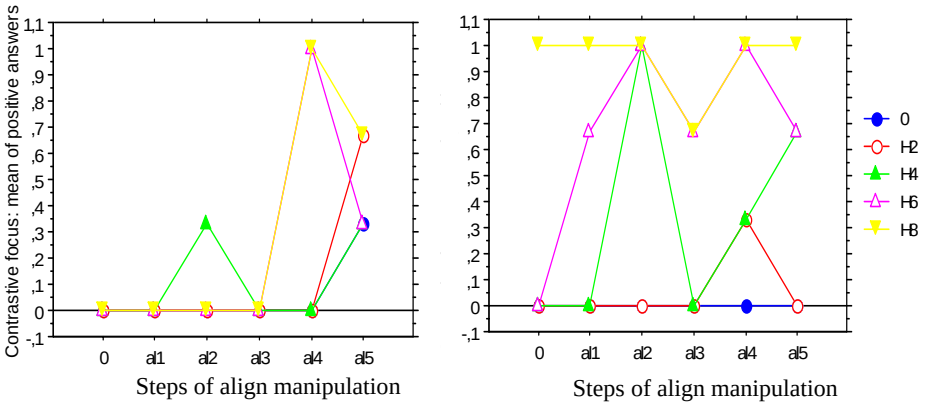

Figura 7: Media delle risposte a favore dell'interpretazione contrastiva in relazione a stimoli realizzati manipolando l'allineamento del picco (e bersaglio basso seguente) per ogni manipolazione di altezza del picco considerata: soggetto FV – grafico di sinistra e soggetto MA – grafico di destra.

Mentre il parlante FV - grafico di sinistra - sembra identificare l'accento contrastivo indipendentemente dalle informazioni fornite dallo *scaling* dei bersagli tonali, ma basandosi principalmente sull'allineamento degli stessi, il parlante MA – grafico di destra – sembra prestare maggiore attenzione alle caratteristiche di *scaling* ed avere difficoltà in relazione alle variazioni di allineamento. Il soggetto, infatti, identifica la configurazione

contrastiva in corrispondenza degli stimoli caratterizzati dall'elevata altezza tonale del picco – stimoli H8 – indipendentemente dalle variazioni di allineamento).

6. DISCUSSIONE

Gli esperimenti descritti in questo lavoro hanno mostrato che i soggetti non percepiscono categoricamente i due accenti, anche se le variazioni di *scaling* mostrano un'influenza sul comportamento linguistico. L'impossibilità di identificare due configurazioni categoricamente distinte potrebbe indicare che i due accenti non rappresentino categorie fonologicamente diverse, anche se rimane il fatto che, in caso di interpretazione contrastiva in fase di produzione, i parlanti realizzano l'accento H*+L, e non quello H+L*, in corrispondenza della parola bersaglio. I dati di produzione mostrano che gli accenti H+L* e H*+L sono prodotti in contesti diversi – neutro e marcato, rispettivamente – ed esprimono funzioni differenti – focalizzazione ampia *vs*. contrastiva. Infatti, H+L* è associato ad enunciati neutri, mentre il contrasto è espresso per mezzo dell'accento H*+L, al massimo in alternativa ad un accento H* che indica focalizzazione non necessariamente contrastiva.

I risultati degli esperimenti, peraltro, potrebbero anche dipendere da limiti del disegno sperimentale. Innanzitutto, possono essere dovuti ad una forte ambiguità – già osservata in letteratura (si veda il paragrafo 3) – tra accenti nucleari neutri e accenti di focalizzazione in posizione finale. L'ambiguità, nei dati qui descritti, è evidente nel caso degli stimoli caratterizzati da un'altezza tonale del bersaglio alto in H+L* che sia maggiore della media per quell'accento, e più simile a quella riscontrata in un accento come H*+L. Altra possibilità è che non sia stato scelto il compito ideale per indurre l'identificazione delle due configurazioni. Possibilità alternative, infatti, consisterebbero nel far ascoltare ai soggetti sequenze domanda-risposta per poi interrogarli relativamente all'adeguatezza degli stimoli nei materiali ascoltati. Un'altra opzione potrebbe consistere nello sfruttare il fatto che, in alcuni contesti, l'accento contrastivo può svolgere la funzione di richiesta di conferma – si veda la discussione in 3 – e poter far riferimento ad una differenziazione tra affermazioni e domande potrebbe permettere l'identificazione delle configurazioni con maggiore facilità. Infine, si potrebbero ottenere risultati diversi ponendo ai soggetti una domanda volta a comprendere l'adeguatezza degli stimoli rispetto a contesti neutri. In questo senso, verificare che gli stimoli possano rappresentare risposte a domande volte ad elicitare risposte neutre, potrebbe far emergere in modo più evidente la peculiarità della funzione associata all'accento marcato, quello contrastivo. Se, infatti, l'accento neutro può adattarsi ad esprimere funzioni specifiche, è possibile che il contrario non sia ammesso, e che l'accento adatto ad esprimere funzioni marcate sia evidentemente inadatto ad un contesto neutro. Quest'ultima modalità sperimentale, oltretutto, permetterebbe anche di escludere un effetto legato alla scelta dell'enunciato neutro come base per la manipolazione nel secondo esperimento. La scelta di utilizzare un enunciato neutro come base per la creazione degli stimoli del secondo esperimento è stata motivata dai risultati del primo esperimento. Tuttavia, questo può aver favorito il prevalere delle risposte a favore dell'interpretazione neutra. Se, modificando il compito, il risultato cambiasse, si avrebbe conferma della fondamentale inadeguatezza di parte del paradigma sperimentale piuttosto che di un effetto legato all'enunciato utilizzato come base per le manipolazioni. Peraltro nel primo esperimento l'allineamento dei bersagli non era stato modificato e non si può quindi escludere a priori che, nel caso in cui fossero stati presenti stimoli con caratteristiche di allineamento manipolate, i risultati sarebbero stati favorevoli all'interpretazione contrastiva. Relativamente ai contesti sperimentali appena descritti, tuttavia, si veda Gili Fivela (in

preparazione), che, grazie ad una serie di esperimenti in cui gli stessi stimoli qui utilizzati sono stati sottoposti ai soggetti in contesti diversi e ponendo domande diverse, conferma l'assenza di categorialità nella percezione dei parlanti. Un ultimo potenziale problema dal punto di vista del disegno sperimentale riguarda invece le scelte relative ai bersagli da manipolare. Infatti la pendenza della porzione discendente del tracciato di f_0 non è stata variata esplicitamente (visto che l'allineamento del secondo bersaglio basso era modificato insieme a quello del bersaglio alto) ma solo indirettamente, negli stimoli nei quali è stato innalzato il livello di frequenza del picco. Un'indagine più sistematica degli effetti della velocità di diminuzione dei valori di frequenza – indicata dalla pendenza della fase discendente – potrebbe essere significativa.

Tornando al presente lavoro, comunque, il risultato meno atteso ha riguardato proprio la mancanza di categorialità della percezione in relazione alla variazione delle caratteristiche di allineamento. Come già osservato questo potrebbe indicare il fatto che H+L* e H*+L non rappresentino categorie fonologicamente diverse. In questo senso, quindi, H*+L sarebbe percepito come una categoria differente rispetto a H* (cfr. Gili Fivela, 2005), ma non rispetto a H+L*. Ulteriori indagini sono richieste per verificare questa possibilità, ma rimane vero che, in fase di produzione e in contesti chiaramente distinti dal punto di vista della forza illocutoria, i parlanti producono tre accenti differenti. Un'ipotesi è che la produzione sia in qualche modo sottoposta a vincoli più rigidi rispetto alla percezione, ad esempio per garantire quella ridondanza di informazioni che può aiutare l'ascoltatore a ricostruire il messaggio veicolato nella produzione verbale. Può darsi che, soprattutto in test percettivi, magari proprio in quelli a risposta forzata, la differenza tra le configurazioni risulti meno definita. Oppure si può ipotizzare che la percezione delle tre configurazioni, effettivamente distinte, sia fortemente influenzata dalle variazioni di tipo pragmatico. Da questo punto di vista, può essere interessante effettuare ancora qualche considerazione sui risultati emersi rispetto alle variazioni di *scaling*.

Anche se in misura minore rispetto a quanto osservato in relazione al confronto tra H* ed H*+L, nel caso degli esperimenti descritti in questo lavoro, cioè del confronto tra H*+L e H+L*, la variazione di altezza tonale incide sulla percezione dei soggetti. Come già osservato per H* ed H*+L, questo può significare sia che lo *scaling* rappresenta un correlato rilevante per l'identificazione dei due accenti – e, nel caso di H*+L e H+L*, si tratterebbe di accenti caratterizzati da toni diversi associati alla sillaba - sia che lo *scaling* incida solo nel senso di rendere esplicita la variazione paralinguistica, amplificando le differenze tra i due accenti. In questo senso, tuttavia, i risultati non sono totalmente coerenti con le predizioni effettuate grazie al riferimento, ad esempio, ai codici biologici di Gussenhoven (2002). Infatti, in base al codice di frequenza, una minor altezza tonale sarebbe prevista in corrispondenza della configurazione interpretata come più assertiva tra H*+L ed H+L*. I dati, però, mostrano che l'accento normalmente realizzato in contesti nei quali si esplicita un contrasto è H*+L, corrispondente a valori di frequenza del picco più elevati rispetto a quelli osservati in H+L*. Relativamente al codice dello sforzo articolatorio, invece, si riscontra maggior coerenza rispetto alle predizioni: la configurazione caratterizzata da maggior escursione tonale è quella più enfatica, contrastiva. Infine, per quanto riguarda l'accento H+L*, l'interpretazione più vicina a quella contrastiva si riscontra quando il bersaglio alto sia effettivamente realizzato a frequenze superiori rispetto alla media riscontrata per l'accento in contesto neutro. Questi ultimi dati relativi al confronto tra H*+L ed H+L* sono quindi coerenti con le predizioni del codice dello sforzo articolatorio di Gussenhoven, ma sembrano contraddire le predizioni fatte in base a quello di frequenza, che invece sembrava confermato dal confronto tra H*+L

e H*. Per poter dare una spiegazione coerente delle due configurazioni in termini paralinguistici, la maggior altezza tonale ed escursione dovrebbe, quindi, essere spiegata ipotizzando un effetto indiretto del codice di frequenza, o, meglio, che il codice sia sensibile all'organizzazione gerarchica dei toni che compongono l'accento. In questo senso, l'innalzamento del bersaglio alto in H+L* può essere visto come strumentale alla percezione del tono L* come 'particolarmente' basso. Il tono associato alla sillaba – L* nel caso di un accento H+L* – sarebbe quindi quello più rilevante, quello in base al quale interpretare i valori legati al codice biologico di frequenza. Se così fosse, si potrebbe ipotizzare che H*+L e H+L* rappresentino soltanto la grammaticalizzazione di una variazione paralinguistica.

7. CONCLUSIONI

Gli esperimenti percettivi discussi in questo contributo mostrano che nel caso di un compito di identificazione di due accenti come H+L* e H*+L i soggetti non sembrano percepire categorie accentuali distinte. Due accenti che in fase di produzione sono tipicamente realizzati in contesti diversi – neutro e marcato – ed esprimono funzioni differenti – focalizzazione ampia e contrastiva, rispettivamente – non sembrano distinti dal punto di vista percettivo. In 5, sono discusse diverse possibilità in relazione all'interpretazione di questi risultati. Sono esaminate sia quelle comunque coerenti con la considerazione dell'esistenza di due categorie fonologiche distinte, che quelle in linea con la considerazione dei due accenti come varianti allofoniche di una stessa categoria. In ogni caso, si potrebbe comunque prevedere che nell'ambito dell'intonazione, l'interpretazione delle configurazioni sia altamente influenzata dalle informazioni pragmatiche. In questo senso, indipendentemente dal considerare i due accenti in esame come fonologicamente distinti o meno, si potrebbe ipotizzare che gli accenti di tono siano associati ad interpretazioni diverse lungo un *continuum* di variazione tra interpretazione neutra e contrastiva. In questo senso, H+L* sarebbe associato ad enunciati spostati verso il polo neutro del *continuum*, mentre H*+L sarebbe associato ad interpretazioni spostate verso il polo contrastivo.

In ogni caso, la variazione di altezza tonale svolge un ruolo nella percezione dei parlanti. Questo ruolo può corrispondere a quello di un parametro di per sé rilevante nell'identificazione delle configurazioni, ma anche a quello di un correlato legato alla variazione paralinguistica che incide solo indirettamente sulla percezione delle configurazioni accentuali. Infine, come già osservato in relazione ad altri esperimenti percettivi, è stata nuovamente osservata una differenza intersoggettiva in relazione ai parametri – di allineamento e di *scaling* – ai quali i soggetti sembrano prestare attenzione nello svolgere il test percettivo. L'osservazione di queste differenze evidenzia ulteriormente l'importanza della considerazione sia delle variazioni di allineamento che di quelle di *scaling* nello studio delle configurazioni accentuali.

8. BIBLIOGRAFIA

Arvaniti, A.; Ladd, D. R.; Mennen, I., 2000. What is a starred tone: Evidence from Greek. In M. Broe e J.Pierrehumbert (a c. d.) *Papers in Laboratory Phonology*, 5, Cambridge, Cambridge University Press, 119-131.

Austin, J. L., 1962. *How to do things with words*. Oxford: Oxford University Press.

Bartels, C.; Kingston, J., 1994. Salient Pitch Cues in the Perception of Contrastive Focus. In P.Bosch e R. van der Sandt (a c. d.) *Focus and Natural Language Processing, Working Papers of the Institute for Logic and Linguistics, 1: Intonation and Syntax*, 1-10.

Beckman, M.; Ayers, G. M.; 1994-97. *Guidelines for ToBI Labelling*. The Ohio State University.

Bruce, G., 1977. *Swedish word accents in sentence perspective*. Lund: CWK, Gleerup.

Cresti, E., 2000. *Corpus dell'italiano parlato*. Firenze: Accademia della Crusca.

Cruttenden, A., 1986. *Intonation, Cambridge Textbooks in Linguistics*. Cambridge: Cambridge University Press.

Firenzuoli, V., 2000. Ordine e istruzione. Espressione di incredulità e contrasto. Descrizione di profili intonativi dal corpus LABLITA. In D. Locchi (a c. d.) *Atti delle X Giornate di Studio del Gruppo di Fonetica Sperimentale*, Napoli 1999, 99-109.

Gili Fivela, B., in preparazione. On the methods for defining categories in intonation.

Gili Fivela, B., in stampa. The coding of target alignment and scaling in pitch accent transcription. In *Italian Journal of Linguistics.*

Gili Fivela, B., 1999. The prosody of left-dislocated topicalized constituents in Italian read speech. In *Proceedings of the European Conference on Speech Communication and Technology*, 1, Budapest, Hungary, 531-534.

Gili Fivela, B., 2004. The phonetics and phonology of intonation: the case of Pisa Italian. Tesi di Perfezionamento, Scuola Normale Superiore, Pisa.

Gili Fivela, B., 2005. La percezione degli accenti: il ruolo dell'allineamento e dello 'scaling' dei bersagli tonali. In *Misura dei parametri. Atti del I Convegno Nazionale dell'Associazione Italiana di Scienze della Voce*, Padova, 30 novembre-1 e 2 dicembre, 313-326.

Gussenhoven, C., 2002. Intonation and interpretation: Phonetics and phonology. In B. Bel e I. Marlien (a c. d.) *Proceedings of the I International Conference on Speech Prosody*, Aix-en-Provence: Laboratoire Parole et Langage, 47-57.

Gussenhoven, C. e Rietveld, T., 2000. The behaviour of H* and L* under variation in pitch range in Dutch rising contours. In *Language and Speech*, 43, 2, 183-203.

Ladd, D. R., 1996. *Intonational Phonology*. Cambridge: Cambridge University Press.

Ladd, D. R. e Schepman, A., 2003. "Sagging transitions" between high pitch accents in English: experimental evidence. In *Journal of Phonetics*, 31, 81-112.

Ohala, J. J., 1983. Cross-language use of pitch: an ethological view. In *Phonetica*, 40, 1-18.

Ohala, J. J., 1994. The frequency code underlines the sound symbolic use of voice pitch. In L. Hilton *et al*. (a c. d.) *Sound symbolism*, Cambridge: Cambridge University Press.

Pierrehumbert, J., 1980. *The phonology and phonetics of English intonation*, PhD thesis, MIT - pubblicata nel 1988 da IULC.

Searle, J. R.; Vanderveken, D., 1985. *Foundations of illocutionary logic*. Cambridge: Cambridge University Press.

't Hart, J.; Collier, R.; Cohen A., 1990. *A perceptual study of intonation. An experimental-phonetic approach to speech melody*. Cambridge: Cambridge University Press.