

CONFINI PROSODICI E VARIAZIONE SEGMENTALE. ANALISI ACUSTICA DELL'ALTERNANZA MONOTTONGO/DITTONGO IN ALCUNI DIALETTI DELL'ITALIA MERIDIONALE

Giovanni Abete, Adrian Simpson
Friedrich-Schiller-Universität Jena
giovanni.abete@libero.it, adrian.simpson@uni-jena.de

1. SOMMARIO

In questo contributo vengono presentati i primi risultati di un'indagine acustica, condotta su un fenomeno di alternanza sincronica tra esiti monottongali e esiti dittongali di alcune variabili vocaliche in quattro dialetti dell'Italia meridionale. L'analisi dei dati, limitata per il momento a Pozzuoli e Torre Annunziata, mette in evidenza due aspetti fondamentali di questo fenomeno, entrambi legati alla posizione delle variabili vocaliche nella struttura prosodica: l'allungamento prepausale e la tendenza delle varianti dittongali a emergere nella posizione finale di sintagma intonativo. L'alternanza monottongo/dittongo in questi dialetti rientra quindi tra i fenomeni di variazione fonetica che dipendono dalla posizione nella struttura prosodica e che forniscono al parlante indici acustici per la segmentazione della catena parlata in costituenti prosodici. Rispetto ad altre ricerche che hanno indagato il rapporto tra confini prosodici e variazione segmentale, utilizzando in genere materiale prodotto *ad hoc* in laboratorio, il presente lavoro si distingue per l'uso di parlato spontaneo relativo a varietà substandard quali i dialetti italiani.

2. INTRODUZIONE

Il presente contributo espone i primi risultati di una ricerca quadriennale che ha riguardato la variabilità del vocalismo tonico in quattro dialetti dell'Italia meridionale (Pozzuoli e Torre Annunziata in Campania, Belvedere Marittimo in Calabria, Trani in Puglia). In particolare, ci si è concentrati su un fenomeno di alternanza sincronica tra esiti monottongali e esiti dittongali, che è tra le caratteristiche più interessanti dei dialetti su menzionati. Come l'analisi evidenzierà, questa alternanza è particolarmente sensibile alla presenza di determinati confini prosodici, con le varianti dittongali che emergono in corrispondenza dei confini di ordine gerarchicamente superiore.¹ Alla questione della variazione tra esiti dittongali e esiti monottongali si aggiunge quella delle variazioni di durata, anch'esse dipendenti in larga misura dalla posizione della variabile nella struttura prosodica. Lo studio delle relazioni tra posizione prosodica, variazioni di durata e alternanza monottongo/dittongo costituisce dunque l'obiettivo primario di questo lavoro. La trattazione sarà limitata ai dati relativi a Pozzuoli e Torre Annunziata.

¹ Il rapporto tra dittongazione e posizione nella frase è stato già messo in luce da Rohlfs (1938 e 1966: § 12), il quale riporta il fenomeno come molto diffuso nei dialetti del versante adriatico, mentre sul versante tirrenico lo registra solo per i dialetti di Pozzuoli e Belvedere Marittimo. Sulla questione si veda anche Lausberg (1939: § 289) e Loporcaro (1988: 159 ss.).

Di seguito si riporta un esempio del fenomeno di alternanza in esame, attraverso alcuni brevi enunciati contenenti il lessema *rete*, tratti dalla produzione spontanea di un parlante del dialetto di Pozzuoli:²

- 1a. a 'k:osir i **rɛtɛs** || a f:ɔ̃ i 'z:ɛtɛsə || (.)
a cucire le reti, a fare le reti (un pescatore deve imparare)
- 1b. tʊ sap:h i **rɛtɛs** || ðə 'βas:əŋ pə vi:'ʃɣɪŋ ||
tu salpi le reti, ti passano (per) vicino (i motoscafi)
- 1c. 'p:ʊrʊ kʷɛnd jɛn æ k:ə'li i **rɛtɛs** || (.)
pure quando andavamo a tirare le reti
- 2a. pe'k:hə 'pɾim:ə ɣʊ 't:hɛndə pjet:s i **rɛtɛs** zə ɣãm'bũ: || (.)
perché prima con trenta pezzi di rete si campava
- 2b. i **rɛtɛs** v m:ʊ'l:ʏt:s ε < ε > (.) < ε > 'nãðu ðip i .ɛtɛ:s ||
le reti a merluzzo è... un altro tipo di rete
- 2c. zãn i mægə'd:zini ɣə nu < u > t:ãn i **rɛtɛs** a:'rɪŋt ||
c'erano i magazzini che noi... buttavamo le reti dentro

Nei primi tre esempi le realizzazioni del lessema *rete* sono in posizione finale di sintagma intonativo, prima di pausa prosodica, e la vocale tonica presenta esiti di tipo dittongale. Negli ultimi tre esempi le realizzazioni sono in posizione interna all'enunciato, posizione che potremmo definire ora interna al sintagma fonologico, ora finale di sintagma fonologico (cfr. § 3), e presentano esiti monotongali del tipo [e]. Perché ci siano esiti dittongali la presenza di una pausa silente non sembra necessaria, come si può vedere negli esempi 1a e 1b.³

La situazione è comunque più complessa di quanto appaia da questi esempi. A parità di posizione prepausale gli esiti non sono sempre dittongali, ma si hanno anche esiti monotongali. Inoltre, gli stessi esiti dittongali in posizione finale sono caratterizzati da una notevole variabilità connessa alle variazioni di durata del dittongo. Il fenomeno pone quindi due problemi: da un lato descrivere correttamente l'alternanza monotongo/dittongo, dall'altro descrivere la variabilità interna agli esiti dittongali. La questione è ulteriormente complicata da un problema metodologico di fondo: distinguere tra esiti dittongali e esiti monotongali è spesso un'operazione arbitraria, se effettuata su base esclusivamente impressionistica. Gli esiti in posizione finale possono presentare talvolta traiettorie dittongali molto accentuate, ma in altri casi tali traiettorie possono essere anche solo accennate; d'altro canto, anche le realizzazioni in posizione interna non presentano andamenti formantici necessariamente piatti.⁴ Pertanto, si è ritenuto necessario procedere a un'analisi acustica del fenomeno, prendendo in esame un indice numerico dell'entità della dittongazione (cfr. § 4) e studiando le variazioni di tale indice in rapporto a diverse

² Tra parentesi uncinate < > vengono trascritte le pause piene, dovute in genere a fenomeni di esitazione.

³ Questa impressione è stata confermata dal confronto su base statistica delle realizzazioni vocaliche prima di pausa prosodica e prima di pausa silente, che non ha mostrato differenze significative. Per i dati si rinvia ad Abete (in preparazione).

⁴ Da tempo si riconosce l'importanza del cosiddetto *Vowel Inherent Spectral Change* anche nell'identificazione di vocali di tipo monotongale (cfr. Neary & Assman, 1986).

posizioni prosodiche, evitando quindi di stabilire una distinzione impressionistica a-priori tra esiti dittongali e esiti monottongali.

Tale studio si inserisce nel filone delle ricerche che negli ultimi anni hanno indagato gli effetti della struttura prosodica non solo a livello soprasegmentale ma anche a livello segmentale (cfr. ad es. Fougeron & Keating, 1997; Keating *et al.*, 2003; Cho, 2004; Cho *et al.*, 2007). Queste ricerche hanno messo in evidenza la mole di variazione fonetica sistematica, di livello anche molto fine, che si accompagna a diverse posizioni nella struttura prosodica, attribuendo in genere a tale variazione una funzione importante nel processo di segmentazione della catena parlata in unità prosodiche di livelli diversi. Rispetto ai lavori citati, generalmente condotti su parlato di laboratorio, la presente ricerca si contraddistingue per l'uso di parlato spontaneo e di varietà substandard quali i dialetti italiani.⁵ Questa scelta ha imposto una riflessione approfondita su diversi problemi metodologici, dalla modalità di elicitazione del parlato, alle tecniche di analisi acustica, ai metodi statistici per un'adeguata interpretazione dei dati. A tali aspetti si potrà in questa sede solo accennare; per una esposizione più estesa si rimanda ad Abete (in preparazione).

3. IL CORPUS

Questo studio si avvale del materiale parlato elicitato da uno stesso raccoglitore (Giovanni Abete) in una campagna di lavoro sul campo tra il 2005 e il 2007. Si tratta complessivamente di circa 28 ore di parlato di 45 informatori. Da questo corpus ampio è stato estrapolato un sotto-corpus di 24 parlanti e 20 ore di registrazioni da sottoporre all'analisi sperimentale. Come accennato, in questa sede si farà riferimento alla parte del corpus relativa a Pozzuoli e Torre Annunziata: 16 parlanti (8 per punto) per circa 12 ore di registrazioni.

Tutti i parlanti intervistati sono maschi, di età compresa tra i 30 e i 60 anni, con poche eccezioni. Sono tutti pescatori, impiegati soprattutto nel campo della piccola pesca, hanno bassa istruzione e sono spesso accomunati da condizioni di disagio economico ed emarginazione sociale.

Le interviste sono state realizzate con una versione adattata dell'intervista libera (cfr. Como, 2006). La grande maggioranza delle registrazioni è stata effettuata all'aperto, nelle baie per il rimessaggio delle barche. Qui i pescatori si intrattengono a compiere piccoli lavori di manutenzione e a rammendare le reti, trascorrendo insieme molte ore della giornata. All'intervista prendevano parte generalmente più persone. Spesso i presenti intervenivano attivamente nella conversazione, anche se non portavano personalmente il microfono.⁶ L'intervistatore si rivolgeva agli intervistati in dialetto napoletano, con slittamenti verso un italiano regionale campano di livello diastraticamente basso. Il dialetto

⁵ Sul rapporto tra fenomeni prosodici e andamenti formantici di un'area (quella livornese) per certi versi comparabile all'area qui in esame cfr. anche Calamai *et al.* (2003) e Marotta *et al.* (2004).

⁶ La presenza all'intervista di più informatori ha attenuato il ruolo potenzialmente inibitore della presenza dell'intervistatore, favorendo dinamiche interazionali più naturali. Studiare gruppi piuttosto che individui è una delle strategie definite da Labov (1972) per attenuare i ruoli sociali di intervistatore e intervistato. Come osserva Milroy (1987: 62), "this has the effect of 'outnumbering' the interviewer and decreasing the likelihood that speakers will simply wait for questions to which they articulate".

usato dall'intervistatore, seppur diverso da quello degli intervistati, favoriva comunque l'elicitazione di parlato dialettale, in quanto veniva interpretato (rispetto all'italiano) come una varietà bassa del repertorio, e sembrava quindi autorizzare l'uso di varietà altrettanto basse. L'intervistatore partecipava attivamente alla conversazione, sia con *feedback* molto brevi (del tipo *ah, mh, eh, ho capito*), sia con interventi più estesi, commentando quanto detto dall'intervistato o proponendo i propri punti di vista. La conversazione veniva orientata il più possibile verso una serie di argomenti predefiniti, quali la pesca, le tecniche utilizzate, i tipi di pesci catturati, i problemi del mercato ittico, le difficoltà della vita del pescatore, eventuali avventure in mare, etc. Quando però il parlante proponeva delle digressioni rispetto a questo canovaccio di base, l'intervistatore lo assecondava, per poi riportare di nuovo il discorso sugli argomenti prestabiliti, non appena la struttura conversazionale lo avesse permesso. L'utilizzo di questo schema fisso di argomenti ha consentito la ricomparsa frequente in tutte le interviste di alcuni *items* lessicali, fornendo un base statisticamente solida per le analisi fonetiche.

Il corpus è stato segmentato ed etichettato manualmente in maniera parziale, limitando questo lavoro a una lista di parole precedentemente selezionate. La procedura è partita da un ascolto impressionistico dei materiali registrati, quindi dalla scelta di determinati *items* lessicali che presentavano variabilità degli esiti nelle vocali toniche, in particolare variabilità tra realizzazioni di tipo dittongale e realizzazioni di tipo monotongale. Sulla base degli *items* scelti sono state individuate tutte le loro realizzazioni nel corpus. Ad esempio, sono state segmentate e etichettate tutte le realizzazioni della parola *rete* nel dialetto di Pozzuoli.

La procedura di segmentazione utilizzata si basa sull'osservazione visiva dell'oscillogramma e dello spettrogramma, e sull'ascolto di porzioni di audio in corrispondenza e a cavallo dei possibili confini fonetici.⁷ Una lista di criteri operativi, corredata da esempi, è stata definita preliminarmente al lavoro di segmentazione.⁸

L'etichettatura ha previsto diversi livelli: segmenti, parole, sintagmi intonativi, enunciati. Altre informazioni prosodiche sono state inserite nella trascrizione al livello segmentale. Questa procedura consente di effettuare un'analisi delle variabili fonetiche all'interno di determinati tipi lessicali, controllando la posizione che le variabili occupano in strutture prosodiche gerarchicamente più alte, come il sintagma intonativo, e fornendo informazioni sul contesto discorsivo più ampio, grazie alla trascrizione del parlato contenuto in uno o più enunciati. I dati del lavoro di etichettatura sono stati quindi trasferiti in un data-base attraverso una serie di procedure automatizzate.

In questo studio si assume una strutturazione gerarchica della prosodia, in cui i costituenti di livello superiore sono composti da costituenti di livello inferiore.⁹ Di particolare importanza per l'analisi del fenomeno in esame sono i costituenti 'sintagma

⁷ Il programma utilizzato per la segmentazione è Wavesurfer 1.8.5 (Sjölander & Beskow, 2006).

⁸ Per le specifiche della procedura di segmentazione si rinvia ad Abete (in preparazione).

⁹ Cfr. Selkirk (1984); Beckman & Pierrehumbert, J. (1986); Nespors & Vogel (1986); Selkirk (1986); Pierrehumbert & Beckman, (1988); Nespors (1993); Beckman (1996); per una rassegna Shattuck-Hufnagel & Turk (1996).

intonativo' e 'sintagma fonologico'.¹⁰ Sulla base di questi due costituenti è possibile distinguere all'interno dei nostri materiali tra 3 diverse posizioni prosodiche (partendo dal livello più basso della gerarchia): 1) posizione interna al sintagma fonologico; 2) posizione finale di sintagma fonologico ma interna al sintagma intonativo; 3) posizione finale di sintagma intonativo. Le tre posizioni saranno indicate rispettivamente dalle sigle *SFa*, *SFb*, *SI*. In figura 1 si riporta un esempio di rappresentazione della struttura prosodica di un sintagma intonativo per un breve enunciato tratto dal corpus di Pozzuoli. Per convenienza espositiva le parole sono glossate in italiano; in basso sono riportate schematicamente le diverse posizioni prosodiche. Nello schema si evidenzia anche che le tre posizioni prosodiche possono essere raggruppate in due categorie fondamentali, opponendo da un lato la posizione finale di *SI*, comunemente definita 'prepausale' e dall'altro le due posizioni *SFa* e *SFb*, racchiuse in una categoria che possiamo qui definire 'interna'.

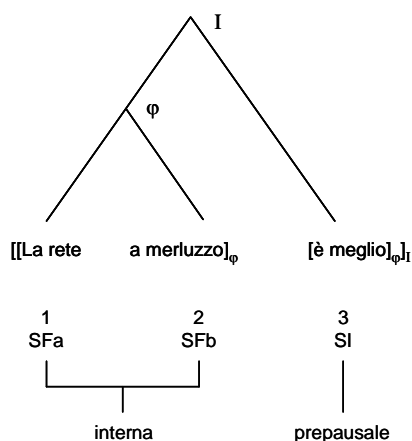


Figura 1: Albero prosodico di un sintagma intonativo tratto dal corpus di Pozzuoli. In basso si riportano le posizioni prosodiche ritenute pertinenti per la presente ricerca

Per la posizione finale di sintagma intonativo, che riveste una particolare importanza per il fenomeno in esame, sono state operate ulteriori distinzioni in base al tipo di andamento melodico di confine e alla funzione pragmatica da esso assolta. Si è quindi definita la categoria *SI_Q* che identifica la posizione finale di sintagma intonativo con intonazione interrogativa; la categoria *SI_C* che identifica la posizione finale di sintagma intonativo con la presenza di un'intonazione di 'continuazione'; la categoria *SI_L* che identifica la posizione finale di sintagma intonativo con la presenza di un'intonazione del tipo 'lista'. Un'etichetta *SI_N* è stata creata ad indicare una categoria 'cestino', che accoglie le posizioni finali di sintagma intonativo prive dei toni ascendenti caratteristici delle categorie

¹⁰ Il sintagma intonativo viene qui definito come la sequenza di parlato inclusa tra due pause prosodiche di livello superiore (cfr. Nespor & Vogel 1986; Nespor 1993). La definizione di sintagma fonologico, invece, si rifà a quella di *intermediate phrase* fornita da Beckman & Pierrehumbert (1986): un contorno intonativo con uno o più *pitch accents*, ma privo di tono di confine finale.

SI_Q, SI_C, SI_L; si tratta quindi di una categoria piuttosto ampia che abbraccia diverse funzioni pragmatiche: può ad esempio essere caratterizzata da un tono di conclusione (soprattutto quando seguita da una pausa silente), ma può avere anche altre funzioni, come quella vocativa o quella imperativa.

È evidente che questa ulteriore suddivisione della posizione finale di sintagma intonativo avrebbe potuto comportare l'individuazione di un numero molto maggiore di sottocategorie. In questa sede però si è preferito separare solo quelle categorie che sembravano avere effetti più macroscopici sulla durata e sulla presenza e l'entità delle dittongazioni, evitando un proliferare di categorie, che avrebbe comportato un'eccessiva frammentazione dei dati e la comparsa di categorie documentate da pochissimi *tokens* e difficilmente utilizzabili nei ragionamenti statistici.

Infine, un'ulteriore categoria SI_V è stata creata per identificare quelle parole che, seppur in posizione finale di sintagma intonativo e caratterizzate da un tono di confine, sono realizzate con una notevole velocità di eloquio, senza il rallentamento tipico della posizione prepausale. Questa peculiarità determina per tale categoria una durata delle toniche in posizione finale molto più bassa della media (cfr. figg. 7-8) e un'incidenza quasi nulla dei fenomeni di dittongazione (cfr. figg. 9-10). In figura 2 si evidenziano tutte le posizioni prosodiche definite per questa ricerca. Sono stati esclusi, invece, dalle rappresentazioni grafiche e dai ragionamenti statistici i *tokens* caratterizzati da allungamenti anomali dovuti a fenomeni di esitazione¹¹.

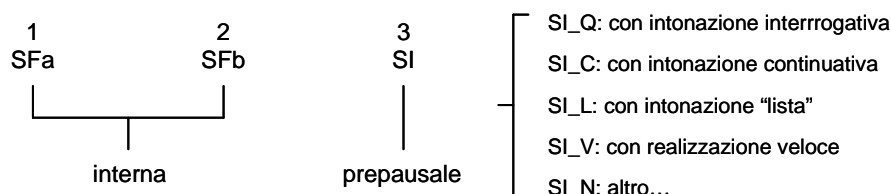


Figura 2: Posizioni nella struttura prosodica e sottocategorie di SI

Sulla base dei *tokens* vocalici etichettati sono state effettuate analisi acustiche della durata e della struttura formantica. Nella tabella 1 si riporta il numero di *tokens* per variabile vocalica e per parlante, analizzati per Pozzuoli e Torre Annunziata. Sono stati etichettati e analizzati 944 *tokens* per il dialetto di Pozzuoli e 799 per Torre Annunziata. Non tutte le variabili vocaliche e non tutti i parlanti sono rappresentati allo stesso modo. In particolare, si rileva un numero nettamente inferiore di realizzazioni di /o/ e /u/, che dipende dalla minore frequenza assoluta di queste vocali nei dialetti esaminati.

Si tenga presente che l'alternanza monotongo/dittongo non riguarda esattamente lo stesso numero di variabili nelle varietà esaminate: mentre a Torre Annunziata sono coinvolte dal fenomeno in questione tutte le vocali a eccezione della /a/, a Pozzuoli sono

¹¹ I dati di durata e del coefficiente di dittongazione di tali *tokens* sono riportati comunque nelle tabelle in appendice sotto l'etichetta SI_H.

coinvolte le alte e le medio-alte anteriori e posteriori.¹² In entrambi i dialetti la presenza delle varianti dittongali non è limitata né dal tipo sillabico, né dalla struttura accentuale della parola, quindi i dittonghi compaiono anche in sillaba chiusa e in parole ossitone e proparossitone.¹³

Pozzuoli							Torre Annunziata							
parlante	/i/	/e/	/ɛ/	/o/	/u/	tot.	parlante	/i/	/e/	/ɛ/	/ɔ/	/o/	/u/	tot.
PZ01	62	62	65	14	22	225	TA01	33	53	42	30	12	11	181
PZ03	13	41	39	13	12	118	TA03	34	37	18	34	9	10	142
PZ05	19	24	42	23	4	112	TA05	5	21	19	11	3	7	66
PZ07	34	6	46	7	4	97	TA07	4	8	26	37	7	1	83
PZ09	20	21	27	0	8	76	TA09	18	24	40	22	5	0	109
PZ11	14	52	51	10	15	142	TA11	4	15	30	22	3	3	77
PZ13	21	23	34	5	5	88	TA13	1	16	20	14	7	4	62
PZ15	26	9	30	6	15	86	TA15	8	17	26	24	2	2	79
tot.	209	238	334	78	85	944	tot.	107	191	221	194	48	38	799

Tabella 1: Numero di tokens analizzati per variabile vocalica e per parlante

4. CARATTERIZZAZIONE DELLA DINAMICA DITTONGALE

Mentre le durate sono state ottenute direttamente dai *files* di etichettatura, per l'analisi della traiettoria dittongale è stata realizzato uno *script* in *Snack* e *tcl/tk* per la stima automatica dei valori formantici. Il metodo di caratterizzazione della traiettoria dittongale utilizzato nella presente ricerca costituisce una evoluzione del metodo di Holbrook & Fairbanks (1962) e segue essenzialmente Simpson (1998), con qualche differenza rispetto all'algoritmo per la stima dei valori formantici. Un esempio di tale procedura è dato in figura 3.

Di ciascun segmento vocalico vengono scartati in automatico i primi 20 ms. e gli ultimi 20 ms.; ciò è necessario per ridurre l'impatto delle transizioni formantiche, che nei criteri di segmentazione utilizzati in questa ricerca vengono incluse nel segmento vocalico. La prima e l'ultima misurazione delle formanti vengono effettuate subito dopo i primi 20 ms. e subito prima degli ultimi 20 ms. I restanti punti in cui effettuare le misurazioni vengono calcolati dividendo in parti uguali la porzione tra la prima e l'ultima misurazione, facendo in modo che la lunghezza dell'unità di segmentazione sia la più vicina possibile al valore nominale di 20 ms. Quindi vengono stimati i valori delle prime tre formanti per ciascuno dei punti precedentemente individuati. Ogni segmento vocalico inferiore ai 60 ms. viene trattato tecnicamente come un monottongo e viene caratterizzato da un singolo insieme di valori di F1, F2 e F3 presi nella porzione centrale della vocale, al centro di una finestra di 20 ms.

Nell'esempio riportato in figura 3 un dittongo di 192 ms. viene rappresentato da 9 misurazioni delle formanti. Esclusi i primi e gli ultimi 20 ms., le misurazioni vengono effettuate nel punto iniziale del segmento, quindi a intervalli regolari della durata nominale di 20 ms. (la durata effettiva in questo caso è di 19 ms.).

¹² In questo dialetto l'alternanza monottongo/dittongo si va estendendo anche alla vocale anteriore medio-bassa nei pescatori al disotto dei 40 anni. Per la questione si rinvia a Abete & Simpson (in stampa).

¹³ Sulla questione dei dittonghi in sillaba chiusa cfr. Abete (2006).

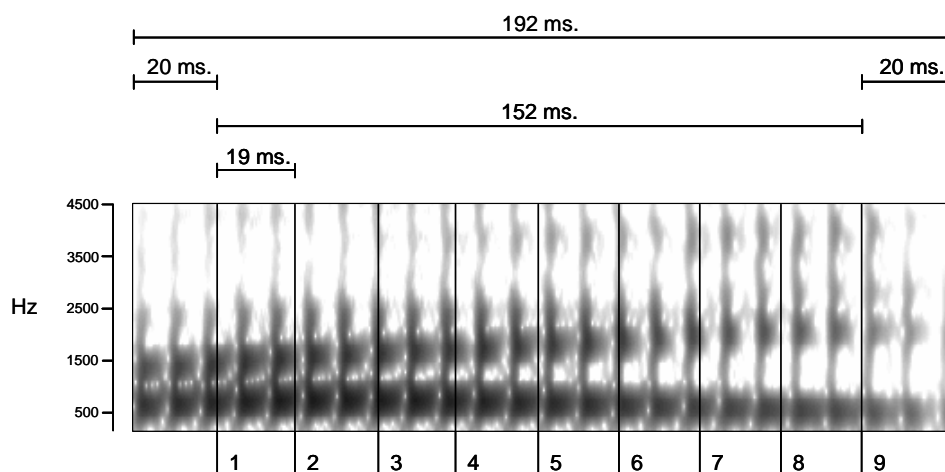


Figura 3: Esempio della procedura di analisi della traiettoria dittongale

Di seguito si riportano le specifiche tecniche per l'analisi formantica:

```

framelength = 0.02 s.
windowlength = 0.02
windowtype = Hamming
preemphasisfactor = 0.9
lpc-order = 12
ds_freq = 10000 Hz
nom_fl_freq = 500 Hz

```

Una volta effettuata la stima automatica dei valori formantici è stato necessario controllare manualmente questi valori e correggerli nel caso di errori. In questa ricerca la presenza di eventuali errori è stata individuata in maniera visiva, attraverso l'osservazione di grafici realizzati per ciascun segmento vocalico. I grafici sono stati ottenuti in automatico, grazie a uno *script* appositamente realizzato con il programma per l'analisi statistica e la rappresentazione grafica dei dati *R*.¹⁴ Questo script realizza grafici delle traiettorie dittongali (basati sull'andamento temporale di F1 e F2), aggiungendo a ciascun grafico una serie di informazioni: codice del file, posizione prosodica, tipo di variabile, durata in ms., coefficiente di dittongazione (cfr. più avanti), la parola da cui il *token* è tratto. Il tutto viene sistemato in un unico file *.pdf, includendo dieci grafici per pagina (v. figura 4).

Se si suddividono i segmenti vocalici in classi di durata di 20 ms., i dittonghi di una stessa classe saranno caratterizzati da un eguale numero di misurazioni formantiche. È quindi possibile calcolare le medie delle misurazioni formantiche di tutti i *tokens* di un dittongo che appartengono a una stessa classe di durata, in modo da poter rappresentare graficamente le traiettorie dittongali medie per le diverse classi di durata dei *tokens* di uno

¹⁴ R Core Development Team (2008).

stesso dittongo. Un esempio di questa modalità di rappresentazione è riportato in figura 5. Questo tipo di grafico consente di visualizzare il rapporto tra struttura del movimento dittongale e variazioni di durata. In questo caso, in particolare, si nota una notevole variabilità del timbro di attacco del dittongo in rapporto alle diverse classi di durata, con le due formanti in attacco che si avvicinano sempre più nei dittonghi di durata maggiore.¹⁵

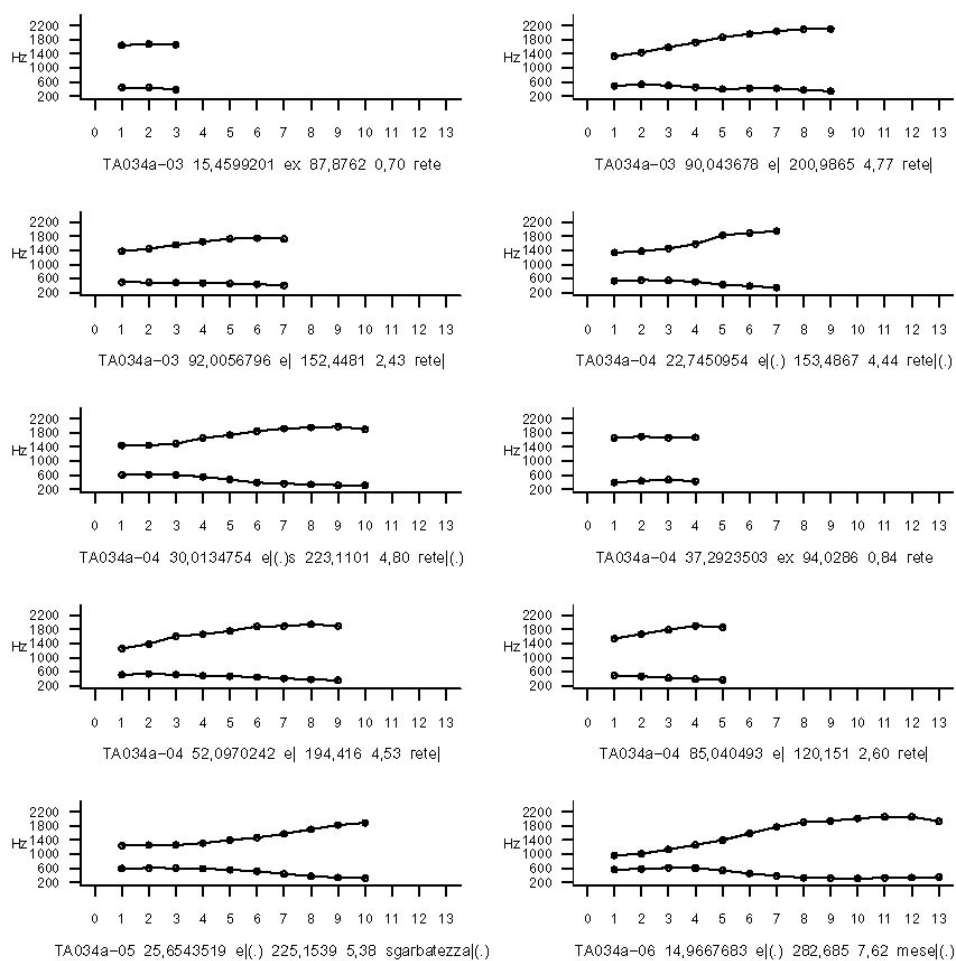


Figura 4: Grafici riassuntivi delle caratteristiche acustiche dei tokens analizzati. Esempio tratto dal corpus di Torre Annunziata e relativo a 10 realizzazioni di /e/

Oltre ad avere una rappresentazione grafica delle traiettorie formantiche, si è avuta la necessità di ottenere un indice numerico relativo all'ampiezza del movimento dittongale, così da poter confrontare l'andamento di tale indice rispetto ad altri parametri linguistici,

¹⁵ Sulla variabilità delle traiettorie dittongali in rapporto a variazioni di durata non possiamo soffermarci in questa sede; per la questione si rinvia ad Abete (in preparazione).

come la posizione della variabile nella struttura prosodica (cfr. § 5.2), o extralinguistici, come l'età dei parlanti (cfr. Abete & Simpson, in stampa). L'indice misura la distanza euclidea nello spazio F1-F2 che intercorre tra i due timbri più distanti raggiunti dal segmento nella transizione dittongale. La base della procedura di calcolo sta nell'individuazione degli scarti tra valore massimo e valore minimo di ciascuna formante (F1 e F2) per le misurazioni effettuate sull'intero dittongo. Si ottengono così due misure separate indicative dell'ampiezza delle transizioni delle prime due formanti.

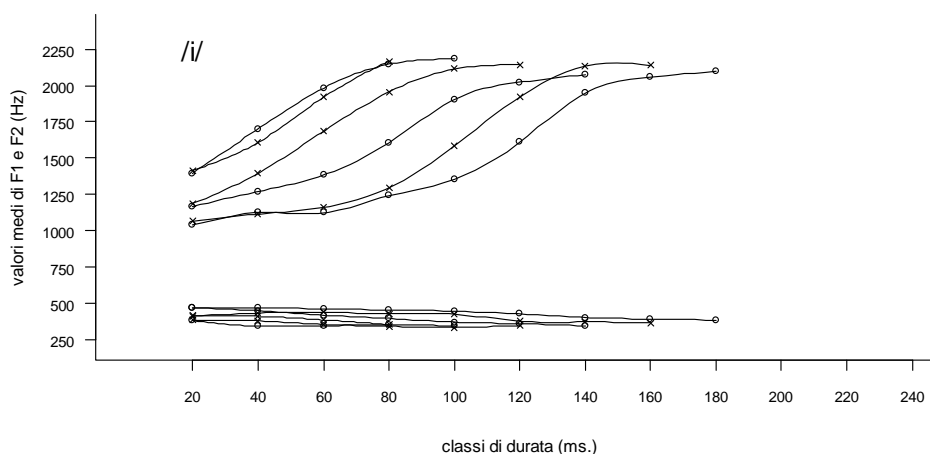


Figura 5: Traiettorie dittongali medie di /i/ distinte per classi di durata (il grafico è stato realizzato con i dati relativi alla parola 'pesce' nel corpus di Pozzuoli)

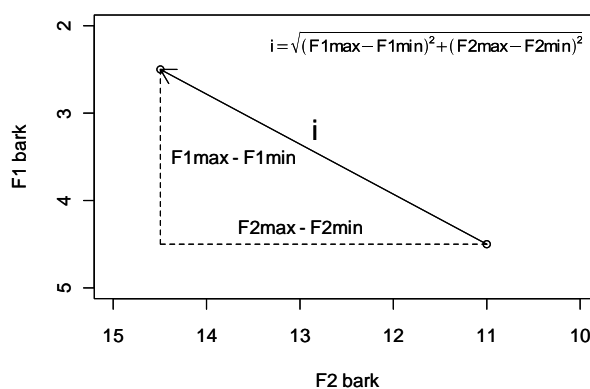


Figura 6: Il coefficiente di dittongazione

Come esemplificato in figura 6, tali misure sono rappresentabili in uno spazio F2-F1 come i cateti di un triangolo rettangolo. Pertanto, è possibile ottenere un unico indice dell'escursione dei valori delle due formanti calcolando l'ipotenusa di questo triangolo attraverso il teorema di Pitagora. La formula per il calcolo dell'indice è dunque la seguente:

$$\text{coeff. ditt.} = \sqrt{(F1_{\text{max}} - F1_{\text{min}})^2 + (F2_{\text{max}} - F2_{\text{min}})^2}$$

Prima di essere immessi nella formula, i valori in Hertz vengono convertiti in Bark secondo la formula di Traunmüller (1990). In tal modo viene valutato in maniera più adeguata l'apporto dei movimenti della prima formante all'ampiezza totale del movimento dittongale. I movimenti della F1, infatti, risultano meno ampi se considerati dal punto di vista della scala acustica in Hertz, mentre vengono rappresentati più adeguatamente se considerati dal punto di vista della scala uditiva in Bark.

A tale indice è stato dato il nome di 'coefficiente di dittongazione'. Il coefficiente ha valore 0 solo nei casi di segmenti vocalici più brevi di 60 ms., giacché questi vengono caratterizzati da una sola misurazione delle formanti in un punto medio. Per quanto riguarda invece vocali più lunghe di 60 ms., il coefficiente di dittongazione si mantiene in genere al disotto del valore 1 per gli esiti che percettivamente risultano di tipo monotongale, mentre supera il valore 1.8 per gli esiti che percettivamente risultano di tipo dittongale. All'ascolto impressionistico dei dati sembra che segmenti caratterizzati da un coefficiente di dittongazione inferiore a 1.8 non producano la percezione di un dittongo.¹⁶

5. ANALISI DEI DATI

5.1 Allungamento prepausale

L'allungamento prepausale, ossia l'allungamento di vocali e consonanti nelle posizioni finali di diversi costituenti prosodici, è un fenomeno noto in moltissime lingue e su di esso esiste una ricca bibliografia.¹⁷ In molti degli studi riportati in letteratura tale fenomeno risulta di tipo incrementale: le durate si fanno progressivamente più lunghe a mano a mano che si sale nella gerarchia dei costituenti prosodici (cfr. ad es. Ladd & Campbell, 1991; Wightman *et al.*, 1992). Le variazioni di durata costituiscono pertanto una delle risorse fondamentali che i parlanti utilizzano per la demarcazione di strutture prosodiche. La natura plausibilmente universale dell'allungamento prepausale è da mettere in relazione con tendenze fisiologiche comuni ad ogni attività motoria di livello superiore (cfr. Vaissière, 1983 e 1995). Tuttavia, nonostante questa base fisiologica, gli effetti di questo fenomeno sono diversi da lingua a lingua, sia per quanto riguarda il dominio dell'allungamento (numero di sillabe coinvolte), sia relativamente ai segmenti interessati (vocali e consonanti), sia per il diverso influsso esercitato dai diversi costituenti prosodici.

Nel presente studio è stata analizzata la durata dei segmenti vocalici tonici in rapporto alla posizione interna o finale di alcuni costituenti prosodici (cfr. § 3). L'analisi acustica delle variazioni di durata ha evidenziato nei dialetti di Pozzuoli e Torre Annunziata gli effetti consistenti dell'allungamento prepausale. Tale allungamento si manifesta con molta

¹⁶ Questa 'soglia critica' per la percezione di dittongazione è stata definita su base impressionistica. In futuro si prevede di approfondire meglio tale questione attraverso lo sviluppo di adeguati test percettivi.

¹⁷ Per l'inglese Lehiste (1972); Oller (1973); Klatt (1975 e 1976); Crystal & House (1990); Edwards *et al.* (1991); Byrd (2000); per il francese Grosjean & Deschamps (1972); Crompton (1980); Rietveld (1980); Fletcher (1991); per lo svedese Lindblom (1968); Lindblom & Rapp (1973); Lyberg (1981); per l'ebraico Berkovits (1984 e 1993); per il tedesco Kohler (1983); per l'italiano o varietà regionali di italiano Vayra & Fowler (1992); Sorianello (1994 e 2006); Dell'Aglio *et al.* (2002); Bertinetto *et al.* (2006).

regolarità soprattutto nelle posizioni che hanno una più rigorosa definizione prosodica, quali la posizione finale di sintagma fonologico e la posizione finale di sintagma intonativo, mentre i dati sono più variabili e specifici di ciascuna varietà per quanto riguarda le sottocategorie della posizione finale di sintagma intonativo (SI_N, SI_Q, SI_C, SI_L), le quali sono state definite invece su base pragmatica e sulla presenza o meno di intonazione ascendente.

I segmenti vocalici tonici del dialetto di Pozzuoli sono particolarmente sensibili alla posizione nella struttura prosodica. La durata delle vocali, infatti, aumenta in corrispondenza dei confini prosodici di ordine gerarchico superiore. Come già documentato per altre lingue, tale allungamento è di tipo incrementale, cioè incide in maniera sempre maggiore quanto più si sale nella gerarchia prosodica. In tabella 2 (appendice) si riportano i dati relativi alla durata di tutti i *tokens* del corpus di Pozzuoli nelle diverse posizioni prosodiche definite per questa ricerca. Per ciascuna classe prosodica vengono riportati il valore minimo, il primo quartile, la mediana, la media, il terzo quartile¹⁸, il valore massimo, il numero dei *tokens* analizzati.

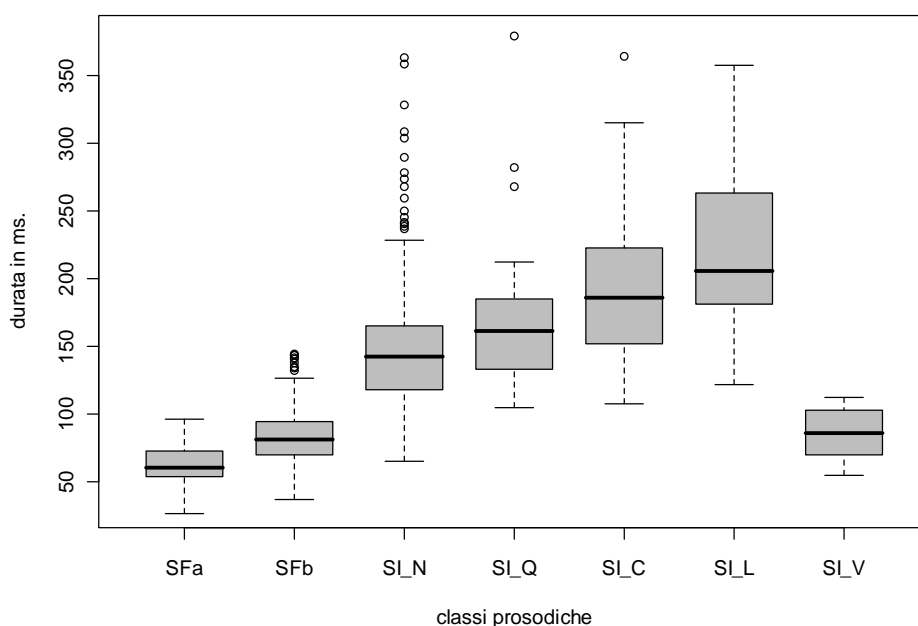


Figura 7: Durate per posizione prosodica nel corpus di Pozzuoli

I dati della tabella 2 sono riportati in un grafico del tipo *box plot* in figura 7. Il grafico evidenzia come i valori di durata in posizione finale di sintagma intonativo (SI_N) siano molto maggiori di quelli in posizione finale di sintagma fonologico (SFb) e come questi ultimi siano a loro volta sensibilmente maggiori dei valori in posizione interna di sintagma fonologico (SFa). Inoltre, a parità di posizione finale di sintagma intonativo, la durata

¹⁸ I quartili (compresa la mediana) e le medie sono stati calcolati escludendo gli *outliers*. Gli *outliers* vengono individuati secondo la formula proposta da Tukey (1977).

continua ad aumentare progressivamente se l'enunciato ha un'intonazione di tipo interrogativo, di tipo 'continuativo', o del tipo 'lista'. Le vocali etichettate come SI_V, che pur essendo percepite in posizione finale di sintagma intonativo sembravano essere realizzate con particolare velocità, hanno in effetti durata sensibilmente inferiore alle vocali in posizione finale di sintagma intonativo e sono comparabili invece alle vocali in posizione finale di sintagma fonologico. Si è deciso di escludere da questa rappresentazione le vocali prodotte con fenomeni di esitazione,¹⁹ in quanto i valori di durata di tali vocali risultano troppo variabili e non costituiscono in effetti una classe prosodica a sé stante.

Per valutare la significatività statistica delle differenze riscontrate tra le mediane nelle diverse classi prosodiche è stata effettuato il test non parametrico di Kruskal-Wallis.²⁰ Il test è stato applicato prima all'insieme dei campioni, per accertare che almeno una delle differenze tra le mediane non fosse dovuta al caso, quindi sono stati effettuati confronti appaiati tra ciascun campione con tutti gli altri. I risultati del test come valori di p. sono riportati nella tabella 3 (appendice), insieme con le differenze in ms. e in percentuale²¹ tra le mediane di ogni classe prosodica con ciascuna delle altre classi. Dal test risulta che quasi tutte le differenze tra le mediane sono statisticamente significative per un valore di p. < 0.01, ad eccezione delle differenze tra SI_N e SI_Q da un lato, e SI_Q e SI_C dall'altro, per le quali p. < 0.05, ma > 0.01. Non risulta esserci una differenza statisticamente significativa tra vocali in posizione finale di sintagma fonologico (SFb) e vocali finali di sintagma intonativo realizzate velocemente (SI_V).

Le differenze di durata tra le varie posizioni prosodiche, pur essendo in genere statisticamente significative (con le eccezioni fatte sopra), non sono tutte di uguale consistenza. Confrontando i valori in tabella 3 con il grafico in figura 7, emerge come la differenza di durata più consistente sia quella tra vocali in posizione interna di sintagma intonativo (SFa e SFb) da un lato e vocali in posizione finale di sintagma intonativo dall'altro. Le vocali in posizione SI_N durano il 75 % in più di quelle in posizione SFb e il 136 % in più di quelle in posizione SFa. A parità di posizione finale di sintagma intonativo, si presenta una certa polarizzazione tra vocali non marcate (SI_N) e vocali con intonazione di 'continuazione' e del tipo 'lista' (SI_C e SI_L). Rispetto alle vocali in posizione SI_N, quelle in posizione SI_C durano il 30 % in più e quelle in posizione SI_L presentano un incremento del 44 %. Meno consistenti, invece, le differenze tra SI_N e SI_Q, SI_Q e SI_C, SI_C e SI_L.

Anche nel dialetto di Torre Annunziata la durata dei segmenti vocalici tonici è fortemente influenzata dalla posizione nella struttura prosodica. La tabella 4 (appendice) e la figura 8 riassumono i dati sulle variazioni di durata. Come per Pozzuoli, anche in questo caso l'allungamento è di tipo incrementale, cioè aumenta progressivamente nel passaggio

¹⁹ Tali vocali sono etichettate come SI_H. Le statistiche riassuntive relative a questa categoria sono riportate nelle tabelle in appendice.

²⁰ Il test di Kruskal-Wallis è un'alternativa non parametrica all'ANOVA. Per una descrizione tecnica del test si veda Hollander & Wolfe (1973: 115-120). La scelta di un test non parametrico è stata dettata dalla natura dei nostri dati, che non soddisfano alcuni requisiti delle statistiche parametriche, come la normalità delle distribuzioni e l'omogeneità delle varianze.

²¹ I valori in ms. e in percentuale vanno letti come incremento del valore nella colonna di destra sul valore della colonna di sinistra.

dalla posizione interna al sintagma fonologico (SFa) a quella finale di sintagma fonologico ma interna al sintagma intonativo (SFb), a quella finale di sintagma intonativo (SI_N). Inoltre, i segmenti in posizione finale di sintagma intonativo caratterizzati da intonazione ascendente (SI_Q, SI_C, SI_L) presentano durate sensibilmente maggiori di quelle dei segmenti nella stessa posizione ma privi di tale contorno intonativo (SI_N). Poco significative, invece, le differenze che si notano all'interno delle diverse classi caratterizzate da intonazione ascendente.

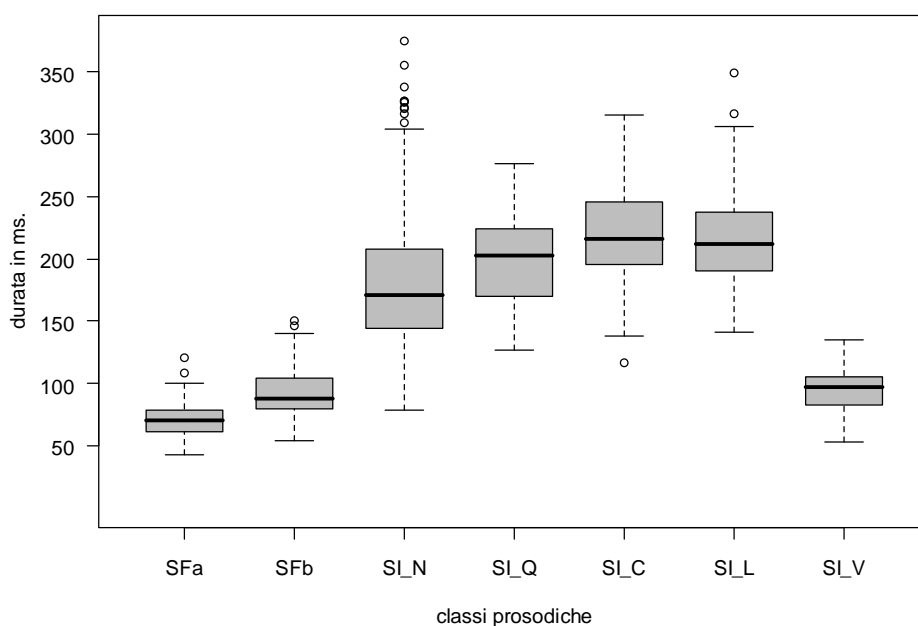


Figura 8: Durate per posizione prosodica nel corpus di Torre Annunziata

La significatività statistica delle differenze notate tra le mediane di ciascuna classe è stata valutata attraverso il test non-parametrico di Kruskal-Wallis. I risultati sono riportati come valori di p. nella tabella 5, insieme con le differenze in millisecondi e in percentuale tra le mediane. Le differenze che non risultano statisticamente significative sono quelle tra SI_N e SI_Q, SI_Q e SI_C, SI_C e SI_L. Non è stata dunque rilevata una differenza significativa tra durate in posizione finale di enuncianti con intonazione di ‘continuazione’ da un lato, e del tipo ‘lista’ dall’altro, mentre nel dialetto di Pozzuoli quest’ultima classe presentava durate sensibilmente maggiori. Come per Pozzuoli, invece, le vocali in posizione finale di sintagma intonativo realizzate con particolare velocità di eloquio (SI_V) sono assimilabili a quelle in posizione finale di sintagma fonologico (SFb).

Le differenze più consistenti sono quelle che intercorrono tra segmenti in posizione non-prepausale e segmenti in posizione prepausale: le vocali in posizione finale di sintagma intonativo non marcata (SI_N) durano il 94 % in più di quelle in posizione finale di sintagma fonologico (SFb) e il 144 % in più di quelle in posizione interna al sintagma fonologico (SFa). Un incremento sensibile, di circa il 26 %, avviene anche nel passaggio dalla posizione interna al sintagma fonologico (SFa) a quella finale di sintagma fonologico

(SFb). A loro volta i segmenti in posizione finale di sintagma intonativo con intonazione di ‘continuazione’ (SI_C) o del tipo ‘lista’ (SI_L) hanno durate sensibilmente superiori a quelle dei segmenti in posizione finale di sintagma intonativo non marcata (SI_N), rispettivamente del 26 % e del 24 %.

I dati sull’allungamento prepausale delle vocali toniche ottenuti in questa ricerca presentano un quadro piuttosto simile a quello già rilevato per le toniche in posizione finale e non finale in studi sperimentali condotti su diverse varietà di italiano (cfr. Albano Leoni *et al.* 1995, Dell’Aglia *et al.* 2002, Soriano & Calamai 2005, Soriano 2006).²²

5.2 Variazioni nel coefficiente di dittongazione

Come descritto in § 4, per ogni *token* vocalico è stato calcolato un coefficiente di dittongazione, che fornisce un’indicazione dell’ampiezza del movimento dittongale. Gli esiti di tipo monotongale presentano valori minimi, al disotto di 1, mentre esiti tipicamente dittongali superano in genere il valore 2 del coefficiente. L’analisi impressionistica dei dati ha indotto a considerare un coefficiente di 1.8 come soglia critica al disotto della quale le traiettorie formantiche non vengono percepite come un movimento dittongale.

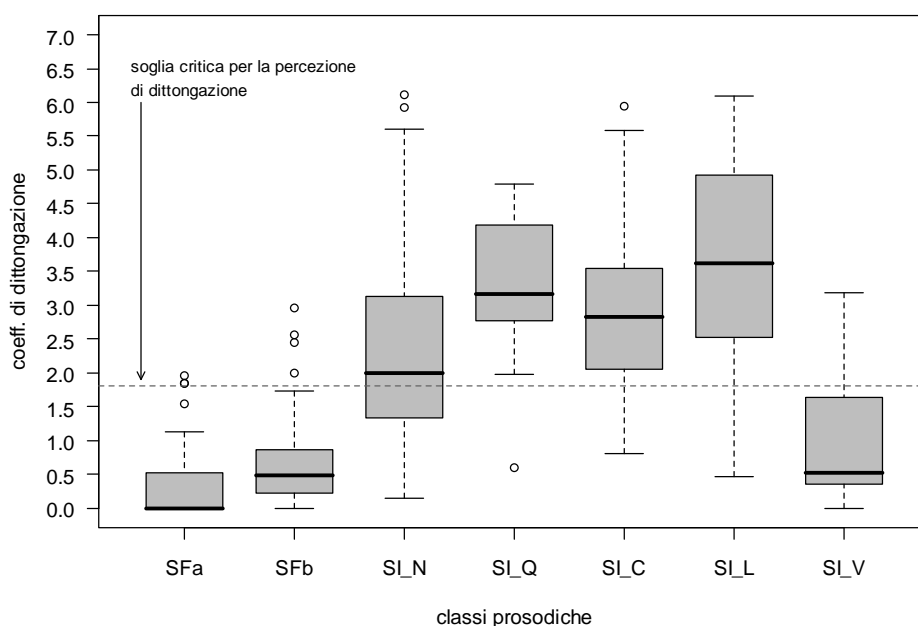


Figura 9: Coefficienti di dittongazione in rapporto alla posizione prosodica nel corpus di Pozzuoli²³

²² Per un confronto più agevole è necessario accorpare in un’unica posizione ‘interna’ i dati relativi alle posizioni SFa e SFb definite per questa ricerca.

²³ Sono stati esclusi i *tokens* di /ε/, che non presentano dittongazione nei parlanti più anziani (cfr. nota 8).

Il coefficiente di dittongazione permette di indagare eventuali correlazioni tra dittongazione e altri parametri di riferimento. Nella tabella 6 (appendice) si riportano dati riassuntivi sul coefficiente di dittongazione nelle diverse classi prosodiche per il dialetto di Pozzuoli. I dati sono rappresentati graficamente in figura 9.

Dalla figura appare chiaramente la stretta correlazione tra dittongazione e posizione della parola nella struttura prosodica. Gli esiti delle variabili esaminate si presentano come esclusivamente monotongali (con sporadiche eccezioni) in posizione interna e finale di sintagma fonologico, mentre sono prevalentemente dittongali in posizione finale di sintagma intonativo,²⁴ con movimenti formantici particolarmente ampi nelle posizioni caratterizzate da intonazione ascendente (SI_Q, SI_C, SI_L).

Come per le variazioni di durata, sono stati effettuati dei test statistici per valutare l'affidabilità delle differenze notate tra le mediane del coefficiente di dittongazione nelle varie classi prosodiche. Nella tabella 7 (appendice) sono riportati i risultati del test di Kruskal-Wallis come valori di p ., insieme con le differenze in Bark e in percentuale tra le mediane di ogni classe prosodica con ciascuna delle altre classi.

Dai test emerge che le differenze nelle mediane del coefficiente di dittongazione tra SI_Q e SI_C, e SI_Q e SI_L non sono statisticamente significative, mentre la differenza per le classi SI_C e SI_L è significativa per un valore di p . < 0.05 ma > 0.01 . Infine, non risulta significativa la differenza delle mediane nei valori del coefficiente di dittongazione tra le classi SFb e SI_V. Dalle differenze in percentuale emerge, come per le durate, una netta contrapposizione tra realizzazioni in posizione interna e finale di sintagma fonologico e le realizzazioni in posizione finale di sintagma intonativo. Gli esiti in posizione finale di sintagma intonativo non marcata (SI_N) presentano un incremento nel coefficiente di dittongazione del 308% rispetto agli esiti in posizione finale di sintagma fonologico (SFb). A parità di posizione finale di Sintagma Intonativo, solo le realizzazioni con intonazione interrogativa (SI_Q) e intonazione del tipo 'lista' (SI_L) sembrano avere coefficienti di dittongazione sensibilmente maggiori.

In tabella 8 (appendice) e in figura 10 si riportano i dati sul coefficiente di dittongazione nel dialetto di Torre Annunziata. Anche per questa varietà si evidenzia una netta polarizzazione tra esiti di tipo monotongale, caratteristici delle posizioni non-prepausali, interna e finale di sintagma fonologico (SFa e SFb), e esiti dittongali, caratteristici delle posizioni finali di sintagma intonativo (SI_N, SI_Q, SI_C, SI_L). I segmenti in posizione SI_V, cioè in posizione finale di sintagma intonativo ma realizzati con particolare velocità di eloquio, non presentano dittongazione, se non piuttosto sporadicamente. In posizione prepausale, invece, gli esiti sono prevalentemente dittongali, ma sono nondimeno presenti esiti monotongali, anche se in percentuale nettamente inferiore.

In posizione finale di sintagma intonativo non si riscontrano grosse differenze nel coefficiente di dittongazione in rapporto al tipo di andamento dell'intonazione, tranne nel caso degli enunciati caratterizzati da intonazione di 'continuazione', che presentano coefficienti di dittongazione sensibilmente maggiori. Per valutare la significatività statistica delle differenze tra le mediane dei coefficienti di dittongazione nelle varie classi prosodiche, sono stati effettuati confronti appaiati tra le mediane con il test di Kruskal-

²⁴ Sulla pertinenza della posizione finale di sintagma intonativo per l'emersione delle varianti dittongali si veda quanto già riportato in Loporcaro (1988: 519 ss.) per il dialetto di Altamura.

Wallis. I risultati dei test come valori di p . sono riportati nella tabella 10, insieme con le differenze tra le mediane in Bark e in percentuale. Dai test statistici risultano significative tutte le differenze riscontrate, ad eccezione delle differenze tra SI_N e SI_Q, SI_N e SI_L, SI_Q e SI_C, SI_Q e SI_L. In altre parole, all'interno della macrocategoria prepausale risulta statisticamente significativo soltanto l'incremento di durata che SI_C presenta rispetto a SI_N, e che SI_L presenta rispetto a SI_C; non sono significative, invece, le altre differenze.

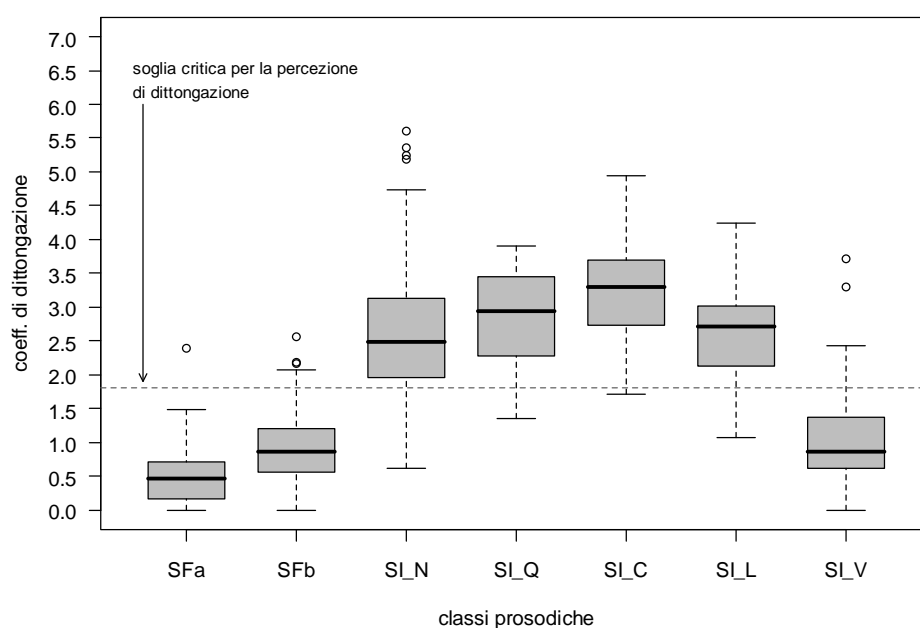


Figura 10: Coefficienti di dittongazione in rapporto alla posizione prosodica nel corpus di Torre Annunziata

Guardando gli incrementi in percentuale, si riscontra un netto aumento del coefficiente di dittongazione (186 %) nel passaggio dalla posizione finale di sintagma fonologico (SFb) a quella finale di sintagma intonativo non marcata (SI_N). Di non molto rilievo le differenze che si riscontrano tra le varie sottoclassi della posizione finale di sintagma intonativo. Il coefficiente di dittongazione in posizione finale con intonazione di 'continuazione' (SI_C) è superiore del 33 % rispetto ai valori in posizione finale non marcata (SI_N); il coefficiente di dittongazione in posizione finale con intonazione del tipo 'lista' (SI_L) è inferiore di circa il 18 % rispetto ai valori in posizione finale con intonazione di 'continuazione' (SI_C).²⁵ Per il resto, le differenze sono di poco conto e, come già evidenziato in precedenza, non risultano statisticamente significative.

²⁵ Per questo ultimo aspetto Torre Annunziata presenta una situazione inversa rispetto a quella di Pozzuoli.

6. CONCLUSIONI

In questo contributo sono stati presentati i primi risultati di un'indagine acustica condotta su un fenomeno di alternanza tra esiti monotongali e esiti dittongali di alcune variabili vocaliche in alcuni dialetti meridionali. L'analisi dei dati, limitata per il momento a Pozzuoli e Torre Annunziata, ha messo in evidenza due aspetti fondamentali di questo fenomeno, entrambi legati alla posizione delle variabili vocaliche nella struttura prosodica: l'allungamento prepausale e la tendenza delle varianti dittongali a emergere nella posizione finale di sintagma intonativo.²⁶

L'analisi delle durate ha mostrato la presenza di un forte allungamento dei segmenti vocalici tonici nella posizione finale di sintagma intonativo rispetto alle durate in posizione interna (l'incremento è del 92% per Pozzuoli e del 116% per Torre Annunziata).²⁷ Un allungamento statisticamente significativo, anche se più modesto, è stato rilevato nella posizione finale di sintagma fonologico rispetto alle durate in posizione interna a tale costituente (l'incremento è del 35% per Pozzuoli e del 26% per Torre Annunziata). I dati relativi alle durate nelle sottocategorie definite per la posizione finale di sintagma intonativo sono meno coerenti, ma si nota comunque un allungamento significativo negli enunciati caratterizzati da intonazione ascendente. Come già documentato in diversi studi (cfr. ad es. Ladd & Campbell, 1991; Wightman *et alii*, 1992), l'allungamento è dunque di tipo incrementale: le durate si fanno progressivamente più lunghe a mano a mano che si sale nella gerarchia dei costituenti prosodici.

L'analisi del coefficiente di dittongazione ha evidenziato una polarizzazione piuttosto netta tra le realizzazioni della variabili vocaliche in posizione interna e quelle in posizione finale di sintagma intonativo. Mentre in posizione interna il coefficiente si mantiene su livelli piuttosto bassi, in posizione finale esso supera tendenzialmente (e spesso in maniera consistente) il valore di 1.8, che su base impressionistica è stato individuato come soglia critica per la percezione di dittongazione. L'alternanza monotongolo/dittongolo in questi dialetti rientra quindi tra i fenomeni di variazione fonetica che dipendono dalla posizione nella struttura prosodica e che forniscono al parlante indici acustici per la segmentazione della catena palata in costituenti prosodici.

Rispetto al quadro delineato sono presenti comunque anche dati in controtendenza. In particolare si notano diversi esiti monotongali in posizione finale, laddove sarebbero stati attesi esiti dittongali. A questi proposito occorre precisare che nei dialetti esaminati l'alternanza monotongolo/dittongolo non è regolata soltanto da fattori strutturali interni, ma assume anche significati sociolinguistici. Assodato che i dittonghi sono in genere esclusi da posizioni interne, nei contesti favorevoli alla dittongazione le varianti dittongali sono spesso in competizione con varianti monotongali. Tale competizione è intrisa di valori extralinguistici, in quanto inevitabilmente le varianti dittongali risultano marcate sia dal punto di vista geografico, perché assenti nella maggioranza delle comunità limitrofe, sia dal

²⁶ Posizione prosodica, durata e dittongazione sono strettamente interrelati. Sugli eventuali rapporti gerarchici tra queste variabili si indagherà in futuri lavori, attraverso modelli di regressione multipla che non è stato possibile esporre in questo contributo (cfr. Abete, in preparazione).

²⁷ Queste percentuali si ottengono includendo nella categoria 'interna' le durate relative alle posizioni SFa (interna al sintagma fonologico) e SFb (finale di sintagma fonologico ma interna al sintagma intonativo).

punto di vista diastratico, perché assenti nella lingua standard. I dittonghi, quindi, non assolvono una funzione esclusivamente ‘grammaticale’, marcando ad esempio la fine del sintagma intonativo, ma veicolano anche significati sociali, come l’appartenenza a un gruppo e la definizione della propria identità e della propria posizione nella società.²⁸ Privata della dimensione sociale, una certa variabilità tra realizzazioni dittongali e realizzazioni monottongali in fine enunciato risulta assolutamente caotica e priva di significato. Una parte dei dati in controtendenza potrebbe spiegarsi, dunque, chiamando in causa fattori sociolinguistici del tipo a cui si è accennato, ma che al momento non sono stati indagati sistematicamente. L’analisi qualitativa di queste ‘eccezioni’ potrebbe essere in futuro un campo di indagine molto fruttuoso, che permetterebbe di integrare nello studio dell’alternanza monottongo/dittongo anche fattori di tipo esterno.

RINGRAZIAMENTI

Si ringraziano i tre revisori anonimi, i cui commenti sono stati molto utili per la stesura finale di questo articolo.

²⁸ Per questi aspetti si veda Abete & Simpson (in stampa).

7. BIBLIOGRAFIA

- Abete, G. (2006), Sulla questione della sillaba superpesante: i dittonghi discendenti in sillaba chiusa nel dialetto di Pozzuoli, in *Analisi prosodica. Teorie, modelli e sistemi di annotazione*, Atti del 2° Convegno Nazionale dell'Associazione Italiana di Scienze della Voce, Salerno, 30 novembre – 2 dicembre 2005 (R. Savy & C. Crocco, editors), Torriana: EDK Editore, 379-398.
- Abete, G. (in preparazione), *I processi di dittongazione nei dialetti dell'Italia meridionale. Un approccio sperimentale*, Tesi di dottorato, Jena: Friedrich-Schiller-Universität.
- Abete, G. & Simpson, A. (in stampa), L'estensione della dittongazione nei giovani pescatori di Pozzuoli (NA). Dati acustici su un cambiamento fonetico in corso, in *La comunicazione parlata*, Atti del 3° Convegno Internazionale, Napoli, 23-25 febbraio 2009.
- Albano Leoni, F., Cutugno, F. & Savy, R. (1995), The vowel system of Italian connected speech, in *Proceedings of the 13th International Congress of Phonetic Sciences* (K. Elenius & P. Branderud, editors), Stockholm, Stockholm University, vol. 4, 396-399.
- Beckman, M. E. (1996), The parsing of prosody, *Language and Cognitive Processes*, 11, 17-67.
- Beckman, M. E. & Pierrehumbert, J. (1986), Intonational structure in Japanese and English, *Phonology Yearbook*, 3, 255-309.
- Berkovits, R. (1984), Duration and fundamental frequency in sentence final intonation, *Journal of Phonetics*, 12, 255-265.
- Berkovits, R. (1993), Utterance-final lengthening and the duration of final-stop closures, *Journal of Phonetics*, 21, 479-489.
- Bertinetto, P.M., Dell'Aglio, M. & Agonigi, M. (2006), Quali fattori influenzano maggiormente la durata vocalica e consonantica in italiano? Un'indagine mediante l'algoritmo di decisione C&RT, in *La Comunicazione Parlata*, Atti del congresso internazionale, Napoli 23-25 febbraio 2006 (M. Pettorino, A. Giannini, M. Vallone & R. Savy, editors), Napoli: Liguori, 13-38.
- Byrd, D. (2000), Articulatory vowel lengthening and coordination at phrasal junctures, *Phonetica*, 57, 3-16.
- Calamai, S., Marotta, G. & Sardelli, E. (2003), La modulazione di frequenza in due varietà toscane (Pisa e Firenze). Una indagine preliminare, *Quaderni del Laboratorio di Linguistica della Scuola Normale Superiore di Pisa*, 4 n.s., 11-25.
- Cho, T. (2004), Prosodically conditioned strengthening and vowel-to-vowel coarticulation in English, *Journal of Phonetics*, 32, 141-176.
- Cho, T., McQueen, J.M. & Cox, E.A. (2007), Prosodically driven phonetic detail in speech processing: the case of domain-initial strengthening in English, *Journal of Phonetics*, 35, 210-243.
- Como, P. (2006), Elicitation techniques for spoken discourse, in *Encyclopedia of language and linguistics* (K. Brown, editor), second edition, vol. 4., Amsterdam: Elsevier, 105-109.
- Crompton, A. (1980), Timing patterns in French, *Phonetica*, 37, 205-234.

- Crystal, T. & House, A. S. (1990), Articulation rate and the duration of syllables and stress groups in connected speech, *Journal of Acoustical Society of America*, 88 (1), 101-112.
- Dell'Aglio, M., Agonigi, M. M. & Bertinetto, P. M. (2002), Le durate dei foni vocalici in rapporto al contesto nel parlato di locutori pisani. Primi risultati, in *La fonetica acustica come strumento di analisi della variazione linguistica in Italia*, Atti delle VIIe Giornate di Studio del Gruppo di Fonetica Sperimentale, Napoli, 14-15 novembre 1996 (a cura di A. Regnicoli), Roma: Il Calamo, 53-58.
- Edwards, J.E., Beckman, M.E. & Fletcher, J. (1991), The articulatory kinematics of final lengthening, *Journal of the Acoustical Society of America*, 89, 369-382.
- Fletcher, J. (1991), Rhythm and final lengthening in French, *Journal of Phonetics*, 19, 193-212.
- Fougeron, C. & Keating, P. A. (1997), Articulatory strengthening at edges of prosodic domains, *Journal of the Acoustical Society of America*, 101 (6), 3728-3740.
- Grosjean, F. & Deschamps, A. (1972), Analyse des variables temporelles du français spontanée, *Phonetica*, 26, 129-156.
- Holbrook, A. & Fairbanks, G. (1962), Diphthong formants and their movements, *Journal of Speech and Hearing Research*, 5, 38-58.
- Hollander, M. & Wolfe, D. A. (1973), *Nonparametric Statistical Methods*, New York: John Wiley & Sons.
- Keating, P., Cho, T., Fougeron, C. & Hsu, C. (2003), Domain-initial articulatory strengthening in four languages, in *Phonetic interpretation* (J. Local, R. Ogden, R. Temple, editors), Papers in Laboratory Phonology 6, Cambridge: Cambridge University Press, 143-161.
- Klatt, D. H. (1975), Vowel lengthening is syntactically determined in a connected discourse, *Journal of Phonetics*, 3, 129-140.
- Klatt, D. H. (1976), Linguistic uses of segmental duration in English: acoustic and perceptual evidence, *Journal of the Acoustical Society of America*, 59, 1208-1221.
- Kohler, K. J. (1983), Prosodic boundary signals in German, *Phonetica*, 40, 89-134.
- Labov, W. (1972), Some principles of linguistic methodology, *Language in Society*, 1, 97-120.
- Ladd, R. & Campbell, N. (1991), Theories of prosodic structure: Evidence from syllable duration, in *Proceedings of the 12th International Congress of Phonetic Sciences*, Aix-en-Provence, II, 290-293.
- Lausberg, H. (1939), *Die Mundarten Südlukaniens* (Beiheft XC zur «Zeitschrift für romanische Philologie»), Halle: Niemeyer.
- Lehiste, I. (1972), The timing of utterance and linguistic boundaries, *Journal of the Acoustical Society of America*, 51, 2018-2024.
- Lindblom, B. (1968), Temporal organization of syllable production, *Speech Transm. Lab., Q. Prog. Status Rep.*, No. 2-3, 1-5.

- Lindblom, B. & Rapp, K. (1973), Some temporal regularities of spoken Swedish, *Papers in Linguistics, University of Stockholm*, 21, 1-59.
- Loporcaro, M. (1988), *Grammatica storica del dialetto di Altamura*, Pisa: Giardini.
- Lyberg, B. (1981), Some observations on the vowel duration and the fundamental frequency contour in Swedish utterances, *Journal of Phonetics*, 9, 261-272.
- Marotta, G., Calamai, S. & Sardelli, E. (2004), Non di sola lunghezza. La modulazione di f₀ come indice sociofonetico, in *Costituzione, gestione e restauro di corpora vocali*, Atti delle XIVe Giornate di Studio del G.F.S., Università della Tuscia (Viterbo), 4-6 dicembre 2003 (A. De Dominicis, L. Mori & M. Stefani, editors), Roma: Esagrafica, 215-220.
- Milroy, L. (1987), *Observing and analysing natural language*, Cambridge: Blackwell.
- Nearey, T. M. & Assman, P. F. (1986), Modeling the role of inherent spectral change in vowel identification, *Journal of the Acoustical Society of America*, 80, 1297-1308.
- Nespor, M. (1993), *Fonologia*, Bologna: Il Mulino.
- Nespor, M. & Vogel, I. (1986), *Prosodic phonology*, Dordrecht: Foris.
- Oller, D. K. (1973), The effect of position in utterance on speech segment duration in English, *Journal of the Acoustical Society of America*, 54 (5), 1235-1247.
- Pierrehumbert, J. & Beckman, M. E. (1988), *Japanese tone structure*, Cambridge (MA): MIT Press.
- R Core Development Team (2008), *R: a language and environment for statistical computing*, <http://www.R-project.org>.
- Rietveld, A. C. M. (1980), Word boundaries in French language, *Language and Speech*, 23(3), 289-296.
- Rohlf's, G. (1938), Der Einfluß des Satzakkentes auf den Lautwandel, *Archiv für das Studium der neueren Sprachen*, CLXXIV: 54-6.
- Rohlf's, G. (1966) [1949], *Grammatica storica della lingua italiana e dei suoi dialetti*, 1: Fonetica, Torino: Einaudi.
- Selkirk, E. (1984), *Phonology and syntax: The relation between sound and structure*, Cambridge (MA): MIT Press.
- Selkirk, E. (1986), On derived domains in sentence phonology, *Phonology Yearbook*, 3, 371-405.
- Shattuck-Hufnagel, S. & Turk, A. E. (1996), A prosody tutorial for investigators of auditory sentence processing, *Journal of Psycholinguistic Research*, 25, 193-247.
- Simpson, A. P. (1998), Characterizing the formant movements of German diphthongs in spontaneous speech, in *Computer Linguistik und Phonetik zwischen Sprache und Sprechen*, Tagungsband der 4. Konferenz zur Verarbeitung natürlicher Sprache – KONVENS – 98, Bonn, 5-7 ottobre 1998 (B. Schröder, W. Lenders, W. Hess & T. Portele, editors), Frankfurt: Lang, 192-200.
- Sjölander, K. & Beskow, J. (2006), *Wavesurfer*, Department of Speech and Music Communication, Stockholm: KTH.

- Sorianello, P. (1994), Il processo dell'allungamento prepausale: dati ed interpretazioni, *Quaderni del dipartimento di Linguistica, Università di Firenze*, 5, 47-73.
- Sorianello, P. (2006), Per una definizione fonetica e fonologica dei confini prosodici, in *La Comunicazione Parlata*, Atti del convegno internazionale di studi, Napoli 23-25 febbraio 2006 (M. Pettorino, A. Giannini, M. Vallone, R. Savy, editors), Napoli: Liguori, 298-318.
- Sorianello, P. & Calamai, S. (2005), Il sistema vocalico romano, in *Italiano parlato. Analisi di un dialogo* (F. Albano Leoni & R. Giordano, editors), Napoli: Liguori, 25-70.
- Trautmüller, H. (1990), Analytical expressions for the tonotopic sensory scale, *Journal of the Acoustical Society of America*, 88, 97-100.
- Tukey, J. W. (1977), *Exploratory data analysis*, Reading: Addison-Wesley .
- Vaissière, J. (1983), Language-independent prosodic features, in *Prosody: models and measurements* (A. Cutler & D. R. Ladd, editors), Berlin: Springer, 53-66.
- Vaissière, J. (1995), Phonetic explanations for cross-linguistic prosodic similarities, *Phonetica*, 52, 123-130.
- Vayra, M. & Fowler, C. (1992), Declination of supralaryngeal gestures in spoken Italian, *Phonetica*, 49(1), 48-60.
- Wightman, C. W., Shattuk-Hufnagel, S., Ostendorf, M. & Price, P. J. (1992), Segmental durations in the vicinity of prosodic phrase boundaries, *Journal of the Acoustical Society of America*, 91, 1707-1717.

1. APPENDICE: DATI RIASSUNTIVI DELLE ANALISI EFFETTUATE E RISULTATI DEI TEST STATISTICI

	SFa	SFb	SI_N	SI_Q	SI_H	SI_C	SI_L	SI_V
Min.	26,00	36,00	65,00	104,00	85,00	107,00	121,00	54,00
I Qu.	53,00	69,00	118,00	133,50	114,20	153,00	181,00	69,50
Mediana	60,00	81,00	142,00	161,00	160,00	186,00	205,00	85,00
Media	60,76	82,30	148,70	172,20	184,30	190,20	220,70	86,52
III Qu.	71,75	94,00	165,00	183,00	260,20	221,50	263,00	102,50
Max.	96,00	144,00	424,00	379,00	330,00	364,00	446,00	112,00
tokens	122	259	352	24	30	80	53	27

Tabella 2 Dati riassuntivi della durata in rapporto alla posizione prosodica per il corpus di Pozzuoli

	SFa						
ms.	21						
%	35	SFb					
p	< 0.01						
ms.	82	61					
%	136,7	75,3	SI_N				
p	< 0.01	< 0.01					
ms.	101	80	19				
%	168,3	98,8	13,4	SI_Q			
p	< 0.01	< 0.01	0.039				
ms.	126	105	44	25			
%	210	129,6	31	15,5	SI_C		
p	< 0.01	< 0.01	< 0.01	0.028			
ms.	145	124	63	44	19		
%	241,7	153,1	44,4	27,3	10,2	SI_L	
p	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01		
ms.	25	4	-57	-76	-101	-120	
%	41,7	4,9	-40,1	-47,2	-54,3	-58,5	SI_V
p	< 0.01	> 0.05	< 0.01	< 0.01	< 0.01	< 0.01	

Tabella 3: Differenze delle mediane delle durate nelle diverse classi prosodiche (Pozzuoli). Valori in ms., in percentuale, e valori di p. nel test di Kruskal-Wallis (*One-Way ANOVA by Ranks*). Il risultato del test applicato all'insieme dei campioni (preliminarmente ai confronti appaiati) è: $\chi^2 = 657.921$, $df = 6$, $p. < 2.2e-16$

	SFa	SFb	SI_N	SI_Q	SI_H	SI_C	SI_L	SI_V
Min.	43,00	54,00	79,00	176,00	127,00	117,00	141,00	53,00
I Qu.	61,50	80,00	144,00	213,50	170,50	195,50	190,50	83,25
Mediana	70,00	88,00	171,00	244,00	203,00	216,00	211,50	97,50
Media	70,86	92,43	180,10	252,40	203,80	221,40	222,70	94,31
III Qu.	79,00	4,00	208,00	279,00	222,80	245,50	237,50	5,00
Max.	21,00	50,00	375,00	362,00	277,00	315,00	402,00	35,00
tokens	135	133	392	12	7	43	42	32

Tabella 4: Dati riassuntivi della durata in rapporto alla posizione prosodica per il corpus di Torre Annunziata

	SFa							
ms.	18							
%	25,7	SFb						
p	< 0.01							
ms.	101	83						
%	144,3	94,3	SI_N					
p	< 0.01	< 0.01						
ms.	174	156	73					
%	248,6	177,3	42,7	SI_Q				
p	< 0.01	< 0.01	0.0501					
ms.	146	128	45	-28				
%	208,6	145,4	26,3	-11,5	SI_C			
p	< 0.01	< 0.01	< 0.01	> 0.05				
ms.	141,5	123,5	40,5	-32,5	-4,5			
%	202,1	140,3	23,7	-13,3	-2,1	SI_L		
p	< 0.01	< 0.01	< 0.01	> 0.05	> 0.05			
ms.	27,5	9,5	-73,5	-146,5	-118,5	-114		
%	39,3	10,8	-43	-60	-54,9	-54	SI_V	
p	< 0.01	> 0.05	< 0.01	< 0.01	< 0.01	< 0.01		

Tabella 5: Differenze delle mediane delle durate nelle diverse classi prosodiche (Torre Annunziata). Valori in ms., in percentuale, e valori di p. nel test di Kruskal-Wallis (*One-Way ANOVA by Ranks*). Il risultato del test applicato all'insieme dei campioni (preliminarmente ai confronti appaiati) è: $\chi^2 = 572.1547$, $df = 6$, $p. < 2.2e-16$.

	SFa	SFb	SI_N	SI_Q	SI_H	SI_C	SI_L	SI_V
Min.	0,00	0,00	0,15	0,60	0,65	0,80	0,47	0,00
I Qu.	0,00	0,22	1,34	2,77	1,05	2,09	2,57	0,35
Mediana	0,00	0,49	2,00	3,16	2,48	2,84	3,62	0,52
Media	0,30	0,59	2,34	3,28	2,34	2,93	3,55	1,04
III Qu.	0,52	0,86	3,13	4,18	3,44	3,52	4,87	1,64
Max.	1,95	2,96	6,11	4,80	4,23	5,95	6,09	3,18
tokens	107	154	183	17	15	62	32	21

Tabella 6: Dati riassuntivi sul coefficiente di dittongazione in rapporto alla posizione prosodica per il corpus di Pozzuoli.

	SFa						
bark	0,49						
%	//	SFb					
p	< 0.01						
bark	2	1,51					
%	//	308	SI_N				
p	< 0.01	< 0.01					
bark	3,16	2,67	1,16				
%	//	545	58	SI_Q			
p	< 0.01	< 0.01	< 0.01				
bark	2,84	2,35	0,84	-0,32			
%	//	480	42	-10	SI_C		
p	< 0.01	< 0.01	< 0.01	> 0.05			
bark	3,62	3,13	1,62	0,46	0,78		
%	//	639	81	15	27	SI_L	
p	< 0.01	< 0.01	< 0.01	> 0.05	< 0.05		
bark	0,52	0,03	-1,48	-2,64	-2,32	-3,1	
%	//	6	-74	-84	-82	-86	SI_V
p	< 0.01	> 0.05	< 0.01	< 0.01	< 0.01	< 0.01	

Tabella 7: Differenze delle mediane del coefficiente di dittongazione nelle diverse classi prosodiche (Pozzuoli). Valori in Bark, in percentuale, e valori di p. nel test di Kruskal-Wallis (*One-Way ANOVA by Ranks*). Il risultato del test applicato all'insieme dei campioni (preliminarmente ai confronti appaiati) è: $\chi^2 = 366.1515$, $df = 6$, $p. < 2.2e-16$

	SFa	SFb	SI_N	SI_Q	SI_H	SI_C	SI_L	SI_V
Min.	0,00	0,00	0,61	1,35	1,65	1,71	1,08	0,00
I Qu.	0,17	0,56	1,96	2,31	1,95	2,73	2,12	0,63
Mediana	0,46	0,87	2,49	2,94	1,99	3,31	2,72	0,86
Media	0,48	0,92	2,58	2,86	2,02	3,24	2,63	1,08
III Qu.	0,71	1,20	3,13	3,41	2,01	3,66	3,01	1,34
Max.	2,39	2,57	5,61	3,90	2,48	4,94	4,25	3,72
tokens	131	121	385	5	12	40	39	32

Tabella 8: Dati riassuntivi sul coefficiente di dittongazione in rapporto alla posizione prosodica per il corpus di Torre Annunziata

	SFa						
bark	0,41						
%	89	SFb					
p	< 0.01						
bark	2,03	1,62					
%	441	186	SI_N				
p	< 0.01	< 0.01					
bark	2,48	2,07	0,45				
%	539	238	18	SI_Q			
p	< 0.01	< 0.01	> 0.05				
bark	2,85	2,44	0,82	0,37			
%	618	280	33	12	SI_C		
p	< 0.01	< 0.01	< 0.01	> 0.05			
bark	2,26	1,85	0,23	-0,22	-0,59		
%	491	213	9	-7	-18	SI_L	
p	< 0.01	< 0.01	> 0.05	> 0.05	< 0.01		
bark	0,40	-0,02	-1,64	-2,09	-2,45	-1,87	
%	86	-2	-66	-71	-74	-69	SI_V
p	< 0.01	> 0.05	< 0.01	< 0.01	< 0.01	< 0.01	

Tabella 9: Differenze delle mediane del coefficiente di dittongazione nelle diverse classi prosodiche (Torre Annunziata). Valori in Bark, in percentuale, e valori di p. nel test di Kruskal-Wallis (*One-Way ANOVA by Ranks*). Il risultato del test applicato all'insieme dei campioni (preliminarmente ai confronti appaiati) è: $\chi^2 = 488.3946$, $df = 6$, $p. < 2.2e-16$.

