

# **XII Ciclo Scuole di Dottorato di Ricerca**

**DOTTORATO DI RICERCA IN STUDI LETTERARI E LINGUISTICI  
UNIVERSITA' DEGLI STUDI DI SALERNO**

## **TESI DI DOTTORATO**

**Validità delle metriche ritmiche:  
un'analisi e verifica su alcune varietà d'italiano  
regionale**

**Tutor  
Prof.ssa Renata Savy**

**Candidato  
Andrea Tarasi  
Matr. 8884900026**

**Anno Accademico 2014/2015**

# Indice

<b>INTRODUZIONE</b>	1
<b>Capitolo I - Il ritmo nelle lingue naturali</b>	3
<b>1. Ritmo e lingua</b>	3
<b>1.1 L'interfaccia prosodica</b>	4
<i>1.1.1 Funzioni grammaticali</i>	4
1.1.1.1 Prosodia e sintassi	5
1.1.1.2 Prosodia e semantica	5
<i>1.1.2 Funzioni testuali</i>	6
1.1.2.1 Prosodia e struttura dell'informazione	6
<i>1.1.3 L'interfaccia ritmica</i>	7
<b>1.2 La dimensione temporale del piano prosodico: parametri e costituenti</b>	7
<i>1.2.1 Accento e prominenzza</i>	7
<i>1.2.2 Durata e intensità</i>	8
1.2.2.1 <i>Speech rate</i>	9
<b>1.3 Fonologia del ritmo</b>	10
<i>1.3.1 Domini e costituenti del ritmo</i>	10
1.3.1.1 La sillaba	10
1.3.1.2 Il piede	12
<i>1.3.2 La Fonologia Prosodica e Metrica</i>	13
<b>2. Ritmo e isocronia</b>	15
<b>2.1 La tipologia ritmica e il concetto di 'classe ritmica'</b>	16
2.1.1 <i>'Stress timed' vs 'Syllable timed'</i>	17
2.1.2 <i>La classe ritmica 'mora-timed'</i>	17
<b>2.2 L'approccio fonologico</b>	18
2.2.1 <i>La proposta di Dauer</i>	18
2.2.2 <i>Isocronia e accorciamento compensatorio</i>	20
2.2.2.1 <i>L'accorciamento compensatorio nelle lingue naturali</i>	20
2.2.3 <i>Isocronia e coarticolazione: la proposta di Fowler</i>	21
<b>2.3 Studi percettivi sull'isocronia</b>	22
2.3.1 <i>Studi linguistici sulla percezione dell'isocronia</i>	22
2.3.2 <i>Il 'P-center'</i>	23
2.3.3 <i>Studi psicologici sull'isocronia</i>	23
<b>2.4 Ritmo e isocronia: confusione terminologica e metodologica</b>	24
<b>Capitolo II - Le metriche ritmiche</b>	26
<b>1. Introduzione</b>	26
<b>2. Le metriche ritmiche</b>	26
<b>2.1 Metodologia generale</b>	27

2.1.1 <i>I delta</i>	27
2.1.2 <i>Il Pairwise Variability Index (PVI)</i>	29
2.1.3 <i>Il varco<math>\Delta C</math></i>	31
2.1.4 <i>Il Compensation and Control Index (CCI)</i>	32
<b>2.2 Rassegna di studi sulle metriche ritmiche</b>	35
2.2.1 <i>La classificazione delle lingue naturali attraverso le metriche ritmiche</i>	35
2.2.1.1 <i>La classificazione di alcune varietà di italiano attraverso le metriche ritmiche</i>	37
2.2.1.2 <i>La classificazione di alcuni dialetti italiani attraverso le metriche Ritmiche</i>	38
<b>2.3 Critiche mosse alle metriche ritmiche</b>	38
<b>Capitolo III – La definizione di un dominio di analisi per il ritmo</b>	40
<b>1. Introduzione</b>	40
<b>2. Le unità di analisi prosodica</b>	42
2.1 <b>Le varie proposte sulle unità di analisi prosodica</b>	42
2.2.1 <i>Corrispondenza tra T-U e SI</i>	43
<b>3. L’approccio fonetico: la T-U</b>	44
3.1 <b>I parametri che definiscono le T-U</b>	45
<b>Capitolo IV – Materiali e metodo</b>	47
<b>1. Introduzione</b>	47
<b>2. Il corpus</b>	47
2.1 <b>Il corpus CLIPS</b>	48
2.2 <b>Selezione dei materiali</b>	50
<b>3 Metodo</b>	51
3.1 <b>Metodo di segmentazione in T-U</b>	51
3.2 <b>Principi di segmentazione in sillabe</b>	54
3.3 <b>Segmentazione in vocali e consonanti</b>	55
3.4 <b>Le metriche considerate per questo lavoro di ricerca</b>	63
<b>V Capitolo – Analisi dei dati</b>	65
<b>1. Sperimentazione delle metriche ritmiche</b>	65
1.1 <b>La metrica di Ramus et al. (1999)</b>	65
1.2 <b>La metrica di Dellwo (2006)</b>	69
1.3 <b>La metrica di Grabe &amp; Low (2002)</b>	72
1.4 <b>Riepilogo dei risultati delle metriche ritmiche</b>	76
<b>2. Variabilità inter-parlatore e intra-parlatore nelle metriche ritmiche</b>	77
2.1 <b>La variabilità inter-parlatore nelle metriche ritmiche</b>	77
2.2 <b>La variabilità intra-parlatore nelle metriche ritmiche</b>	82

<b>Conclusioni</b>	92
<b>Bibliografia</b>	94
<b>Sitografia</b>	105

## Introduzione

Lo scopo di questa ricerca è verificare la capacità classificatoria delle cosiddette metriche ritmiche su un corpus di parlato semi-spontaneo.

Negli ultimi quindici anni, lo studio del ritmo delle lingue naturali è avvenuto esclusivamente attraverso le metriche ritmiche. Questi modelli effettuano una classificazione ritmico-tipologica delle lingue naturali nelle classi ritmiche tradizionali *syllable- stress- mora-timed* che rappresentano il fondamento delle ricerche fonetiche sul ritmo (cfr. Abercrombie 1967; Ladefoged 1975). Le metriche realizzano la classificazione ritmico-tipologica delle lingue naturali tramite formule matematiche che misurano la variabilità della durata consonantica e vocalica. L'uso della durata degli intervalli vocalici e consonantici da parte di questi modelli come parametro centrale per la classificazione delle lingue naturali si fonda sulla osservazione di Bertinetto (1977) e di Dauer (1983), secondo cui l'impressione *stress- syllable-timed* si manifesta attraverso specifiche proprietà strutturali delle lingue.

I capitoli di questa ricerca affrontano il tema del ritmo come fenomeno linguistico, inquadrandolo tra i vari aspetti della prosodia. Nel primo capitolo del lavoro (cfr. §1), si accenna alle funzioni di interfaccia della prosodia, si definiscono i parametri e i costituenti delle funzioni ritmiche, gli approcci fonologici che hanno prodotto una sistematizzazione delle strutture ritmiche e infine si esamina la confusione teorica e metodologica tra ritmo, come complessa e variabile organizzazione del piano prosodico-temporale, e isocronia, come eventuale tendenza tipologica di tale organizzazione.

Questa confusione ha come conseguenza l'introduzione delle metriche ritmiche tra gli approcci sviluppati per lo studio del ritmo: una descrizione approfondita di queste metriche è nel secondo capitolo della ricerca (cfr. §2), in cui si affronta anche il tema della tipologia ritmica, si analizza la classificazione resa dalle metriche ritmiche per alcune lingue naturali, varietà di italiano regionale e dialetti italiani e, infine, si esaminano le critiche mosse a questi modelli.

La maggior parte delle metriche ritmiche (cfr. Ramus *et al.* 1999; Grabe & Low 2002; Dellwo 2006), sono sperimentate su unità linguistiche molto diverse tra loro. Così facendo, l'analisi ritmica avviene su unità linguistiche eterogenee per dimensione, struttura e parametri. Questo *modus operandi*, però, comporta la perdita di informazioni linguistiche, e specificatamente ritmiche, che l'unità linguistica riflette, perché essa è il 'mezzo' attraverso cui si esplicitano relazioni tra i vari livelli della grammatica. Si rende necessario, dunque, individuare un'unità ritmicamente pertinente, che funga da dominio per lo studio del fenomeno ritmo. Questo tema si affronta nel terzo capitolo della ricerca (cfr. §3).

Accanto al problema della definizione dell'unità di analisi, un forte dibattito si è sviluppato in merito ai corpora su cui si sperimentano le metriche ritmiche. Nella maggior parte delle ricerche, le metriche ritmiche sono applicate a parlato letto, mentre pochissimi lavori arrivano a testare questi modelli su parlato spontaneo e semi-spontaneo: come sottolinea Arvaniti (2009, 2013), in tal caso i risultati sono altamente contrastanti. Sulla scia del lavoro di Arvaniti, in questa ricerca si verifica la validità classificatoria di alcune metriche su parlato semi-spontaneo che compone il corpus di questa ricerca. Tale corpus comprende otto varietà di italiano regionale, che danno una rappresentazione geolinguistica dell'intero territorio italiano. La descrizione del corpus della ricerca si trova nel quarto capitolo (cfr. §4). Nel capitolo successivo (cfr. §5), si sperimentano le suddette metriche sul corpus della ricerca e si presentano i risultati della sperimentazione.

# Capitolo I - Il ritmo nelle lingue naturali

## 1. Ritmo e lingua

Il ritmo è parte fondamentale della nostra vita, perché caratterizza vari aspetti biologici come i battiti del cuore, l'alternanza tra il giorno e la notte ma anche numerose attività come la musica, la danza e la lingua. La lingua è fortemente caratterizzata dal ritmo e la sua natura si palesa attraverso la ricorrenza ciclica di specifici elementi. Infatti, in modo generale, il ritmo linguistico può essere definito come l'alternanza regolare tra elementi deboli ed elementi forti. Gli elementi deboli e forti corrispondono a specifici elementi linguistici, la cui forza dipende da un concetto relativo, cioè dalla differenziazione tra elementi percettivamente prominenti e non prominenti.

La definizione generica di ritmo linguistico sopra data, comporta la necessità di individuare gli elementi linguistici su cui misurare il concetto di forza o debolezza, i parametri che veicolano tale funzione e le strutture ritmico-linguistiche che inglobano queste dinamiche. A tali propositi, sono state avviate numerose ricerche che seguono varie direzioni. Tali ricerche introducono nuovi elementi, nuovi parametri e si posizionano sui versanti più disparati: da quello fonetico-fonologico a quello psicologico-percettivo a quello più propriamente tipologico che ha come obiettivo la realizzazione di modelli di classificazione delle lingue naturali su basi ritmiche. Tali ricerche raggiungono spesso risultati eterogenei, che mettono in discussione la definizione stessa di ritmo linguistico. A questo punto, è necessario chiedersi: qual è il dominio del ritmo linguistico? Quale funzione svolge? Come vedremo in §1.3.1 e §1.3.2, corrispondono a sillabe e piedi. Questi elementi guidano la produzione, percezione e sincronizzazione tra l'emittente e il ricevente, allo scopo di consentire la decodifica del messaggio verbale. Quindi, il ritmo appartiene all'ambito della codifica 'prosodica' del messaggio, cioè a ciò che viene definita 'interfaccia prosodica'. Tale interfaccia lega e specifica il piano della produzione fonica con il piano della significazione.

Nello sviluppo di questo capitolo, quindi: accenneremo alle funzioni di interfaccia (cfr. §1.1), definiremo i parametri (cfr. §1.2) e poi i costituenti (cfr. §1.3) delle funzioni ritmiche, delineremo gli approcci fonologici che hanno prodotto una sistematizzazione delle strutture ritmiche (cfr. §1.3.2; §1.3.3) per arrivare infine alla continua contrapposizione e confusione tra ritmo e isocronia (cfr. §2). Tale contrapposizione è il punto centrale di questa tesi, perché rappresenta l'origine di visioni e di ricerche interessanti sebbene molto controverse sul piano teorico, metodologico e analitico.

## **1.1 L'interfaccia prosodica**

Un aspetto importante delle ricerche condotte sulla prosodia è lo studio delle sue funzioni e delle sue relazioni con gli altri livelli della grammatica (cfr. §1.1.1, §1.1.2 e rispettivi sotto paragrafi). Le funzioni linguistiche della prosodia si distinguono in testuali e grammaticali. Tra le funzioni testuali menzioniamo la capacità della prosodia di porre in rilievo gli elementi prominenti, mentre tra le funzioni grammaticali ricordiamo la capacità della prosodia di esprimere sia contrasti sistematici come quelli tra modalità differenti (interrogativa vs. dichiarativa) sia di suddividere il parlato in unità prosodiche (su questo tema torneremo più avanti).

Le ricerche sull'interfaccia prosodica hanno spesso ristretto il proprio campo di azione verso le funzioni dell'intonazione. Pochissimi studi, invece, hanno esaminato le funzioni del piano ritmico-temporale e le sue relazioni con gli altri livelli della grammatica (cfr. §1.1.3). La preponderanza assegnata al piano melodico è dovuta soprattutto al metodo di valutazione adottato nei vari studi che consiste, in molti casi, in un'analisi prevalentemente uditiva degli andamenti prosodici, attraverso la quale è facilmente valutabile la curva melodica. Diversa, invece, è la situazione del piano ritmico-temporale, perché i parametri presentano una forte complessità sia nella misura che nella valutazione. Tuttavia, dal momento che il ritmo è un componente della codifica prosodica, possiamo assumere con una certa sicurezza che anch'esso manifesta una forte correlazione con gli altri livelli della grammatica. La prominenza, infatti, non è solo un concetto di natura intonativa (cfr. Halliday 1985) ma anche di natura sintattica e semantica (cfr. Soriano 2006).

Nei paragrafi e sotto paragrafi successivi, descriviamo i rapporti che intercorrono tra il piano melodico e quello ritmico-temporale con alcuni livelli della grammatica.

## **1.1.1 Funzioni grammaticali**

### **1.1.1.1 Prosodia e sintassi**

L'aspetto più studiato della relazione prosodia-sintassi è il rapporto che intercorre tra intonazione e sintassi. Le prime ricerche che studiano la relazione tra questi due livelli della grammatica postulano una completa subordinazione della prosodia alla sintassi (cfr. Chomsky & Halle 1968). Con il passare degli anni, viene smontata l'ipotesi del completo isomorfismo tra i due livelli della grammatica, avanzando quella di assenza di coestensione tra una particolare struttura sintattica e l'unità prosodica. Si evidenziano, piuttosto, delle correlazioni più probabili, anche se altamente variabili, che coinvolgono un'ampia gamma di strutture sintattiche (cfr. Crystal 1969; Halliday 1976; Cruttenden 1986; Voghera 1990), anche e soprattutto in funzione della velocità d'eloquio (cfr. Sornicola 1981).

Esiste, però, un'altra tematica di confine tra sintassi e prosodia e riguarda l'analisi dei meccanismi che inducono l'alterazione dell'ordine basilare degli elementi di una frase. Questi studi dimostrano che lo spostamento degli elementi all'interno di una frase produce una diversa segmentazione del parlato e, di conseguenza, una corrispondenza variabile tra unità sintattiche e unità prosodiche (cfr. Sorianello 2006).

### **1.1.1.2 Prosodia e semantica**

L'area di interfaccia prosodia/semantica è molto vasta, perché riguarda lo studio dell'ampia gamma di funzioni di significazione della prosodia. Sebbene lo studio di tali funzioni esuli dagli scopi di questo lavoro, vale la pena accennarle molto brevemente.

Anche in questo caso, l'aspetto più trattato è la relazione tra intonazione e semantica, attraversata da due importanti questioni: 1) quali sono le unità significative dell'intonazione; 2) quali significati sono associati a queste unità.

La prima questione viene affrontata attraverso due approcci: quello olistico e quello compositivo. Secondo l'approccio olistico (Bolinger 1964, Liberman 1975, Liberman & Sag 1974, cfr. von Stechow 1999: 90), il contorno intonativo esprime una funzione che non può essere scomposta in elementi significativi, quindi le piccole unità che compongono il contorno intonativo non contribuiscono a determinarne il significato. L'approccio compositivo, invece, assume che gli elementi che compongono il contorno intonativo convogliano il significato. Quindi, il significato del contorno intonativo è estratto dai suoi elementi (cfr. Pierrehumbert & Hirschberg 1990; Hobbs 1990; Selkirk 1995).

La seconda questione, invece, viene affrontata sia attraverso approcci che assegnano una funzione molto vaga alla peculiarità intonativa (Bolinger 1964, cfr. von Heusinger 1999), sia tramite approcci che attribuiscono un significato piuttosto specifico alla peculiarità intonativa (Lieberman & Sag 1974, cfr. von Heusinger 1999).

## **1.1.2 Funzioni testuali**

### **1.1.2.1 Prosodia e struttura dell'informazione**

L'interfaccia tra intonazione e struttura informativa riguarda la relazione tra unità tonale (cfr. Cap.3) e unità di informazione. Infatti, l'interfaccia intonazione/struttura informativa sancisce che l'unità tonale venga considerata come espressione di una unità di informazione:

<<*The information unit is what its name implies: a unit of information. Information is a process of interaction between what is already known or predictable and what is new or unpredictable. It is this interplay of new and not new that generates information in the linguistic sense*>> (cfr. Halliday 1985: 274-275).

Quindi, la natura della unità tonale è quella di segnalatore dell'organizzazione del contenuto del messaggio in una parte 'nuova' (*new o unpredictable*), che è anche l'unica obbligatoria, e una parte 'non nuova' o 'data' (*already known o predictable*), che è facoltativa. L'informazione 'data' corrisponde all'informazione che il locutore ritiene in qualche modo presente nella consapevolezza del proprio interlocutore e per questo motivo ritiene opportuno non sottolinearla. La differenziazione tra le due parti avviene attraverso un punto di rilievo, che coincide con un elemento prominente: la parte dell'unità tonale che contiene l'elemento prominente indica l'informazione nuova, mentre il resto della unità tonale denota quella data.

La prominenza serve anche a marcare il centro di interesse del discorso, detto anche *focus*. Sulla relazione focus/prominenze esiste una ampia letteratura, perché tale aspetto in linguistica è stato affrontato da diverse prospettive. Ancora una volta, però, nessuna ricerca ha analizzato l'apporto del ritmo in questa relazione. Lo studio di tali dinamiche esula dagli scopi di questo lavoro, quindi ci limitiamo a dire che la collocazione della prominenza è una delle strategie che consente nelle lingue di porre il *focus* su alcune parti della unità tonale. Nella letteratura dedicata al *focus*, si distingue tradizionalmente tra: *focus* esteso (*broad focus*), *focus* ristretto (*narrow focus*) e *focus* contrastivo (*contrastive focus*), le cui differenze sono veicolate, per l'appunto, anche e soprattutto da fatti prosodici.

### **1.1.3 Interfaccia ritmica**

Come accennato in §1.1, pochissimi studi hanno analizzato il rapporto tra ritmo e altri livelli della grammatica e le uniche relazioni esaminate sono: ritmo/sintassi e ritmo/pragmatica.

Gli studi condotti sulla interfaccia ritmo/sintassi dimostrano sia un condizionamento reciproco tra i due livelli della grammatica, dove la struttura temporale risulta condizionata dalla sintassi e la struttura sintattica subisce a sua volta l'influsso del ritmo (cfr. Marotta 1985), sia la capacità del ritmo di intervenire nella scelta di specifiche costruzioni sintattiche (cfr. Shih *et al.* 2009). Le ricerche condotte sulla relazione ritmo/pragmatica, invece, rivelano sia una continua modifica del ritmo di elocuzione con cui vengono prodotti i turni, perché i locutori si adattano a specifiche condizioni comunicative (cfr. Licoppe & Smoreda 2005), sia un ritmo di eloquio variabile che dipende dalla modalità (assertiva vs. esclamativa) della frase (cfr. Sorianello 2011).

## **1.2 La dimensione temporale del piano prosodico: parametri e costituenti**

L'analisi fonetica e prosodica del parlato si basa sulla individuazione di variazioni attraverso la misurazione di parametri che hanno la caratteristica della relatività: sul piano temporale sono pertinenti le variazioni di durata dei segmenti, mentre sul piano intonativo i valori di intensità, altezza e i loro andamenti nel tempo. Questi parametri interagiscono tra loro e concorrono a segnalare i costituenti e le parti prominenti a diversi livelli nella stringa fonetica. Nei paragrafi e sotto paragrafi successivi, definiremo innanzitutto i parametri delle funzioni ritmiche.

### **1.2.1 Accento e prominenzza**

Accento e prominenzza sono i parametri su cui gravita l'organizzazione ritmica di una lingua. L'accento, da un punto di vista articolatorio, è correlato alla quantità di energia impiegata per articolare una sillaba. L'aumento della pressione dell'aria all'interno del condotto fonatorio è dovuta a una maggiore attività dei muscoli espiratori. Sul piano fonetico-acustico ciò implica una maggiore durata, intensità e dinamica di  $f_0$  della sillaba accentata, che può essere avvertita a livello uditivo. Il ruolo di questi parametri fisici non è facilmente individuabile, perché la loro azione spesso si interseca e si sovrappone.

La prominenzza, invece, tradizionalmente definita '*a local degree of stress or emphasis*' (cfr. Liberman & Pierrehumbert 1984: 157) è il risultato, sul piano percettivo, dell'accentazione: essa riguarda il grado di salienza assegnato a sillabe e parole e svolge una funzione sul piano fonologico, vale a dire 'linguistico'.

Va sottolineato, però, che secondo la fonologia in un eloquio non esistono solo sillabe con accento primario<sup>1</sup> e sillabe non accentate, ma esistono anche sillabe con accenti secondari<sup>2</sup>. Le sillabe che portano tali accenti sono meno prominenti delle sillabe con accento primario, ma più prominenti delle sillabe senza accento.

Nel sotto paragrafo successivo, esaminiamo i parametri che caratterizzano le funzioni ritmiche.

### 1.2.2 Durata e intensità

La durata e l'intensità, come descritto in §1.2.1, sono parametri fisici che partecipano alla realizzazione dell'accento e della prominente, elemento cardine per la descrizione del ritmo delle lingue naturali.

La durata è l'estensione temporale relativa con cui sono prodotti i suoni della lingua considerati nella propria realtà fisica. E' un parametro molto variabile in funzione del contesto enunciativo e del contesto linguistico. Infatti, vari fattori influenzano la durata e di seguito ne citiamo alcuni: il numero dei foni circostanti, lo stile utilizzato durante la comunicazione e soprattutto la velocità di elocuzione o *speech rate* (tale parametro è descritto in §1.2.2.1).

La durata è un parametro che contribuisce alla realizzazione della prominente: le sillabe o i foni prominenti hanno una durata maggiore rispetto a quelli non prominenti. A tale proposito, gli studi effettuati sulla variazione temporale dei componenti sillabici in diverse condizioni (cfr. Greenberg *et al.* 2003: 13) dimostrano che la durata interessa soprattutto i nuclei vocalici, perché i nuclei rivelano una maggiore 'elasticità' rispetto agli altri componenti della sillaba (ricordiamo che la descrizione della sillaba e dei suoi componenti è affrontata in §1.3.1.1. Questo significa che il nucleo assorbe la maggior parte dell'impatto della prominente. Le sillabe prominenti sono pronunciate con una maggiore intensità, cioè con una prominente dovuta a un aumento del volume della voce. L'aumento del volume della voce consente di individuare sul piano uditivo le sillabe prominenti e, di conseguenza, permette di determinare e descrivere l'andamento ritmico degli enunciati. Sul piano fisiologico, l'intensità è correlata a più fattori: la pressione dell'aria proveniente dai polmoni, la quantità di energia articolatoria, il grado di tensione delle pliche vocali. Questo correlato acustico acquista significato solo se viene confrontato con valori di intensità rilevati nello stesso enunciato per i segmenti precedenti o successivi.

Un altro parametro delle funzioni ritmiche centrale per lo studio del ritmo delle lingue naturali è lo *Speech Rate*.

---

<sup>1</sup> Per una definizione di accento primario e secondario (cfr. Nespor 1993).

<sup>2</sup> Per una definizione di accento secondario (cfr. Nespor 1993).

### 1.2.2.1 *Speech Rate*

La definizione di *Speech Rate* o Velocità di Elocuzione (d'ora in poi VDE) ancora oggi è una questione aperta. Vari autori hanno cercato di definire tale fenomeno, ma spesso con una certa sovrapposizione rispetto a un altro fenomeno, cioè alla Velocità di Articolazione o *Articulation Rate* (d'ora in poi VDA). La sovrapposizione tra i due fenomeni può essere notata nelle seguenti definizioni:

VDE: <<[...] espresso come numero di sillabe al secondo della sequenza articolata, risultante dal rapporto tra il numero delle sillabe della sequenza articolata e la durata dell'intera catena fonica>> (Zmarich *et al.* 1996, cfr. Romito *et al.* 2005: 4).

VDA: <<il numero delle sillabe diviso per la durata della catena fonica>> (Magno Caldognetto & Vagges 1993, cfr. Romito *et al.* 2005: 4).

Tutto ciò produce forti problematiche, in quanto il calcolo dei due parametri avviene attraverso la stessa formula, cioè dividendo il numero delle sillabe per la durata della catena fonica, che corrisponde alla porzione di segnale compresa tra due pause. La soluzione a questa sovrapposizione viene da Magno Caldognetto & Vagges:

VDE: <<definita come numero delle sillabe diviso per la durata dell'enunciato>> (cfr. Magno Caldognetto & Vagges 1991: 424).

A differenza delle definizioni precedenti che prevedevano la divisione del numero delle sillabe per la durata della catena fonica, gli autori dividono le sillabe per la durata dell'enunciato che comprende al proprio interno anche la durata della pausa che lo delimita.

Sul piano temporale, come accennato in §1.2.2, la VDE ha un forte impatto sulla durata degli elementi linguistici: la durata delle vocali, delle consonanti e delle sillabe è fortemente influenzata dalla VDE adottata dal locutore. In condizioni 'normali' di eloquio, la VDE viene scelta dal locutore in relazione a vari fattori come il contesto e lo stile o registro. Pertanto, differenze di VDE affiorano quando questo parametro viene testato su tipi diversi di parlato: il parlato spontaneo è caratterizzato da una VDE diversa rispetto al parlato letto. Alcuni studi (cfr. Bertini & Bertinetto 2007: 3; Taranto *et al.* 2011: 1) individuano variazioni di durata consonantica e vocalica tra i due tipi di parlato strettamente connesse a variazioni di VDE da parte dei locutori. Queste fluttuazioni di

VDE che modificano le durate vocaliche e consonantiche realizzano tangibili conseguenze sul ritmo dell'eloquio.

Dopo la descrizione dei parametri delle funzioni ritmiche, affrontiamo la fonologia del ritmo (cfr. §1.3) che determina e descrive i domini e i costituenti del ritmo.

## 1.3 Fonologia del ritmo

### 1.3.1 Domini e costituenti del ritmo

#### 1.3.1.1 La sillaba

La sillaba è nozione intuitivamente presente nella competenza del parlante nativo. L'esistenza della sillaba nel comportamento dei parlanti delle diverse lingue è un fatto riscontrabile empiricamente: la sillabazione si manifesta in varie pratiche come il canto, la marcia ed emerge anche dal comportamento dei parlanti al momento dei primi stadi dell'acquisizione linguistica: le lallazioni coinvolgono sillabe e non suoni isolati. In più, il riconoscimento della sillaba sul piano fonologico permette di spiegare in modo più efficace il comportamento di molti processi tra cui lapsus linguistici, fenomeni di rafforzamento o indebolimento ecc.

In ambito fonologico, la definizione della sillaba gode di un accordo generale. La sillaba è una unità fonologica dotata di una struttura interna. Questa struttura è composta da un *nucleo* o *apice* che in genere è occupato da un elemento vocalico obbligatorio<sup>3</sup> e da elementi consonantici opzionali che prendono il nome di *incipit* o *attacco* quando precedono il nucleo, di *coda* quando invece lo seguono. Il nucleo e la coda formano la *rima*. La composizione della sillaba può essere schematizzata nel modo seguente:

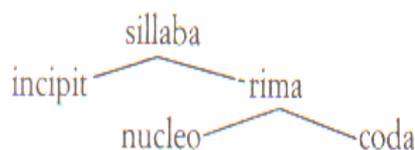


Figura 1.1: struttura della sillaba (da Nespor 1993: 156).

Una sillaba può essere aperta o chiusa: è aperta quando la sillaba termina per vocale, cioè è priva di una coda, mentre è chiusa o implicata quando contiene una consonante in coda.

L'identificazione, invece, di tratti fonetici a cui affidare il riconoscimento della sillaba e dei suoi confini risulta problematica. Varie teorie articolatorie e motorie e diversi approcci uditivi o

---

<sup>3</sup> Esistono lingue come l'inglese dove il nucleo della sillaba può essere occupato da alcune consonanti.

propriamente acustici<sup>4</sup> hanno cercato di fornire una definizione di sillaba che fosse non solo giustificata sul piano fonologico, ma anche ancorata a realtà di natura fisica. Attualmente sembra esserci accordo generale nel considerare la sillaba come un insieme di fonemi raggruppati intorno a un picco di intensità sonora, che di norma coincide con la vocale, il fono più sonoro e quindi più forte anche sul piano percettivo. La disposizione dei fonemi intorno al nucleo è dettata sia da principi fonologici, sia da criteri di ordine fonotattico e fonetico. Le consonanti non possono disporsi in maniera casuale intorno al nucleo e ogni lingua ha specifiche restrizioni fonotattiche e precise regole di sillabazione. Tali restrizioni e regole si uniformano sia a principi di sonorità universalmente validi sia a principi linguo-specifici. In linea generale, i fonemi si dispongono lungo il continuum fonico secondo il loro grado di sonorità intrinseco. Il grado di sonorità intrinseco dei fonemi è stabilito sulla base delle cosiddette scale di sonorità<sup>5</sup> che ordinano gerarchicamente i segmenti che possono occupare varie posizioni all'interno della sillaba e tale gerarchia è di tipo monotratto: un unico parametro, grado di sonorità o di forza consonantica, determina la gradazione lungo la scala di sonorità per definire i vari stadi della gerarchia. In genere, le vocali sono caratterizzate dal massimo grado di sonorità e pertanto costituiscono l'elemento intorno al quale si raggruppano i fonemi consonantici in un ordine di sonorità che cresce progressivamente verso il nucleo e decresce nel momento in cui ci si allontana dal nucleo.

Esistono casi però in cui la buona formazione sillabica consente due sillabificazioni possibili ma solo una è quella reale, quella cioè che corrisponde alle intuizioni dei parlanti nativi di una certa lingua e che può predire delle regolarità nel sistema fonologico. Per risolvere questo problema, Kurylowicz (Kurylowicz 1947, cfr. Nespor 1993) propone che l'inizio e la fine possibili di una sillaba in una data lingua 'x' siano simili all'inizio e alla fine possibili di una parola nella stessa lingua, ma questo principio di sillabificazione lascia alcuni casi irrisolti, per cui è stato affiancato da due principi ausiliari. Uno di questi principi, detto Principio dell'Incipit Massimo, stabilisce che nel caso in cui esistano due possibilità per una consonante mediana, nel senso che sia una coda oppure un incipit possibile, allora si sceglie per l'incipit della sillaba successiva. Per esempio, questo principio stabilisce che 'caro' sia sillabificato [ka.ro] e non [kar.o]. Le consonanti che non sono state assegnate attraverso questo principio, vengono inserite nella coda della sillaba precedente.

Un altro costituente rilevante per la descrizione del ritmo delle lingue naturali è il piede.

---

<sup>4</sup> Per una rassegna su questi studi (cfr. Bertinetto 1981).

<sup>5</sup> Per una descrizione delle scale di sonorità (cfr. De Dominicis 2003).

### 1.3.1.2 Il piede

Il piede inteso come una unità ritmica ha origine nella metrica classica, dove indica un modulo ritmico formato da sillabe brevi o lunghe tenute insieme da un accento che cade costantemente sulla sillaba lunga. In un piede si distingue un tempo forte, che coincide con la sillaba su cui cade l'accento, e un tempo debole, concetti musicali strettamente connessi al tipo di accento (forte o debole) che caratterizza un tempo. Anche la fonologia utilizza il piede come costituente imprescindibile per la descrizione del ritmo delle lingue naturali. Come nella metrica classica, in fonologia il piede può dominare una o più sillabe e le sillabe contenute in un piede non hanno uguale forza, cioè prominente: una sola è prominente e le altre sono deboli. Per maggiore chiarezza, di seguito riportiamo un esempio di struttura del piede:

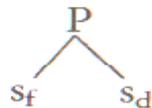


Figura 1.2: struttura del piede (da Nespor 1993: 164).

Studi effettuati in ambito generativo delineano alcuni principi ‘universali’ che dettano la costruzione del piede:

1. posizione dell'elemento prominente;
2. sensibilità o meno alla quantità sillabica;
3. complessità del piede, vale a dire numero limitato o illimitato di sillabe all'interno del piede.

Secondo il primo principio universale, l'elemento prominente è sempre situato a uno dei due limiti estremi, il sinistro oppure il destro. Risulta rara, invece, la possibilità che la sillaba prominente sia seguita e preceduta da sillabe non prominenti. Nel caso in cui ciò si manifesti, il piede che ne risulta è definito ‘anfibraco’. Se il piede anfibraco contiene una sola sillaba è definito ‘degenerato’. In più, se la sillaba di un piede anfibraco degenerato non è prominente, allora è altamente marcato. Le lingue che possiedono piedi anfibrachi sono definite marcate.

Il secondo principio universale sancisce che il piede può essere sensibile alla quantità sillabica. Le lingue sensibili alla quantità sillabica consentono lunghe sequenze di sillabe non prominenti, quindi ritmi irregolari. Pertanto, in base a questo principio si possono classificare lingue che presentano sequenze ritmiche più regolari e lingue con sequenze aritmiche, cioè sequenze con ritmo irregolare.

Infine, il terzo principio universale stabilisce il numero di sillabe che può contenere un piede, limitato o illimitato. Le lingue che consentono piedi con numero illimitato di sillabe sono caratterizzate da un ritmo irregolare.

Nel sotto paragrafo successivo, esponiamo i modelli fonologici che perseguono l'obiettivo di descrivere il ritmo delle lingue naturali.

### **1.3.3 La Fonologia Prosodica e Metrica**

La Fonologia Metrica e la Fonologia Prosodica, nate in seno al modello Autosegmentale, producono una sistematizzazione delle strutture ritmiche.

Nel modello della Fonologia Metrica (cfr. Liberman & Prince 1977), il piano prosodico della lingua è descritto attraverso una struttura ad albero con nodi forti e deboli, organizzata in costituenti gerarchici. La prominente, che corrisponde ai nodi forti, è assegnata dentro ciascun costituente attraverso regole linguo-specifiche. L'albero metrico poggia sulla griglia metrica. La costruzione della griglia è soggetta a due regole: 1) al primo livello della griglia metrica, tutte le sillabe ricevono un asterisco; 2) a ogni costituente prosodico superiore alla sillaba corrisponde un livello sulla griglia metrica e a ogni elemento prominente viene assegnato un asterisco al livello corrispondente. L'elemento prominente è definito elemento terminale designato (*Designated Terminal Element* o DTE). Quindi, la regola 1) assegna gli asterischi alle sillabe, che rappresentano le unità temporali in una concezione astratta del tempo; la regola 2) assegna le prominente sull'asse verticale sulla base dei nodi forti dell'albero prosodico, in modo che ogni nodo contrassegnato dall'etichetta forte, cioè il DTE di un costituente prosodico, riceve un asterisco al relativo livello della griglia metrica. Quindi, la griglia metrica assorbe dall'albero prosodico le informazioni sulla distribuzione delle prominente e le esplicita nella propria rappresentazione. Di seguito riportiamo un esempio:

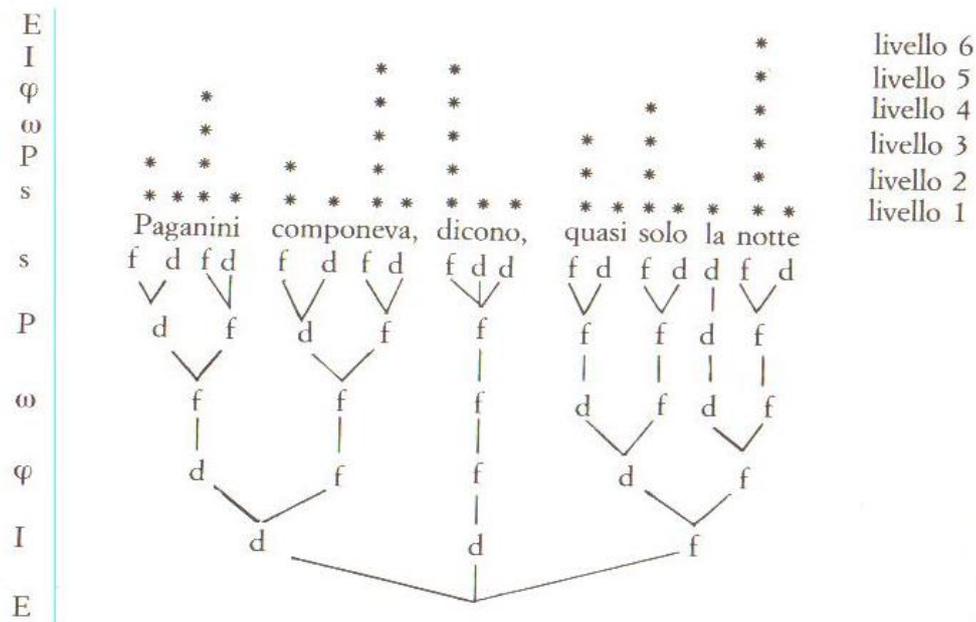


Figura 1.4: rappresentazione albero prosodico/griglia metrica (da Nespor 1993: 240).

La parte superiore della figura 1.4 contiene la rappresentazione della griglia metrica, la parte inferiore mostra l'albero prosodico e nella parte sinistra sono riportati i costituenti.

Nella Fonologia Prosodica (cfr. Nespor & Vogel 1986; Nespor 1993), il livello fonologico o prosodico e quello sintattico sono organizzati gerarchicamente in diversi costituenti. Tra questi costituenti esiste una connessione, anche se non fissata in modo deterministico, che riguarda soprattutto i costituenti prosodici di livello più alto. Allo scopo di verificare proprio tali relazioni, la Fonologia Prosodica postula la costruzione della gerarchia prosodica e della gerarchia sintattica. Di seguito, per maggiore chiarezza, riportiamo un esempio:

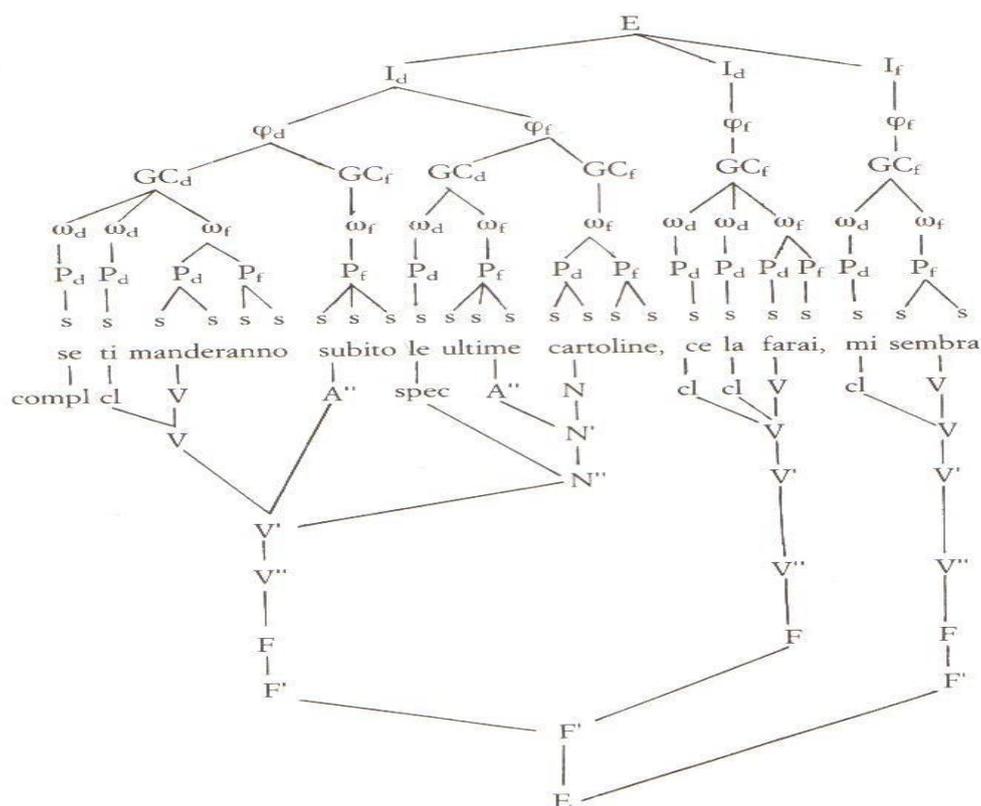


Figura 1.3: rappresentazione albero prosodico/albero sintattico (da Nespor 1993: 210)..

Nella figura 1.3, la parte inferiore raffigura l'albero sintattico e i costituenti<sup>6</sup> che lo compongono, la parte superiore illustra l'albero prosodico e i propri costituenti<sup>7</sup>.

## 2. Ritmo e isocronia

In linguistica, il ritmo è stato studiato in diversi ambiti e in questa ricerca poniamo l'attenzione verso gli studi fonetici (cfr. §2.1), fonologici (cfr. §2.2) e percettivi (cfr. §2.3). Questi studi, come vedremo più avanti, sviluppano metodologie e modelli che sono attraversati da una continua confusione tra due fenomeni ben distinti: ritmo e isocronia (cfr. §2.4). Il ritmo come fenomeno linguistico è stato definito nel paragrafo precedente (cfr. §1), mentre per isocronia, sempre in linguistica, si intende la ripetizione costante di specifici elementi linguistici, cioè sillabe o piedi. Nelle varie ricerche sul ritmo, la metodologia adottata per descrivere questo fenomeno prosodico prevede la misura della distanza, espressa attraverso la durata, tra due sillabe accentate o due piedi accentati. Così facendo, si passa dallo studio dell'alternanza tra elementi deboli ed

<sup>6</sup> F = Frase; V = Verbo; N = Nome; A = Avverbio; cl = Clitico; compl = Complementatore.

<sup>7</sup> s = Sillaba; P = Piede;  $\omega$  = Parola Fonologica; GC = Gruppo Clitico;  $\Phi$  = Sintagma Fonologico; I = Sintagma Intonativo; E = Enunciato.

elementi forti (ritmo) alla misura della distanza espressa attraverso la durata (isocronia). Tutto ciò produce come risultato una descrizione delle lingue naturali su basi isocrone e non ritmiche.

## 2.1 La tipologia ritmica e il concetto di ‘classe ritmica’

Attraverso la tipologia ritmica o ‘*rhythm typology*’ le lingue naturali vengono classificate ritmico-tipologicamente in base a comportamenti isocroni in specifiche classi ritmiche: *syllable-timed*, *stress-timed* e *mora-timed*.

I primi studi fonetici sul ritmo delle lingue naturali, come accennato in §2, descrivono tale fenomeno, cercando di individuare nei segnali vocali la ricorrenza costante, attraverso misure di durata (isocronia), di specifici elementi linguistici, cioè sillabe e piedi. Questi primi studi fonetici però raggiungono risultati in parte divergenti: Scripture (1902), Russelot (1924), Classe (1939) e Jones (1956) (Scripture 1902; Russelot 1924; Classe 1939; Jones 1956, cfr. Cumming 2010: 4) dimostrano che nei segnali vocali esiste una ricorrenza isocrona di sillabe e di piedi, mentre Scripture (Scripture 1902, cfr. Cumming 2010: 4) non individua alcuna ricorrenza isocrona a livello di piede.

Negli anni quaranta, anche se le ricerche appena citate mostrano risultati in parte divergenti, vari autori imputano le differenze ritmiche che esistono tra le lingue naturali alla ricorrenza isocrona di sillabe e piedi. L’idea che le lingue possiedano un ritmo differente è avanzata da Lloyd James (Lloyd James 1940, cfr. Cumming 2010: 5). L’autore osserva che il ritmo dello spagnolo è simile a quello di una ‘mitragliatrice’ (*Machine Gun*), mentre quello dell’inglese a un messaggio in ‘codice morse’ (*Morse Code*). Successivamente Pike (Pike 1945, cfr. Mairano 2010: 10), nel tentativo di fornire un supporto empirico a questa dicotomia, imputa tale differenza al modo in cui agisce l’isocronia nelle varie lingue: l’isocronia incide sulla durata sillabica in lingue come lo spagnolo, mentre influenza gli intervalli interaccettuali in lingue come l’inglese. Partendo da questa considerazione, l’autore postula l’esistenza di due classi ritmiche, che possiamo anche definire classi ‘isocrone’(in questo lavoro, consideriamo l’espressione classe ritmica sinonima di classe isocrona, perché tali classi si fondano su principi isocroni e non ritmici): *syllable-timed* (italiano, spagnolo e francese) e *stress-timed* (inglese e tedesco). Nelle lingue *syllable-timed* (isocronia sillabica o isosillabiche), l’isocronia si manifesta a livello di sillaba, nelle lingue *stress-timed* (isocronia accentuale o isoaccettuali), l’isocronia si manifesta a livello di piede. Nelle lingue *stress-timed*, la distanza tra un accento e l’altro è relativamente costante. Questo processo provoca un adeguamento della durata sillabica, cioè le sillabe devono rientrare entro i confini stabiliti da due accenti. Nelle lingue *syllable-timed*, invece, la durata delle sillabe tende a rimanere relativamente costante. Nonostante Pike (1945) postulasse differenze di isocronia in tutte le lingue naturali, sostiene

che i due tipi di isocronia possano anche coesistere in una stessa lingua: ad esempio, una lingua come l'inglese che è considerata *stress-timed* in alcune circostanze può anche assumere le caratteristiche di una lingua *syllable-timed* (ad esempio nel canto) (cfr. Cumming 2010: 5).

Sebbene l'autore effettui questa importante precisazione, come vedremo in §2.1.1, le differenze di isocronia vengono adottate per distinguere le lingue naturali in modo dicotomico, cioè le lingue devono necessariamente appartenere a una delle due classi 'isocrone' menzionate. Questo passaggio avvia la cosiddetta *rhythm typology* o tipologia ritmica.

### 2.1.1 'Stress timed' vs 'Syllable timed'

Nel 1967, Abercrombie modifica l'idea originaria di ritmo, o meglio di isocronia, proposta da Pike. Secondo Abercrombie, le lingue naturali presentano un solo tipo di isocronia e postula l'appartenenza obbligatoria delle lingue naturali a una delle due classi ritmiche *syllable-timed/stress-timed*:

<<usually a language has one or the other type of rhythm but not both since the two types are incompatible'. [...] English has a stress-timed rhythm which manifests itself in all modes of spoken expression>> (cfr. Abercrombie 1967: 97).

La posizione di Abercrombie avvia la cosiddetta 'tipologia ritmica', che persegue l'obiettivo di classificare ritmico-tipologicamente le lingue naturali in una delle due classi ritmiche *syllable-timed/stress-timed* su basi isocrone. Successivamente, come vedremo in §2.1.2, solo per il giapponese si postula la classe ritmica *mora-timed*. Secondo Abercrombie, infatti, nelle lingue naturali il ritmo è reso dalla ricorrenza isocrona di specifici impulsi: *chest-pulses* (impulso toracico) e *stress-pulses* (impulso addominale). Nelle lingue *syllable-timed* gli impulsi toracici hanno una sequenza isocrona, mentre nelle lingue *stress-timed* la sequenza isocrona degli impulsi addominali rafforza quelli toracici.

### 2.1.2 La classe ritmica 'mora-timed'

Ladefoged (1975) postula una terza classe ritmica, definita *mora-timed*, a cui appartiene solo il giapponese. La scelta di istituire questa terza classe ritmica è data dalla centralità che questa unità occupa nel sistema ritmico del giapponese<sup>8</sup>. Anche in questo caso, però, va sottolineato che

---

<sup>8</sup> L'importanza della mora nel sistema ritmico del giapponese era stata già discussa in altri lavori (Jinbo 1927, Bloch 1950, Trubetzkoy 1958, cfr. Cumming 2010) a cui rimandiamo anche per una sua definizione e descrizione.

Ladefoged pone la mora al centro del sistema ritmico del giapponese, utilizzando l'isocronia come fenomeno di riferimento. Infatti, secondo l'autore, le more sono caratterizzate da durate costanti (isocronia). Dal punto di vista isocrono, la classe ritmica *mora-timed* presenta caratteristiche affini a quelle delle lingue classificate *syllable-timed*, perché la durata isocrona delle rispettive unità è il parametro che le caratterizza e le distingue dalle lingue classificate *stress-timed*. Quindi, lingue come il giapponese (*mora-timed*) e l'italiano (*syllable-timed*) hanno simili caratteristiche ritmiche.

## 2.2 L'approccio fonologico

Numerosi fonetisti hanno accettato la tripartizione ritmico-tipologica appena esposta, ma vari studi hanno anche provato l'inesistenza dell'isocronia nei segnali vocali (per una rassegna su questi studi cfr. Lehiste 1977; Bertinetto 1989). In questi studi è stato dimostrato che gli intervalli interaccentuali variano in proporzione al numero delle sillabe contenute, quindi la durata degli intervalli tra accenti consecutivi non è costante, mentre la durata sillabica varia in base ai segmenti contenuti, quindi le sillabe non presentano sempre una durata isocrona.

Questa incertezza che caratterizza gli studi fonetici sul ritmo ha avviato la ricerca di una 'strada fonologica' che superi il paradosso dell'isocronia fisica. Tali studi, che descriviamo di seguito, hanno cercato di individuare i fattori fonologici responsabili delle differenze ritmico-tipologiche tra le lingue naturali.

### 2.2.1 La proposta di Dauer

Dauer (1983) avvia più di ogni altro autore un approccio innovativo sullo studio del ritmo. Innanzitutto, in base agli esperimenti effettuati, l'autrice sostiene che le differenze tra lingue *stress-timed* e *syllable-timed* non vanno imputate alla durata degli intervalli interaccentuali, perché gli accenti si manifestano in modo regolare sia nelle lingue *stress-timed* sia in quelle *syllable-timed* ma a fatti fonologici, fonetici, lessicali e sintattici, che concorrono a rendere costanti gli intervalli interaccentuali e infrasillabici. Nonostante l'autrice correli la ricorrenza costante degli intervalli interaccentuali e infrasillabici tra le lingue ai fatti appena citati, Dauer affronta lo studio dell'isocronia seguendo esclusivamente un approccio fonologico e indica tre parametri come responsabili delle differenze isocrone tra le lingue naturali: struttura sillabica, riduzione vocalica e accento.

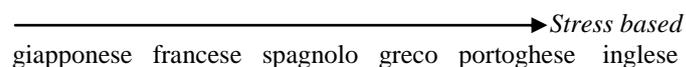
Per quanto riguarda la struttura sillabica, le variazioni isocrone vengono imputate ai tipi sillabici, cioè alle strutture sillabiche consentite dalle lingue: le lingue a isocronia sillabica (italiano, francese e spagnolo) hanno soprattutto strutture CV (una consonante (C) che precede una vocale (V)), mentre le lingue a isocronia accentuale (inglese) presentano tipi sillabici più complessi, cioè

che contengono un numero maggiore di segmenti consonantici. Questo significa che le lingue a isocronia sillabica hanno una maggiore impressione di regolarità dovuta alla ripetizione continua di sillabe con struttura CV. Una ulteriore differenza individuata tra le lingue *syllable-stress-timed* è la tendenza di accentare le sillabe pesanti, cioè che contengono un numero maggiore di segmenti (es. CCVC) e a deaccentare le sillabe leggere, quelle che includono pochi segmenti (es. CV). Dato che le lingue *stress-timed* consentono un numero maggiore di segmenti consonantici all'interno della sillaba, allora in queste lingue la struttura sillabica e l'accento hanno una forte correlazione.

Il secondo parametro indicato da Dauer è la centralizzazione delle vocali atone. Questo processo provoca il passaggio di tutte o di alcune vocali non accentate a vocale media centrale<sup>9</sup>. Questo processo è ritenuto proprio delle lingue *stress-timed*, perché caratterizzate da una forte riduzione delle vocali atone.

Infine, secondo Dauer, l'accento contribuisce a distinguere le lingue naturali nelle classi ritmiche, perché comporta una notevole differenza di durata delle sillabe toniche tra le lingue *stress-timed* e *syllable-timed*. Le lingue *stress-timed* presentano una durata delle sillabe accentate molto superiore rispetto a quelle delle lingue *syllable-timed*.

L'autrice, dopo avere dimostrato che nelle lingue *syllable-stress-timed* i tre parametri individuati sono responsabili delle regolarità isocrone, propone un approccio innovativo per la classificazione delle lingue naturali. Tale approccio prevede di utilizzare i tre parametri per disporre le lingue lungo un continuum *stress-based*, abbandonando così la distinzione dicotomica delle lingue naturali nelle classi ritmiche tradizionali. Quindi, le lingue naturali hanno un 'ritmo' più o meno *stress-based* e si dispongono lungo un continuum dove agli estremi si collocano le lingue più/meno *stress-based*. Di seguito riportiamo il continuum proposto da Dauer (cfr. Dauer 1983: 60):



La proposta di Dauer (1983) è seguita da altri autori (per una rassegna di questi studi cfr. Bertinetto 1981). Questi autori però istituiscono, in modo continuo, vari parametri fonologici come fattori per disporre le lingue lungo il continuum *stress-based*. Tra questi parametri ricordiamo: la quantità vocalica con valore fonologico (Brakel 1985, cfr. Bertinetto 1989) e la tendenza verso la dittongazione delle sillabe accentate (Major 1985, cfr. Bertinetto 1989).

Tuttavia, questa procedura genera una certa circolarità che non consente di arrivare a una classificazione definitiva delle lingue naturali. In questo clima di incertezza, come vedremo nel

<sup>9</sup> Per una descrizione più approfondita della centralizzazione rimandiamo ai seguenti studi (cfr. Bertinetto 1981; Vayra 1989; Nespor 1993).

sotto paragrafo successivo (cfr. §2.2.2), nasce la consapevolezza che sia le caratteristiche prosodiche sia quelle segmentali contribuiscono all'organizzazione ritmica del parlato. Questo non significa che i fattori connessi alla durata perdono la loro importanza ma, al contrario, il dominio della durata diventa la caratteristica centrale da indagare per classificare le lingue naturali.

## **2.2.2 Isocronia e accorciamento compensatorio**

### **2.2.2.1 L'accorciamento compensatorio nelle lingue naturali**

La ricerca di Lindblom & Rapp (1973) apre un nuovo scenario sullo studio del ritmo delle lingue naturali, perché considera il fenomeno dell'accorciamento compensatorio come centrale per la descrizione del ritmo delle lingue naturali. Il termine accorciamento compensatorio si riferisce sia al fenomeno fonologico con cui, in alcune lingue, le sillabe accentate di un piede o di una parola tendono a essere compresse in relazione al numero delle sillabe atone successive (compensazione inter-sillabica) sia al fenomeno con cui i fonemi contenuti nella sillaba vengono compressi in relazione al numero di fonemi che la compongono (compensazione intra-sillabica). Il primo tipo di accorciamento compensatorio ha come dominio il piede, mentre il secondo la parola.

La ricerca dimostra che la compensazione inter-sillabica caratterizza le lingue *stress-timed*, mentre la compensazione intra-sillabica si manifesta nelle lingue *syllable-timed*. Gli autori indagano il modo in cui l'accorciamento compensatorio determini il ritmo delle lingue naturali, ponendo nuovamente al centro dell'attenzione fatti di durata, quindi isocroni, e utilizza tale fenomeno per classificare dicotomicamente le lingue naturali nelle classi ritmiche *syllable- stress-timed*, tornando all'obiettivo della tipologia ritmica.

Un'altra ricerca di assoluto rilievo che indaga il rapporto isocronia/accorciamento compensatorio è quella di Fowler (1977). L'ipotesi di base è che le lingue naturali possano differire nel modo in cui i gesti vocalici e consonantici si associano nel flusso articolatorio. Le lingue tradizionalmente considerate *stress-timed* permettono un maggior livello di compressione rispetto alle lingue *syllable-timed*. Queste ultime mantengono un maggiore controllo sulla durata dei segmenti vocalici e consonantici. Partendo da questa distinzione, l'autrice propone di classificare ritmico-tipologicamente le lingue naturali in lingue a compensazione e lingue a controllo '*control vs. compensation*'.

Tale schema di compressione temporale è stato descritto ed esaminato per varie lingue<sup>10</sup>, tra cui l'italiano e alcune sue varietà (cfr. Vayra *et al.* 1984; Romito & Trumper 1993), ma i risultati dimostrano deboli tracce di accorciamento compensatorio sia nel dominio della parola sia in quello

---

<sup>10</sup> Per una rassegna su questi studi (cfr. Bertinetto 1989).

della sillaba. L'inconsistenza dei risultati spinge a cercare nuovi approcci fonologici che consentano di studiare e di classificare ritmico-tipologicamente le lingue naturali ma, come vedremo nel sotto paragrafo successivo, ancora una volta il ritmo viene confuso e sostituito dall'isocronia.

### **2.2.3 Isocronia e coarticolazione: la proposta di Fowler**

Fowler (1981) avanza una nuova ipotesi sullo studio del ritmo, che pone in stretta relazione il fenomeno della coarticolazione con quello dell'accorciamento compensatorio.

La coarticolazione è un fenomeno che produce la reciproca interferenza tra fonemi adiacenti. In letteratura, si distingue tra coarticolazione anticipatoria, perseverativa e bidirezionale (cfr. Nespor 1993; Romito 2003). Nel caso di coarticolazione anticipatoria, durante la produzione di un fonema alcuni organi dell'apparato vocale anticipano l'articolazione di un fonema che segue, nel caso di coarticolazione perseverativa viene prolungata l'articolazione di un fonema precedente, nel caso della coarticolazione bidirezionale un fonema cambia sotto l'influenza congiunta di due fonemi, quello che lo precede e quello che lo segue.

Nello studio di Fowler (1981), le aspettative prevedono che, durante la produzione dell'eloquio, la coarticolazione abbia effetti perseverativi, mentre l'accorciamento compensatorio anticipatori. Proprio sulla base di queste aspettative vengono postulati i due punti attraverso cui distinguere le lingue naturali nei due gruppi ritmici *syllable- stress-timed*: a) nelle lingue *stress-timed* (inglese e tedesco), gli effetti della coarticolazione sono perseverativi e scarsi mentre, gli effetti dell'accorciamento compensatorio sono anticipatori; b) nelle lingue *syllable-timed* (italiano e spagnolo), gli effetti della coarticolazione e dell'accorciamento compensatorio vanno nella direzione opposta, cioè sono anticipatori. Quindi, le lingue isosillabiche sono presumibilmente caratterizzate da uno sforzo articolatorio più distribuito che tende a non essere concentrato sulla sillaba accenta del piede.

Ancora una volta, però, queste predizioni sono mantenute solo in parte e varie ricerche dimostrano che nelle lingue naturali gli effetti previsti per i due fenomeni non si manifestano in modo così chiaro da consentire una netta classificazione nelle classi ritmiche tradizionali (cfr. Vayra 1985; Vayra *et al.* 1987; per una rassegna su questi studi cfr. Bertinetto 1989).

## 2.3 Studi percettivi sull'isocronia

Gli studi percettivi sull'isocronia sono tutt'ora molto pochi (cfr. §2.3.1; §2.3.2). Come vedremo in §2.3.3, una maggiore attenzione su questo tema si è diffusa negli ultimi venti anni grazie all'interesse di alcuni psicologi che hanno studiato la capacità di neonati e di primati non umani di classificare le lingue naturali nelle classi ritmiche *syllable- stress-timed*.

### 2.3.1 Studi linguistici sulla percezione dell'isocronia

Le prime ricerche linguistiche che studiano il ritmo dal punto di vista percettivo sono fortemente caratterizzate da una descrizione isocrona del fenomeno ritmo. Infatti, la ricerca di Classe (Classe 1939, cfr. Cumming 2010: 5), che rappresenta il punto di partenza degli studi linguistici sulla percezione del ritmo delle lingue naturali, pone l'attenzione sui seguenti aspetti: l'*isocronia* dipende 1) sia dal numero delle sillabe atone comprese tra due sillabe toniche; 2) sia dal numero di segmenti che compongono le sillabe. Questi aspetti dell'isocronia, secondo l'autore, si palesano in maniera evidente nella percezione, mentre si manifestano in modo debole nelle misure acustiche. Tale risultato viene confermato successivamente anche da Lehiste (1977) e Donovan & Darwin (Donovan & Darwin 1979, cfr. Cumming 2010: 7). Proprio sulla base dei risultati raggiunti, questi autori sostengono che l'isocronia sia un fenomeno percettivo confinato nella lingua.

Questi autori dimostrano che l'isocronia è un fenomeno prettamente percettivo, ma le ricerche linguistiche sulla percezione dell'isocronia spostano l'attenzione verso la possibilità di classificare percettivamente le lingue naturali nelle classi ritmiche tradizionali, basata sulla capacità di percepire la ricorrenza isocrona di specifici elementi linguistici, cioè sillabe e piedi (cfr. Miller 1984; Scott *et al.* 1985; Ramus 2003; Arvaniti & Ross 2010). I risultati di queste ricerche però dimostrano l'impossibilità di adempiere percettivamente a questo task.

Vari autori hanno cercato di capire quali variabili inficiano la capacità di classificare percettivamente le lingue naturali nelle classi ritmiche tradizionali, dimostrando che la variabilità tra parlanti (variabilità inter-parlatore), la variabilità all'interno del parlante (variabilità intra-parlatore) (Benguerel & D'arcy cfr. Cumming 2010, cfr. Barry & Koreman 2009) e la velocità di elocuzione (cfr. Deelwo 2008) condizionano fortemente la percezione dell'isocronia.

Infine, White *et al.* (2012) ottengono risultati completamente opposti rispetto a quanto visto finora. Gli autori testano l'ipotesi secondo cui sia possibile distinguere le lingue naturali in base a specifiche caratteristiche correlate alla durata, indipendentemente dalla appartenenza di una lingua a una specifica classe ritmica, raggiungendo risultati positivi.

I primi studi linguistici sulla percezione del ritmo (Classe 1939 e Donovan & Darwin 1979, cfr. Cumming 2010; cfr. Lehiste 1977), come vedremo in §2.3.2, stimolarono alcuni studiosi a

ricercare percettivamente quale unità prosodica o parte della unità prosodica sia responsabile della percezione regolare o irregolare degli eventi isocroni. Tale unità prosodica o parte di unità prosodica, come vedremo in §2.3.2, è definita *Perceptual center* o *P-center*. In più, vari autori (cfr. §2.3.2) cercarono di individuare i correlati articolatori o acustici che veicolano il *P-center*, ma con risultati poco incoraggianti.

### **2.3.2 Il ‘P-center’**

Marcus *et al.* (1976) (Marcus *et al.* 1976, cfr. Marcus 1981) inaugurano questo nuovo filone di ricerca che ha l’obiettivo di individuare e descrivere quale unità prosodica oppure quale parte dell’unità prosodica sia responsabile della trasmissione percettiva dell’isocronia e quali siano i parametri acustici o articolatori che veicolano percettivamente l’evento isocrono. Diverse ricerche (Marcus *et al.* 1976, cfr. Marcus 1981; cfr. Marcus 1981; Hoeqvist 1983; Fox 1987; Fowler 1979; Whalen *et al.* 1991) osservano che esiste un ipotetico punto definito *P-center* o *Perceptual center* responsabile della trasmissione isocrona di un certo elemento.

Le prime ricerche (Marcus *et al.* 1976, cfr. Marcus 1981; cfr. Marcus) definiscono in maniera generale un ipotetico punto (*‘P-center* o *Perceptual center’*) in cui si manifesta percettivamente la ricorrenza isocrona di un certo elemento. Tale punto, secondo gli studi sopra citati, non è caratterizzato da specifici correlati acustici o articolatori, perché la sua collocazione è il risultato di numerosi ‘eventi’ che si manifestano nell’intero segnale.

Nonostante questi risultati negativi, la ricerca ha continuato a perseguire affannosamente questo obiettivo, dando vita a un panorama molto confuso, ma anche caratterizzato da risultati poco incoraggianti. Alcuni studiosi (cfr. Hoeqvist 1983; Fox 1987) cercano di caratterizzare acusticamente il *P-center* e associano la sua ricorrenza isocrona ad alcune parti specifiche della sillaba. Questi autori estendono i risultati ottenuti a tutte le lingue naturali e considerano il *P-center* un fenomeno universale. Una descrizione articolatoria del *P-center* viene elaborata, invece, da Fowler (1979), ma con il passare degli anni si assiste al ritorno verso le prime teorie elaborate sul *P-center* (cfr. Whalen *et al.* 1991) che dimostrano l’impossibilità di descrivere il *P-center* sia articolatoriamente che acusticamente.

I risultati raggiunti da tali ricerche sono molto problematici, per cui gli studiosi abbandonano la strada intrapresa.

### **2.3.3 Studi psicologici sull’isocronia**

Negli ultimi trent’anni, la percezione del ritmo è stata oggetto di studio anche della psicologia. I primi studi psicologici, attraverso test percettivi, hanno concentrato le proprie ricerche

sulla capacità di neonati (cfr. Mehler *et al.* 1988; Nazzi *et al.* 1998; Nazzi *et al.* 2000; Nazzi *et al.* 2003), di adulti (Ramus *et al.* 2003) e di primati non umani (cfr. Tincoff *et al.* 2005) di classificare le lingue nelle categorie ritmiche tradizionali. I risultati estratti da questi test percettivi dimostrano che i neonati tradizionali hanno questa capacità quando lingue non conosciute sono poste a confronto con la propria lingua nativa. Quindi, il ritmo della propria lingua diventa un parametro classificatorio: le lingue con ritmo simile alla propria lingua nativa sono poste nella stessa classe ritmica, mentre quelle ritmicamente distanti dalla propria lingua nativa sono raggruppate e distinte nelle altre classi ritmiche tradizionali (cfr. Mehler *et al.* 1988; Nazzi *et al.* 1998; Nazzi *et al.* 2000; Nazzi *et al.* 2003). Inoltre, i neonati riescono a differenziare il ritmo della propria lingua nativa rispetto a quello delle lingue classificate nella stessa classe ritmica della propria lingua nativa.

Negli ultimi dieci anni, invece, alcuni autori hanno testato attraverso test percettivi la capacità classificatoria di adulti (cfr. Ramus *et al.* 2003) e di primati non umani (cfr. Tincoff *et al.* 2005). A differenza dei neonati, sia gli adulti che i primati non umani classificano ritmicamente le lingue che appartengono a classi differenti mentre non distinguono il ritmo della propria lingua nativa rispetto a quello delle lingue che rientrano nella stessa classe ritmica.

In linea generale, tutto ciò concorre a dimostrare che il ritmo è un importante fenomeno prosodico presente nella competenza linguistica di ciascun individuo.

## **2.4 Ritmo e isocronia: confusione terminologica e metodologica**

L'esistenza delle due classi ritmiche *stress- syllable-timed* è un fondamento delle ricerche fonetiche sul ritmo, esteso successivamente anche alle ricerche fonologiche e percettive (cfr. §2 di questo capitolo). Ancora oggi, le ricerche fonetiche, percettive e fonologiche condotte sul ritmo delle lingue naturali seguono la strada indicata da Abercrombie. Tuttavia, questi studi non sono riusciti a individuare prove certe che supportino la presenza della isocronia nei segnali vocali, cioè la durata equa di piedi e di sillabe rispettivamente nelle lingue *stress- syllable-timed*, che rappresenta il fondamento della dicotomia *stress- syllable-timed* (cfr. Arvaniti 2009). Come conseguenza, la nozione di tipi ritmici ha cominciato a perdere il suo fascino e la confusione tra i concetti di ritmo e isocronia diventa sempre più straripante.

La maggior parte degli studi che si occupano di isocronia testano i vari parametri e modelli su corpora di parlato letto, di esigua ampiezza: brani di varia lunghezza, frasi, parole monosillabiche, bisillabiche e trisillabiche utilizzate per testare le caratteristiche metriche dei piedi, sillabe, ecc. (cfr. Ramus *et al.* 1999; Grabe & Low 2002; Dellwo & Wagner 2003; Rouas & Farinas 2004; Dellwo 2006; Mairano & Romano 2008). Tali corpora mancano di una stratificazione diafasica e diatopica: le analisi sono effettuate spesso su parlato letto, solo Bertinetto (cfr. Bertinetto

& Bertini 2008) sperimentano la propria metrica su parlato semi-spontaneo, e l'eloquio è reso da soggetti con diversa provenienza geografica. Questa metodologia comporta due conseguenze: i risultati del parlato letto sono considerati come modello ed estesi al parlato spontaneo; la diversa provenienza dei soggetti che compongono il corpus, incide sui risultati delle classificazioni.

Infine, un altro problema rilevante che colpisce l'aspetto metodologico delle ricerche sull'isocronia è la mancanza di accordo tra gli studiosi sull'unità prosodica (sillabe, piedi, unità tonali, frasi) che deve fungere da metro di analisi. Da questo fatto ne consegue che gli studi sull'isocronia utilizzando diverse unità prosodiche. Rifiutando questa impostazione metodologica, riteniamo che le unità prosodiche svolgano un ruolo attivo nello studio del comportamento ritmico, e non isocrono, di una lingua naturale. La scelta dell'unità prosodica è un passaggio molto delicato, essa rappresenta l'unità di riferimento che consente di studiare e di capire le dinamiche ritmiche di una lingua.

## Capitolo II - Le metriche ritmiche

### 1. Introduzione

Come abbiamo visto nel capitolo precedente (§2.1) la tipologia ritmica, basandosi sugli studi di isocronia, propone l'esistenza di classi ritmiche *syllable-timed* e *stress-timed* che, a seconda degli approcci, delineano tendenze cui le lingue si conformano (Pike 1945) o distinzioni che operano una classificazione dicotomica (Abercrombie 1967).

Vari studi condotti negli anni sessanta (cfr. §1) contrastano con i risultati di Abercrombie. Le misure di durata degli intervalli interaccentuali effettuate in vari studi (cfr. §1), infatti, dimostrano l'assenza oppure una lievissima presenza del fenomeno della isocronia e suggeriscono l'impossibilità di classificare ritmico-tipologicamente le lingue naturali attraverso questo principio. In seguito a questi risultati, diversi autori avanzano nuove proposte che perseguono l'obiettivo di classificare ritmico-tipologicamente le lingue naturali. La proposta più recente è quella delle metriche ritmiche, cioè modelli che attraverso specifiche formule matematiche perseguono l'obiettivo di classificare ritmico-tipologicamente le lingue naturali. Nel paragrafo successivo, si descrivono proprio questi modelli.

### 2. Le metriche ritmiche

Negli ultimi quindici anni, le ricerche sul ritmo hanno concentrato i propri sforzi verso la realizzazione di modelli, definiti metriche ritmiche, che classificano ritmico-tipologicamente le lingue naturali nelle tradizionali classi ritmiche. Le formule matematiche alla base di questi modelli si fondano sostanzialmente sulla misura delle durate vocaliche e consonantiche.

Questo nuovo filone di ricerca sul ritmo è avviato da Ramus *et al.* (1999), che propongono tre correlati fonetici (cfr. §2.1.1) ritenuti in grado di classificare le lingue naturali nelle classi ritmiche *stress-timed*, *syllable-timed* e *mora-timed*. Un approccio simile è stato sviluppato da Grabe & Low (2002), mentre altri dimostrato l'instabilità delle metriche rispetto ad alcuni fattori, soprattutto nei confronti dello *speech rate*. Questo problema ha indotto a proporre modifiche alle metriche ritmiche così da ridurre la sensibilità a tale parametro (cfr. Dellwo & Wagner 2003; Dellwo 2006). Di recente, invece, Bertinetto & Bertini (2008) hanno realizzato una nuova metrica ritmica diversa rispetto alle precedenti, che basa la propria logica sulla compensazione.

Nel paragrafo successivo descriviamo le metriche ritmiche sviluppate dagli autori appena citati (cfr. §2.1.1; §2.1.2; 2.1.3; §2.1.4), perché saranno sperimentate sul corpus di questa ricerca. Inoltre,

sempre nel paragrafo seguente, presentiamo una rassegna di studi sulle metriche ritmiche (cfr. §2.2) e le critiche mosse a questi modelli (cfr. §2.3).

## 2.1 Metodologia generale

### 2.1.1 I *delta*

Ramus *et al.* (1999), allo scopo di descrivere le caratteristiche ritmiche delle lingue naturali, propongono tre correlati fonetici ritenuti in grado di catturare alcune proprietà fonologiche (cfr. Dauer 1983) responsabili della percezione *stress-timed* o *syllable-timed* di una lingua. Le proprietà fonologiche responsabili in maniera maggiore rispetto ad altre sono: 1) presenza vs. assenza di riduzione vocalica; 2) struttura sillabica complessa vs. struttura sillabica semplice. I correlati acustici proposti da Ramus *et al.* (2009) inglobano queste proprietà fonologiche in specifiche formule matematiche, che hanno come parametro essenziale la durata degli intervalli vocalici e consonantici.

Da queste misure, gli autori ricavano tre variabili:  $\Delta V$ ,  $\Delta C$  e %V. Il  $\Delta V$  rappresenta la deviazione standard degli intervalli vocalici e indica la presenza/assenza di un alto grado di riduzione vocalica nelle sillabe atone: le lingue *stress-timed*, che consentono un elevato grado di riduzione vocalica, presentano ipoteticamente una maggiore variabilità di durata della vocale tonica rispetto a quella atona, breve e ridotta. Pertanto, da queste lingue si attendono valori alti di  $\Delta V$ . Il  $\Delta C$  indica la deviazione standard degli intervalli consonantici e descrive la complessità sillabica: più il valore è alto, maggiore è la complessità della struttura sillabica. Questa affermazione si basa sulla osservazione che le lingue con struttura sillabica semplice, quelle *syllable timed*, consentono esclusivamente clusters consonantici semplici, mentre le lingue *stress-timed* presentano clusters consonantici semplici e complessi con conseguenti valori alti di  $\Delta C$ . La deviazione standard misura il grado di variabilità su una lista di valori, comparando ogni possibile coppia attraverso la formula seguente:

$$\sigma_x = \sqrt{\frac{\sum_{i=1}^N (x_i - \langle x \rangle)^2}{N - 1}}$$

Il %V va inteso come un correlato acustico che misura sia la complessità della struttura sillabica sia la presenza/assenza di un elevato grado di riduzione vocalica. Il fatto che %V presenti un valore alto nelle lingue con elevato grado di riduzione vocalica (*stress-timed*) è abbastanza intuitivo, mentre l'ipotesi che %V mostri un valore basso anche nelle lingue con struttura sillabica

complessa è da imputare alla percentuale alta di elementi consonantici che compongono la struttura sillabica complessa.

Le aspettative degli autori sulla classificazione delle lingue sono illustrate nella figura 2.1 (cfr. Mairano 2010: 38), dove il cerchio A rappresenta le lingue classificate tradizionalmente come *syllable-timed* (sono previsti valori bassi di  $\Delta V$  e di  $\Delta C$ ), mentre il cerchio B racchiude le lingue tradizionalmente classificate *stress-timed* (sono previsti valori alti di  $\Delta V$  e  $\Delta C$ ).

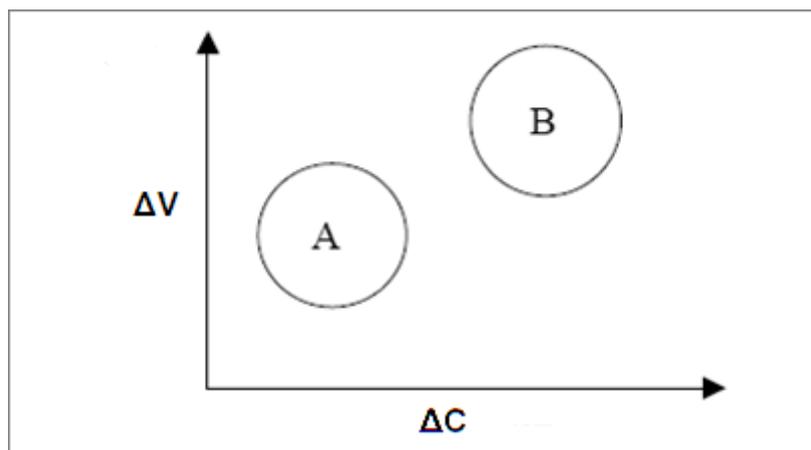


Figura 2.1: figura che mostra le presunte differenze dei valori di  $\Delta V$  e  $\Delta C$  per le lingue *syllable timed* (A) e *stress-timed* (B) (da Mairano 2010: 38).

Ramus *et al.* (1999), allo scopo di testare la classificazione ritmico-tipologica resa dai delta, compiono un esperimento su otto lingue: inglese, tedesco e polacco considerate *stress-timed*, italiano, francese, catalano e spagnolo ritenute *syllable-timed* e il giapponese definito *mora-timed*. Il corpus è composto da quattro parlanti nativi per ciascuna lingua che producono cinque frasi. Su questo corpus, gli autori calcolano i tre correlati acustici e presentano i risultati su tre grafici, combinando i valori dei correlati ottenuti per ciascuna lingua:

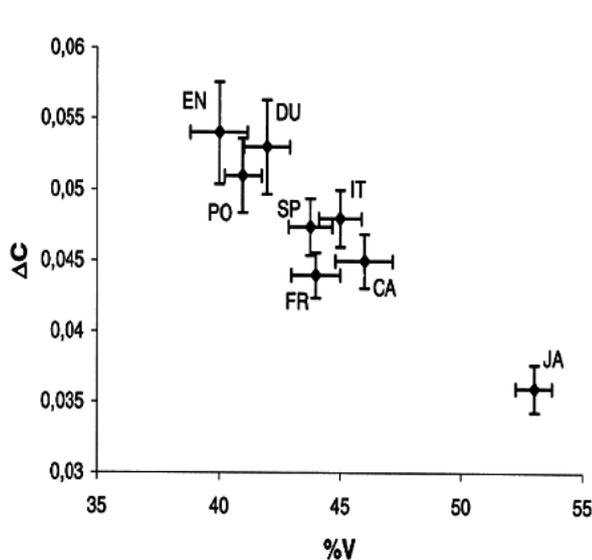


Figura 2.2: %V/ΔC (da Ramus *et al.* 1999: 273).

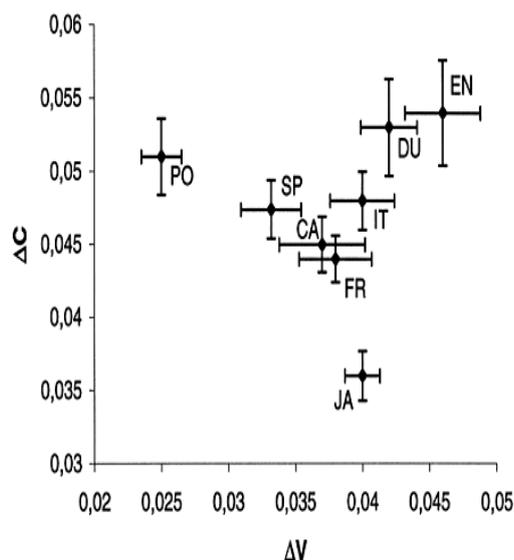


Figura 2.3: ΔV/ΔC (da Ramus *et al.* 1999: 273).

Questi grafici sembrano confermare le ipotesi di partenza: a) l'inglese e il tedesco formano un gruppo con valori alti di  $\Delta V$  e  $\Delta C$  e valori bassi di %V; b) l'italiano, il francese, il catalano e lo spagnolo compongono un gruppo con valori opposti; c) il giapponese occupa una posizione isolata e presenta valori di  $\Delta V$  e  $\Delta C$  più bassi rispetto a quelli delle lingue *syllable-timed*, ma valori più alti di %V; d) il polacco condivide con il tedesco e l'inglese i valori di %V e  $\Delta C$ , ma ha valori molto bassi di  $\Delta V$ . Gli autori concludono che il grafico %V/ΔC consente di distinguere le lingue naturali nelle classi ritmiche tradizionali, mentre  $\Delta V$  risulta influenzato da altri fattori.

### 2.1.2 Il *Pairwise Variability Index* (PVI)

Grabe & Low (2002) presentano un approccio leggermente diverso rispetto a quello in (cfr. §2.3.1.1). Tale approccio si basa sul *Pairwise Variability Index* (d'ora in poi PVI), che fornisce indicazioni sulla variabilità degli intervalli vocalici e consonantici. Il PVI, proprio come i delta, viene applicato alla durata degli intervalli vocalici e consonantici, ma il suo vantaggio consiste nel considerare la successione temporale dei segmenti ( $m$  è il numero degli intervalli, mentre  $d_k$  è la durata dell'intervallo  $k^{th}$ ):

$$rPVI = \left[ \sum_{k=1}^{m-1} |d_k - d_{k+1}| / (m - 1) \right]$$

In altre parole, la formula del *raw* PVI (d'ora in poi  $rPVI$ ) calcola la differenza di durata di ogni intervallo con quello successivo e, infine, viene calcolata la media delle differenze. Questo

procedimento contrasta con la logica della deviazione standard, la cui formula considera tutti gli intervalli possibili (successivi e non successivi). La sensibilità della durata dei segmenti vocalici alla velocità di eloquio ha indotto le autrici a utilizzare il *normalised* PVI (d'ora in poi *nPVI*) per i segmenti vocalici. L'*nPVI* in sostanza è la stessa formula dell'*rPVI*, ma calcola la differenza di durata di ogni intervallo consonantico con quello successivo e la divide per la durata media. Infine, viene calcolata la media dei valori ottenuti e moltiplicata per cento:

$$nPVI = 100 \times \left[ \sum_{k=1}^{m-1} \left| \frac{d_k - d_{k+1}}{(d_k + d_{k+1})/2} \right| / (m - 1) \right]$$

Grabe & Low (2002) confrontano la propria metrica ritmica con quella proposta da Ramus *et al.* (1999). Le autrici conducono l'esperimento sulle lingue considerate da Ramus *et al.* (1999), tranne l'italiano, e aggiungono altre lingue: estone, tedesco, greco, lussemburghese, malese, cinese mandarino, rumeno, inglese parlato a Singapore, tamil<sup>11</sup>, thailandese e gallese. Il corpus è composto da un parlante nativo per ciascuna lingua e ogni locutore legge il testo '*The North Wind and the Sun*<sup>12</sup>'. Dalle analisi condotte attraverso la propria metrica, le autrici individuano quattro lingue *stress-timed* (inglese britannico, tedesco, olandese e thailandese), quattro lingue *syllable-timed* (francese, spagnolo, tamil e inglese parlato a Singapore), una lingua *mora-timed* (giapponese), sette lingue non classificate (estone, greco, lussemburghese, malese, cinese mandarino, rumeno e gallese) e due lingue a ritmo 'misto' (polacco e catalano):

---

<sup>11</sup> Lingua parlata in India.

<sup>12</sup> Fiaba scritta da Aesop e usata come testo standard nelle ricerche fonetiche.

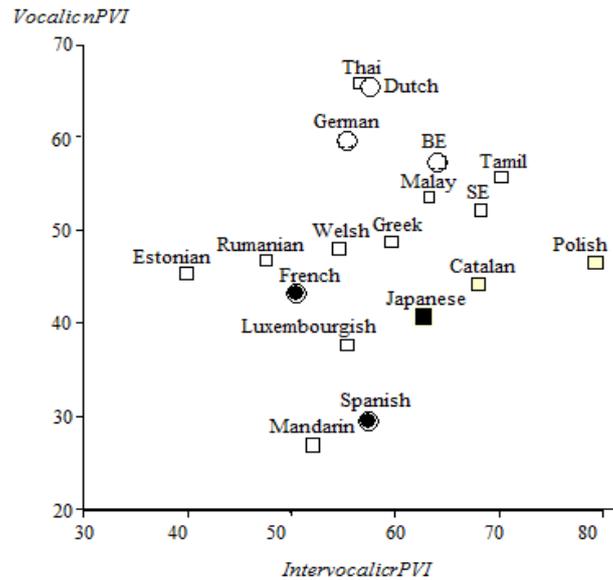


Figura 2.4: *nPVI/rPVI* (da Grabe & Low 2002: 7).

Secondo le autrici, i risultati del PVI confermano le aspettative: le lingue *stress-timed* presentano valori vocalici alti di *nPVI*, invece, l'*rPVI* sembra avere un potere discriminativo minore, perché le lingue *syllable- stress-timed* mostrano valori affini. In più, la metrica sembra rendere un buon risultato per le lingue con ritmo ‘misto’ come il polacco, dove il valore alto di *rPVI* riflette la complessità della sua struttura sillabica, mentre il valore basso di *nPVI* indica l’assenza del fenomeno fonologico della riduzione vocalica. Successivamente, le autrici calcolano sullo stesso corpus i correlati %V e  $\Delta C$ , così da confrontare i risultati con quelli del PVI. Grabe & Low (2002) trovano disposizioni simili in alcuni casi, mentre in altri alcune lingue si posizionano in aree diverse del grafico, contrastando con le aspettative. Le autrici imputano queste differenze alla poca ‘attenzione’ dei correlati proposti da Ramus *et al.* (1999) nei confronti di alcune variabili, soprattutto dello *speech rate*, e affermano che il PVI realizza una migliore descrizione del ritmo delle lingue rispetto a quella dei delta. Grabe & Low (2002) concludono il proprio lavoro affermando che la disposizione delle lingue ottenuta nel grafico *rPVI/nPVI* dimostra la impossibilità di distinguere categoricamente le lingue nei tipi *stress-timed* e *syllable-timed*.

### 2.1.3 Il varco $\Delta C$

L'applicazione dei correlati proposti da Ramus *et al.* (1999) a corpora più vasti e vari ha fatto emergere la sensibilità di  $\Delta C$  ma soprattutto di  $\Delta V$  a molte variabili, in particolare allo *speech rate*. Gli effetti di questa variabile sui valori dei delta è stata studiata attentamente da Dellwo & Wagner (2003). In questo lavoro, gli autori esaminano l’influenza dello *speech rate* sui correlati %V e  $\Delta C$  e le lingue considerate sono l’inglese, il francese e il tedesco. I risultati riflettono i gruppi

ritmici di Ramus *et al.* (1999), ma gli autori individuano una forte relazione tra la velocità di eloquio e i valori dei correlati (soprattutto del  $\Delta C$ ). Tale relazione però non è così forte da produrre spostamenti nella classificazione resa per le lingue oggetto di studio.

Allo scopo di correggere la sensibilità mostrata da  $\Delta C$ , gli autori propongono di dividere i valori di  $\Delta C$  per la durata media degli intervalli consonantici e moltiplicare per cento, ottenendo così il *varco* $\Delta C$ :

$$\text{varco}\Delta C = \Delta C / \text{mean}C * 100$$

Questa idea segue la stessa base logica della normalizzazione proposta nella formula dell'*nPVI*, dove le durate degli intervalli sono divise con la durata media.

Il *varco* $\Delta C$  viene testato successivamente in Dellwo (2006). In questo lavoro, l'autore calcola tale parametro sui dati estratti dal corpus *BonnTempo* e i risultati sembrano incoraggianti:

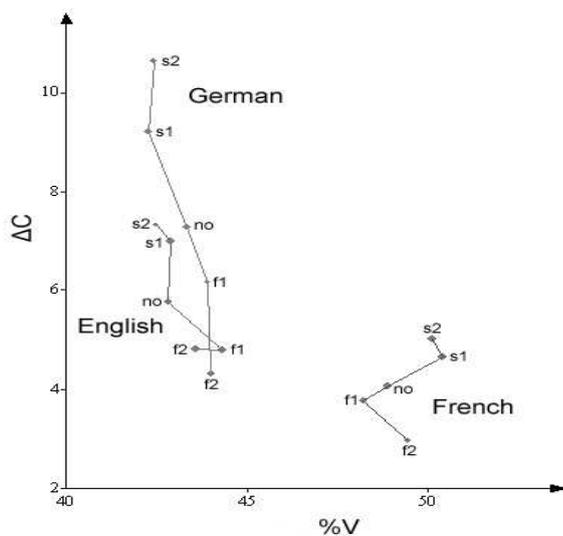


Figura 2.5:  $\Delta C / \%V$  (da Dellwo 2006: 2).

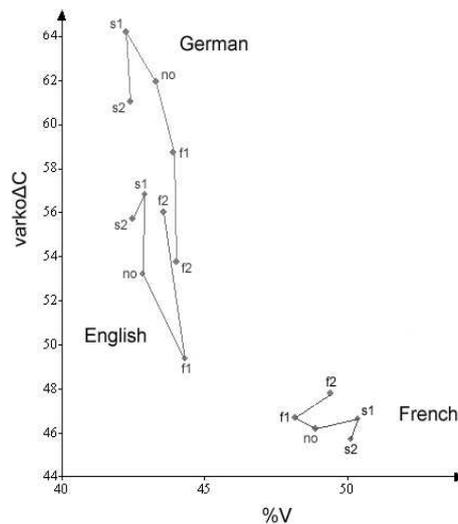


Figura 2.6: *varco* $\Delta C / \%V$  (da Dellwo 2006: 5).

Il *varco* $\Delta C$  differenzia i gruppi *syllable-stress-timed* in modo più chiaro rispetto al correlato  $\Delta C$ . Quindi, l'uso del coefficiente di variazione di  $\Delta C$  migliora la differenziazione delle classi ritmiche.

#### 2.1.4 Il *Compensation and Control Index* (CCI)

Una nuova metrica, il *Compensation and Control Index* (d'ora in poi CCI), viene proposta da Bertinetto & Bertini (2008). Tale formula è ispirata all'*rPVI* e come le altre metriche viene applicata, anche se con differenze importanti, alle misure vocaliche e consonantiche. Diversamente

da quanto visto per le metriche precedenti, il CCI ha una logica differente che si basa sugli studi effettuati sulla compensazione (cfr. Fowler 1977). Il CCI misura il livello di compressione consentito da una lingua, cioè quanto i segmenti di una lingua possono dilatarsi o restringersi a seconda del contesto in cui si trovano.

Questa idea parte dal lavoro di Bertinetto (1989), in cui viene rivisitata la dicotomia tradizionale *stress-timed* vs. *syllable-timed* in termini di lingue a controllo vs. lingue a compensazione: le lingue a controllo, che corrispondono alle lingue *syllable-timed*, dovrebbero presentare bassi livelli di compensazione intra- e inter-sillabica, mentre le lingue a compensazione dovrebbero mostrare alti livelli di compensazione intra e inter-sillabica.

Come dichiarato da Bertinetto & Bertini (2008), il CCI punta a descrivere il comportamento intra-sillabico che interessa l'alternanza accentuale. Allo scopo di descrivere il comportamento intra-sillabico, il CCI considera i segmenti che compongono ogni intervallo vocalico e consonantico: la formula del CCI è una modifica dell'*rPVI* e prevede la divisione delle durate degli intervalli per il numero dei segmenti che lo compongono:

$$CCI = \frac{100}{m-1} \sum_{k=1}^{m-1} \left| \frac{d_k}{n_k} - \frac{d_{k+1}}{n_{k+1}} \right|$$

Le aspettative degli autori sono che le lingue a controllo dovrebbero posizionarsi lungo la bisettrice, perché in queste lingue le fluttuazioni vocaliche e consonantiche dovrebbero essere più o meno le stesse. Invece, le lingue a compensazione tenderebbero a posizionarsi al di sotto della bisettrice, perché in queste lingue le fluttuazioni vocaliche sono maggiori rispetto a quelle consonantiche. Tale atteggiamento è dovuto alla differenza tra le vocali toniche completamente articolate e le vocali atone ridotte. Gli autori non prevedono una collocazione delle lingue nella zona al di sopra della bisettrice, perché questo atteggiamento implicherebbe un livello di compensazione consonantica maggiore rispetto a quella vocalica:

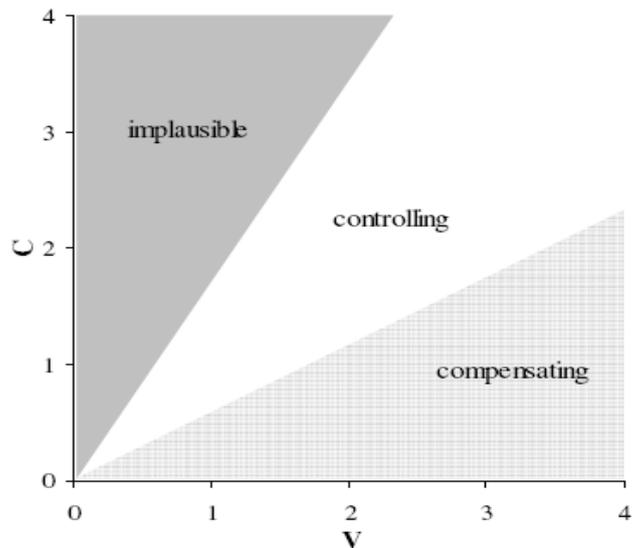


Figura 2.7: aspettative di classificazione del CCI (da Bertinetto & Bertini 2008: 428).

Nel lavoro di Bertinetto & Bertini (2008), il CCI viene sperimentato su un corpus composto da dieci locutori pisani e i dialoghi sono stati acquisiti attraverso la tecnica del *map-task*<sup>13</sup>. Successivamente, i dialoghi sono stati suddivisi in unità tonali e in base a certi criteri stabiliti dagli autori, per gli scopi dello studio, sono state selezionate solo alcune unità tonali su cui è stato applicato il CCI. I risultati della sperimentazione sembrano riflettere le aspettative delle lingue a controllo, perché i dieci locutori pisani si collocano lungo la bisettrice:

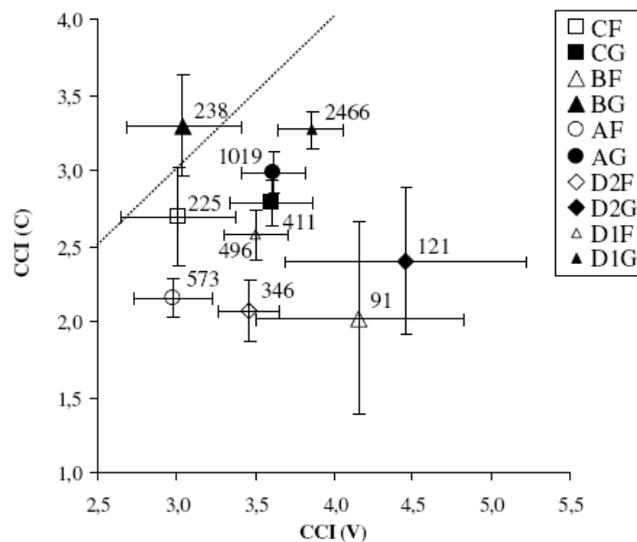


Figura 2.8: risultati CCI (da Bertinetto & Bertini 2008: 429).

<sup>13</sup> Tecnica di elicitazione ideata da Brown (1983) ma sviluppata successivamente dal gruppo di ricerca HCRC.

Successivamente, in base ai valori di *speech rate*, i locutori vengono suddivisi in tre gruppi e, in linea con le aspettative, al decrescere della velocità di eloquio, le consonanti e le vocali vengono compresse fino a una certa soglia più o meno allo stesso modo (anche se una maggiore compressione riguarda le consonanti).

## 2.2 Rassegna di studi sulle metriche ritmiche

### 2.2.1 La classificazione delle lingue naturali attraverso le metriche ritmiche

Dopo l'analisi di ciascuna metrica ritmica, presentiamo una rassegna di studi che hanno utilizzato questi modelli per classificare ritmico-tipologicamente varie lingue naturali.

Gibbon & Gut (2001) classificano ritmico-tipologicamente l'ibibio<sup>14</sup> e l'inglese britannico attraverso la *Rhythm Ratio*, una misura che prende spunto dal PVI. Il corpus della lingua ibibio è costituito da due locutori che leggono dieci frasi che contengono almeno dodici sillabe, mentre il corpus dell'inglese britannico è composto da un locutore che legge una fiaba. Secondo gli autori, la *Rhythm Ratio* restituisce una classificazione molto chiara: l'ibibio ha un ritmo più *syllable-timed* rispetto all'inglese britannico.

Ghazali *et al.* (2002) studiano la variazione ritmica in sei dialetti arabi: marocchino, tunisino e algerino (dialetti occidentali), giordano, siriano ed egiziano (dialetti orientali). Gli autori calcolano i correlati %V e  $\Delta C$  sulle versioni tradotte nei vari dialetti di *'The North Wind and the Sun'*. I risultati dimostrano valori bassi di %V e valori alti di  $\Delta C$  per i dialetti occidentali. Tale tendenza conferma le aspettative secondo cui i dialetti arabi occidentali hanno preferenze isoaccentuali maggiori rispetto ai dialetti dell'area orientale.

Mok & Dellwo (2008) testano diverse metriche ritmiche allo scopo di classificare ritmico-tipologicamente alcune varietà cinesi: inglese mandarino, cantonese, inglese cantonese, mandarino pechinese. L'esperimento si divide in tre fasi: nella prima fase, i locutori leggono in cantonese o mandarino la fiaba *'The North Wind and the Sun'*, nella seconda fase, i locutori ripetono quanto letto senza consultare il testo così da produrre parlato semi-spontaneo; infine, i partecipanti leggono la versione inglese della fiaba. I risultati rendono una classificazione *syllable-timed* per il cantonese e il mandarino pechinese, nonostante l'alto grado di variazione del %V per i due diversi stili di eloquio (letto vs. semi-spontaneo). Una situazione poco chiara si manifesta per l'inglese cantonese e per l'inglese mandarino che percettivamente rivelano un ritmo a isocronia sillabica, ma i risultati mostrano una classificazione che varia in base alle metriche ritmiche. Gli autori concludono affermando che questi risultati mettono a dura prova le misure acustiche.

---

<sup>14</sup> L'ibibio è una lingua africana parlata in Nigeria.

Mairano & Romano (2008) raggiungono risultati simili a quelli di Mok & Dellwo (2008). Gli autori cercano di classificare ritmico-tipologicamente alcuni dialetti italiani e romeni attraverso varie metriche: i delta, i varco, il PVI e il CCI. I risultati mostrano che le metriche ritmiche classificano le varietà dialettali in maniera discordante. Risultati ancora più critici sono riportati in Mairano *et al.* (2010). In questo lavoro, gli autori sperimentano su un corpus di numerose lingue, le metriche ritmiche proposte da Ramus *et al.* (1999), Grabe & Low (2002), Dellwo (2006) e Bertinetto & Bertini (2008) ottenendo risultati molto eterogenei:

- la metrica di Ramus *et al.* (1999), con la proiezione  $\Delta C\text{-}\%V$ , restituisce risultati incerti sul moldavo, munteno e spagnolo (valori inaspettatamente bassi di  $\%V$  e medi di  $\Delta C$ ), mentre la proiezione  $\Delta C\text{-}\Delta V$  evidenzia risultati inaspettati per il giapponese, il finlandese ed il transilvano (valori alti di  $\Delta V$  e bassi o medio-bassi di  $\Delta C$ );
- la metrica di Grabe & Low (2002) rende risultati incerti per due dei campioni di italiano, romeno olteno, ceco, russo, turco ed islandese che si collocano al centro del grafico, quindi a metà strada tra i due gruppi;
- la metrica di Dellwo (2006) consegna dati molto confusi. Il giapponese e le tre varietà di romeno risulterebbero tra le lingue accentuali, mentre il finlandese si situerebbe in una posizione intermedia tra i due gruppi.
- la metrica di Bertinetto & Bertini (2008) restituisce risultati incerti per il russo, l'arabo, su uno dei due parlanti di italiano e, in parte, sul portoghese europeo, l'inglese australiano e l'islandese che risulterebbero a controllo.

Tepperman & Nava (2011) classificano il ritmo di varie lingue attraverso il *Generalized Variability Index* (GVI), una metrica che prende spunto ancora una volta dal PVI. In questa ricerca viene utilizzato il corpus *Globalphone* (cfr. Tepperman & Nava 2011), che contiene circa cento soggetti registrati per ciascuna lingua. Dai risultati emerge, in sintonia con le aspettative, che il francese, lo spagnolo e il turco sono lingue *syllable-timed*, mentre lo svedese, l'inglese, il portoghese e il tedesco sono lingue *stress-timed*. A tale visione generica però esistono delle eccezioni, come dimostrano i risultati del russo (lingua considerata *stress-timed*) e del giapponese (lingua considerata *mora-timed*): il primo assume una posizione *syllable-timed*, mentre il secondo *stress-timed*.

Beňuš & Simko (2012) effettuano una classificazione ritmico-tipologica dello slovacco attraverso i delta, i varco e il PVI. Il corpus è composto da parlato letto e si articola in due task: 1) un parlante nativo slovacco legge centodue frasi a una velocità di eloquio naturale; 2) due frasi simili prodotte da tre locutori slovacchi con una velocità di elocuzione continua. Le metriche ritmiche classificano in modo unanime lo slovacco tra le lingue *syllable-timed*.

In letteratura, pochi studi hanno sperimentato le metriche ritmiche sulle varietà e sui dialetti italiani allo scopo rendere una loro classificazione ritmico-tipologica. Dato che l'obiettivo di questa ricerca è lo studio del ritmo di alcune varietà di italiano, riteniamo opportuno presentare, nei sotto paragrafi successivi, una rassegna sia dei pochi studi che hanno investigato il ritmo di alcune varietà di italiano (§2.2.1.1) sia delle pochissime ricerche che hanno sperimentato le metriche ritmiche su alcuni dialetti italiani (§2.2.1.2).

### **2.2.1.1 La classificazione di alcune varietà di italiano attraverso le metriche ritmiche**

White *et al.* (2009) testano il varco  $\Delta V$  e il %V allo scopo di classificare ritmicotipologicamente alcune varietà regionali italiane (veneta e siciliana). I risultati mostrano una classificazione *syllable-timed* per entrambe le varietà. Gli autori sottolineano il risultato della varietà siciliana che, secondo le aspettative, doveva posizionarsi nel polo a isocronia accentuale. Giordano & D'Anna (2010) classificano quindici varietà di italiano regionale (Bari, Bergamo, Cagliari, Catanzaro, Firenze, Genova, Lecce, Milano, Napoli, Palermo, Parma, Perugia, Roma e Torino) attraverso il %V, i delta, i varco e il PVI. Il corpus, estratto da CLIPS<sup>15</sup> (cfr. Albano Leoni 2006, Savy & Cutugno 2009), è costituito da trentaquattro locutori che provengono dalle regioni sopra menzionate. Gli autori classificano le varietà di italiano in relazione a tre stili di eloquio: parlato monologico pre-programmato, parlato dialogico spontaneo e parlato letto. I risultati mostrano che: a) il valore di  $\Delta C$  è basso nel parlato letto, ma aumenta nel parlato dialogico e pre-programmato (quest'ultimo presenta il valore più alto). Al contrario, il valore  $\Delta V$  decresce dal parlato pre-programmato agli altri due stili di eloquio; b) i valori consonantici di *rPVI* sono uguali a quelli di  $\Delta C$ , mentre quelli di *nPVI* sono stabili. La variazione delle metriche ritmiche nella classificazione delle quindici varietà di italiano è stata calcolata solo per il parlato letto: i valori di %V sono coerenti con quelli delle lingue a isocronia sillabica, anche se in alcuni casi sono vicini alle aspettative delle lingue isoaccentuali. I valori di  $\Delta V$  e di  $\Delta C$  mostrano una forte variabilità: il valore di  $\Delta V$  di Perugia e di Catanzaro è vicino alle lingue isoaccentuali, mentre quello di Napoli è simile alle lingue moraiche. Le stesse oscillazioni si manifestano per i valori di *nPVI* e *rPVI*.

Risultati più eterogenei sono riportati in Romano *et al.* (2010). In questo studio, gli autori esaminano le caratteristiche ritmiche del parlato letto prodotto da un singolo locutore per ciascuna delle seguenti località: Roccaforte Ligure, Briga Alta, Exilles, Capanne di Marcarolo, Campertogno e Bagnolo Piemontese. I risultati collocano le due varietà liguri agli estremi opposti: Roccaforte

---

<sup>15</sup> [www.clips.unina.it](http://www.clips.unina.it).

Ligure si situa tra le lingue a isocronia accentuale, Capanne di Marcarolo tra quelle a isocronia sillabica. Le varietà di Bagnolo Piemontese e Campertogno assumono una posizione a isocronia accentuale, mentre Exilles e Briga Alta sono a isocronia sillabica.

### **2.2.1.2 La classificazione di alcuni dialetti italiani attraverso le metriche ritmiche**

Nel lavoro di Russo & Barry (2008) vengono testate le metriche ritmiche proposte da Ramus *et al.* (1999), Grabe & Low (2002) e Barry (2003) (Barry 2003, cfr. Russo & Barry 2008) allo scopo di descrivere e classificare ritmico-tipologicamente il dialetto di Ischia. In questo studio viene analizzato parlato spontaneo prodotto da dieci locutori e i risultati classificano il dialetto di Ischia tra le lingue *stress-timed*.

In un lavoro successivo (cfr. Russo & Barry 2010), gli autori cercano di classificare ritmico-tipologicamente i dialetti di Ischia, Capri e Pozzuoli attraverso il PVI. Per ogni località sono stati registrati due locutori, che producono dieci enunciati spontanei. I risultati mostrano che i locutori campani hanno un valore di PVI-V elevato, simile a quello di parlanti tedeschi, mentre il PVI-C assume valori bassi vicini a quelli di locutori francesi. In più, l'autrice riscontra una ampia dispersione dei valori di PVI per enunciato individuale: quattro enunciati presentano valori più vicini al polo isoaccentuale. Tale risultato indica che la struttura sillabica è caratterizzata da un alto livello di variabilità.

## **2.3 Critiche mosse alle metriche ritmiche**

Le incertezze classificatorie e la metodologia su cui si basano le metriche ritmiche hanno prodotto forti critiche nei confronti di questi modelli. La critica più pressante avanzata alle metriche riguarda il fatto di tenere conto solo di criteri che rendono informazioni sullo *speech timing* (cfr. Arvaniti 2009): le metriche effettuano una descrizione del fenomeno dell'isocronia, perché suppongono una relazione semplice e lineare tra la durata e alcune categorie fonologiche astratte, come la struttura sillabica e la riduzione vocalica.

Kohler (2009) effettua una affermazione ancora più forte su questo punto. Secondo l'autore, la relazione tra strutture fonologiche e misure di durata può riflettere fino a una certa misura i modelli della durata segmentale, ma essa non è in grado di catturare i modelli del movimento ritmico che evolvono nel tempo con una certa regolarità. Così facendo, le metriche ritmiche si pongono sull'asse paradigmatico, perché classificano le lingue naturali in una delle tradizionali classi ritmiche, ponendo la durata dei segmenti vocalici e consonantici come parametro cardine ai fini classificatori. Tale approccio esclude l'asse sintagmatico, cioè lo studio delle relazioni tra gli elementi che realizzano e descrivono le potenzialità ritmiche delle lingue naturali. Va anche

sottolineato che, nonostante le metriche ritmiche siano delle pure formule matematiche, ancora oggi non esiste un parametro capace di misurare i loro risultati e, nello stesso tempo, un metodo chiaro in grado di interpretare le loro differenze e similarità: solitamente, la similarità o la differenza dei risultati viene giudicata sulla base di classificazioni precedenti.

Una ulteriore critica mossa alle metriche ritmiche riguarda l'attenzione posta esclusivamente verso la natura fisica del segnale. Tali modelli non riflettono il ruolo che ciascun tratto acustico svolge nella resa percettiva del ritmo. Questa osservazione è sollevata da Barry & Koreman (2009). Gli autori affermano che il legame tra percezione e produzione del ritmo è stato studiato raramente nelle ricerche condotte sulla classificazione ritmico-tipologica delle lingue naturali effettuata dalle metriche ritmiche. Barry & Koreman (2009) dimostrano che valori simili di PVI indicano ritmi percepiti diversamente, mentre valori contrastanti di PVI specificano ritmi percepiti in maniera simile. Questi risultati portano gli autori a criticare il PVI e, nello stesso tempo, ad aumentare lo scetticismo nei confronti dell'utilità delle metriche ritmiche.

Forti critiche riguardano anche i corpora su cui si testano le metriche. Arvaniti (2009) dimostra che la tecnica di elicitazione può turbare significativamente i risultati delle metriche ritmiche, perché le classificazioni restituite da questi modelli variano in base al tipo di materiale su cui sono sperimentate. In più, l'autrice nota che le differenze di classificazione scompaiono quando le metriche sono testate su parlato spontaneo. Infatti, nella maggior parte delle ricerche sul ritmo<sup>16</sup>, le metriche ritmiche sono adoperate su parlato letto (§ 3.2 di questo capitolo). A supporto di questa critica, riportiamo i corpora su cui sono state costruite e testate le metriche ritmiche:

1. Ramus *et al.* (1999) considerano otto lingue: inglese, tedesco, francese, polacco, spagnolo, catalano, italiano e giapponese. Per ciascuna lingua sono stati registrati quattro parlanti che leggono cinque frasi;
2. Grabe & Low (2002) esaminano diciotto lingue *timed* (inglese britannico, tedesco, olandese, thailandese, tamil, francese, spagnolo, inglese parlato a Singapore, giapponese, polacco, catalano, estone, greco, lussemburghese, malese, cinese mandarino, rumeno e gallese) e per ognuna di esse è stato registrato un soggetto che legge il testo '*The North Wind and the Sun*';
3. Dellwo (2006) analizza il corpus *BonnTempo* composto da lingue considerate *stress-timed* (inglese e tedesco) *syllable-timed* (francese ed italiano) e non classificate (polacco e ceco). Per ciascuna lingua sono stati registrati un numero differente di parlanti che leggono brevi testi composti da circa ottanta sillabe.

---

<sup>16</sup> Solo Bertinetto & Bertini (2008) sperimentano la propria metrica su parlato semi-spontaneo.

Da questi corpora emerge con forza una ulteriore critica, che riguarda il numero dei locutori: in genere, i corpora contengono un numero molto esiguo di locutori per cui sembra difficile potere affermare che una certa lingua mostri un determinato tipo di ritmo.

Infine, un grande problema riguarda le unità linguistiche che fungono da dominio di analisi del ritmo. Come vedremo nel capitolo successivo (cfr. §3), negli studi sul ritmo si testano le metriche sulle unità linguistiche più disparate. Tali unità linguistiche presentano differenze sotto vari aspetti: estensione della unità linguistica, struttura della unità linguistica, variazioni sul piano temporale e su quello intonativo, ecc. Da questo scenario emerge con chiarezza la necessità di individuare una sola unità linguistica che sia ritmicamente pertinente e linguisticamente omogenea, cosicché possa fungere da dominio di analisi del ritmo. Per raggiungere tale scopo, facciamo un passo indietro. Nel primo capitolo di questa ricerca (cfr. §1), abbiamo visto che il ritmo appartiene all'ambito della codifica prosodica del messaggio, cioè a ciò che è definita interfaccia prosodica. Tale interfaccia lega e specifica il piano della produzione fonica con il piano della significazione. Quindi, dato che il ritmo fa parte della codifica prosodica, la unità linguistica che può fungere da dominio di analisi del ritmo va ricercata tra le unità linguistiche che racchiudono ed esprimono le relazioni tra il ritmo e gli altri livelli della grammatica. In questo modo, la unità che rappresenterà il dominio delle analisi ritmiche sarà una unità linguisticamente omogenea e ritmicamente pertinente. A questo punto è lecito chiedersi: quale unità può rappresentare il dominio di analisi del ritmo? Questa domanda racchiude il tema del capitolo successivo.

## Capitolo III – La definizione di un dominio di analisi per il ritmo

### 1. Introduzione

Come abbiamo già visto nel secondo capitolo negli ultimi quindici anni, lo studio del ritmo avviene attraverso specifici modelli che basano la classificazione/descrizione delle lingue naturali esclusivamente sulla durata di segmenti vocalici e consonantici racchiusi in unità linguistiche eterogenee per dimensione, struttura e parametri; nel migliore dei casi, si tratta di frasi o enunciati (cfr. Ramus *et al.* 1999; Grabe & Low 2001; Dellwo 2006; Bertini & Bertinetto 2007) definiti su base sintattica o semantica che, com'è noto, si adattano solo parzialmente alla realtà del parlato (cfr. per esempio, Sornicola 1981; Voghera 1992); in altri (la maggioranza), sono costituite da stringhe indivise e produzioni casuali, vale a dire qualsiasi porzione di parlato indipendentemente dalla sua estensione e dalla sua struttura.

Tutto ciò produce una visione non corretta del fenomeno ritmo. A questo punto, è evidente che bisogna individuare una unità ritmicamente pertinente. La scelta di questa unità è un passaggio molto delicato. Nel primo capitolo di questa ricerca (cfr. §1), si è visto come il ritmo appartenga alla interfaccia prosodica, cioè come tale fenomeno intrattenga relazioni con gli altri livelli della grammatica e come diversi parametri operino in co-variazione insieme a questi livelli. Pertanto, un'analisi corretta del ritmo implica la scelta di un'unità che rispecchi tali dinamiche, solo in questo modo si evita il rischio di ridurre lo studio del ritmo esclusivamente alla misura della durata e all'uso di unità linguistiche poco pertinenti a un fenomeno così complesso.

E' necessario, dunque, partire da una funzione fondamentale della prosodia, che è quella di facilitare la scomposizione o *parsing* del flusso parlato in raggruppamenti acusticamente coerenti, che guidano l'ascoltatore nella decisione di quali suoni oppure insieme di suoni compongano le singole unità, passibili di ulteriori analisi linguistiche. Gli studi psicolinguistici, tra cui (Schafer & Speer 1998, cfr. Avesani & Vayra 2005), dimostrano come sia l'intera struttura prosodica ad avere effetti sulle decisioni di *processing* a tutti i livelli della rappresentazione linguistica: sintattica, semantica, pragmatica e anche lessicale. Pertanto, secondo il nostro punto di vista, la unità prosodica deputata a rappresentare il dominio per le analisi ritmiche è la Unità Tonale (d'ora in poi T-U), perché esprime le relazioni con vari livelli della grammatica che costituiscono l'interfaccia tra le condizioni pragmatiche, il piano di progettazione semantica, l'organizzazione della informazione, con la struttura sintattica e il materiale segmentale (fonetico, fonologico, morfologico) di cui è composto il messaggio e, nello stesso tempo, assicura la omogeneità dei parametri ritmico-acustici che operano in co-variazione insieme ad altri livelli prosodici (cfr. Savy 2001).

Negli studi linguistici, la unità di analisi prosodica è variamente definita secondo gli approcci teorici: Gruppo Tonale (*Tone Group*), Gruppo di Respiro (*Breath Group*), Sintagma Intonativo (*Intonational Phrase*) e T-U (*Tone Unit*). La descrizione di questi approcci teorici è il tema del paragrafo successivo.

## 2. Le unità di analisi prosodica

### 2.1 Le varie proposte sulle unità di analisi prosodica

Un dibattito che per lungo tempo ha diviso l'opinione dei linguisti è il riconoscimento del carattere discreto o continuo della intonazione. Questo dibattito ha come tema centrale la possibilità o meno di segmentare un contorno melodico in unità linguisticamente discrete e formalmente riconoscibili. Tale dibattito ha prodotto due schieramenti teorici opposti: approccio per configurazioni e approccio per livelli.

L'approccio per configurazioni nasce nell'ambito della tradizione intonativa inglese (cfr. Armstrong & Ward 1926; Jones 1956) e nel corso del tempo subisce numerose variazioni (cfr. Crystal 1969; Halliday 1963, 1967). Infatti, la prima elaborazione di questo approccio prevedeva che la curva intonativa non potesse essere scomposta in unità discrete. Questo significa che un contorno melodico (*Tune*) costituisce un blocco non scomponibile, quindi una unità funzionale nel suo complesso. In questo approccio, la unità che funge da dominio delle analisi intonative è definita *Breath Group* (Gruppo di Respiro) (cfr. Jones 1956).

Secondo Stetson (1951), il Gruppo di Respiro corrisponde, grosso modo, a una unità intonazionale originata dalla compressione del diaframma da parte dei muscoli addominali: il loro compito è di articolare internamente l'enunciato, creando una serie di unità ritmiche, corrispondenti alle sillabe e delimitate da accenti più o meno forti. Questa unità prosodica è rilevante per la descrizione dei fatti ritmici, perché è strettamente connessa all'approccio di studio sul ritmo che propone Abercrombie (1967) Secondo l'autore, infatti (cfr. capitolo 1 §2.1.1), il ritmo è reso dalla ricorrenza isocrona di specifici impulsi, che dipendono da fatti respiratori: *chest-pulses* (impulso toracico) e *stress-pulses* (impulso addominale). Quindi, il Gruppo di Respiro può essere vista come l'unità più ampia che racchiude l'alternanza tra gli impulsi.

Dagli anni sessanta in poi, l'approccio per configurazione subisce delle revisioni (cfr. Crystal 1969, Halliday 1963, 1967). Queste revisioni prevedono che la curva intonativa non sia più analizzata in termini di *Tunes*, ma come una sequenza di unità definite *Tone Units* (T-U). Come abbiamo già visto, due motivi rendono questa unità prosodica pertinente per le analisi ritmiche: 1) esprime le relazioni con vari livelli della grammatica; 2) assicura la omogeneità dei parametri ritmico-acustici che operano in co-variazione insieme ad altri livelli prosodici (cfr. Savy 2001).

L'approccio per livelli, invece, si sviluppa e si diffonde negli Stati Uniti. Tale approccio assume che la intonazione possa essere descritta come una sequenza di unità discrete (*pitch levels* o *pitch phonemes*), con dimensione variabile, definite da fratture intonative. Il primo autore che sviluppa una metodologia per lo studio della intonazione su livelli è Pike (cfr. Soriano 2006).

L'evoluzione più recente di questo approccio è il modello Autosegmentale e Metrico (d'ora in poi AM: cfr. cap1, §1.3.3) proposto da Pierrehumbert (cfr. Pierrehumbert 1980) in cui si considerano rappresentazioni gerarchiche attraverso cui si formalizzano i rapporti tra unità fonologiche non adiacenti nella catena parlata lineare. Il dominio privilegiato di applicazione in tal senso è la prominza, elemento che si individua tramite la griglia metrica (cfr. cap1, §1.3.3) come riflesso dei nodi forti dell'albero prosodico. Le informazioni ritmiche trasmesse dalle sillabe e dai piedi confluiscono all'interno del costituente fonologico più alto dell'albero prosodico, il Sintagma Intermedio (SI).

Nella definizione (piuttosto incerta) di questo costituente concorrono fattori sintattici (cfr. Nespor & Vogel 1986; Selkirk 2005) e fattori semantici correlati alla prominza, che possono determinare il numero di contorni intonativi di un enunciato. La natura di queste informazioni è molto variabile e dipendente da fatti di produzione: tra questi, la VDE, definita nel primo capitolo (cfr. §1.2.2.1), è il parametro che più di ogni altro determina la cosiddetta 'ristrutturazione del SI' in segmenti più piccoli. Infatti, più la stringa è prodotta velocemente, maggiori sono le possibilità che il SI non venga suddiviso in SI più piccoli. Quindi, se la VDE è alta, allora il SI tende a essere lungo; al contrario, se la VDE è bassa, il SI tende a essere più breve. Questi 'atteggiamenti' dimostrano la forte correlazione tra VDE e concezione isocrona del ritmo, perché la estensione del SI rispetta le nozioni postulate sulla durata in termini di isocronia ritmica.

## **2.2.1 Corrispondenza tra T-U e SI**

Come vedremo in questo paragrafo, vari aspetti dimostrano una forte relazione tra T-U (unità fonetica) e SI (unità fonologica).

Ricordiamo, in primo luogo, che per la T-U vengono chiamate in causa le correlazioni probabili con diverse strutture sintattiche (cfr. capitolo1 §1.1.1.1), così come svolge un ruolo nella sua strutturazione la collocazione della prominza come punto di rilievo per marcare il contrasto tra informazione data e nuova (cfr. capitolo1 §1.1.2.1).

Inoltre, gli stessi parametri concorrono sostanzialmente alla loro definizione e individuazione. Anche se il SI è una unità fonologica, i parametri che determinano i suoi confini sono di tipo fonetico: presenza di pause, variazioni del contorno intonativo che realizzano fratture del suo contorno e la VDE (cfr. Crystal 1969; Pierrehumbert 1980; Pierrehumbert & Beckman

1986; Ladd 1996). Infatti, l'estensione del SI dipende da fattori fonetici, come la VDE adottata dal locutore. Tale processo produce lo studio fonologico di vari fenomeni linguistici attraverso la codifica di informazioni di tipo fonetico. Questo significa che la fonetica costituisce la interfaccia tra la rappresentazione cognitiva della fonologia e i fenomeni acustici. Allora, dato che le informazioni fonetiche determinano il dominio su cui si estendono le unità prosodiche in esame, la T-U nella sua definizione fonetica rappresenta, in questa ricerca, il dominio di analisi per il ritmo. Tale unità prosodica garantisce, come sopra descritto, relazioni con altri livelli della grammatica (interfaccia) e omogeneità dei parametri ritmici e intonativi. Inoltre, l'obiettivo di questa ricerca è verificare la validità delle classificazioni ritmico-tipologiche di alcune metriche ritmiche. Le metriche che si esploreranno seguono l'approccio fonetico sul ritmo e oggi questi modelli rappresentano l'unico 'mezzo' attraverso cui avviene la classificazione ritmico-tipologica delle lingue naturali. Pertanto, tale ricerca si inserisce nel filone degli studi fonetici sul ritmo delle lingue naturali. Dato che questa ricerca si inserisce negli studi fonetici sul ritmo, nel paragrafo successivo (cfr. §3) si definiscono le unità di analisi dal punto di vista fonetico.

### **3. L'approccio fonetico: la T-U**

La definizione di T-U come una unità melodica che demarca blocchi informativi del messaggio si trova nel primo capitolo di questa ricerca (cfr. §1). Nell'approccio fonetico, le unità prosodiche che fungono da dominio di analisi sono individuate attraverso vari criteri fonetico-acustici.

La T-U è unità prosodica strettamente correlata alla intonazione. Infatti, la funzione primaria della intonazione, variamente definita nel corso degli anni come strutturale, linguistica o distintiva, consiste nella segmentazione del parlato in gruppi di parole coesi e coerenti sul piano testuale. A ciascuna unità discreta, individuata per via di eventi demarcativi, si assegna il nome di T-U.

Vari parametri, come vedremo nel paragrafo successivo (cfr. §3.2), contribuiscono alla segmentazione del parlato nelle unità definite T-U. Proprio la relazione tra questi parametri ha un ruolo cardine nella definizione di tale unità prosodica:

*<<una unità intonazionalmente e temporalmente integrata: cioè un raggruppamento tenuto insieme da andamenti globali della  $F_0$  e da una dinamica temporale amalgamante>>* (cfr. Caputo 1992: 364).

Da questa definizione emerge che la T-U è una unità data dalla integrazione tra un piano 'intonazionale', dove sono rilevanti i parametri e le variazioni di frequenza fondamentale ( $f_0$ ) e di

intensità (I) e un piano ritmico-temporale, in cui è rilevante il parametro durata (d). Questa definizione ha il pregio di superare la concezione secondo cui sul piano fonetico la descrizione e l'analisi della T-U si traduce essenzialmente nella descrizione e nella analisi degli andamenti del parametro associato al tono, cioè alla frequenza fondamentale (F0).

### 3.1 I parametri che definiscono le T-U

Uno dei punti più complessi degli studi sull'intonazione è l'individuazione di un criterio che consenta la delimitazione della T-U, cioè un criterio capace di individuare le marche di confine o *boundary markers*. Vari criteri sono stati testati allo scopo di rendere semplice e generale la delimitazione delle T-U, ma ancora oggi tale questione resta aperta.

Una parte di studi condotti su questo tema fa riferimento alla presenza di una pausa come principale segnale di demarcazione (cfr. Brown 1990), ma è stato verificato che non c'è sempre corrispondenza regolare tra una pausa e il confine della T-U. Infatti, negli enunciati prodotti spontaneamente può capitare di incontrare pause all'interno di una T-U (connesse a fenomeni di esitazione o di giuntura) e di non incontrarne al confine di T-U. Alcuni autori (cfr. Nespor & Vogel 1986), invece, ricorrono al criterio di 'pausa potenziale', cioè la presenza di un punto in cui potrebbe essere inserita una pausa. L'inserimento nella stringa di una pausa potenziale non crea perturbazioni al contorno intonativo.

Un'altra marca prosodica di confine considerata rilevante è il brusco calo, nella parte finale della T-U, dei valori di F0 o della coppia parametrica F0-I (cfr. Caputo 1994) al di sotto del *range* di variazione utilizzato durante la produzione del resto della T-U. In questo caso, l'inizio di una nuova T-U è segnalato dal rientro degli indici all'interno del *range*. Questo fenomeno è spesso correlato con il fenomeno della *declination* intonativa<sup>17</sup> (cfr. Caputo 1994).

Un altro segnale di confine della T-U, riconosciuto nella maggioranza degli studi, è la presenza di un forte rallentamento della VDE verso la parte finale della stringa, fenomeno noto come *prepausal lengthening*, che produce un significativo allungamento dell'ultima vocale tonica della sequenza.

Infine, alla realizzazione del confine prosodico concorrono anche alcuni fenomeni connessi alla qualità della voce. Infatti, le sillabe laringalizzate o desonorizzate sono più frequenti alla fine oppure all'inizio di una T-U, ma non al suo interno.

Nella realizzazione del confine prosodico, questi tipi di *boundary markers* possono cooccorrere oppure essere presenti separatamente.

---

<sup>17</sup> Per una definizione della declinazione intonativa (cfr. Sorianello 2006).

Dopo la definizione delle marche che delimitano i confini delle T-U, nel capitolo successivo si descrivono i materiali e il metodo utilizzati in questa ricerca.

## Capitolo IV – Materiali e metodo

### 1. Introduzione

Come già detto nel secondo capitolo della ricerca (cfr. §2), la unità che rappresenta il dominio delle analisi ritmiche non è l'unico problema che attanaglia gli studi sul ritmo. Esiste anche un forte dibattito che riguarda i corpora su cui si sperimentano le metriche ritmiche.

La maggior parte delle ricerche sul ritmo sperimentano le metriche su corpora di parlato letto (cfr. Ramus *et al.* 1999; Grabe & Low 2002; Dellwo 2003, 2006), invece, pochissimi lavori si basano su corpora di parlato spontaneo o semi-spontaneo (cfr. Bertini & Bertinetto 2007; Bertinetto & Bertini 2008; Arvaniti 2009): queste sperimentazioni restituiscono classificazioni ritmico-tipologiche contrastanti (cfr. Arvaniti 2009).

Come ben sappiamo, il parlato letto è un tipo di parlato poco rappresentativo della realtà linguistica di una popolazione, perché la sua acquisizione avviene in condizioni 'ideali' e 'guidate' in cui le caratteristiche fonetiche, fonologiche, prosodiche e le informazioni sintattiche e semantiche sono sotto controllo. È ben noto, invece, che nel parlato spontaneo e semi-spontaneo lo stile di elocuzione produce effetti macroscopici a livello di comportamenti ritmici (cfr. Bertini & Bertinetto 2007; Bertinetto & Bertini 2008). Soprattutto la VDE ha un forte impatto sulla durata dei segmenti vocalici e consonantici che provoca come conseguenza una maggiore variabilità della durata (cfr. cap.1 §1.2.2.1).

La scelta del corpus è una fase fondamentale, delicata e fondante per lo sviluppo dei lavori di ricerca. Questa scelta è correlata al grado di generalità del corpus che dipende dalla misura in cui i materiali sono stati selezionati rispetto alle diverse varietà di una lingua. Questo significa che un corpus deve fornire un modello della lingua analizzata capace di restituire, nel modo più accurato possibile, un quadro delle varietà e delle tendenze linguistiche rispettandone le proporzioni e permettendo, in questo modo, di generalizzare le proprietà linguistiche del corpus all'intera popolazione. Nel paragrafo successivo si descrivono le motivazioni che hanno guidato la scelta del corpus di questa ricerca che segue la direzione appena tracciata.

### 2. Il corpus

Il corpus analizzato in questa ricerca consiste in una selezione di quindici dialoghi semispontanei di tipo *map-task* selezionati tra i materiali del corpus CLIPS (*Corpora e Lessici di Italiano Parlato e Scritto*, cfr. Albano Leoni 2006; Savy & Cutugno 2009), di cui si dà qui di seguito una breve descrizione.

## 2.1 Il corpus CLIPS

CLIPS è un progetto di raccolta di testi scritti e parlati di italiano, stratificato dal punto di vista diatopico, diafasico e diamesico (cfr. Savy & Cutugno 2009). Agli scopi di questa ricerca risulta fondamentale la stratificazione diatopica e diafasica: la prima perché consente un'esplorazione di diverse varietà di italiano e la verifica (o eventuale rettifica) delle caratteristiche ritmiche di tali varietà; la seconda, perché consente la scelta di materiali non predisposti *ad hoc* per misure metriche, ma spontanei (almeno sotto il profilo fonetico-prosodico).

La stratificazione diatopica di CLIPS, quindi la scelta delle località che compongono il corpus, tiene conto, innanzitutto, di specifiche esigenze di rappresentatività sociolinguistica, analizzate tenendo conto di aspetti complessi e molteplici<sup>18</sup>. Tali esigenze hanno anche delineato le caratteristiche degli informatori che compongono il corpus: Infatti, per ridurre al minimo la interferenza di variabili sociolinguistiche non controllabili e nello stesso tempo rispondere ai requisiti di rapidità e di relativa semplicità delle operazioni di raccolta del materiale, è stato ritenuto indispensabile adottare un campione sostanzialmente omogeneo, quanto meno per le variabili fondamentali: status socioeconomico, classe di età, livello di istruzione e residenza in centri medi e grandi. I parlanti che costituiscono il corpus sono di entrambi i sessi, con percentuali equivalenti. Tali locutori hanno necessariamente caratteristiche biografiche che li qualificano come rappresentativi del punto linguistico al quale appartengono.

Sulla base della rappresentatività sociolinguistica, si è puntato alla rappresentazione geolinguistica dell'intero territorio italiano, facendo riferimento alla partizione dialettale dell'Italia tradizionalmente descritta in base ai raggruppamenti delle varietà dialettali in famiglie linguistiche, individuando infine i seguenti punti di raccolta (cfr. Savy & Cutugno 2009: 5):

- a) Torino, Milano, Bergamo, Genova e Parma (area gallo-italica);
- b) Venezia (area veneta);
- c) Firenze (area toscana);
- d) Perugia e Roma (area mediana);
- e) Napoli e Bari (area meridionale);
- f) Lecce, Catanzaro e Palermo (area meridionale estrema);
- g) Cagliari (area sarda).

---

<sup>18</sup> (per i quali si rimanda a Sobrero & Tempesta 2007: [http://www.clips.unina.it/it/documenti/1\\_scelta\\_informatori\\_e\\_localita.pdf](http://www.clips.unina.it/it/documenti/1_scelta_informatori_e_localita.pdf)).

In merito alla variazione diafasica, il corpus CLIPS è stratificato in 5 sottocorpora (cfr. Savy & Cutugno 2009: 3):

- 1) parlato radiotelevisivo;
- 2) parlato dialogico;
- 3) parlato letto;
- 4) parlato telefonico,
- 5) parlato ortofonico.

I primi 4 seguono la distribuzione regionale (diatopica) sopra descritta, mentre l'ultimo consiste in registrazioni effettuate da parlatori professionisti secondo norme di pronuncia 'standard' (cfr. Falcone *et al.* 2007<sup>19</sup>; Bonomi *et al.* 2007<sup>20</sup>).

Come anticipato, agli scopi della ricerca è stato preso in considerazione il sottocorpus dialogico, composto, nel complesso, di 240 dialoghi, prodotti da 120 coppie di parlatori (116 maschi e 124 femmine, cfr. Savy & Cutugno, 2009: 7), per un totale di circa 50 ore di parlato (idem: 9).

I dialoghi sono stati raccolti attraverso due diverse tecniche di elicitazione, appartenenti alle cosiddette 'situazioni-gioco' (cfr. Cerrato 2007<sup>21</sup>): la tecnica del *map-task* (sviluppata dal HCRC di Edimburgo, cfr. Carletta *et al.* 1996) e la tecnica del 'test delle differenze' ('*Spot the differences*', Pean *et al.* 1993). Entrambe le situazioni producono conversazioni semi-spontanee di tipo *task-oriented*, volte cioè a dare e a ricevere indicazioni utili per raggiungere un 'task', che hanno il pregio di coinvolgere il parlante e distrarlo dalla condizione di registrazione; ne risulta così uno scambio comunicativo abbastanza 'naturale' tra i due partecipanti (sebbene fortemente vincolato dal punto di vista pragmatico-lessicale), ma allo stesso tempo si consente un buon controllo delle condizioni di registrazione. Per questo motivo, si tratta di due tecniche frequentemente utilizzate nella raccolta di corpora ad uso fonetico.

Un ulteriore vantaggio nella scelta del corpus Clips, infine, è costituito dal fatto che una percentuale intorno al 10% del corpus (cfr. Savy & Cutugno 2009: 9) è stata etichettata a diversi livelli segmentali, ivi compreso un livello fonetico-acustico riutilizzabile ai fini di questa ricerca<sup>22</sup>.

---

<sup>19</sup> [http://www.clips.unina.it/it/documenti/9\\_descrizione\\_del\\_corpus\\_ortofonico.pdf](http://www.clips.unina.it/it/documenti/9_descrizione_del_corpus_ortofonico.pdf).

<sup>20</sup> [http://www.clips.unina.it/it/documenti/8\\_definizione\\_database\\_ortofonico.pdf](http://www.clips.unina.it/it/documenti/8_definizione_database_ortofonico.pdf).

<sup>21</sup> [http://www.clips.unina.it/it/documenti/2\\_tecniche\\_di\\_elicitazione\\_dialogica.pdf](http://www.clips.unina.it/it/documenti/2_tecniche_di_elicitazione_dialogica.pdf).

<sup>22</sup> Le norme di etichettatura sono consultabili in Savy (2007: cfr. [http://www.clips.unina.it/it/documenti/11\\_specifiche\\_trascrizione\\_ortografica.pdf](http://www.clips.unina.it/it/documenti/11_specifiche_trascrizione_ortografica.pdf)).

## 2.2 Selezione dei materiali

Come si è detto, i materiali oggetto di questa ricerca costituiscono un estratto selezionato del corpus CLIPS.

La scelta delle varietà analizzate rispecchia la distribuzione geolinguistica di partenza in percentuale ridotta, contemplando i seguenti raggruppamenti:

- a) area gallo-italica: Torino e Milano;
- b) area veneta: Venezia;
- c) area toscana: Firenze;
- d) area mediana: Roma;
- e) area meridionale: Napoli;
- f) area meridionale estrema: Lecce e Palermo.

Per ogni località è stato selezionato 1 dialogo di tipo *map-task* o test delle differenze (per un totale di 8 dialoghi e 15 locutori).

A valle di una prima scrematura dei materiali, sono stati considerati per l'analisi ritmica solo turni di lunghezza superiore a tre sillabe, per un totale di 414 turni, suddivisi come in tabella 4.1. Successivamente i turni sono stati analizzati (secondo la procedura che verrà descritta nel successivo §3) prima in unità tonali (T-U, totale 830) e poi in sillabe (totale 8276), ottenendo la seguente distribuzione:

<b>Varietà di italiano regionale</b>	<b>n. turni</b>	<b>n. T-U</b>	<b>n. sillabe</b>
<b>Firenze</b>	98	129	1531
<b>Lecce</b>	54	96	851
<b>Milano</b>	32	65	617
<b>Napoli</b>	62	164	1668
<b>Palermo</b>	32	96	956
<b>Roma</b>	49	127	1200
<b>Torino</b>	25	51	404
<b>Venezia</b>	62	102	1049

Tabella 4.1: numero di turni, di T-U e di sillabe che costituiscono corpus della ricerca.

Nel paragrafo successivo, si espone il metodo di segmentazione del materiale sonoro in T-U (cfr. §3.1), sillabe (cfr. §3.2), consonanti e vocali (cfr. 3.3).



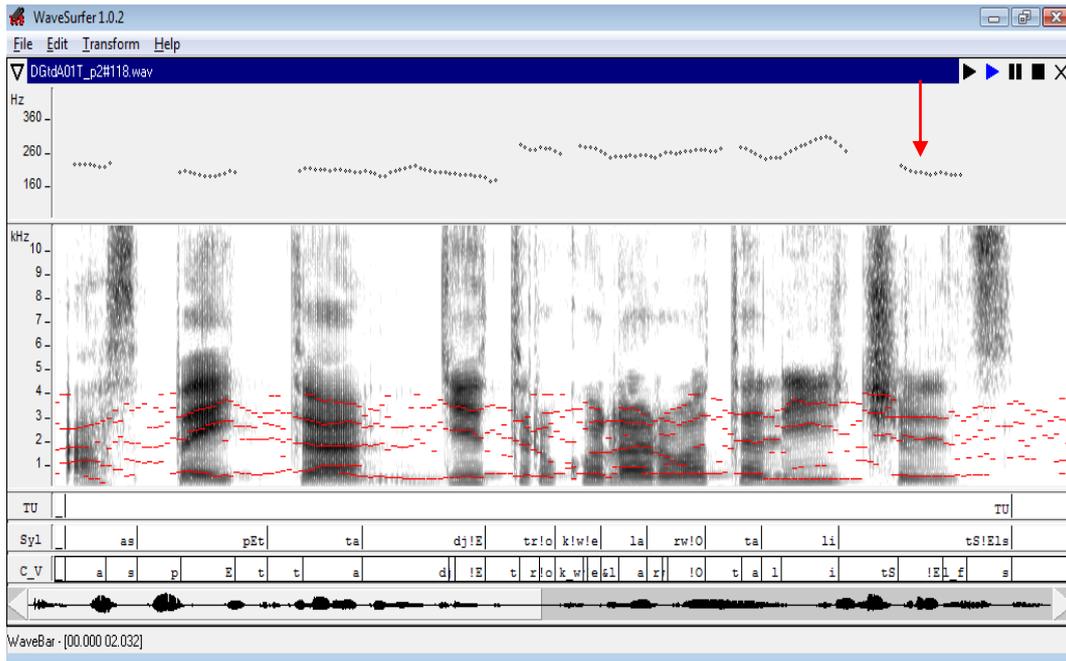


Figura 4.2: parlante di sesso femminile della varietà di Torino.

In questa figura, la freccia specifica il brusco calo della F0 nella parte finale della T-U. Questo brusco calo delimita il confine della T-U.

La figura 4.3 specifica il fenomeno del *prepausal lengthening*. Il testo della T-U è: ‘e poi due normali’:

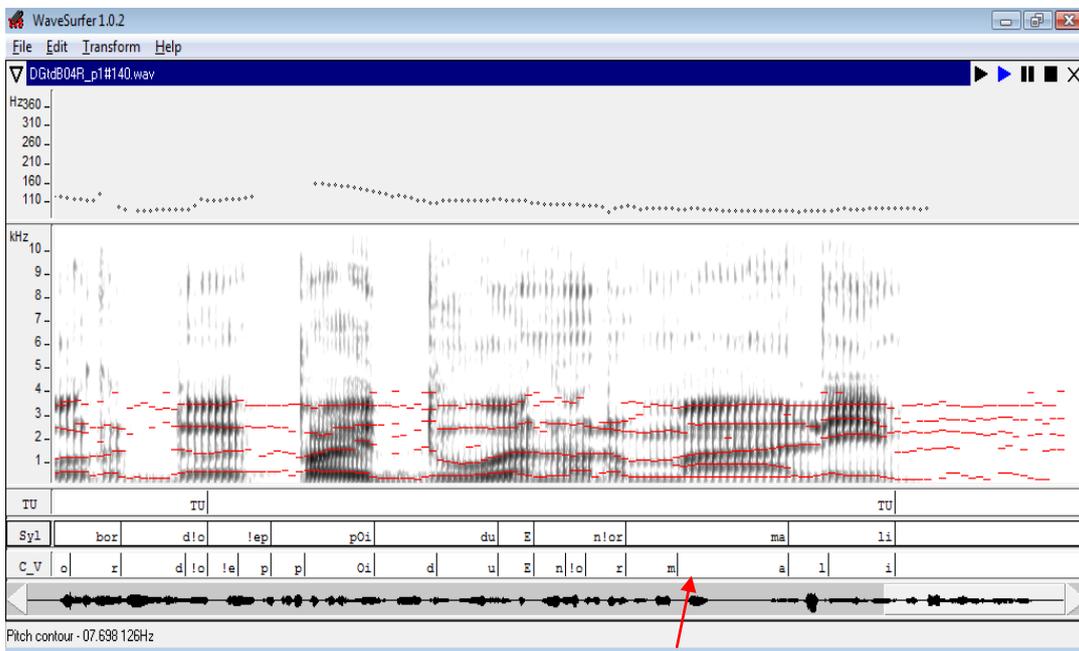


Figura 4.3: parlante di sesso maschile della varietà di Roma.

Il *prepausal lengthening* consiste in un forte rallentamento della VDE verso la parte finale della stringa, che produce un significativo allungamento dell'ultima vocale tonica della sequenza.

Nella figura 4.3, la freccia rossa designa la vocale tonica allungata per via del *prepausal lengthening*.

Accanto a questi criteri oggettivi e misurabili, come detto in cap.3 §3.1, si può considerare quello più interpretativo di ‘pausa potenziale’ con cui si individua un punto di possibile inserimento di pausa nella curva intonativa (così come nella struttura sintattico-semantiche dell’enunciato), seguito sempre da un *reset* (anche parziale) dei parametri acustici, come avviene, per esempio, nella sequenza che segue:

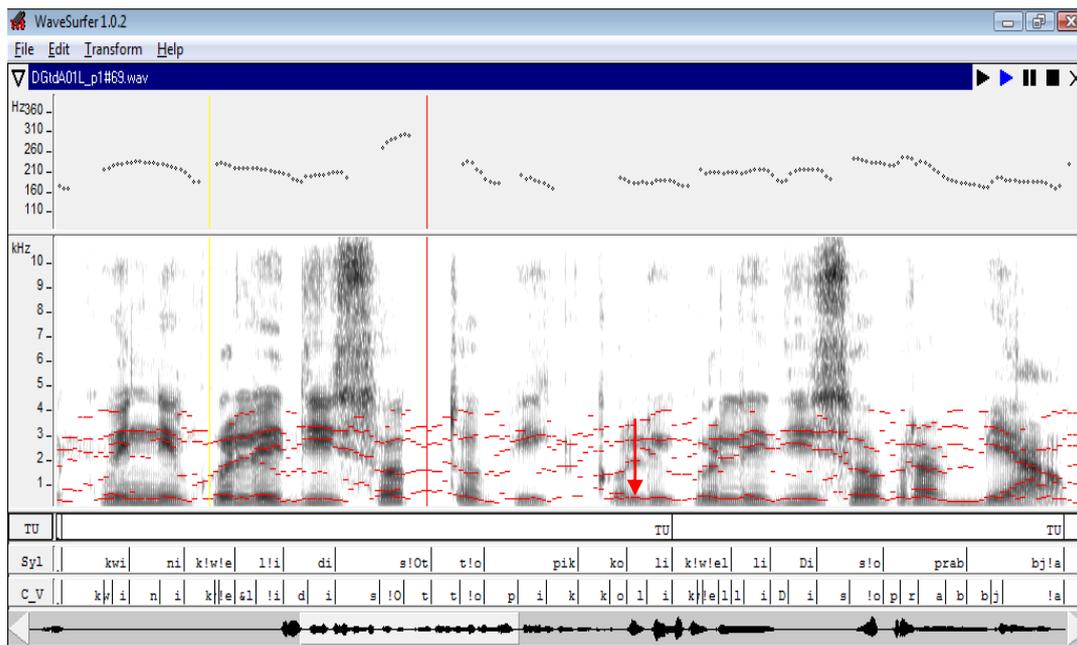


Figura 4.4: parlante di sesso femminile della varietà di Lecce.

In questo caso, la presenza della pausa potenziale è stata rilevata sia attraverso l’analisi percettiva del segnale, perché durante l’eloquio si percepisce uno ‘stacco’ che comporta la suddivisione del turno in T-U differenti, sia mediante l’analisi acustica, con la presenza del reset dei valori di F0. Nella figura 4.4, la freccia indica il punto in cui si inserisce la pausa potenziale.

Dopo la segmentazione del corpus in T-U, si è proceduto con la segmentazione delle T-U individuate in sillabe.

### 3.2 Principi di segmentazione in sillabe

La segmentazione delle T-U in sillabe avviene su base fonetica, cioè le sillabe sono individuate sul parlato effettivamente realizzato. Quindi, non si considerano le sillabe ‘potenziali’, cioè fonologiche, ma quelle reali, cioè fonetiche.

In questo lavoro, la individuazione dei confini sillabici e la segmentazione dei gruppi consonantici e degli incontri vocalici sono trattati nel modo seguente:

- a) muta + liquida (es: [pr], [tr]) sono considerati nessi tautosillabici (es. ‘centimetri’ < ‘cen.ti.me.tri’ < [tSen"ti:metri]);
- b) per quanto riguarda il nesso [s] + consonante, seguendo le scale di sonorità, la fricativa dentale sorda è posta sempre nella coda della sillaba precedente. Un esempio tratto dal corpus di questa tesi è: ‘una sporgenza’ che diviso in sillabe diventa ‘uas.por.gen.tsa’ > [%uaspor"dZentsa]. Il nesso [s] + consonante viene collocato nell’incipit della prima sillaba solo nel caso in cui si trovi nella posizione di attacco della T-U;
- c) le consonanti lunghe (es. ‘possano’ < ‘pos.sa.no’ < ["possano]), rafforzate (es. ‘lasciamo’ < ‘laS.Sa.mo’ < [la"SSamo]) e quelle che subiscono raddoppiamento fonosintattico (es. ‘può muovere’ < ‘pwom.mwo.ve.re’ [pwo"mmwo:vere]) hanno lo stesso trattamento, cioè si suddividono tra la coda della sillaba precedente e l’incipit di quella successiva;
- d) in merito ai dittonghi ascendenti (es. ‘muovere’ < ‘mwo.ve.re’ < ["mwo:vere]), l’approssimante è considerata come consonante. Quindi, l’approssimante va nell’incipit della sillaba, mentre la vocale costituisce il nucleo;
- e) i dittonghi discendenti (es. ‘ai’ in ‘mai’), invece, rappresentano un caso più problematico rispetto a quelli ascendenti. Innanzitutto, si sottolinea che, in questa tesi, le semivocali sono considerate come allofoni delle vocali (cfr. Muljačić 1972). Quindi, nel momento in cui le vocali di questo tipo di dittongo sono percepite come staccate, esse formano il nucleo di due sillabe differenti. Quindi, per la suddivisione in sillabe di parole che contengono dittonghi discendenti si ricorre al criterio percettivo: se il dittongo presenta percettivamente uno stacco, allora si separano le vocali che generano così due sillabe differenti. Si ricorre a questo principio, perché nel parlato semi-spontaneo una VDE può provocare fenomeni di riduzione fonico-acustica che fonde o rende opachi i nuclei vocalici, rendendo spesso impossibile la separazione delle due vocali di un dittongo ascendente. Quindi, con una VDE bassa si può assistere alla separazione delle due vocali in due sillabe differenti (es. ‘**ma.i**’), invece, con una VDE alta può manifestarsi una mancata separazione delle due vocali, che genera una sillaba con nucleo complesso (es. ‘**mai**’);

- f) la sillabificazione degli incontri vocalici rappresenta un altro problema rilevante. Le difficoltà sia di segmentazione dei foni vocalici adiacenti, sia di definizione e attribuzione di parte di segnale a categorie definite, porta a considerare queste sequenze come costituenti di un nucleo vocalico complesso. Se all'interno del nucleo vocalico complesso, però, si individua percettivamente uno stacco, allora tale nucleo viene suddiviso tra due sillabe e il confine è posto arbitrariamente in prossimità della vocale che produce lo iato. Un esempio è: 'molto meno io'. La parte finale di questa stringa presenta un incontro vocalico che si risolve nel modo seguente: 'mol.to.me.no.io' < [%moltomeno"io];
- g) infine, gli allungamenti vocalici e consonantici, le pause di esitazione e le interiezioni non sono state considerate ai fini delle analisi ritmiche, quindi questo materiale è stato eliminato dai segnali vocali. Nel caso di allungamenti consonantici o vocalici, allo scopo di bilanciare il corpus, è stata eliminata la intera sillaba che contiene tali allungamenti.

Le sillabe individuate attraverso questi principi sono state a loro volta suddivise in foni vocalici e consonantici.

### 3.3 Segmentazione in vocali e consonanti

La segmentazione dei segnali sonori in vocali e consonanti rappresenta una delle fasi più delicate e importanti di questo lavoro, perché è il passo essenziale per la sperimentazione delle metriche ritmiche.

Un problema particolarmente delicato è costituito dalla misurazione delle componenti segmentali dei dittonghi. Infatti, come ben sa chiunque si sia cimentato nella materia, la divisione di questi dittonghi è impresa difficile e delicata, a maggior ragione in un corpus di parlato semi-spontaneo, con fasi di elocuzione accelerate o addirittura ipoarticolate. È dunque essenziale, per il funzionamento delle metriche calcolare esattamente le durate delle vocali e delle semiconsonanti. In questa ricerca, come descritto (cfr. §3.2), nei dittonghi ascendenti l'approssimante è considerata come consonante (es. 'pjede' < CCV.CV), mentre nei dittonghi discendenti le semivocali sono considerate come allofoni delle vocali (es. 'mai' < 'CV.V').

Nel primo caso, allo scopo di separare la vocale dalla consonante, si è proceduto con un'analisi percettiva della sequenza. Allo scopo di individuare percettivamente il confine 'fisico' tra i due foni, si parte dalla vocale, l'elemento più stabile e di maggiore durata in questi dittonghi, e si spostano i confini verso la semiconsonante (cfr. Trumper *et al.* 1993). Di seguito, si riporta un esempio estratto dal corpus della ricerca:

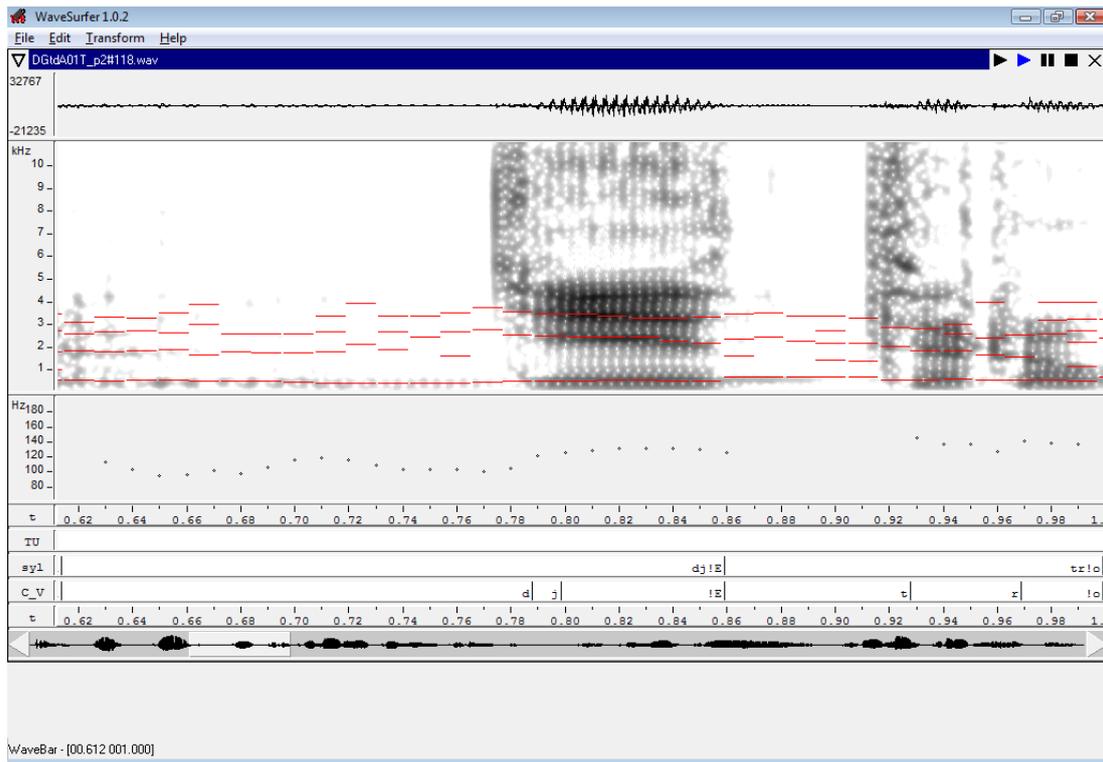


Figura 4.5: esempio dittongo ascendente ["djetto].

Nel secondo caso, si è proceduto innanzitutto con un'analisi percettiva del dittongo. Se percettivamente il dittongo presenta uno stacco, allora le vocali sono separate in due nuclei sillabici differenti. L'analisi percettiva è anche supportata da informazioni acustiche. Il confine che separa le due vocali è specificato acusticamente dal punto sullo spettrogramma in cui si manifesta la variazione dell'andamento delle formanti, che indica il passaggio da una vocale all'altra. Di seguito, si rende un esempio estratto dal corpus:

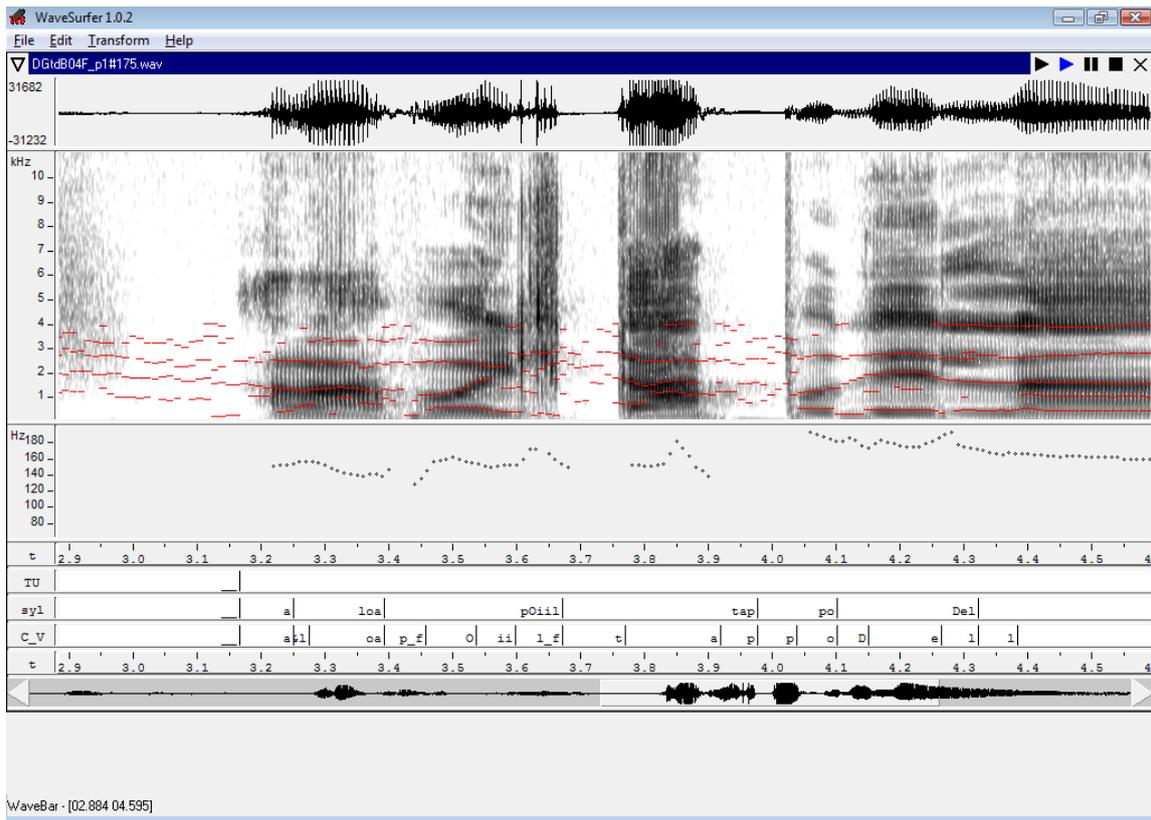


Figura 4.6: esempio dittongo discendente [oi].

Lo stesso procedimento vale per lo iato: se due vocali sono in iato, allora si manifesta lo stacco percettivo e il punto in cui si palesa la variazione dell'andamento delle formanti rappresenta il confine tra i due foni vocalici. Anche per questo caso si riporta un esempio estratto dal corpus della ricerca:

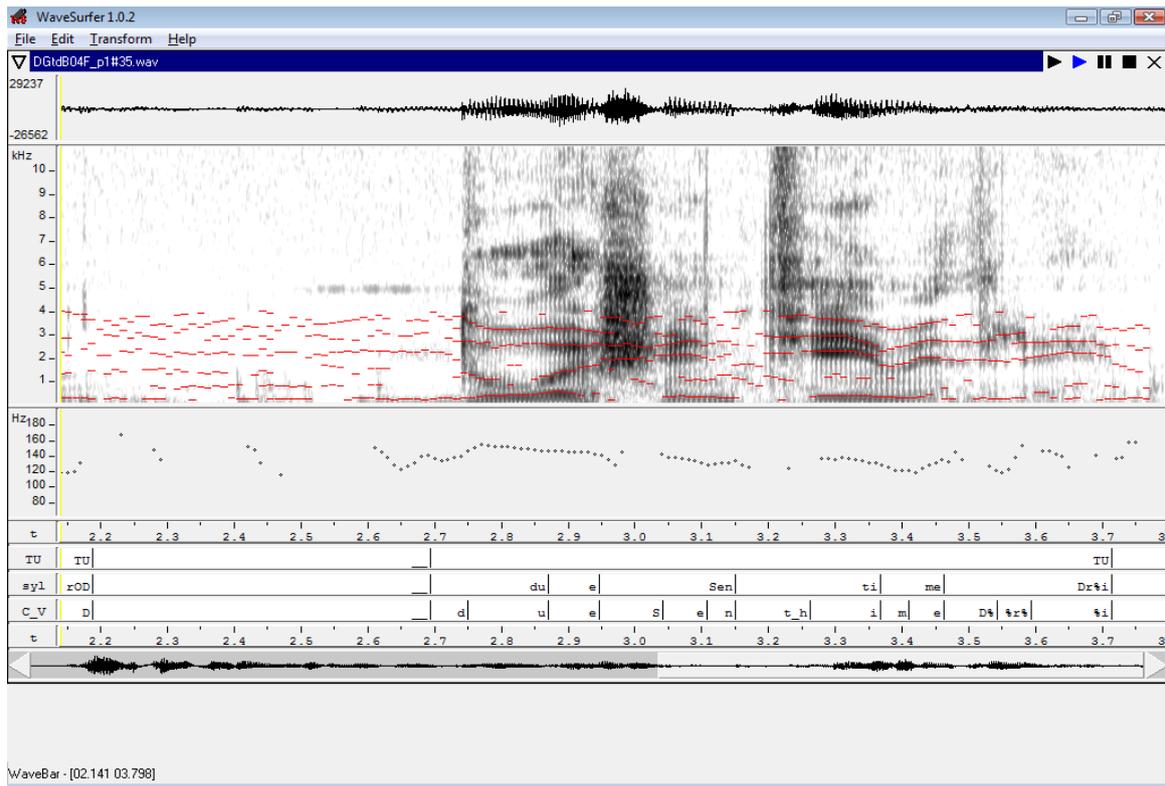


Figura 4.7: esempio di iato ["due].

Nel caso in cui le vocali che compongono il dittongo discendente, però, siano soggette a riduzione fonico-acustica che rende indistinguibile i due foni vocalici, allora entrambi i foni compongono il nucleo della sillaba:

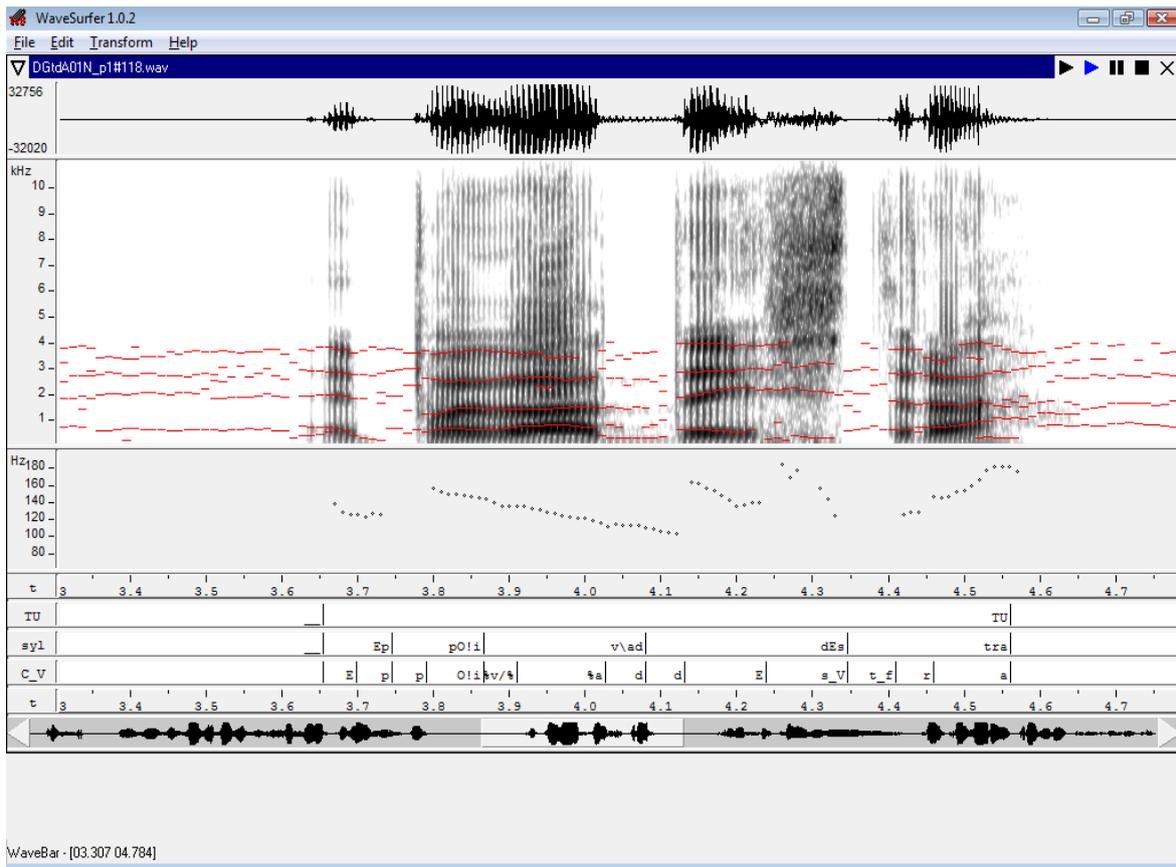


Figura 4.8: esempio di dittongo discendente [pɔi] in cui la riduzione fonico-acustica rende le due vocali indistinguibili.

Questa procedura è utilizzata anche nel caso in cui le vocali di uno iato subiscano riduzione fonico-acustica che le rende indistinguibili:

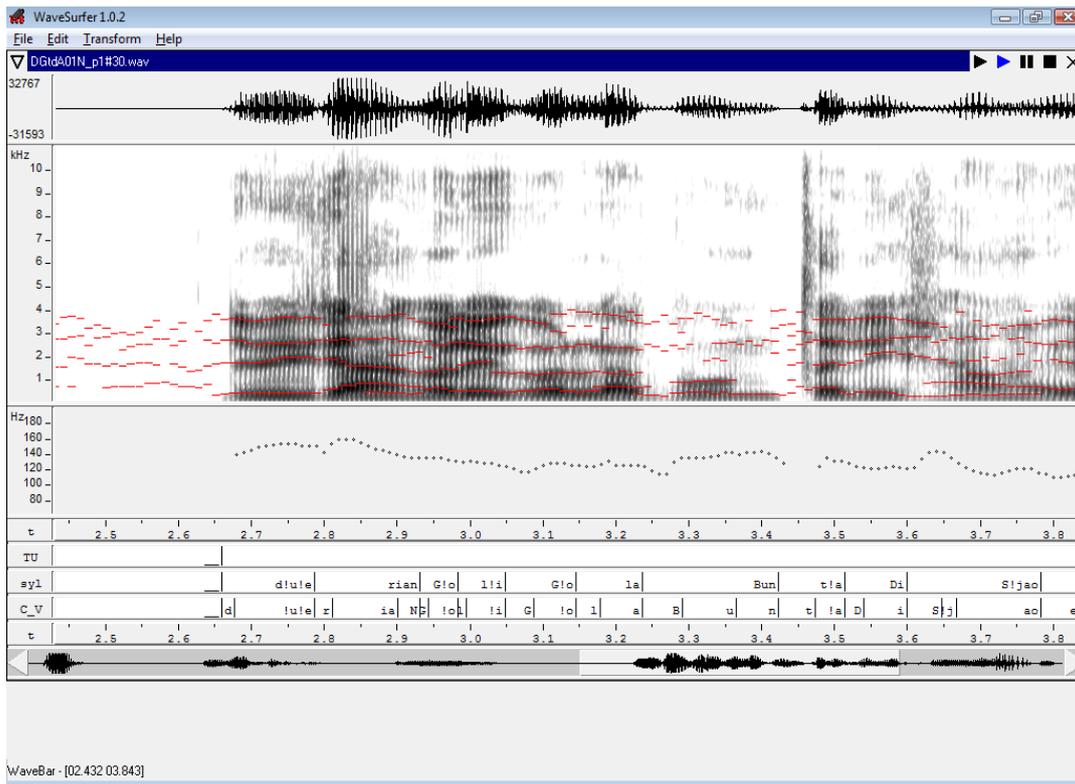


Figura 4.9: esempio di iato [ "due] in cui la riduzione fonico-acustica rende le due vocali indistinguibili.

Un altro problema di segmentazione è rappresentato dagli incontri vocalici. In questi casi si effettua, innanzitutto, un'analisi percettiva delle sequenze. Se dall'analisi percettiva si riscontra uno stacco tra le vocali, il confine è posto nel punto in cui si manifesta la variazione dell'andamento formantico. Un esempio estratto dal corpus è il testo di parte di una T-U 'la mia è un dito' di cui si riporta lo spettrogramma:

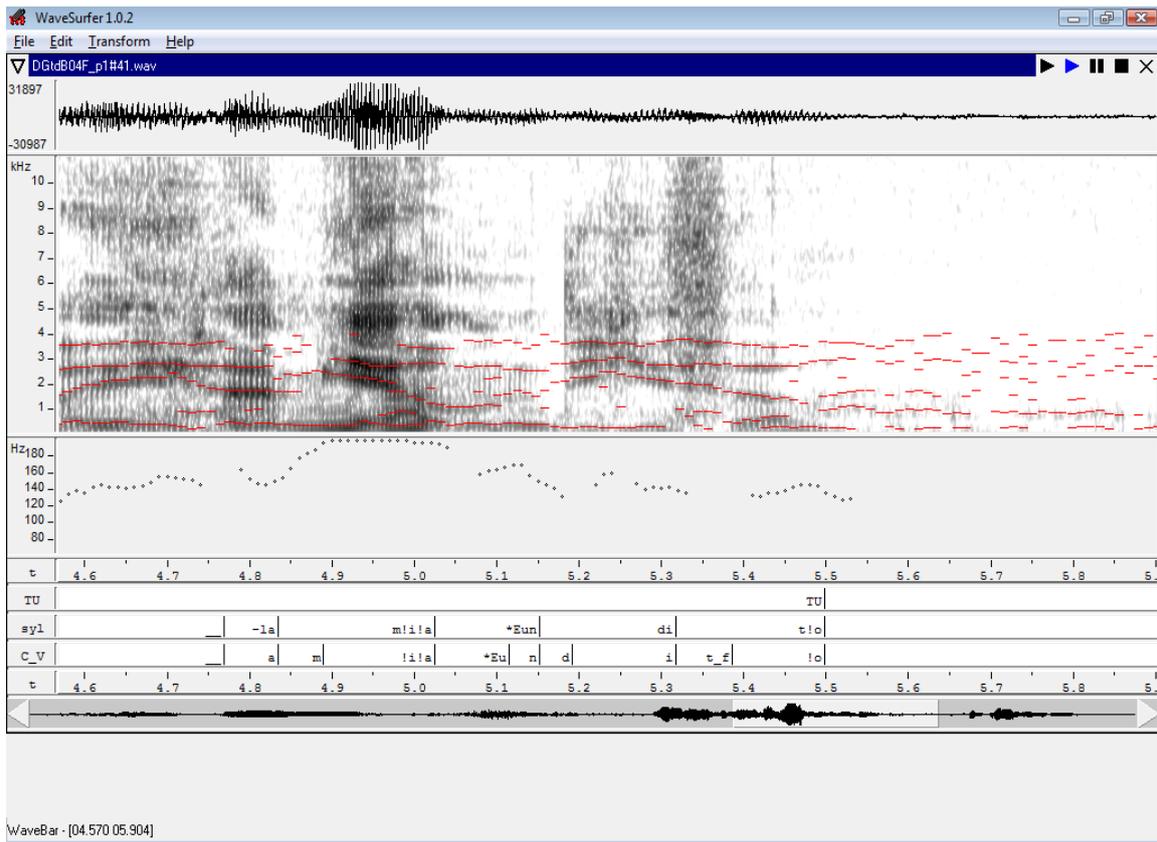


Figura 4.10: esempio di segmentazione di incontri vocalici, quando dall'analisi percettiva si individua uno stacco tra le vocali.

Se non si registra uno stacco, invece, gli incontri vocalici compongono il nucleo di una sola sillaba per via della sinalefe, cioè la fusione in una unica sillaba della vocale (o dittongo) finale di una parola con la vocale (o dittongo) iniziale della successiva. Un esempio che si manifesta nel corpus è la parte di una T-U 'passiamo alla' di cui si rende lo spettrogramma:

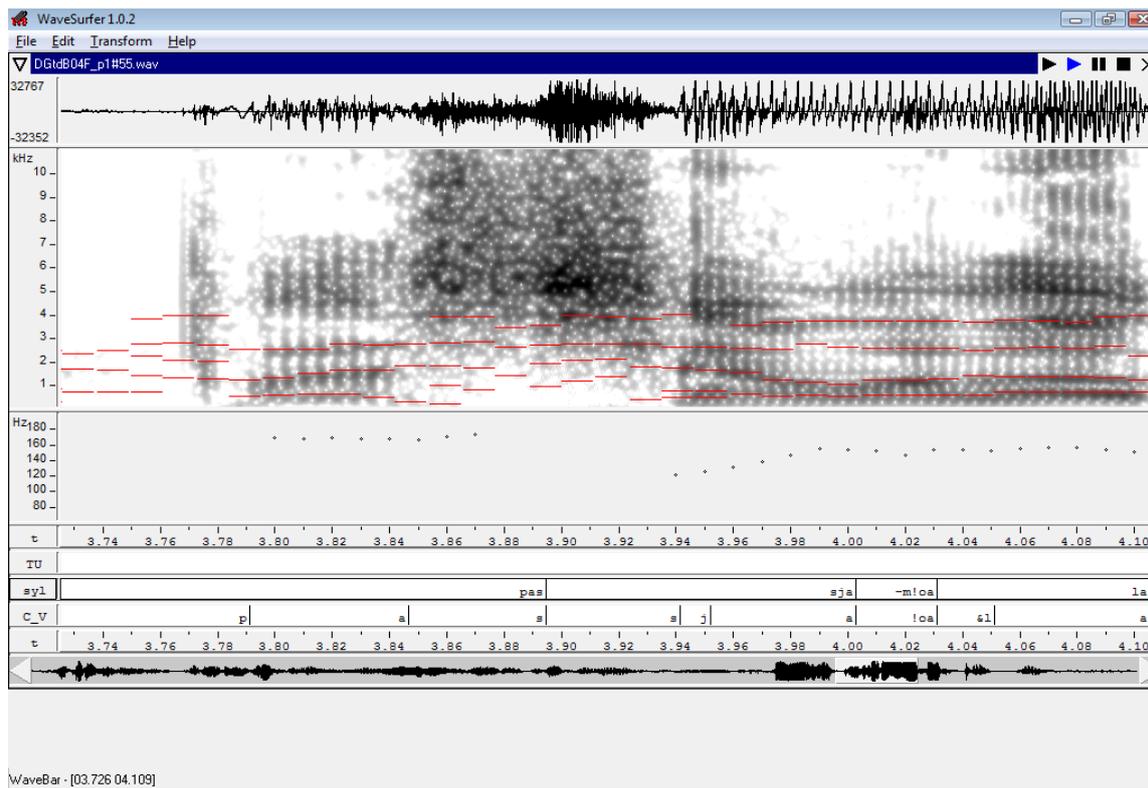


Figura 4.11: esempio di segmentazione di incontri vocalici, quando le due vocali compongono il nucleo di una sola sillaba per via della sinalefe.

Tale procedura ha il vantaggio di non assegnare una durata fissa ai segmenti vocalici e consonantici, scongiurando il rischio di introdurre un livellamento artificiale della varianza.

Nel caso di consonanti lunghe, rafforzate e raddoppiate si è adottato, invece, un criterio a base fonologica. L'intervallo consonantico corrispondente a una consonante lunga, raddoppiata e rafforzata è stato pertanto interpretato come composto da due segmenti fonemati, la cui durata è stata convenzionalmente considerata pari alla metà dell'intero intervallo.

Dopo la descrizione del metodo di segmentazione in T-U, dei principi per la segmentazione sillabica, consonantica e vocalica del materiale sonoro, si riporta un esempio di segmentazione del materiale sonoro:

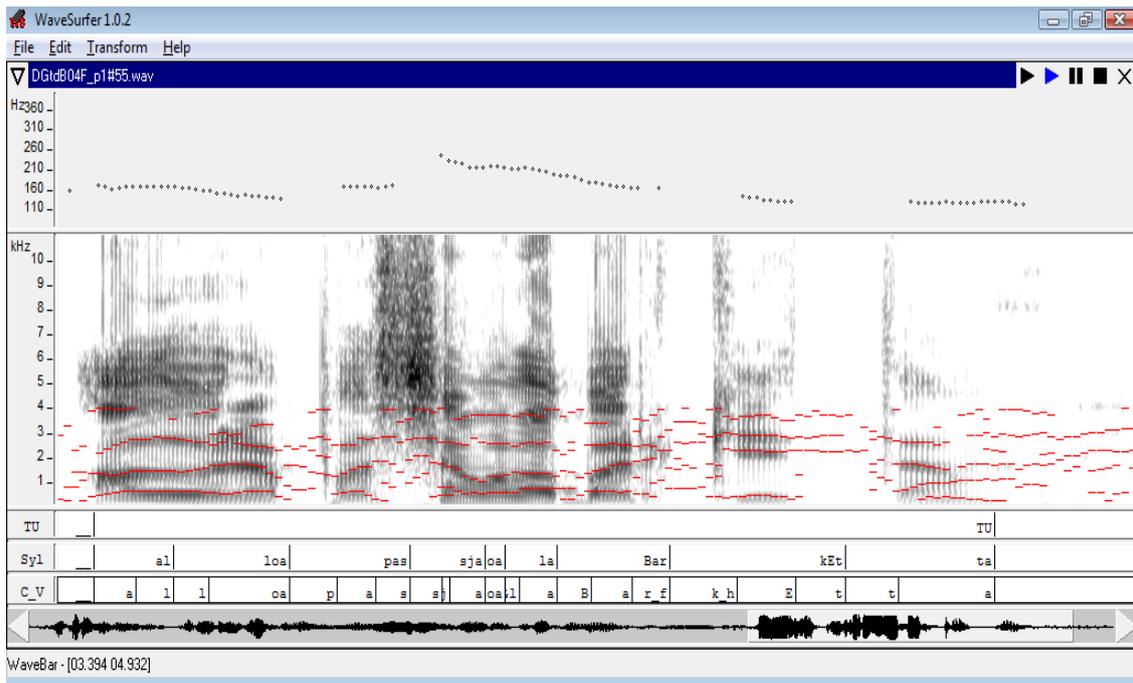


Figura 4.12: esempio di segmentazione in T-U, sillabe, consonanti e vocali estratto dal corpus della ricerca.

Descritti i casi critici che riguardano la segmentazione in vocali e consonanti della catena parlata, l'analisi percettiva e spettrale rappresenta la procedura generale che guida la segmentazione dell'eloquio in foni vocalici e consonantici. Questi principi consentono una corretta segmentazione in consonanti e vocali dell'eloquio e tale segmentazione, come vedremo nel paragrafo successivo (cfr. §3.4), è il materiale su cui avviene la sperimentazione delle metriche ritmiche.

### 3.4 Le metriche considerate per questo lavoro di ricerca

Dopo la segmentazione del materiale sonoro in consonanti e vocali, il corpus è pronto per la sperimentazione delle metriche ritmiche. In questa ricerca, si sperimentano le metriche proposte da Ramus *et al.* (1999), Grabe & Low (2002) e Dellwo (2006). La scelta di queste metriche dipende dal fatto che il calcolo dei correlati avviene su materiale segmentato foneticamente in consonanti e vocali. Quindi, tali metriche seguono un approccio fonetico che corrisponde a quello in cui si inserisce questa ricerca sul ritmo. Il calcolo dei correlati delle metriche avviene attraverso Correlatore<sup>23</sup> (cfr. Mairano 2010).

La metrica di Bertinetto & Bertini (2008), invece, segue un approccio *phonologically driven*, a voler utilizzare le parole degli autori, cioè questo modello prevede che la durata di ogni intervallo consonantico o vocalico venga diviso per il numero di segmenti fonologici che lo compongono. Pertanto, dato che tale modello segue un approccio fonologico, non è sperimentato in questa ricerca.

<sup>23</sup> <http://www.lfsag.unito.it/correlatore/index.html>.

Una descrizione completa delle metriche ritmiche appena citate si trova nel secondo capitolo del presente lavoro (cfr. §2).

Come sopra specificato, questo lavoro si inserisce nel filone di ricerca fonetico sul ritmo. Pertanto, le analisi ritmiche sul materiale sonoro avvengono attraverso le suddette metriche che seguono l'approccio fonetico: Ramus *et al.* (1999), Grabe & Low (2002) e Dellwo (2006). Proprio la sperimentazione di queste metriche ritmiche sul corpus della ricerca esplicita l'obiettivo del lavoro: verificare la validità di alcune metriche ritmiche. In questa ricerca, per validità si intende la capacità delle metriche ritmiche di classificare ritmico-tipologicamente le varietà di italiano che compongono il corpus.

Nella introduzione di questo capitolo, è stato specificato che la capacità classificatoria delle metriche si testa spesso su parlato letto, pochi lavori utilizzano le metriche su parlato spontaneo, il tipo di parlato che realmente riflette le proprietà ritmiche di una lingua naturale, o semi-spontaneo che ha caratteristiche fonetiche e morfosintattiche vicine al parlato spontaneo. La sperimentazione delle metriche ritmiche su parlato letto ha prodotto una estensione o generalizzazione dei risultati classificatori al parlato spontaneo e semi-spontaneo. Oggi, però, questa generalizzazione è posta in discussione (cfr. Arvaniti 2009): se le classificazioni ritmico-tipologiche effettuate dalle metriche ritmiche sul parlato letto rispecchiano la realtà ritmica delle lingue naturali, allora le differenze ritmiche tra le lingue naturali si massimizzano nel parlato spontaneo, che è in grado di catturare i diversi modelli ritmici. Allora, la scelta in questa ricerca del parlato semi-spontaneo consente di raggiungere un duplice obiettivo: 1) verificare la capacità classificatoria delle metriche ritmiche; 2) colmare la enorme lacuna sul tipo di parlato su cui si sperimentano le metriche ritmiche.

Nel capitolo successivo (cfr. §5) verrà avanzata la sperimentazione delle metriche citate sul corpus di riferimento della ricerca.

## V Capitolo – Analisi dei dati

### 1. Sperimentazione delle metriche ritmiche

#### 1.1 La metrica di Ramus *et al.* (1999)

La metrica di Ramus *et al.* (1999) è la prima di cui si mostrano i risultati della sperimentazione sul corpus di questa ricerca. Dopo avere sperimentato diverse combinazioni di parametri metrici, (cfr. §2) gli autori considerano la proiezione %V- $\Delta$ C come quella che rende in modo migliore le differenze ritmico-tipologiche tra le lingue naturali (cfr. Ramus *et al.* 1999), pertanto questa sarà la coppia di parametri che sarà impiegata anche in questo capitolo. Di seguito, si riporta sia la figura 5.1 che contiene i risultati della sperimentazione sul corpus della ricerca, sia la figura 5.2 che ritrae i risultati di Ramus *et al.* (1999) con la proiezione %V- $\Delta$ C:

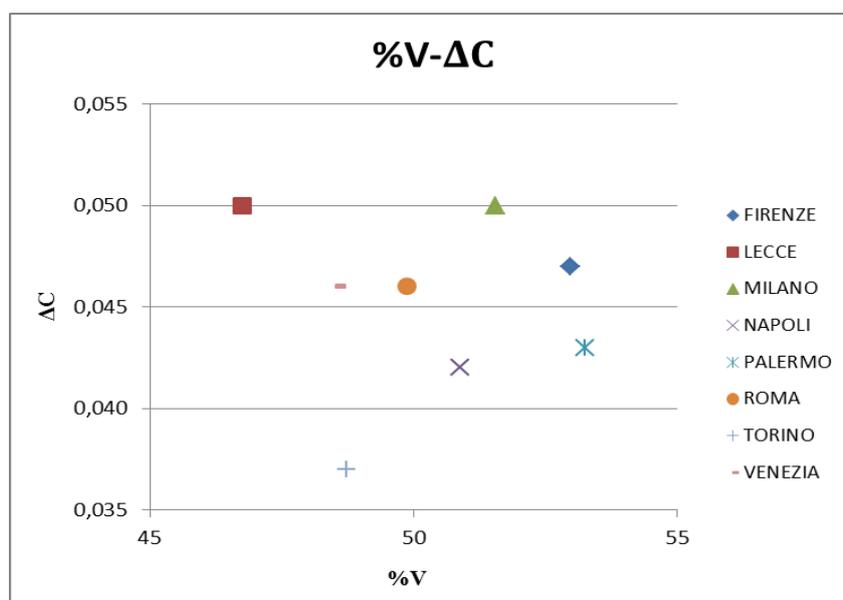


Figura 5.1: risultati metrica %V- $\Delta$ C.

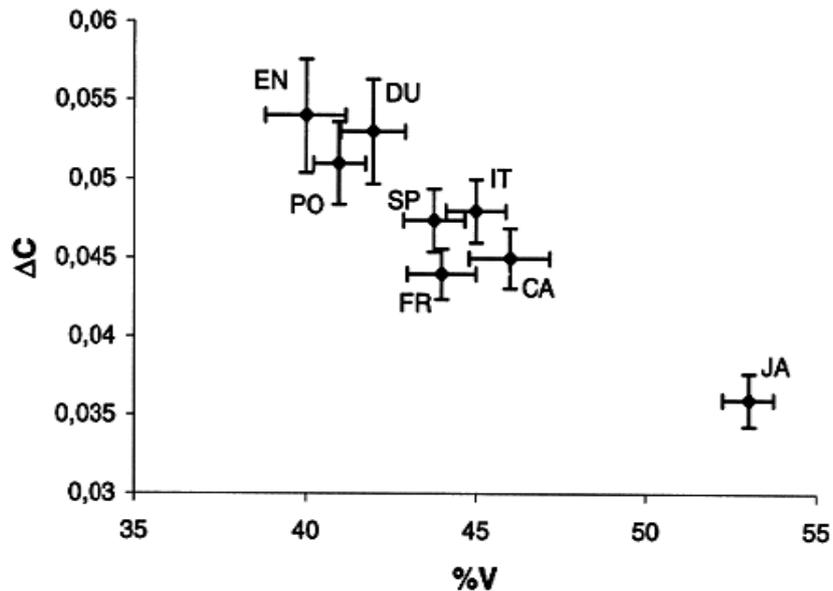


Figura 5.2: risultati metrica %V- $\Delta C$  (da Ramus *et al.* 1999).

I punti della figura 5.1 ritraggono i valori medi dei correlati calcolati dalle metriche per ciascuna varietà, quindi rendono la classificazione delle varietà di italiano regionale.

Entrambe le figure hanno sulle ascisse i valori di %V in secondi, mentre sull'asse delle ordinate i valori di  $\Delta C$  in millisecondi. Questa rappresentazione consente di confrontare in modo omogeneo i risultati di questa ricerca con quelli di Ramus *et al.* (1999). Si nota subito che l'immagine prodotta con i dati delle varietà regionali dell'Italiano non riporta gli assi delle deviazioni standard: ciò è dovuto alla enorme variabilità della durata consonantica e vocalica che caratterizza i segnali vocali del corpus di questa ricerca. L'eventuale inserimento della deviazione standard nella figura 5.1 non avrebbe consentito una discriminazione ritmico-tipologica delle varietà di italiano regionale, facendole collassare tutte al centro del grafico. In ogni caso, si riportano a solo scopo esemplificativo le varianze vocaliche e consonantiche nella tabella 5.1. I risultati della deviazione standard sono espressi in millisecondi e in secondi rispettivamente per la durata consonantica e vocalica così come per il corrispondente valore medio delle due grandezze:

Varietà di italiano regionale	Media Dur. Vocale	Dev.st. Dur. Vocale	Media Dur. Consonante	Dev.st. Dur. Consonante
Firenze	91,50	56,97	0,083	0,049
Lecce	84,39	48,71	0,097	0,051
Milano	91,28	58,30	0,088	0,051
Napoli	77,11	47,20	0,079	0,044
Palermo	86,74	47,20	0,080	0,044
Roma	90,92	56,40	0,093	0,049
Torino	79,17	51,05	0,084	0,040
Venezia	84,93	49,35	0,096	0,048

Tabella 5.1: deviazione standard durata vocali e consonanti, durata media delle vocali e delle consonanti.

La tabella 5.1 riflette una differenza molto importante tra il lavoro di Ramus *et al.* (1999) e questa ricerca. La differenza principale dipende dai corpora che si utilizzano nelle due ricerche (cfr. Arvaniti 2012): questo lavoro si basa su un corpus di parlato semi-spontaneo, dove diverse variabili rendono le durate vocaliche e consonantiche molto instabili (cfr. §1), mentre Ramus *et al.* (1999) usano un corpus di parlato letto, in cui le durate vocaliche e consonantiche hanno una notevole stabilità (cfr. Bertini & Bertinetto 2007). Infatti, dal confronto delle due figure (cfr. figura 5.1 e figura 5.2), nel parlato semi-spontaneo le varietà Firenze, Milano, Napoli e Palermo hanno valori di %V più alti rispetto a quelli del parlato letto. La varietà Torino, invece, nel parlato semi-spontaneo presenta valori di %V e  $\Delta C$  molto bassi. Quindi, questo dato sottolinea una prima importante differenza tra le due ricerche: nel parlato letto (cfr. Ramus *et al.* 1999) l'eloquio è più controllato per cui le durate vocaliche e consonantiche sono stabili, mentre nel parlato semi-spontaneo, che si usa in questa ricerca, la durata delle vocali e delle consonanti è molto variabile.

Le due figure che ritraggono i risultati delle metriche ritmiche presentano anche differenze dal punto di vista classificatorio. Innanzitutto, si specifica che la proiezione %V- $\Delta C$  della metrica di Ramus *et al.* (1999) assume che le lingue *stress-timed* hanno valori alti di  $\Delta C$ , perché presentano una maggiore variabilità di durata dei segmenti consonantici, e valori bassi di %V, cioè una residua occorrenza di elementi vocalici. Il contrario avviene per le lingue *syllable-timed*, cioè valori più bassi di  $\Delta C$  e più alti di %V rispetto alle lingue isoaccentuali. Sulla base di queste informazioni, tale metrica prevede che le lingue tradizionalmente considerate isoaccentuali (inglese, tedesco) si collocano in alto a sinistra, con valori alti di  $\Delta C$  e bassi di %V, mentre le lingue considerate tradizionalmente isosillabiche (italiano, francese, spagnolo) si posizionano più in basso rispetto alle lingue isoaccentuali, con valori più alti di %V e più bassi di  $\Delta C$  rispetto alle lingue *stress-timed*.

Dalla sperimentazione sul corpus della ricerca e dal confronto con i risultati di Ramus *et al.* (1999), la metrica %V- $\Delta C$  restituisce la seguente classificazione ritmico-tipologica: a) la varietà Lecce, che

nella figura 5.1 è in alto a sinistra, sarebbe isoaccentuale; b) le varietà Roma e Venezia apparterrebbero al polo isosillabico; c) la varietà Firenze, che si posizione in alto a destra con valori molto alti di %V e  $\Delta C$ , risulterebbe non classificabile; d) le varietà Milano e Palermo, situate nella parte di centro-destra del grafico, si staccano dalle lingue isosillabiche, perché hanno valori molto alti di %V. Quindi, tali varietà sarebbero inclassificabili; e) la varietà Torino, in basso a sinistra nella figura 5.1, risulterebbe non classificabile, perché ha valori molto bassi di %V e  $\Delta C$ ; f) la varietà Napoli si distanzia dal gruppo *syllable-timed* per valori di  $\Delta C$  più bassi, pertanto sarebbe non classificabile. La metrica di Ramus *et al.* (1999) restituisce chiaramente un panorama classificatorio molto complicato: la metrica classificherebbe tre varietà di italiano regionale nelle classi ritmiche *syllable- stress-timed*, mentre le altre cinque risulterebbero inclassificabili.

Questo risultato delinea uno scenario che pone in discussione la validità classificatoria della metrica di Ramus *et al.* (1999). In accordo con Arvaniti (cfr. Arvaniti 2009), il confronto dei risultati di questa ricerca con quelli di Ramus *et al.* (1999) dimostra che il metodo di elicitazione e di conseguenza il tipo di corpus su cui si sperimentano le metriche possono significativamente colpire i risultati di questi modelli. L'impatto del tipo di parlato sulla metrica si manifesta con la dispersione della maggior parte delle varietà di italiano regionale nella figura 5.1, ma anche dalla distanza tra i valori di %V nelle varietà classificate *syllable-timed*. Quindi, la classificazione ritmico-tipologica può variare in base al metodo di elicitazione e al corpus su cui si sperimentano le metriche: pertanto, sia la mancata classificazione della maggior parte delle varietà di italiano regionale che la distanza nei valori di %V tra le lingue del gruppo isosillabico dipende dal tipo di parlato su cui si sperimenta la metrica. Nel parlato semi-spontaneo, i parametri su cui si fondano i correlati della metrica subiscono l'azione di vari parametri (cfr. §1) che, di conseguenza, disturbano la classificazione. Infatti, i risultati non conformi alle aspettative di questa metrica, cioè la maggior parte delle varietà non si posiziona nelle zone della figura che circoscriverebbero i poli *syllable- stress-timed*. Quindi, il parlato semi-spontaneo mette in crisi l'impianto teorico proposto nella metrica. Il valore del correlato %V è dato dalla somma del numero degli intervalli vocalici che compongono la T-U diviso la durata della T-U stessa (cfr. Ramus *et al.* 1999: 272). Quindi, tale correlato dipende dal numero di segmenti vocalici che compongono la T-U e dalla durata che caratterizza ciascuna T-U. Nelle varietà Firenze, Milano e Palermo questo correlato ha un valore più alto rispetto a tutte le altre varietà di italiano regionale, con la conseguente mancata classificazione ritmico-tipologica delle tre varietà. Il numero di segmenti vocalici e la durata delle T-U, però, dipendono da numerosi fattori. Innanzitutto, i locutori delle varietà di italiano regionale producono all'interno delle T-U un numero diverso di vocali. Nello stesso tempo, la durata delle T-U dipende da vari fattori (cfr. §1). Quindi, un valore alto di %V dipende da fattori elocutivi e non ritmici, che

possono variare da locutore a locutore. Il valore di  $\Delta C$ , invece, descrive la deviazione standard degli intervalli consonantici, che si calcola sulla base della durata dei segmenti consonantici. Nel parlato semi-spontaneo, la durata dei segmenti consonantici subisce l'azione di numerose variabili, soprattutto della VDE (cfr. §1). Quindi, in questa metrica, la minore o maggiore isoaccidentalità delle lingue naturali, cioè il valore alto o basso di  $\Delta C$ , è strettamente connesso alla gestione delle durate consonantiche. Anche la gestione delle durate consonantiche dipende da fattori produttivi, per cui il suo valore può variare da locutore a locutore, ma anche all'interno del locutore stesso (cfr. Arvaniti 2009; Mairano 2010).

Alla luce di questa analisi, i correlati %V e  $\Delta C$  sembrano connessi a fattori produttivi che risultano influenzati da numerose variabili. Tali variabili si manifestano nel parlato semi-spontaneo in tutta la loro interezza, incidendo sulla classificazione ritmico-tipologica delle varietà di italiano regionale. Pertanto, la metrica di Ramus *et al.* (1999) sembra inopportuna per la classificazione ritmico-tipologica delle varietà di italiano che costituiscono il corpus. Inoltre, il fatto che la metrica in esame sia influenzata da fatti produttivi e non propriamente ritmici conferma la critica più pressante mossa in letteratura verso questi modelli (cfr. Arvaniti 2009; Kohler 2009). I correlati %V e  $\Delta C$  riflettono informazioni che riguardano la produzione dell'eloquio: il numero delle vocali e la gestione delle durate possono variare da locutore a locutore, ma anche all'interno del locutore stesso (cfr. Arvaniti 2009; Mairano 2010). In più, la misura della durata consonantica come parametro fondamentale per la classificazione ritmico-tipologica delle lingue realizza uno 'shift' importante, cioè il passaggio dallo studio del ritmo a quello della isocronia. Quindi, questa metrica realizza la classificazione ritmico-tipologica delle varietà di italiano regionale sulla base di informazioni che riguardano la produzione dell'eloquio e la isocronia, ma non pone al centro dell'attenzione i parametri che realizzano il ritmo linguistico inteso come l'alternanza tra sillabe prominenti e non prominenti (cfr. §1).

## 1.2 La metrica di Dellwo (2006)

La incidenza della VDE sulla durata dei segmenti consonantici spinge Dellwo (2006) a realizzare un nuovo correlato per la metrica di Ramus *et al.* (1999), il  $\text{varko}\Delta C$  (cfr. Dellwo 2006). Questo nuovo correlato normalizza le durate consonantiche proprio rispetto alla VDE (cfr. §2). Di seguito, si riporta sia la figura 5.3 che contiene i risultati della sperimentazione della metrica sul corpus della ricerca, sia la figura 5.4 che raffigura i risultati di Dellwo (2006) con la proiezione %V- $\text{varko}\Delta C$ :

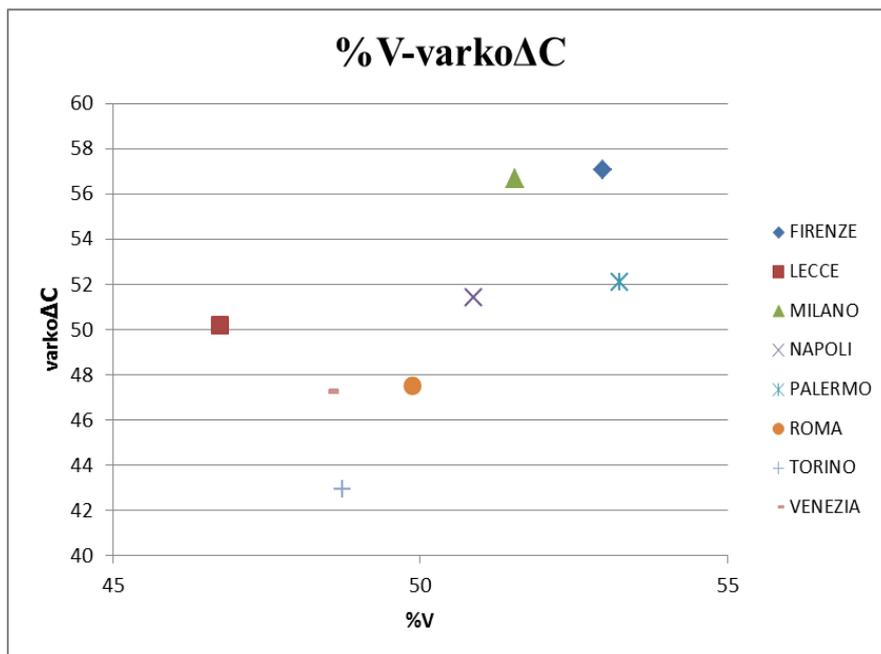


Figura 5.3: risultati metrica % V-varkoΔC.

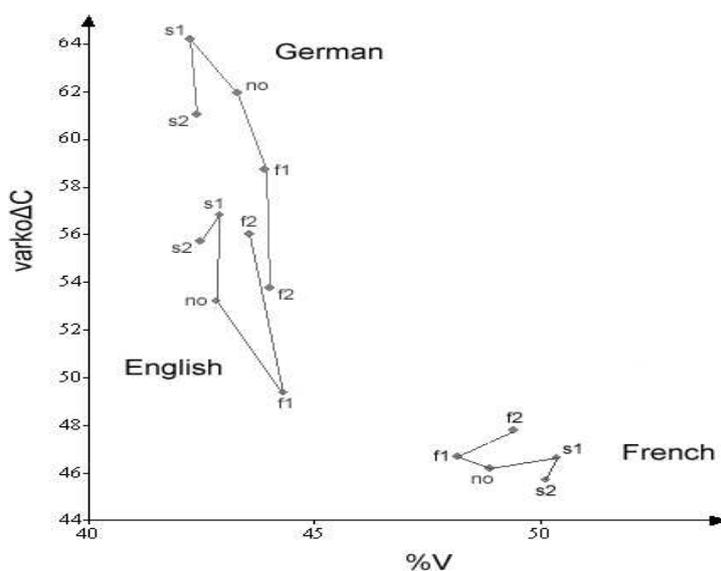


Figura 5.4: risultati metrica % V-varkoΔC (da Dellwo 2006).

I punti della figura 5.3 raffigurano i valori medi dei correlati calcolati dalle metriche per ciascuna varietà, quindi rispecchiano la classificazione delle varietà di italiano regionale. Le due figure hanno rispettivamente sull'asse delle ascisse i valori di %V e sull'asse delle ordinate i valori di varkoΔC espressi in secondi. Questa rappresentazione consente di confrontare in modo omogeneo i risultati di questa ricerca con quelli di Dellwo (2006).

La metrica di Dellwo (2006) ha le stesse aspettative di Ramus *et al.* (cfr. §2), però, dato che le varietà di italiano hanno valori alti di %V e varkoΔC, allora le lingue tradizionalmente considerate isosillabiche (italiano, francese, spagnolo) si collocano generalmente in basso a sinistra, quindi con valori bassi di varkoΔC e di %V, mentre le lingue tradizionalmente considerate

isoaccentuali (inglese, tedesco) si collocano in linea di massima in alto a destra, cioè con valori alti di  $\text{varko}\Delta C$  e di %V (cfr. Mairano 2010).

Dalla sperimentazione sul corpus della ricerca e dal confronto con i risultati di Dellwo (2006), la metrica restituisce la seguente classificazione ritmico-tipologica: a) le varietà Firenze e Milano, che nella figura 5.3 sono in alto a destra, sarebbero isoaccentuali; b) la varietà Torino, che nella figura 5.3 si posiziona in basso a destra, apparterebbe al polo isosillabico; c) le varietà Napoli, Roma e Venezia, che si disperdono più o meno nella zona centrale della figura 5.3, sarebbero non classificabili; d) la varietà Palermo, che ha valori di %V e  $\text{varko}\Delta C$  simili, assumerebbe una posizione inclassificabile; e) la varietà Lecce, con valori medi di  $\text{varko}\Delta C$  e bassi di %V, risulterebbe non classificabile. Quindi, anche la metrica di Dellwo (2006) offre un panorama classificatorio molto complicato: la metrica classificherebbe tre varietà di italiano regionale nelle classi ritmiche *syllable-stress-timed*, mentre le altre cinque risulterebbero inclassificabili. Tutto ciò, produce una classificazione ritmico-tipologica delle varietà difforme dalle attese della metrica, cioè la maggior parte delle varietà non si posiziona nelle zone della figura che delimiterebbero i poli *syllable-stress-timed*. Pertanto, anche la metrica di Dellwo (2006) sembra inopportuna alla classificazione ritmico-tipologica delle varietà di italiano che costituiscono il corpus.

Questo risultato pone in discussione la validità classificatoria della metrica di Dellwo (2006). Anche in questo caso, il confronto dei risultati di questa ricerca con quelli di Dellwo (2006) conferma che la tecnica di elicitazione e di conseguenza il tipo di corpus su cui si sperimentano le metriche determinano i risultati di questo modello. Le Figure 5.3 e 5.4 mostrano, infatti, come la classificazione ritmico-tipologica di tale metrica possa cambiare in base al tipo di metodo di elicitazione e di corpus su cui si calcolano i correlati (cfr. Arvaniti 2012): Dellwo (2006) utilizza un corpus di parlato letto e la metrica restituisce una classificazione molto chiara delle lingue, mentre in questa ricerca si usa un corpus di parlato semi-spontaneo e la metrica restituisce una classificazione confusa delle varietà di italiano regionale. Infatti, come si nota dalle Figure 5.3 e 5.4, nel parlato semi-spontaneo le varietà Firenze, Lecce, Milano, Napoli e Palermo hanno valori di %V e di  $\text{varko}\Delta C$  molto diversi rispetto a quelli del parlato letto. Proprio per questo motivo, buona parte delle varietà non possono essere classificate nelle classi ritmiche tradizionali. Inoltre, sempre dalla figura 5.3, si nota una certa distanza tra le lingue che compongono il gruppo a isocronia accentuale. In questo gruppo ritmico, la varietà regionale Firenze ha valori di %V più alti rispetto alla varietà Milano. Anche in questo caso, il parlato semi-spontaneo influenza i parametri dei correlati che la metrica di Dellwo (2006) utilizza per classificare ritmico-tipologicamente le lingue naturali. Il correlato  $\text{varko}\Delta C$ , lasciando da parte %V che presenta gli stessi problemi della metrica precedente (cfr. §1.1), nasce con l'obiettivo di normalizzare la durata dei segmenti consonantici rispetto alla

VDE. Nonostante la metrica preveda tale normalizzazione, la classificazione della metrica di Dellwo (2006) mostra esclusivamente una diversa collocazione di alcune varietà rispetto alla metrica di Ramus *et al.* (1999), ma il potere classificatorio resta uguale. Quindi, la normalizzazione della durata consonantica sembra non migliorare la capacità classificatoria di questa metrica. Anzi, come dimostra la tabella 5.2, il nuovo correlato non cattura la notevole variabilità consonantica:

<b>Varietà di italiano regionale</b>	<b>Media Dur. Consonante</b>	<b>Dev.st. Dur. Consonante</b>
Firenze	83,78	49,52
Lecce	97,34	51,42
Milano	88,05	51,46
Napoli	79,15	44,73
Palermo	80,24	44,05
Roma	93,98	49,06
Torino	84,98	40,07
Venezia	96,12	48,20

Tabella 5.2: deviazione standard durata consonanti espressa in secondi.

Anche per questa metrica si sottolinea il problema dei correlati che usa per classificare ritmico tipologicamente le lingue naturali. Il correlato %V, come visto (cfr. §1.1), e il correlato  $\text{varko}\Delta C$  esprimono rispettivamente informazioni sulle caratteristiche elocutive e isocrone dei segnali sonori, che possono variare da locutore a locutore, ma anche all'interno del locutore stesso (cfr. Arvaniti 2009; Mairano 2010). Quindi, i due correlati non riferiscono alcuna indicazione sul fenomeno ritmo inteso come alternanza tra sillabe prominenti e non prominenti. Allora, anche questa metrica segue un approccio che pone in serio dubbio la capacità della metrica di descrivere le caratteristiche ritmiche delle lingue naturali.

### **1.3 La metrica di Grabe & Low (2002)**

La metrica di Grabe & Low (2002) è l'ultimo modello di cui si mostrano i risultati della sperimentazione sul corpus di questa ricerca. Di seguito, si mostra sia la figura 5.5 che contiene i risultati della sperimentazione sul corpus della ricerca, sia la figura 5.6 che raffigura i risultati di Grabe & Low (2002) con la proiezione nPVI-rPVI:

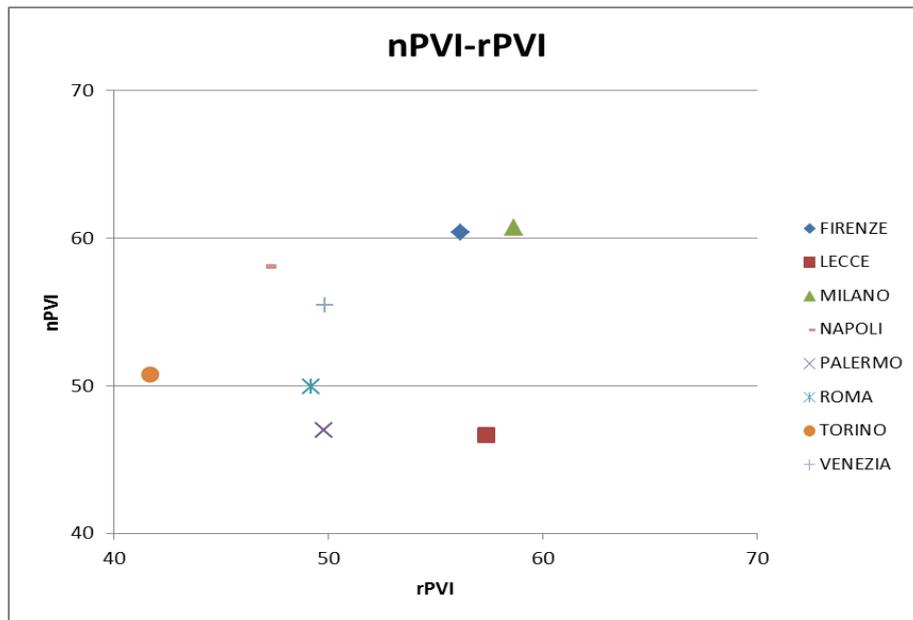


Figura 5.5: risultati metrica nPVI-rPVI.

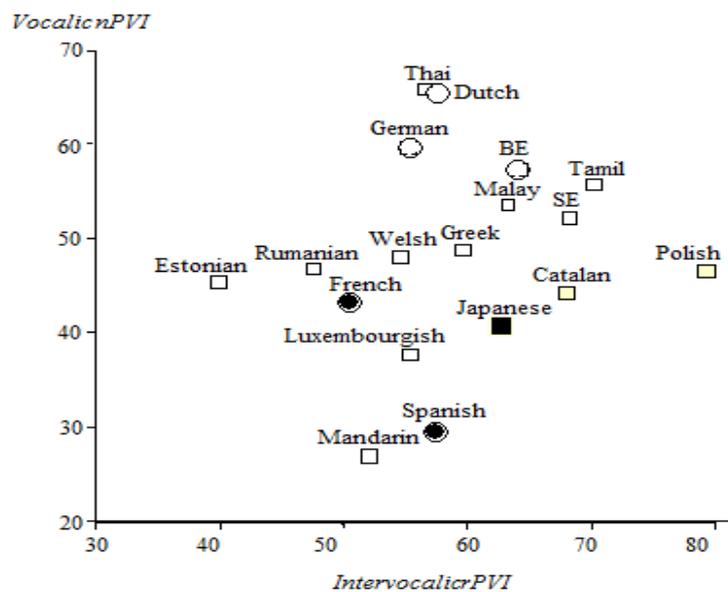


Figura 5.6: risultati metrica nPVI-rPVI (da Grabe & Low 2002).

Nella figura 5.5, i punti rappresentano i valori medi dei correlati calcolati dalle metriche per ciascuna varietà, quindi riflettono la disposizione delle varietà di italiano regionale. Le due figure hanno rispettivamente sull'asse delle ascisse i valori di rPVI e sull'asse delle ordinate i valori di nPVI espressi in secondi. Questa rappresentazione consente di confrontare in modo omogeneo i risultati di questa ricerca con quelli di Grabe & Low (2002).

Tale metrica assume che le lingue tradizionalmente considerate isosillabiche (italiano, francese, spagnolo) si collocano generalmente in basso a sinistra, quindi con valori bassi di rPVI e

di nPVI, mentre le lingue tradizionalmente considerate isoaccentuali (inglese, tedesco) si collocano in linea di massima in alto a destra, quindi con valori alti di rPVI e di nPVI.

Dalla sperimentazione sul corpus della ricerca e dal confronto con i risultati di Grabe & Low (2002), la metrica restituisce la seguente classificazione ritmico-tipologica: a) le varietà Firenze e Milano, che nella figura 5.5 sono in alto a destra, sarebbero isoaccentuali; b) le varietà Palermo e Roma, che nella figura 5.5 si posizionano in basso a destra, apparterrebbero al polo a isocronia sillabica; c) le varietà Napoli e Venezia, che si disperdono nella zona centrale della figura 5.5, sarebbero non classificabili; d) la varietà Torino, con valori molto bassi di rPVI e di nPVI risulterebbe inclassificabile; e) la varietà Lecce, con valori alti di rPVI e bassi nPVI apparirebbe non classificabile. Anche la metrica di Grabe & Low (2002), nonostante ponga una varietà di italiano regionale in più nei due poli ritmici tradizionali rispetto alle metriche precedenti (cfr. §1.1; §1.2), restituisce un panorama classificatorio poco chiaro: la metrica classificherebbe quattro varietà di italiano regionale nelle classi ritmiche *syllable- stress-timed*, mentre le restanti risulterebbero inclassificabili. Tutto ciò, produce una classificazione ritmico-tipologica delle varietà che non rispetta le aspettative della metrica, perché parte delle varietà non si posizionano nelle zone della figura che definirebbero i poli *syllable- stress-timed*. Pertanto, anche la metrica di Grabe & Low (2002) sembra poco opportuna per la classificazione ritmico-tipologica delle varietà di italiano che costituiscono il corpus.

Quindi, anche questa ultima classificazione ritmico-tipologica pone in discussione la validità classificatoria della metrica di Grabe & Low (2002).

La mancata classificazione della metrica di parte delle varietà di italiano regionale dipende, ancora una volta, dal parlato semi-spontaneo che compone il corpus della ricerca (cfr. Arvaniti 2012): rPVI e nPVI non sono altro che il risultato della misura della durata degli intervalli vocalici e consonantici (cfr. §2). Come descritto per le metriche precedenti (cfr. §1.1; §1.2) e nel primo capitolo della ricerca (cfr. §1), nel parlato semi-spontaneo la durata degli elementi vocalici e consonantici dipende da numerose variabili. Infatti, come si nota dalle Figure 5.5 e 5.6, nel parlato semi-spontaneo le varietà Lecce, Napoli, Torino e Venezia hanno valori di nPVI e di rPVI molto diversi rispetto a quelli del parlato letto. Quindi, a differenza della ricerca di Grabe & Low (2002) che sperimentano la propria metrica su un corpus parlato letto, in questa ricerca la metrica si sperimenta su un corpus di parlato semi-spontaneo, dove la variabilità della durata vocalica e consonantica ha le seguenti proporzioni:

<b>Varietà di italiano regionale</b>	<b>Media Dur. Vocale</b>	<b>Dev.st. Dur. Vocale</b>	<b>Media Dur. Consonante</b>	<b>Dev.st. Dur. Consonante</b>
Firenze	91,50	56,97	83,78	49,52
Lecce	84,39	48,71	97,34	51,42
Milano	91,28	58,30	88,05	51,46
Napoli	77,11	47,20	79,15	44,73
Palermo	86,74	47,20	80,24	44,05
Roma	90,92	56,40	93,98	49,06
Torino	79,17	51,05	84,98	40,07
Venezia	84,93	49,35	96,12	48,20

Tabella 5.3: valore medio della durata vocalica e consonantica e rispettive deviazioni standard espresse in secondi.

Come si nota dalla tabella 5.3, la deviazione standard delle vocali e delle consonanti ha un valore molto alto in tutte le varietà di italiano regionale e tale valore indica la notevole variabilità della durata vocalica e consonantica. Nonostante questa metrica preveda la normalizzazione della durata vocalica rispetto alla VDE, la variabile che incide di più sulla durata dei segmenti (cfr. §1), la tabella 5.3 mostra per la durata vocalica valori molto elevati di deviazione standard. Tale deviazione impedisce, ancora una volta, la classificazione ritmico-tipologica delle varietà di italiano.

Proprio l'uso da parte di questa metrica della durata come parametro fondante della classificazione ritmico-tipologica evidenzia due forti critiche. La prima critica riguarda le caratteristiche intrinseche della durata. Tale parametro, come dimostra anche questo studio (cfr. tabella 5.3), è molto instabile (cfr. §1) e la considerazione nelle ricerche sul ritmo della durata come perno classificatore può causare, come succede in questa ricerca, classificazioni ritmico-tipologiche poco chiare. Questo dato sottolinea che la gestione della durata non segue un modello uniforme in tutte le lingue naturali, così da consentire una classificazione ritmico-tipologica in linea con le aspettative della metrica (cfr. Arvaniti 2009), (cfr. Arvaniti 2009; Mairano 2010) e dipendono dal modo in cui le gestiscono i locutori. La seconda critica, come per alcuni correlati delle metriche precedenti (cfr. §1.1; §1.2), investe l'approccio che questo modello usa per descrivere ritmicotipologicamente le lingue naturali. I correlati rPVI e nPVI esprimono informazioni sulle caratteristiche isocrone dei segnali sonori, perché si basano rispettivamente sulla misura della durata delle consonanti e delle vocali, e non trasmettono alcuna indicazione sul fenomeno ritmo inteso come alternanza tra sillabe prominenti e non prominenti. Quindi, anche questa metrica segue un approccio che pone seri dubbi sulla sua funzionalità e sulla sua capacità di descrivere le caratteristiche ritmiche delle lingue naturali.

Allo scopo di rendere i risultati di questa ricerca fruibili in maniera immediata, nel paragrafo successivo si effettua un riepilogo dei risultati.

## 1.4 Riepilogo dei risultati delle metriche ritmiche

Dalla sperimentazione delle metriche sul corpus di questa ricerca emerge la seguente classificazione ritmico-tipologica delle varietà di italiano:

Metriche	Varietà di italiano regionale							
	Firenze	Lecce	Milano	Napoli	Palermo	Roma	Torino	Venezia
%V- $\Delta$ C	<i>stress-timed</i>	non classificata	<i>stress-timed</i>	non classificata	<i>syllable-timed</i>	<i>syllable-timed</i>	non classificata	non classificata
%V- varko $\Delta$ C	<i>stress-timed</i>	non classificata	<i>stress-timed</i>	non classificata	non classificata	non classificata	<i>syllable-timed</i>	non classificata
nPVI-rPVI	<i>stress-timed</i>	non classificata	<i>stress-timed</i>	non classificata	<i>syllable-timed</i>	<i>syllable-timed</i>	non classificata	<i>syllable-timed</i>

Tabella 5.4: riepilogo dei risultati della classificazione ritmico-tipologica resa dalle metriche per le varietà di italiano regionale.

Come si nota dalla tabella 5.4, i risultati delle metriche ritmiche disegnano per le varietà di italiano regionale una classificazione ritmico-tipologica molto frastaglia e poco soddisfacente. Particolare è il risultato classificatorio della varietà Napoli, perché nessuna delle metriche sperimentate sul corpus realizza una classificazione di questa varietà di italiano regionale. Probabilmente, la mancata classificazione della varietà dipende dalla VDE dei locutori, perché la durata media delle vocali e delle consonanti (cfr. tabella 5.4) risultano le più basse tra le varietà di italiano in esame.

Va sottolineato, però, che alcune metriche classificano allo stesso modo le varietà Firenze, Milano e Roma (cfr. tabella 5.4). Quindi, tra queste metriche esiste una correlazione (cfr. Arvaniti 2012). Tale correlazione si manifesta tra la metrica di Ramus *et al.* (1999) e di Dellwo (2006) per la classificazione delle varietà Firenze e Milano e tra la metrica di Ramus *et al.* (1999) e di Grabe & Low (2002) per la classificazione della varietà Roma. Questa correlazione tra le metriche, però, è poco consistente, perché la classificazione unanime delle varietà in esame si manifesta nei pochissimi casi appena citati (cfr. Arvaniti 2012). Questa scarsa consistenza di correlazione si manifesta anche attraverso il rapporto tra metriche e classificazione delle varietà di italiano regionale in esame (cfr. Arvaniti 2012): la correlazione tra la metrica di Ramus *et al.* (1999) e di Dellwo (2006) per le varietà Firenze e Milano si manifesta sull'asse delle vocali (%V), perché questo correlato presenta lo stesso valore (52,98, il valore è espresso in secondi); la correlazione tra la metrica di Ramus *et al.* (1999) e di Grabe & Low (2002) per la varietà Roma, invece, si esprime sia sugli assi delle vocali (%V = 49,88; nPVI = 49,96. I valori sono espressi in secondi) sia sugli assi delle consonanti ( $\Delta$ C = 46,00; rPVI = 49,19. I valori sono espressi in secondi). Questi risultati

dimostrano che le differenze classificatorie tra le varietà di italiano regionali si attenuano solo in pochissime metriche (cfr. Arvaniti 2012).

I risultati della presente ricerca, in linea con quelli di altri lavori (cfr. Arvaniti 2009; Kohler 2009; Arvaniti 2012), dimostrano la poca capacità di questi modelli di classificare ritmico-tipologicamente le lingue naturali nelle classi ritmiche tradizionali. I lavori appena citati e questa ricerca provano che questi modelli hanno una residua capacità di classificare le lingue naturali sulla base di parlato spontaneo e semi-spontaneo. Questa scarsa capacità classificatoria si manifesta perché, come descritto nei paragrafi precedenti di questo capitolo (cfr. §1.1; §1.2; §1.3), nel parlato semi-spontaneo e spontaneo i parametri su cui le metriche fondano la classificazione delle lingue naturali sono fortemente instabili (cfr. §1). Inoltre, nel parlato spontaneo e semi-spontaneo sussistono significative variabilità inter-parlatore variabilità intra-parlatore (cfr. Arvaniti 2009, Mairano 2010, Arvaniti 2012) come andremo a mostrare nel prossimo paragrafo.

## **2. Variabilità inter-parlatore e intra-parlatore nelle metriche ritmiche**

### **2.1 La variabilità inter-parlatore nelle metriche ritmiche**

Come descritto nel paragrafo precedente (cfr. §1.4), i risultati delle metriche sperimentate sul corpus sono affetti dalla variabilità inter- e intra-parlatore (cfr. §2.2), che disturba la classificazione ritmico-tipologica delle varietà di italiano regionale. In *primis* si mostrano i risultati della variabilità inter-parlatore che interessa la metrica di Ramus *et al.* (1999). I punti nella figura 5.7 ritraggono il valore medio dei correlati di ciascun locutore delle varietà in esame calcolati per la proiezione %V- $\Delta C$ <sup>24</sup>:

---

<sup>24</sup> Per l'analisi della variabilità inter-parlatore, non si esamina la varietà Milano, perché il corpus è composto da un solo locutore.

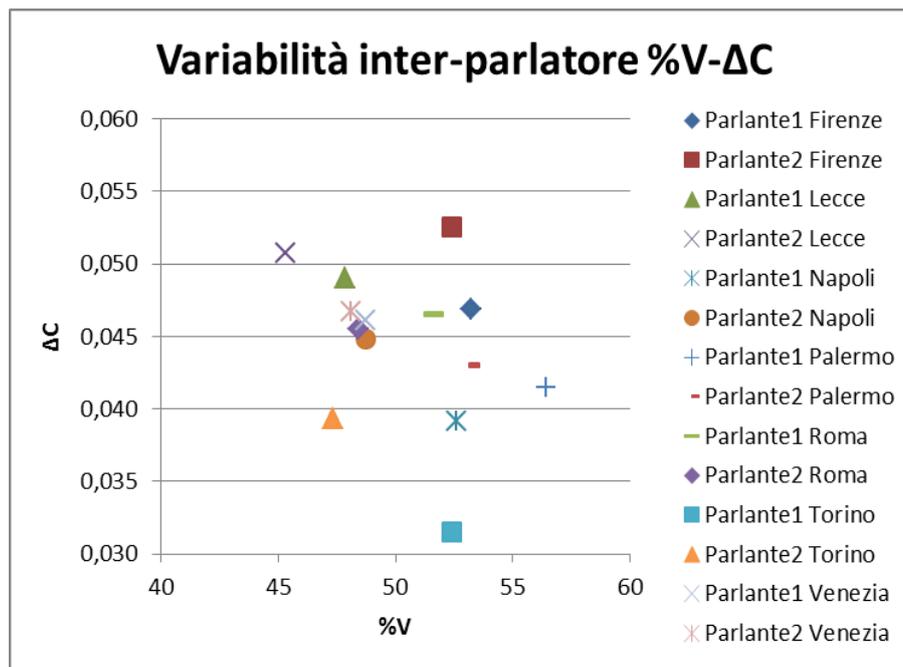


Figura 5.7: variabilità inter-parlatore metrica %V-ΔC.

La figura 5.7 mostra una distribuzione molto confusa dei locutori di ciascuna varietà e sottolinea, in accordo con altre ricerche (cfr. Arvaniti 2009; Kohler 2009; Arvaniti 2012), una notevole variabilità inter-parlatore tra i locutori delle singole varietà. Infatti, dalla figura 5.7 emerge che il Parlante2 Lecce avrebbe caratteristiche del polo *stress-timed*, mentre il Parlante1 Lecce, i Parlanti 1 e 2 Venezia, il Parlante1 Firenze e il Parlante2 Napoli apparterebbero al polo a isocronia sillabica. Il resto dei parlanti, invece, si disperde all'interno del perimetro della figura. La figura accerta che solo i locutori della varietà Venezia hanno valori simili di %V e ΔC, mentre il resto dei parlanti hanno notevoli differenze sia sull'asse %V sia su quello ΔC. Tali differenze nei correlati si manifestano anche tra i locutori che compongono le varietà classificate dalla metrica (cfr. figura 5.1). Allo scopo di catturare quantitativamente questa variabilità che si manifesta sugli assi della figura 5.7, si procede con il calcolo delle distanze relative tra due punti misurate sui correlati %V e ΔC per i parlanti di ciascuna varietà di italiano regionale:

	%V	ΔC
<b>Firenze</b>	0,86	0,006
<b>Lecce</b>	2,57	0,002
<b>Napoli</b>	1,09	0,006
<b>Palermo</b>	3,26	0,002
<b>Roma</b>	3,20	0,001
<b>Torino</b>	5,07	0,008
<b>Venezia</b>	0,67	0,001

Tabella 5.3: variabilità inter-parlatore metrica %V-ΔC misurata attraverso il calcolo delle distanze relative tra due punti.

La tabella 5.3 riporta le distanze relative tra due punti calcolata tra i locutori per i correlati %V e  $\Delta C$ . I risultati della tabella 5.3 confermano una enorme variabilità inter-parlatore sia sull'asse %V tra i locutori delle varietà Lecce, Napoli, Palermo, Roma e Torino sia sull'asse  $\Delta C$  tra i parlanti delle varietà Firenze, Lecce, Napoli e Torino. Quindi, la classificazione ritmico-tipologica delle singole varietà di italiano regionale è affetta da una notevole variabilità che produce due *output*: a) differenze di %V- $\Delta C$  tra i locutori delle varietà Lecce e Roma classificate rispettivamente *stress-* e *syllable-timed* dalla metrica di Ramus *et al.* (1999) (cfr. figura 5.1); b) mancata classificazione delle varietà da parte della metrica di Ramus *et al.* (1999), perché la variabilità inter-parlatore incide in modo efferato su entrambi i correlati oppure su uno dei due correlati della metrica. Quindi, non solo la variabilità inter-parlatore influenza la classificazione delle metriche ritmiche e rende poco soddisfacente il panorama ritmico-tipologico delle varietà di italiano regionale in esame, ma mette a nudo il limite di questi modelli, cioè l'utilizzo di informazioni ritmicamente non pertinenti per descrivere il ritmo delle lingue naturali.

Dopo aver visto la incidenza della variabilità inter-parlatore sulla metrica di Ramus *et al.* (1999), si riportano i risultati di tale variabilità per la metrica di Dellwo (2006):

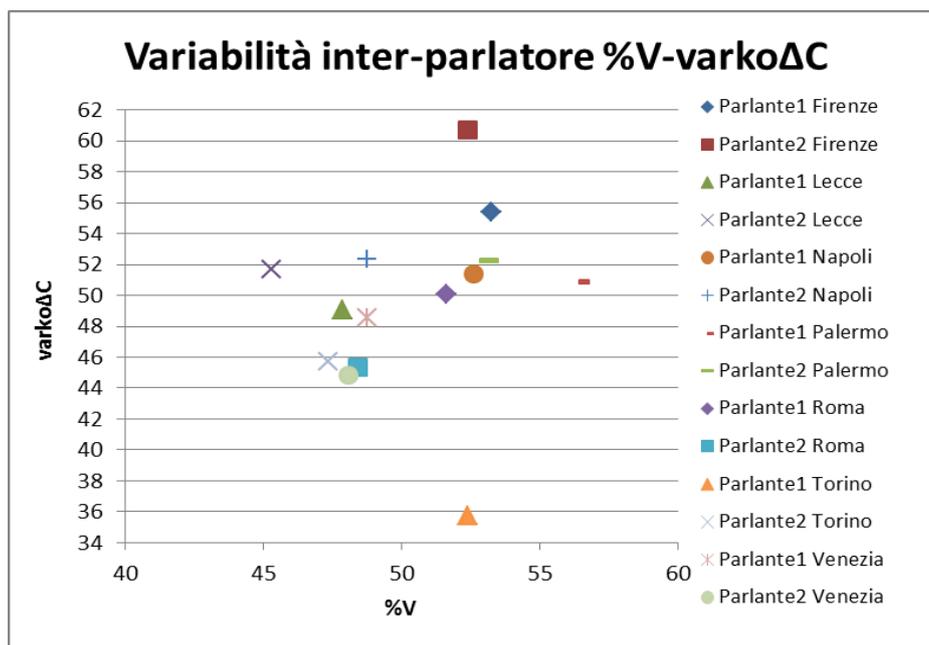


Figura 5.8: variabilità inter-parlatore metrica %V-varko $\Delta C$ .

I punti nella figura 5.8 raffigurano il valore medio dei correlati di ciascun locutore delle varietà in esame calcolati per la proiezione %V-varko $\Delta C$ .

La figura 5.8 rende una situazione molto confusa. Da questa figura, infatti, affiora che i Parlanti1 e 2 Firenze appartenerebbero al polo *stress-timed*, mentre il Parlanti 2 Torino, il Parlante2 Roma e il Parlante2 Venezia avrebbero caratteristiche del polo a isocronia sillabica. Il resto dei

parlanti si posiziona in modo confuso all'interno della figura, quindi è impossibile classificarli ritmico-tipologicamente. Allo scopo di descrivere quantitativamente la variabilità che si manifesta sugli assi della figura 5.8, anche in questo caso si calcolano le distanze relative tra due punti misurate sui correlati %V e  $\text{varko}\Delta\text{C}$  dei parlanti di ciascuna varietà di italiano regionale:

	<b>%V</b>	<b>varko<math>\Delta</math>C</b>
<b>Firenze</b>	0,86	5,31
<b>Lecce</b>	2,57	2,66
<b>Napoli</b>	3,88	0,93
<b>Palermo</b>	3,26	1,40
<b>Roma</b>	3,20	4,72
<b>Torino</b>	5,07	9,98
<b>Venezia</b>	0,67	3,73

Tabella 5.4: variabilità inter-parlatore metrica %V-varko $\Delta$ C misurata attraverso il calcolo delle distanze relative tra due punti.

I risultati della tabella mostrano una enorme variabilità inter-parlatore sia sull'asse %V tra i locutori delle varietà Lecce, Napoli, Palermo, Roma e Torino sia sull'asse  $\text{varko}\Delta\text{C}$  tra i parlanti delle varietà Firenze, Lecce, Roma, Torino e Venezia. Allora, anche in questo caso, la variabilità inter-parlatore produce due risultati: a) differenze di %V-varko $\Delta$ C tra i locutori delle varietà Firenze e Torino classificate rispettivamente *stress-* e *syllable-timed* dalla metrica di Dellwo (2006) (cfr. figura 5.3); b) mancata classificazione della maggior parte delle varietà da parte della metrica di Dellwo (2006). Questo risultato dimostra, come in altri studi (cfr. Arvaniti 2009; Kohler 2009; Arvaniti 2012), ancora una volta che i correlati delle metriche si basano su parametri isocroni e produttivi, che nel parlato semi-spontaneo hanno una notevole variabilità.

Infine, si mostra la influenza della variabilità inter-parlatore sulla metrica di Grabe & Low (1999):

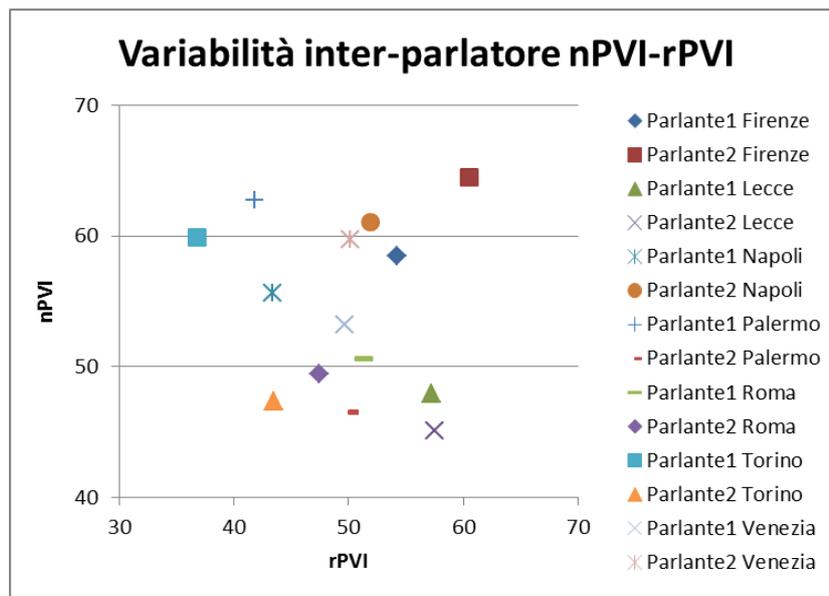


Figura 5.9: variabilità inter-parlatore metrica nPVI-rPVI.

Anche in questa figura, i punti rappresentano il valore medio dei correlati di ciascun locutore delle varietà in esame calcolati per la proiezione nPVI-rPVI.

Gli effetti di una notevole variabilità inter-parlatore tra i locutori delle varietà di italiano regionale si palesa anche nella figura 5.9. Infatti, tale figura mostra che il Parlanti2 Firenze apparterebbe al polo *stress-timed*, mentre il Parlante2 Palermo, il Parlanti1 e 2 Roma, il Parlante2 Torino e il Parlante1 Venezia avrebbero caratteristiche del polo a isocronia sillabica. Il resto dei parlanti assume una posizione che è impossibile classificare ritmico-tipologicamente. Anche per questa metrica, per descrivere quantitativamente la variabilità che si manifesta sugli assi della figura 5.9, si effettua il calcolo delle distanze relative tra due punti sui correlati nPVI e rPVI dei parlanti di ciascuna varietà di italiano regionale:

	nPVI	rPVI
<b>Firenze</b>	5,99	6,31
<b>Lecce</b>	2,82	0,26
<b>Napoli</b>	5,37	8,51
<b>Palermo</b>	16,22	8,23
<b>Roma</b>	1,16	3,90
<b>Torino</b>	12,56	6,69
<b>Venezia</b>	6,56	0,50

Tabella 5.5: variabilità inter-parlatore metrica nPVI-rPVI misurata attraverso attraverso il calcolo delle distanze relative tra due punti.

I risultati della tabella mostrano una enorme variabilità inter-parlatore sull'asse nPVI tra i locutori di tutte le varietà di italiano regionale del corpus e una variabilità inter-parlatore sull'asse

rPVI per la maggior parte delle varietà (cfr. Arvaniti 2012): Firenze, Napoli, Palermo, Roma e Torino. Pertanto, anche la classificazione ritmico-tipologica delle singole varietà di italiano regionale della metrica di Grabe & Low (2002) risulta affetta da una notevole variabilità inter-parlatore che produce gli stessi risultati descritti in precedenza: a) differenze di nPVI-rPVI sia tra i locutori della varietà Firenze classificata *stress-timed* dalla metrica di Dellwo (2006) (cfr. figura 5.5) sia tra i parlanti delle varietà Palermo e Roma classificate *syllable-timed* dalla metrica di Dellwo (2006) (cfr. figura 5.5); b) mancata classificazione della metà delle varietà di italiano regionale da parte della metrica di Grabe & Low (2002).

In conclusione, i risultati di questo paragrafo, in accordo con altri studi (cfr. Arvaniti 2009, Mairano 2010, Arvaniti 2012), provano che la variabilità inter-parlatore incide sulla classificazione ritmico-tipologica delle varietà di italiano. Le figure e le tabelle di questo paragrafo dimostrano che i locutori presentano variabilità inter-parlatore nel calcolo di tutti i correlati delle metriche ritmiche. Questo significa che i locutori delle diverse varietà di italiano regionale gestiscono in maniera differente i parametri isocroni e produttivi su cui si basano i correlati delle metriche ritmiche, disturbando la classificazione finale delle varietà in esame. Pertanto, questi risultati confermano in modo netto che le metriche effettuano lo studio del ritmo delle lingue naturali attraverso parametri ritmicamente non pertinenti, che si discostano dalla definizione linguistica di ritmo.

Le classificazioni confuse che le metriche restituiscono per le varietà di italiano regionale non dipendono solo dalla incidenza della variabilità inter-parlatore, ma anche da differenze che riguardano il parlante stesso, cioè dalla variabilità intra-parlatore. Nel paragrafo successivo, si mostra proprio come la variabilità intra-parlatore disturbi la classificazione delle metriche ritmiche.

## **2.2 La variabilità intra-parlatore nelle metriche ritmiche**

Alcuni studi (cfr. Arvaniti 2009; Mairano 2010) dimostrano che la variabilità intra-parlatore è una delle variabili che mette in discussione la classificazione delle metriche ritmiche. Questo dato emerge anche dalle classificazioni che le metriche restituiscono per le varietà di italiano regionale in esame. A conferma di questo dato, si mostrano i risultati della variabilità intra-parlatore che interessa la metrica di Ramus *et al.* (1999):

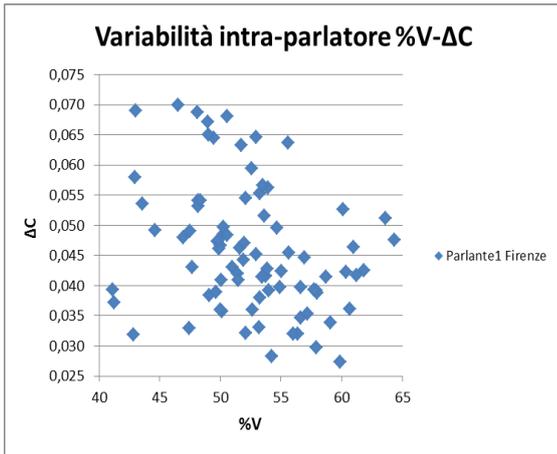


Figura 5.10: variabilità intra-parlatore metrica %V-ΔC.

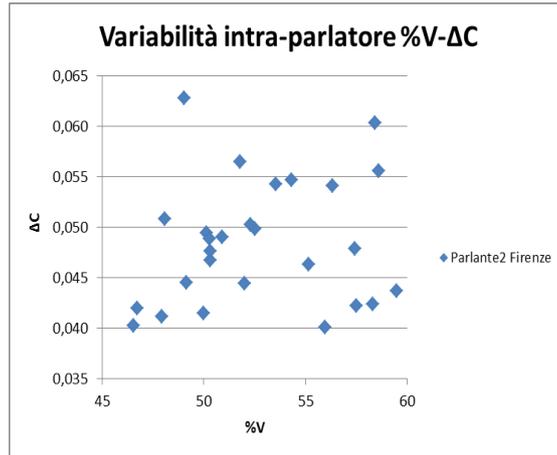


Figura 5.11: variabilità intra-parlatore metrica %V-ΔC.

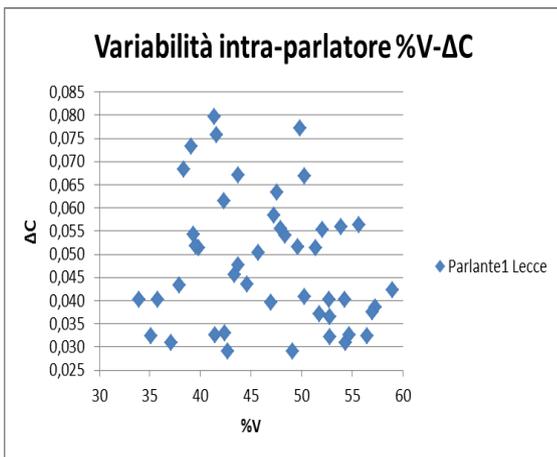


Figura 5.12: variabilità intra-parlatore metrica %V-ΔC.

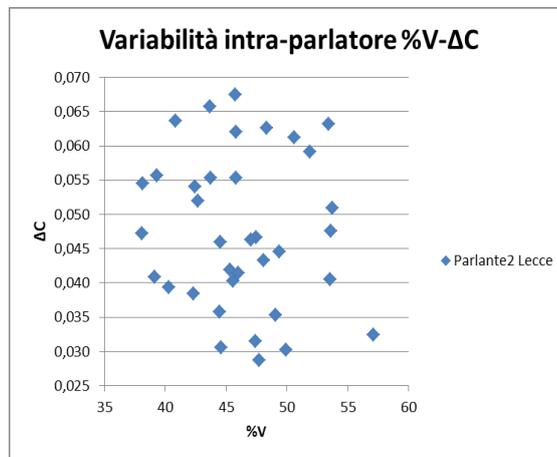


Figura 5.13: variabilità intra-parlatore metrica %V-ΔC.

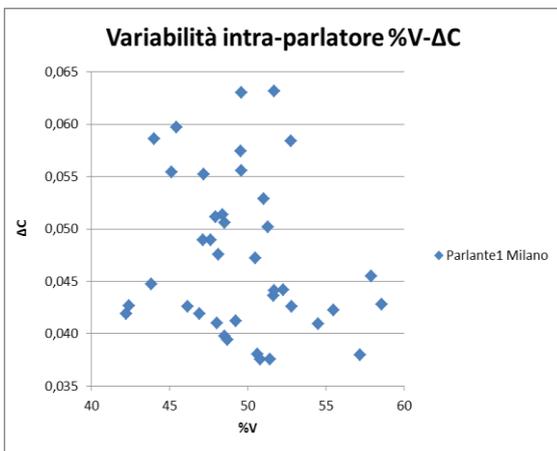


Figura 5.14: variabilità intra-parlatore metrica %V-ΔC.

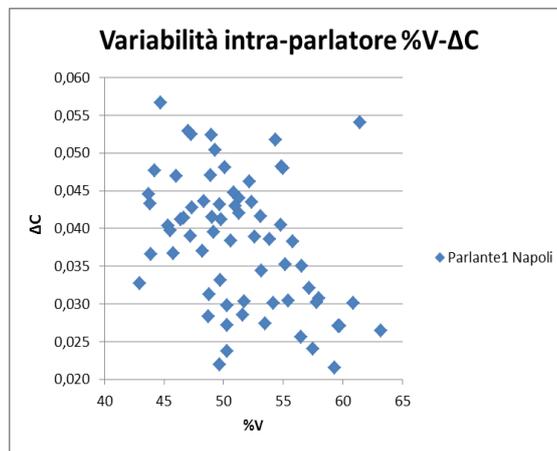


Figura 5.15: variabilità intra-parlatore metrica %V-ΔC.

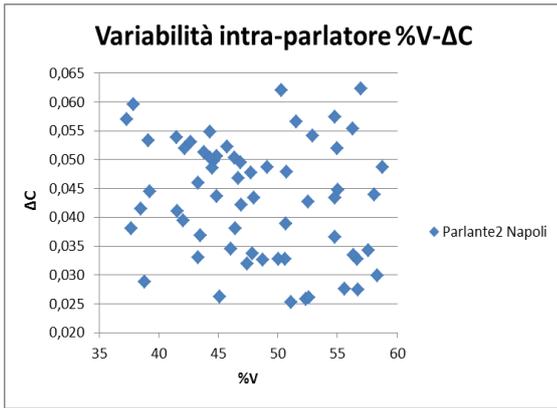


Figura 5.16: variabilità intra-parlatore metrica %V-ΔC.

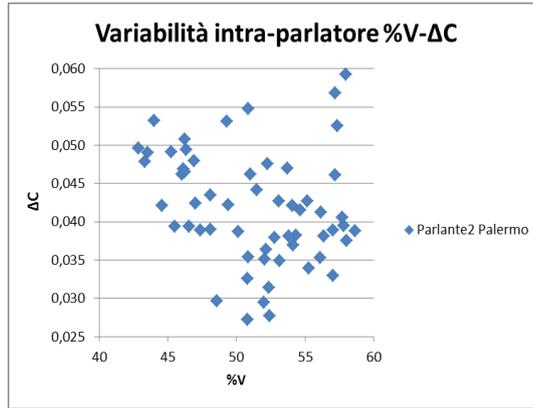


Figura 5.17: variabilità intra-parlatore metrica %V-ΔC.

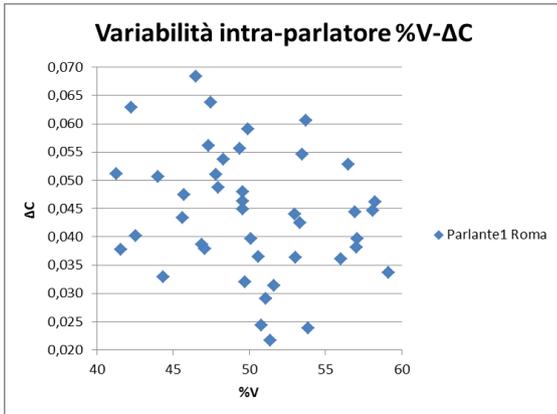


Figura 5.17: variabilità intra-parlatore metrica %V-ΔC.

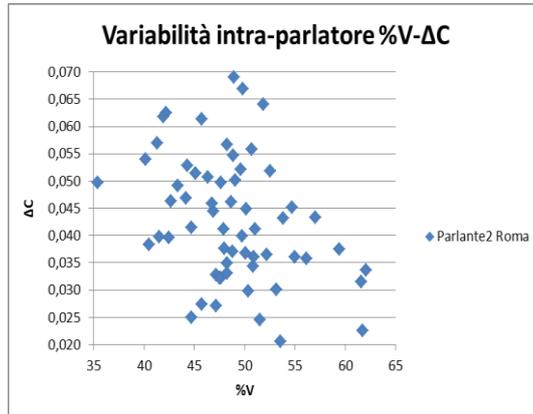


Figura 5.18: variabilità intra-parlatore metrica %V-ΔC.

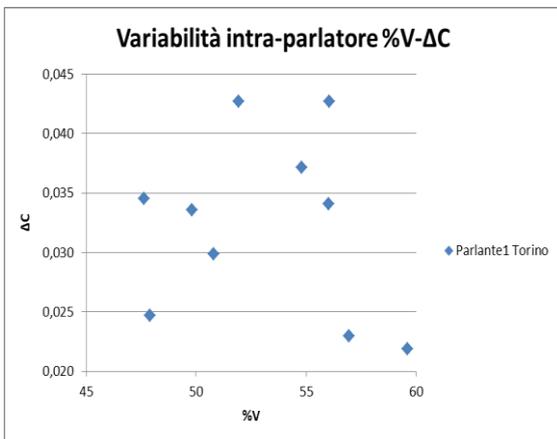


Figura 5.19: variabilità intra-parlatore metrica %V-ΔC.

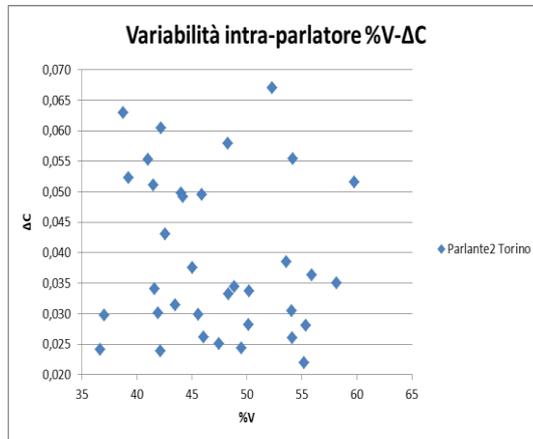


Figura 5.20: variabilità intra-parlatore metrica %V-ΔC.

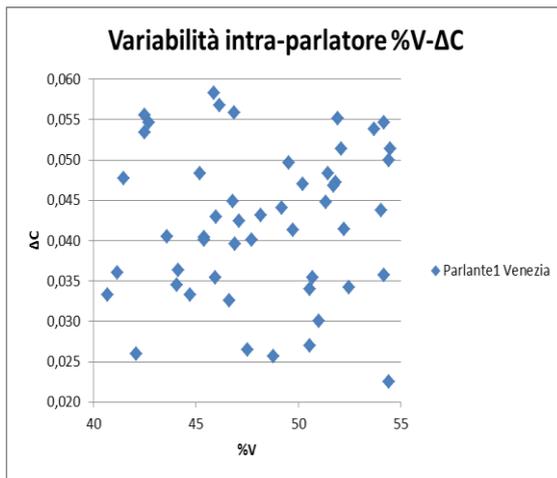


Figura 5.21: variabilità intra-parlatore metrica %V- $\Delta C$ .

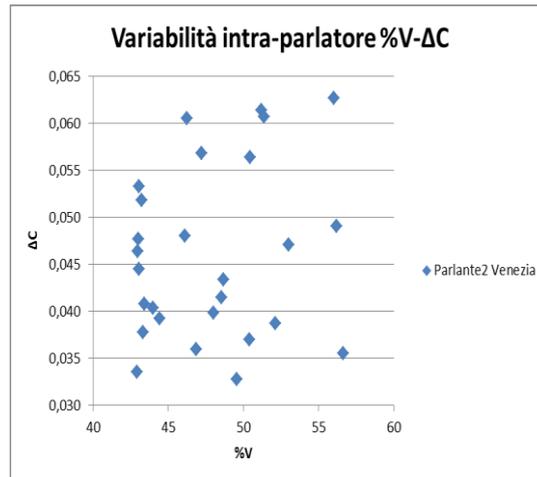


Figura 5.22: variabilità intra-parlatore metrica %V- $\Delta C$ .

Nelle figure, i punti ritraggono il risultato dei correlati delle metriche ritmiche calcolati per ciascuna T-U prodotta dai locutori delle varietà di italiano regionale che compongono il corpus<sup>25</sup>.

Come dimostrano le figure, i punti si distribuiscono in modo molto confuso. Infatti, si ottengono tre risultati: 1) alcuni punti si posizionano in alto a sinistra e rappresentano il polo isocaccentuale, con valori alti di  $\Delta C$  e bassi di %V; 2) una parte dei punti si pone nella zona isosillabica, con valori di  $\Delta C$  più bassi rispetto alle lingue isoaccentuali e valori di %V più alti rispetto alla lingue *stress-timed*; 3) infine, un numero elevato di punti si disperdono all'interno delle figure. Tali punti non rispettano le aspettative della metrica, perché non si posizionano nelle aree che delimitano i poli *syllable- stress timed*, per cui risultano inclassificabili. Questi risultati attestano la notevole variabilità intra-parlatore, perché dimostrano come durante la produzione delle T-U ciascun locutore abbia una propria variabilità che caratterizza sia il valore di  $\Delta C$  che quello di %V. Quindi, i locutori delle diverse varietà di italiano regionale gestiscono in maniera autonoma i parametri isocroni e produttivi su cui si basano i correlati delle metriche ritmiche, disturbando la classificazione delle varietà in esame.

Gli stessi risultati si ottengono con la metrica di Dellwo (2006):

<sup>25</sup> Per l'analisi della variabilità intra-parlatore, non si considerano i risultati del Parlante1 Palermo, perché produce un numero basso di T-U.

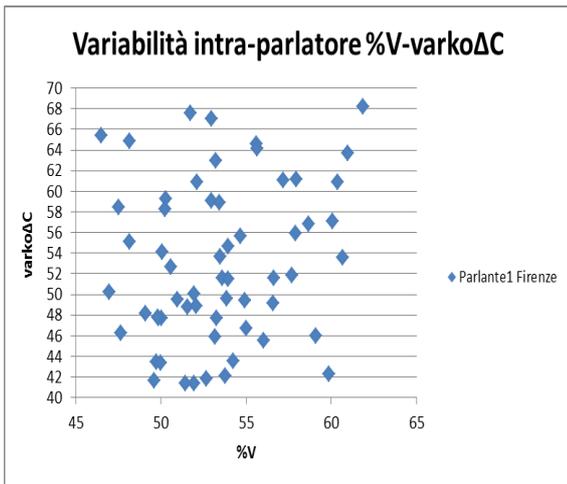


Figura 5.23: variabilità intra-parlatore metrica %V-varkoΔC.

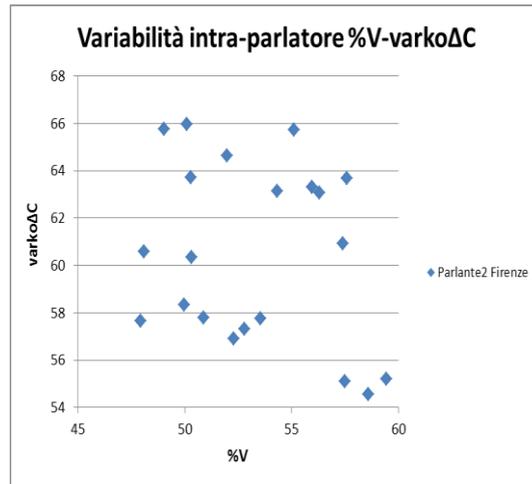


Figura 5.24: variabilità intra-parlatore metrica %V-varkoΔC.

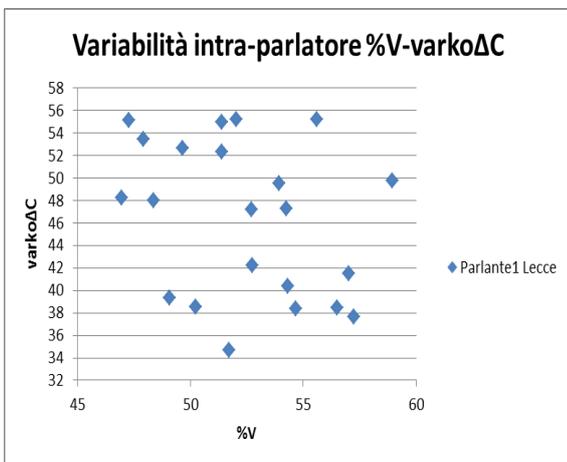


Figura 5.25: variabilità intra-parlatore metrica %V-varkoΔC.

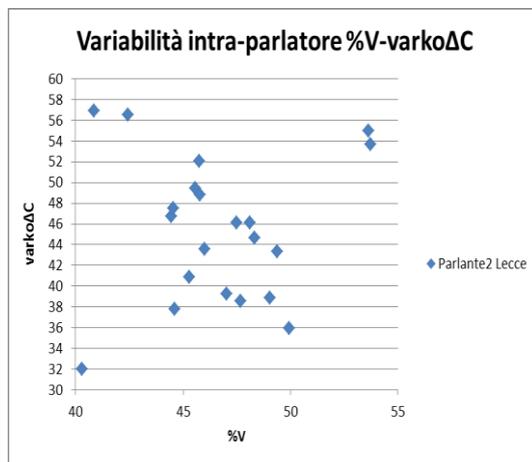


Figura 5.26: variabilità intra-parlatore metrica %V-varkoΔC.

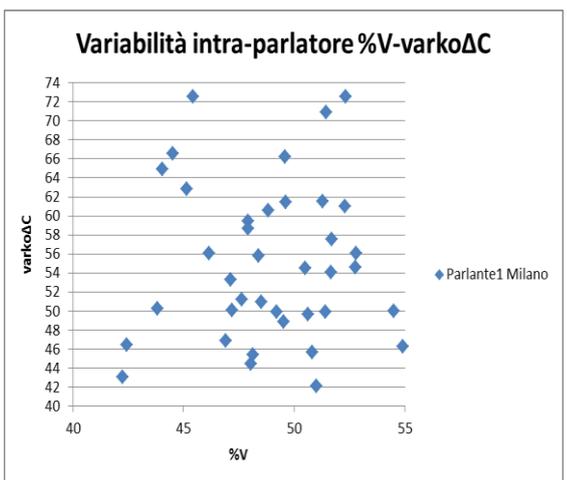


Figura 5.27: variabilità intra-parlatore metrica %V-varkoΔC.

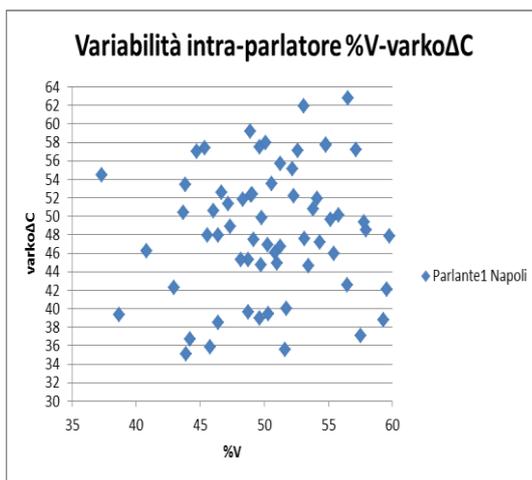


Figura 5.28: variabilità intra-parlatore metrica %V-varkoΔC.

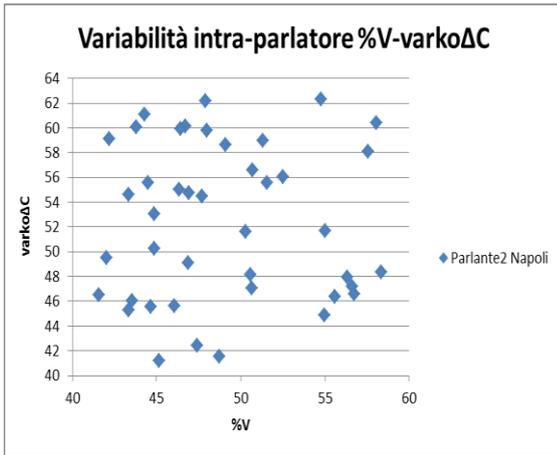


Figura 5.29: variabilità intra-parlatore metrica %V-varkoΔC.

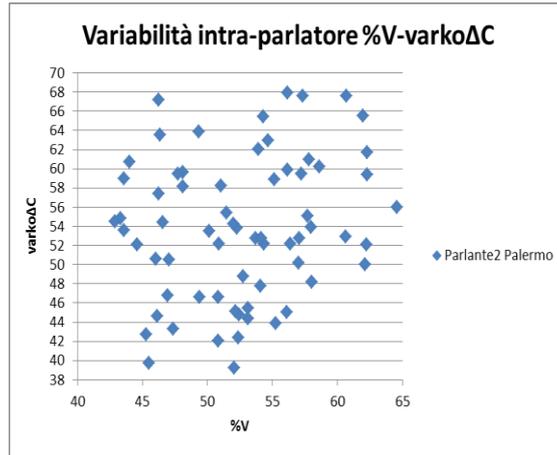


Figura 5.30: variabilità intra-parlatore metrica %V-varkoΔC.

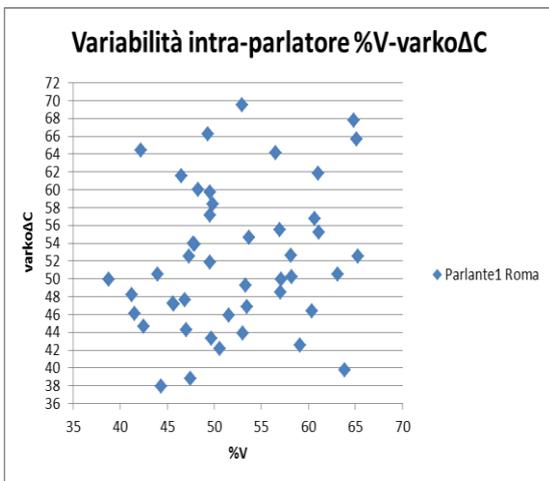


Figura 5.31: variabilità intra-parlatore metrica %V-varkoΔC.

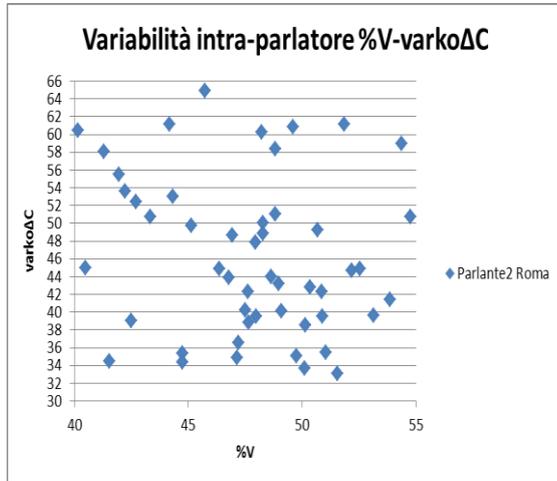


Figura 5.32: variabilità intra-parlatore metrica %V-varkoΔC.

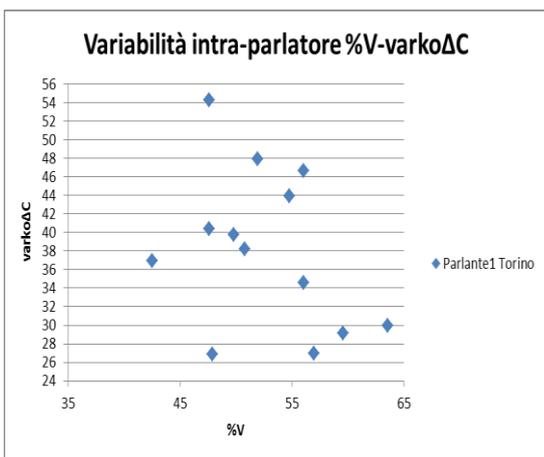


Figura 5.33: variabilità intra-parlatore metrica %V-varkoΔC.

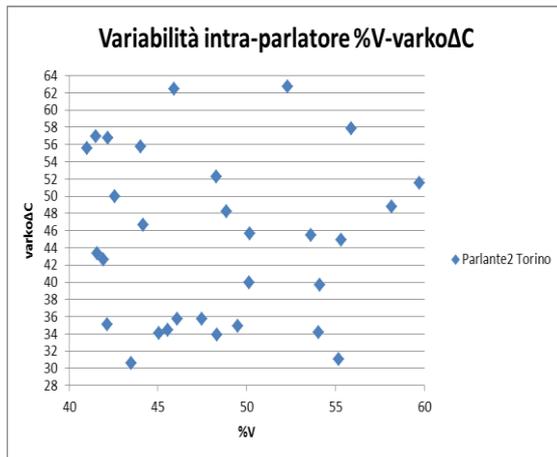


Figura 5.34: variabilità intra-parlatore metrica %V-varkoΔC.

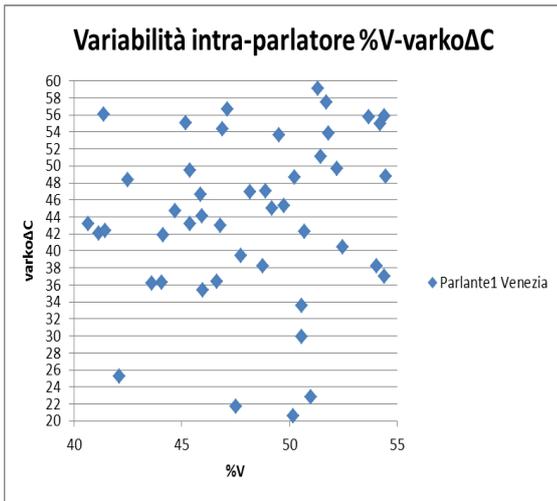


Figura 5.35: variabilità intra-parlatore metrica %V-varkoΔC.

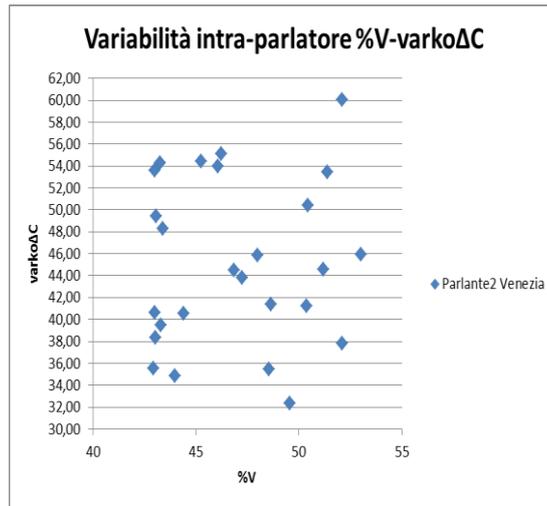


Figura 5.36: variabilità intra-parlatore metrica %V-varkoΔC.

Nelle figure, i punti si distribuiscono in modo confuso: alcuni punti si pongono in alto a destra e rappresentano il polo isocaccentuale, con valori alti di  $\Delta C$  e di %V; parte dei punti si pone in basso a destra nella zona isosillabica, con valori bassi di  $\Delta C$  e di %V; infine, un numero elevato di punti si disperdono all'interno delle figure e risultano inclassificabili. Quindi, anche in questo caso, i risultati accertano una notevole variabilità intra-parlatore, che appura come ciascun locutore abbia una propria variabilità sia nei valori di varko $\Delta C$  e sia in quelli di %V. Pertanto, tale variabilità disturba le classificazioni delle varietà di italiano regionale.

Infine, la incidenza della variabilità intra-parlatore emerge anche dalla metrica di Grabe & Low (2002):

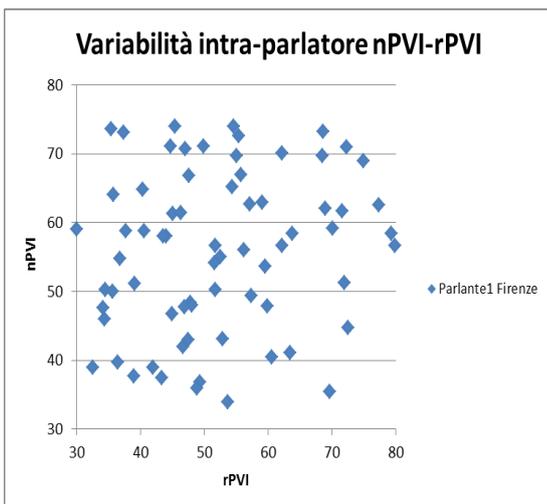


Figura 5.37: variabilità intra-parlatore metrica nPVI-rPVI.

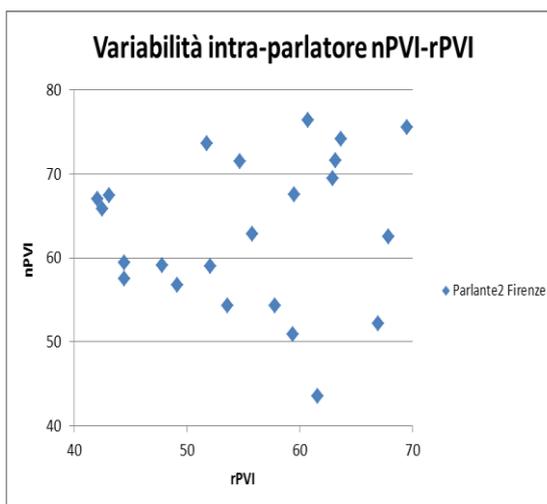


Figura 5.38: variabilità intra-parlatore metrica nPVI-rPVI.

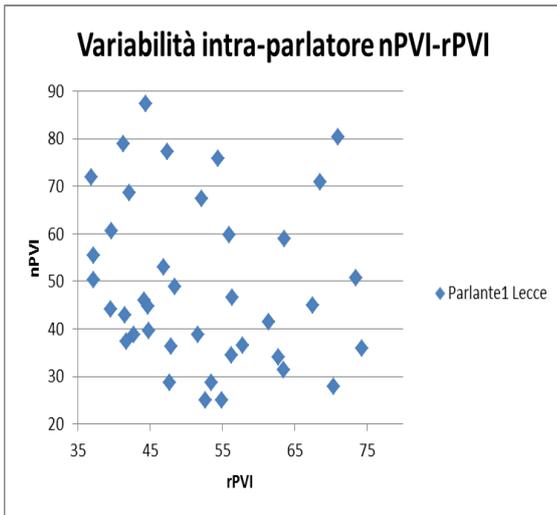


Figura 5.39: variabilità intra-parlatore metrica nPVI-rPVI.

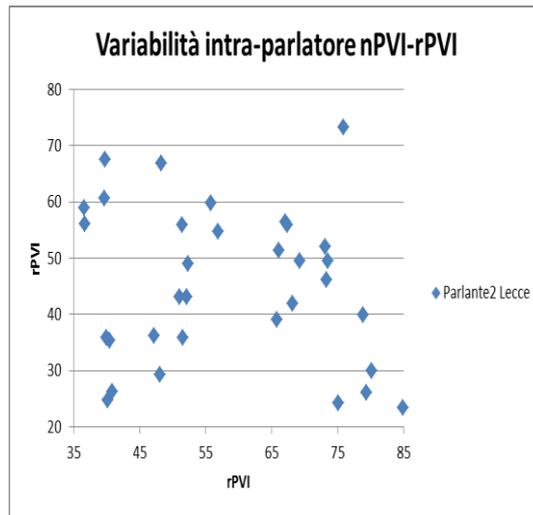


Figura 5.40: variabilità intra-parlatore metrica nPVI-rPVI.

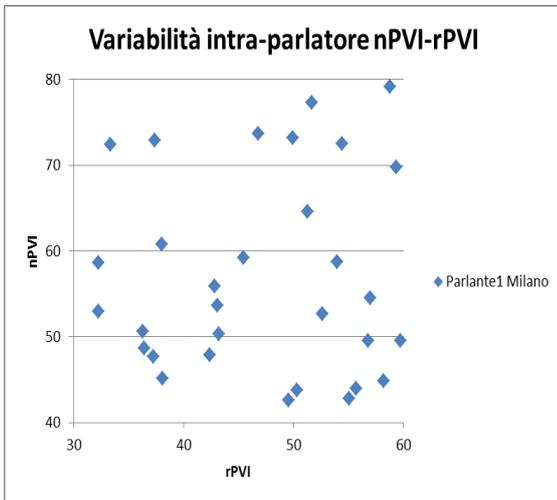


Figura 5.41: variabilità intra-parlatore metrica nPVI-rPVI.

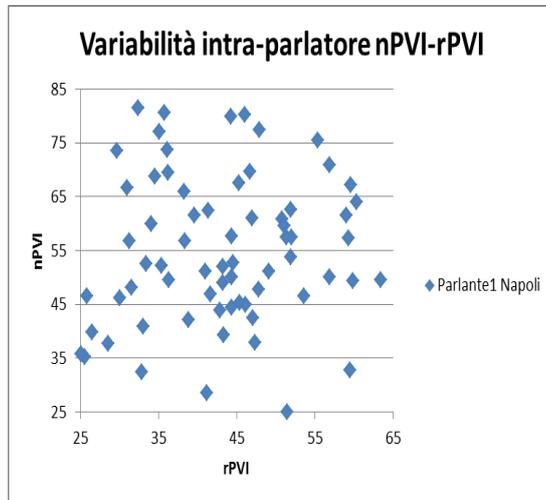


Figura 5.42: variabilità intra-parlatore metrica nPVI-rPVI.

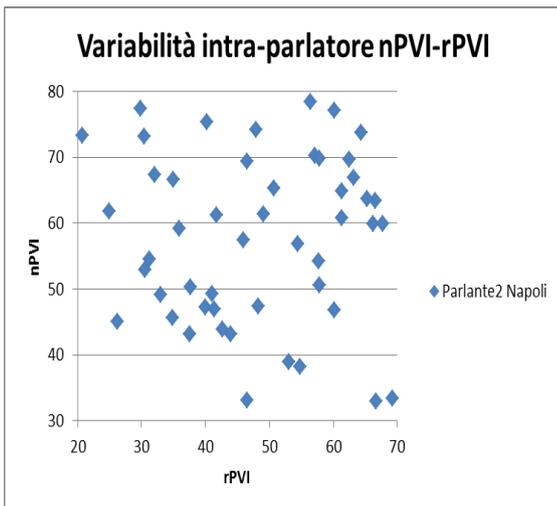


Figura 5.43: variabilità intra-parlatore metrica nPVI-rPVI.

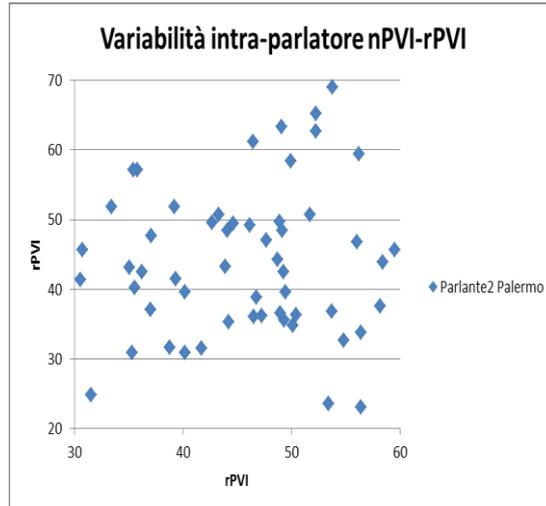


Figura 5.44: variabilità intra-parlatore metrica nPVI-rPVI.

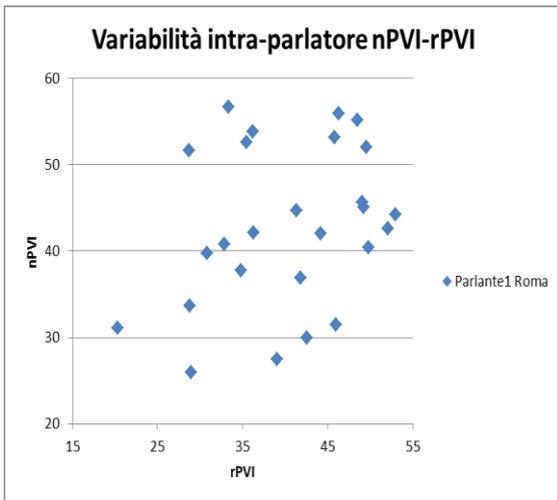


Figura 5.45: variabilità intra-parlatore metrica nPVI-rPVI.

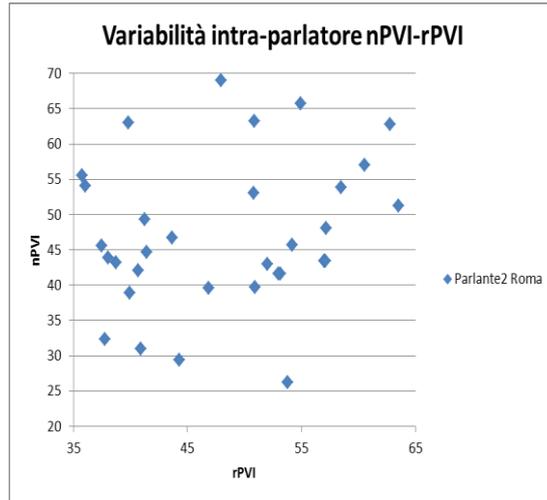


Figura 5.46: variabilità intra-parlatore metrica nPVI-rPVI.

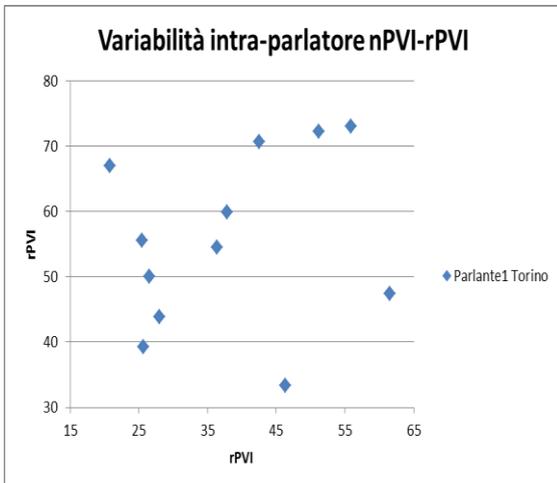


Figura 5.47: variabilità intra-parlatore metrica nPVI-rPVI.

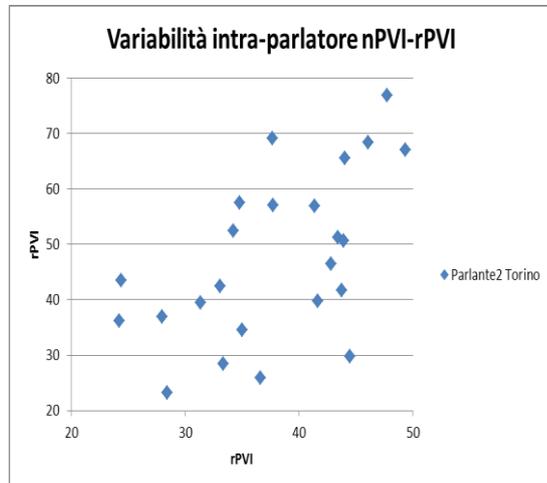


Figura 5.48: variabilità intra-parlatore metrica nPVI-rPVI.

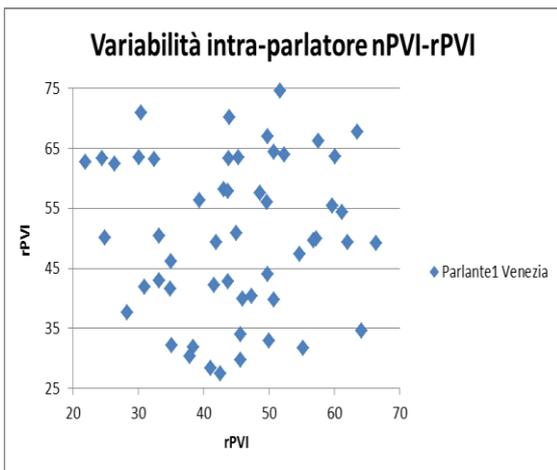


Figura 5.49: variabilità intra-parlatore metrica nPVI-rPVI.

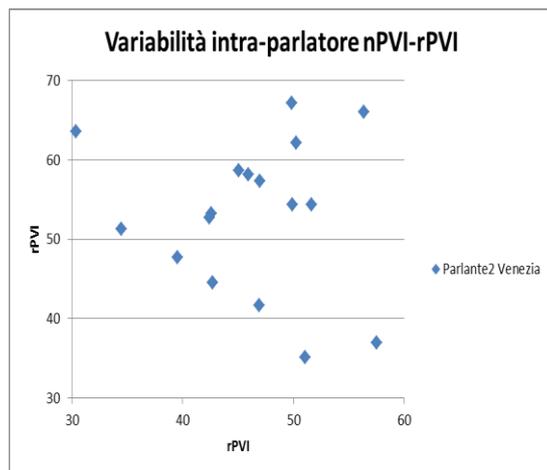


Figura 5.50: variabilità intra-parlatore metrica nPVI-rPVI.

Le figure, dunque, mostrano come i punti si distribuiscano in modo molto confuso: 1) alcuni punti si posizionano in alto a sinistra, cioè nel polo a isocaccentuale, con valori alti di  $\Delta C$  e di  $\%V$ ; 2) una parte dei punti si pone in basso a destra nella zona isosillabica, con valori bassi di  $\Delta C$  e di

%V; 3) infine, un numero elevato di punti si disperdono all'interno delle figure. Tali punti non rispettano le aspettative della metrica, perché non si posizionano nelle aree che delimitano i poli isocronia sillabica e accentuale, per cui risultano inclassificabili. Anche questi risultati dimostrano la notevole variabilità intra-parlatore, perché sottolineano come durante la produzione delle T-U ciascun locutore presenta una propria variabilità che caratterizza sia il valore di nPVI sia quello di rPVI.

I risultati di questo paragrafo dimostrano che nel parlato semi-spontaneo i correlati delle metriche ritmiche sono sottoposti a una notevole variabilità intra-parlatore. Quindi, i locutori delle diverse varietà di italiano regionale gestiscono i parametri su cui si basano i correlati delle metriche ritmiche diversamente nelle varie TU prodotte, laddove era atteso che, in riferimento all'uso di una specifica varietà ritmica, conservassero, almeno in media, un comportamento coerente. Chiaramente, tutto ciò si ripercuote sulle classificazioni ritmico-tipologiche rese per le varietà di italiano regionale, che infatti presentano un panorama molto confuso (cfr. figura 5.1; figura 5.3; figura 5.5). Quindi, anche i risultati della variabilità intra-parlatore confermano due importanti dati: a) le metriche effettuano lo studio del ritmo delle lingue naturali attraverso parametri isocroni e produttivi molto variabili, che nel parlato semi-spontaneo dipendono da numerose variabili; b) le metriche studiano il ritmo delle lingue naturali attraverso parametri ritmicamente non pertinenti, che si discostano dalla definizione linguistica di ritmo (cfr. §1).

## Conclusioni

La sperimentazione delle metriche ritmiche sul corpus della ricerca pone serie questioni agli studiosi di ritmo. I risultati dimostrano, infatti, che il parlato semi-spontaneo disturba la classificazione delle metriche ritmiche, perché tali modelli restituiscono classificazioni ritmico-tipologiche poco chiare delle varietà di italiano regionale (cfr. tabella 5.4). Questo scenario pone seri dubbi sulla utilità e sulla capacità classificatoria di questi modelli e avvia delle argomentazioni critiche nei confronti delle metriche ritmiche, che oggi nelle ricerche sul ritmo rappresentano gli unici modelli adottati per descrivere e classificare ritmico-tipologicamente le lingue naturali.

Dai risultati della ricerca emergono dati molto importanti. Innanzitutto, si assiste a una mancata classificazione da parte delle metriche ritmiche della maggior parte delle varietà di italiano regionale. In accordo con Arvaniti (cfr. Arvaniti 2009), la scarsa capacità di classificazione dipende dal tipo di parlato su cui sono state sperimentate le metriche ritmiche: parlato semi-spontaneo. Questo tipo di parlato, infatti, ha un forte impatto sui parametri che tali modelli utilizzano per la classificazione ritmico-tipologica delle lingue naturali: la durata dei segmenti vocalici e consonantici ( $\Delta C$ ,  $\text{varko}\Delta C$ ,  $nPVI$  e  $rPVI$ ) e il numero dei segmenti vocalici (%V). Come ben assodato, la durata degli elementi vocalici e consonantici nel parlato semi-spontaneo dipende da numerose variabili (cfr. §1), che rendono questo parametro molto instabile. Tale instabilità realizza notevoli variazioni nei correlati  $\Delta C$ ,  $\text{varko}\Delta C$ ,  $nPVI$  e  $rPVI$  e, di conseguenza, nei risultati delle metriche ritmiche. Anche il numero dei segmenti vocalici nel parlato semi-spontaneo dipende da numerosi fattori come il diverso numero di vocali che ricorre nelle T-U, il valore alto della VDE che può produrre sia la cancellazione di segmenti vocalici sia la riduzione fonico-acustica delle vocali, per cui gli incontri vocalici si fondono in un unico nucleo vocalico (sinalefe). Tutto ciò, chiaramente, produce una riduzione del numero delle vocali che ricorrono all'interno dei segnali sonori, influenzando il valore di %V. Pertanto, questi correlati si basano su parametri che nel parlato semi-spontaneo sono molto variabili, per cui un valore alto o basso dei correlati %V e di  $\Delta C$ ,  $\text{varko}\Delta C$ ,  $nPVI$  e  $rPVI$  dipende da fattori elocutivi e isocroni. La scarsa stabilità di questi correlati nel parlato semi-spontaneo si ripercuote sul funzionamento delle metriche ritmiche che, infatti, restituiscono per le varietà di italiano regionale un quadro classificatorio incerto e confuso. Quindi, la sperimentazione delle metriche ritmiche sul corpus della ricerca sottolinea che le metriche ritmiche restituiscono classificazioni ritmico-tipologiche insoddisfacenti.

Il fatto che i correlati delle metriche siano connessi a fattori elocutivi e isocroni conferma la critica più pressante mossa alle metriche ritmiche, ovvero questi modelli restituiscono una cruda misura di fatti produttivi e isocroni (cfr. Arvaniti 2009). Infatti, le metriche analizzano e descrivono il ritmo delle lingue naturali attraverso la durata dei segmenti vocalici e consonantici e il numero dei

segmenti vocalici, che non sono i parametri fondanti della definizione linguistica del fenomeno ritmo (cfr. §1). Pertanto, la isoaccidentalità o isosillabicità delle lingue naturali è esclusivamente connessa a parametri isocroni e produttivi. Questi parametri, come dimostrano i risultati del quinto capitolo (cfr. §5), variano da locutore a locutore (variabilità inter-parlatore) e all'interno dello stesso locutore (variabilità intra-parlatore). In merito alla variabilità inter-parlatore, i risultati della ricerca provano che i locutori presentano notevoli differenze nel calcolo di tutti i correlati delle metriche ritmiche. Questo significa che i locutori delle diverse varietà di italiano regionale gestiscono in maniera differente i parametri isocroni e produttivi su cui si basano i correlati delle metriche ritmiche, disturbando la classificazione finale delle varietà in esame. Gli stessi risultati si hanno per la variabilità intra-parlatore. I locutori delle diverse varietà di italiano regionale gestiscono in modo autonomo e variabile i parametri su cui si basano i correlati delle metriche ritmiche, influenzando i risultati delle metriche.

Infine, nonostante, in questa ricerca, il dominio delle analisi ritmiche sia rappresentato dalla T-U, unità prosodica che cattura ed esprime le relazioni tra ritmo e gli altri livelli della grammatica e in cui i parametri ritmici presentano caratteristiche omogenee (cfr. §3), le metriche rendono per le varietà di italiano regionale classificazioni ritmico-tipologiche insoddisfacenti. Tale insuccesso delle metriche non va imputato alla unità prosodica che in questo lavoro funge da dominio per le analisi ritmiche ma, anche in questo caso, ai parametri che i correlati misurano per effettuare la classificazione ritmico-tipologica delle varietà di italiano regionale. Come sopra specificato, tali parametri non sono ritmicamente pertinenti (cfr. §1), perché sono la espressione di fatti elocutivi e isocroni, ma non ritmici, che nel parlato semi-spontaneo sono soggetti a una variabilità difficile da controllare, indipendentemente dalla unità linguistica che funge da dominio di analisi. Quindi, in conclusione, emerge con chiarezza che lo studio del ritmo delle lingue naturali richiede la realizzazione di nuovi modelli che descrivano il fenomeno ritmo basandosi su parametri ritmicamente pertinenti. Sicuramente, per raggiungere questo scopo è necessario partire dalla definizione linguistica di ritmo (cfr. §1). Tale definizione fornisce gli elementi chiave, cioè i parametri corretti che consentono una reale descrizione del fenomeno ritmo. Pertanto, questi parametri devono costituire il fulcro dei nuovi modelli che si svilupperanno per descrivere il ritmo delle lingue naturali.

## Bibliografia

- Abercrombie, D. (1967), *Elements of general phonetics*, Edinburgh: Edinburgh University.
- Armstrong, L. & Ward, I.C. (1926), *Handbook of English Intonation*, Cambridge: Heffner.
- Aronoff, M. & Oehrle, R. (1984), *Language Sound Structure*, Cambridge: MIT Press.
- Arvaniti, A. (2009), “Rhythm, timing and the timing of rhythm”, *Phonetica*, 66: 46-63.
- Arvaniti, A. (2012), “The usefulness of metrics in the quantification of speech rhythm”, *Journal of Phonetics*, 40: 351-373.
- Arvaniti, A. & Ross, T. (2010), Rhythm classes and speech perception, in *Proceedings of the 5<sup>th</sup> conference of Speech Prosody*, Chicago, U.S.A.: ISCA, pp. 11-14.
- Avesani, C. & Vayra, M. (2005), “Quale informazione codificare per la sintesi dell’intonazione?”, in M. Biffi, L. Salibra & O. Calabrese (eds.), *Italia linguistica. Discorsi di scritto e di parlato*, Siena: Protagon Editori Toscani, pp. 235-248.
- Barry, W. & Koreman, J. (2009), “Do rhythm measures reflect perceived rhythm?”, *Phonetica*, 66: 1-17.
- Beňuš, S & Simko, J. (2012), “Rhythm and tempo in Slovak”, in *Proceedings of the 6th conference on Speech Prosody*, Shanghai, China: ISCA, pp. 502-505.
- Bertinetto, P. M. (1977) “‘Syllabic Blood’, ovvero l’italiano come lingua ad isocronismo sillabico”, *Studi di Grammatica Italiana*, 6: 69-96.
- Bertinetto, P.M. (1981), *Strutture prosodiche dell’italiano*, Firenze: Accademia della Crusca.
- Bertinetto, P.M. (1989), “Reflections on the dichotomy ‘Stress’ vs. ‘Syllable-timing’”, *Revue de Phonétique Appliquée* 91/93: 99-130.

Bertinetto, P.M. & Bertini, C. (2008), "On modeling the rhythm of natural languages", in *Proceedings of the 4th conference on Speech Prosody*, Campinas, Brazil: ISCA, pp. 427-430.

Bertini, C. & Bertinetto, P.M. (2007), "Prospezioni sulla struttura ritmica dell'italiano basate sul corpus semispontaneo AVIP/API", in L. Romito, V. Galatà & R. Lio (eds.), "La fonetica sperimentale: metodo e applicazioni", *Atti del IV Convegno Nazionale dell'Associazione Italiana di Scienze della Voce*, Arcavacata di Rende (Cosenza), Italia, Torriana (RN): EDK Editore, pp. 3-21.

Biffi M., Salibra L., & Calabrese O. (2005), *Italia linguistica. Discorsi di scritto e di parlato*, Siena: Protagon Editori Toscani.

Brown, G. (1990), *Listening to spoken English*, London: Longman.

Caputo, M.R. (1994), "L'intonazione delle domande sì-no in un campione di italiano parlato", in P. Salza (ed.), "Gli aspetti prosodici dell'italiano", *Atti delle IV giornate di Studio del Gruppo di Fonetica Sperimentale*, Roma: Esagrafica, pp. 9-18.

Carletta J., Isard A., Isard S., Kowtko J., Doherty-Sneddon G., & Anderson A. (1996), "*HCRC dialogue structure coding manual. Technical Report HCRC/TR-82*", Human Communication Research Centre: University of Edinburgh.

Channon, R. & Shockey, L. (1987), *A Festschrift for Ilse Lehiste*, Columbus, U.S.A.: The Ohio State University, Department of Linguistics.

Chomsky, N. & Hall, M. (1968), *The sound pattern of English*, New York: Harper e Row.

Cohen P.R., Morgan, J.L. & Pollack, M.E. (1990), *Intentions in Communication*, Cambridge: MIT Press.

Colin, E & Anderson, J. (1986), *Phonology Yearbook*, 3, Cambridge: Cambridge University Press.

Crystal, D. (1969), *Prosodic systems and intonation in English*, London: The Hague.

Cruttenden A. (1986), *Intonation*, Cambridge: Cambridge University Press.

Cumming, R.E. (2010), *A Cross-linguistic Study of Rhythm: French and Swiss German*, Tesi di Dottorato non pubblicata, Cambridge: University of Cambridge.

Dauer, R.M. (1983), "Stress-timing and syllable-timing reanalyzed", *Journal of Phonetics*, 11: 51-62.

De Dominicis, A. (2003), *Fonologia. Modelli e tecniche di rappresentazione*, Roma: Carocci.

Dellwo, V. (2006) "Rhythm and Speech Rate: A Variation Coefficient for deltaC", in P. Karnowski & I. Sziget (eds.), *Language and language-processing*, Frankfurt am Main: Peter Lang, pp. 231-241.

Dellwo, V. (2008), "The role of speech rate in perceiving speech rhythm", in *Proceedings of the 4<sup>th</sup> conference on Speech Prosody*, Campinas, Brazil: ISCA, pp. 375-378.

Dellwo, V. & Wagner, P. (2003), "Relations between language rhythm and speech rate", in D. Recasens, M.J. Solé & J. Romero (eds.), *Proceedings of the 15<sup>th</sup> International Congress of Phonetic Sciences*, Barcelona, Spain, Barcelona: Futurgraphic, pp. 471-474.

Fowler, C.A. (1977), *Timing control in speech production*, Bloomington: Indiana University Linguistics Club.

Fowler, C.A. (1979), "Perceptual centers, speech production and perception", *Perception & Psychophysics*, 25: 375-388.

Fowler, C. A. (1981), "Production and perception of coarticulation among stressed and unstressed vowels", *Journal of Speech and Hearing Research*, 46: 127-139.

Fox, R.A. (1987), "Perceived P-center location in English and Japanese", in R. Channon & L. Shockey (eds.), *A Festschrift for Ilse Lehiste*, Columbus, U.S.A.: The Ohio State University, Department of Linguistics.

Frota S., Vigário, M. & Freitas, M.J. (2005), *Prosodies*, Berlin: Mouton de Gruyter.

- Ghazali, S., Hamdi, R. & Barkat M. (2002), “Speech Rhythm Variation in Arabic Dialects”, in *Proceedings of the 1<sup>st</sup> conference on Speech Prosody*, Aix-en-Provence, France: ISCA, pp. 331-334.
- Gibbon, D. & Gut, U. (2001), “Measuring speech rhythm”, in P. Dalsgaard, B. Lindberg, H. Benner & Z.H. Tan (eds.), *Proceedings of Eurospeech 2001*, Aalborg, Denmark: ISCA, pp. 95-98.
- Giordano, R. & D’Anna, L. (2010), “A comparison of rhythm metrics in different speaking styles and in fifteen regional varieties of Italian”, in *Proceedings of the 5<sup>th</sup> conference on Speech Prosody*, Chicago, U.S.A.: ISCA, pp. 11-14.
- Goldsmith J.A., J. Riggle & A.C.L. Yu (1995), *The Handbook of Phonological Theory*, Oxford: Blackwell.
- Grabe, E. & Low, E.L. (2002), “Durational variability in speech and the rhythm class hypothesis”, *Laboratory Phonology*, 7: 515-546.
- Greenberg S., Carvey H., Hitchcock L. & Chang S. (2003), “Temporal properties of spontaneous speech – A syllable-centric perspective”, *Journal of Phonetics*, 31: 465-485.
- Halliday, M.A.K. (1963), “Intonation in English Grammar”, *Transactions of the Philological Society*, 62: 143-169.
- Halliday, M.A.K. (1967), *Intonation and Grammar in British English*, The Hague: Mouton.
- Halliday, M.A.K. (1976), *System and function in language*, London: Oxford University Press.
- Halliday, M.A.K. (1985), *An Introduction to functional grammar*, London: Edward Arnold.
- Heusinger, K. v. (1999), *Intonation and Information Structure*, Tesi di Dottorato non pubblicata, Konstanz: University of Konstanz.

- Hobbs, Jerry R. (1990), "The Pierrehumbert-Hirschberg Theory of Intonational Meaning Made Simple", in P.R. Cohen, J.L. Morgan & M.E. Pollack (eds.), *Intentions in Communication*, Cambridge: MIT Press, pp. 313-324.
- Hoeqvist, C.E. (1983), "The perceptual centre and rhythm categories", *Language and Speech*, 26: 367-376.
- Jones, D. (1956), *An Outline of English Phonetics*, Cambridge: Heffer.
- Kohler, K.J. (2009), "Rhythm in speech and language. A new research paradigm", *Phonetica*, 66: 29-45.
- Karnowski, P. & Szigeti, I. (2006), *Language and language-processing*, Frankfurt am Main: Peter Lang.
- Kraut R., Brynin M. & S. Kiesler, *Domesticating Information Technologies*, Oxford: University Press.
- Ladd, D.R. (1996), *Intonational Phonology*, Cambridge: Cambridge University Press.
- Ladefoged, P. (1975), *A course in phonetics*, New York: Harcourt Brace Jovanovich.
- Lehiste, I. (1977), "Isochrony reconsidered", *Journal of Phonetics*, 5: 253-263.
- Lieberman, M. & Prince, A. (1977), "On stress and linguistic rhythm", *Linguistic Inquiry*, 8: 249-336.
- Lieberman, M. & Pierrehumbert, J. (1984) "Intonational Invariance under Changes in Pitch Range and Length", in M. Aronoff & R. Oehrle (eds.), *Language Sound Structure*, Cambridge: MIT Press, pp. 157-233.
- Licoppe, C. & Smoreda, Z. (2005), "Rhythms and ties: towards a pragmatics of technologically-mediated sociability", in R. Kraut, M. Brynin & S. Kiesler, *Domesticating Information Technologies*, Oxford: University Press, pp. 911-961.

Lindblom, B. & Rapp, K. (1973), "Some temporal regularities of spoken Swedish", *Papers in Linguistics from the University of Stockholm*, 21: 1-59.

Magno Caldognetto, E. & Vaggies, K. (1991), "Indici di fluenza, tipologia e distribuzione delle sillabe nel parlato spontaneo", in *Atti del XIX Convegno Nazionale dell'Associazione Italiana di Acustica*, Roma: Esagrafica, pp. 423-429.

Mairano, P. (2010), *Rhythm typology: acoustic and perceptive studies*, Tesi di Dottorato non pubblicata, Torino: Università degli Studi di Torino.

Mairano, P. & Romano, A. (2008) "Distances rythmiques entre variétés romanes", in A. Turculen (ed.), *La variation diatopique de l'intonation dans le domaine roumain et roman*, Iasi: University Press A.I. Cuza, pp. 251-272.

Marcus, S.M. (1981), "Acoustic determinants of perceptual center (P-center) location", *Perception & Psychophysics*, 30: 247-256.

Marotta, G. (1985), *Modelli e misure ritmiche. La durata vocalica in italiano*, Bologna: Zanichelli.

Mehler, J., Jusczyk, P. & Lambertz, G. (1988), "A precursor of language acquisition in young infants", *Cognition*, 29: 143-178.

Miller, M. (1984), "On the perception of rhythm", *Journal of Phonetics*, 12: 75-83.

Mok, P.P.K. & Dellwo, V. (2008), "Comparing native and non-native speech rhythm using acoustic rhythmic measures: Cantonese, Beijing Mandarin and English", in *Proceedings of the 4<sup>th</sup> conference on Speech Prosody*, Campinas, Brazil: ISCA, pp. 63-66.

Muljačić, Z. (1972), *Fonologia della lingua italiana*, Bologna: il Mulino.

Nazzi, T., Bertocini, J. & Mehler, J. (1998), "Language discrimination by newborns: toward an understanding of the role of rhythm", *Journal of Experimental Psychology: Human Perception and Performance*, 24: 756-766.

Nazzi, T., Jusczyk, P.W. & Johnson, E.K. (2000), "Language discrimination by English-learning 5-month-olds: effects of rhythm and familiarity", *Journal of Memory and Language*, 43: 1–19.

Nazzi, T. & Ramus, F. (2003), "Perception and acquisition of linguistic rhythm by infants", *Speech Communication*, 41: 233-243.

Nespor, M. & Vogel, I (1986), *Prosodic Phonology*, Dordrecht: Foris.

Nespor, M. (1993), *Fonologia*, Bologna: il Mulino.

Pean V., Williams S. & Eskenazy M. (1993), "The design and recording of ICY, a corpus for the study of intraspeaker variability and the characterisation of speaking styles", in *Proceedings of Eurospeech*, Berlin, Germany: European Speech Communication Association, pp. 627-630.

Pierrehumbert, J. (1980), *The phonology and phonetics of English intonation*, Bloomington: Indiana University Linguistics Club.

Pierrehumbert, J. & Beckman, M. (1986), "Japanese Tone Structure", in E. Colin & J. Anderson (eds.), *Phonology Yearbook*, 3, Cambridge: Cambridge University Press, pp. 255-309.

Pierrehumbert, J. & Hirschberg, J. (1990), "The meaning of intonational contours in the interpretation of discourse", in P.R. Cohen, J.L. Morgan & M.E. Pollack (eds.), *Intentions in Communication*, Cambridge: MIT Press, pp. 271-311.

Ramus, F., Nespor, M. & Mehler, J. (1999), "Correlates of linguistic rhythm in the speech signal", *Cognition*, 73: 265-292.

Ramus, F., Dupoux, E. & Mehler, J. (2003), "The psychological reality of rhythm classes: perceptual studies", in D. Recasens, M.J. Solé & J. Romero (eds.), *Proceedings of the 15th International Congress of Phonetic Sciences*, Barcelona, Spain, Barcelona: Futurgraphic, pp. 337-342.

Romito, L. (2003), *Manuale di fonetica articolatoria, acustica e forense*, Rende: Centro Editoriale e Librario, Università della Calabria.

Romito, L. & Trumper, J. (1993), “Problemi teorici e sperimentali posti dall'isocronia”, in Trumper J. (ed.), *Quaderni del Dipartimento di Linguistica dell'Università della Calabria*, 10: 89-118.

Romito L., Galata' V. & Lio R. (2005), “Fluency Articulation and Speech Rate as new parameters in the Speaker Recognition”, in M.M. Gonzalez Gonzalez, E. Fernandez Rei & B. Gonzalez Rei (eds.), *Actas del III Congreso de Fonética Experimental*, Santiago de Compostela: Xunta de Galicia, pp. 537-549.

Rouas, J.L. & Farinas, J. (2004), “Comparaison de méthodes de caractérisation du rythme des langues”, in *Workshop MIDL (Modélisation pour l'Identification Des Langues)*, Paris, France, pp. 45-50.

Russo, M. (2010), *Prosodic Universals: Comparative Studies in Rhythmic Modeling and Rhythm Typology*, Roma: Aracne.

Russo, M. & Barry, W.J. (2008), “Measuring rhythm. A quantified analysis of Southern Italian Dialects Stress Time Parameters”, *Language Design, Special Issue*, 2: 315-322.

Russo, M & Barry, J. (2010), “Il pairwise Variability Index (PVI e PVIs): valori ritmici per i dialetti italiani e per l'italiano regionale. Implicazioni tipologiche”, in M. Russo (ed.), *Prosodic Universals: Comparative Studies in Rhythmic Modeling and Rhythm Typology*, Roma: Aracne, pp. 185-226.

Savy, R. (2001), “L'interfaccia tra livelli di analisi del parlato: rapporti tra riduzioni segmentali e schemi prosodici”, in F. Albano Leoni (ed.), “Dati Empirici e Teorie Linguistiche”, *Atti del XXXIII Congresso internazionale di studi della Società di Linguistica Italiana*, Roma: Bulzoni editore, pp. 309-328.

Savy, R. & Cutugno, F. (2009), “CLIPS. Diatopic, diamesic and diaphasic variations in spoken Italian”, in M. Mahlberg, V. González-Díaz, C. Smith (eds.), *Proceedings of the 5<sup>th</sup> Corpus Linguistic Conference*, pp. 1-24.

Scott, D.R., Isard, S.D. & Boysson-Bardies, de B. (1985), “Perceptual isochrony in English and in French”, *Journal of Phonetics*, 13: 155-162.

Selkirk, E. (1995), “Sentence prosody: intonation, stress and phrasing”, in J.A. Goldsmith, J. Riggle & A.C.L. Yu (eds.), *The Handbook of Phonological Theory*, Oxford: Blackwell, pp. 550-569.

Selkirk, E. (2005), “Comments on Intonational Phrasing in English”, in S. Frota, M. Vigário & M.J. Freitas (eds.), *Prosodies*, Berlin: Mouton de Gruyter, pp. 11-58.

Shih, S., Grafmiller, J., Futrell, R. & Bresnan, J. (2009), “Rhythm’s role in genitive construction choice in spoken English”, in R. Vogel & R. van de Vijver (eds.), “Rhythm Beyond the Word”, Paris: Mouton.

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

Sorianello P. (2011), “Aspetti pragmatici e prosodici dell’atto esclamativo”, *Studi Linguistici e Filologici Online (SLIFO)*, 9, 287-332. Disponibile da <http://www.humnet.unipi.it/slifo/> [ultimo accesso: 20 novembre 2014].

Sornicola R. (1981), *Sul parlato*, Bologna: Il Mulino.

Speer S.R., Kjelgaard M.M. & Dobroth K.M. (1996), “The influence of prosodic structure on the resolution of temporary syntactic closure ambiguities”, *Journal of Psycholinguistic Research*, 25: 247-268.

Stetson, R. H. (1951), *Motor Phonetics*, Amsterdam: North-Holland.

Taranto M., Bertini C. & Bertinetto, P.M.B. (2011), “Rhythmic Index Elaborator (RIE) come strumento di indagine della struttura ritmica. Un’applicazione al pisano semi-spontaneo vs. letto”, in B. Gili Fivela, A. Stella, L. Garrapa & M. Grimaldi (eds.), *Atti del VII Convegno Nazionale dell’Associazione Italiana di Scienze della Voce*, Roma: Bulzoni, pp. 357-368.

Tepperman, J. & Nava, E. (2011), “Long-distance rhythmic dependencies and their application to automatic language identification”, in *Proceedings of Interspeech 2011*, Florence, Italy: ISCA, pp. 1068-1072.

Tincoff, R., Hauser, M., Tsao, F., Spaepen, G., Ramus, F. & Mehler, J. (2005), “The role of speech rhythm in language discrimination: further tests with a non-human primate”, *Developmental Science*, 8: 26-35.

Trumper J., Romito L., Maddalon M., Mendicino A. & Belluscio G. M. G. (1993),  
Stime manuali: un esperimento, in J. Trumper & L. Romito (eds.), *Atti del Convegno Teoria e Sperimentazione: Parametri, tratti e segmento*, Roma: Esagrafica, pp. 61-79.

Turculen, A. (2008), *La variation diatopique de l'intonation dans le domaine roumain et roman*, Iasi: University Press A.I. Cuza.

Vayra, M (1985), “Effets transsyllabiques de coarticulation voyelle-voyelle en Italien parlé”, in *Actes des IVèmes Journées d'Etudes sur la Parole du Groupe Communication Parlée*, Paris, France, pp. 39-42.

Vayra, M. (1989), “Aspetti della riduzione vocalica in italiano”, in *Atti del XVII Convegno Nazionale Associazione Italiana di Acustica*, Roma: Esagrafica, pp. 421-426.

Vayra, M., Avesani, C. & Fowler, C.A. (1984), Patterns of temporal compression in spoken Italian, in M.P.R. Van den Broecke & A. Cohen (eds.), *Proceedings of the 10<sup>th</sup> International Congress of Phonetic Sciences*, Holland: Foris Publications, pp. 540-546.

Vayra, M., Fowler, C.A. & Avesani, C. (1987), “Word level coarticulation and shortening in Italian and English speech”, *Studi di Grammatica Italiana*, 18: 249-269.

Vigario M., Frota S. & Freitas M.J. (2009), *Phonetics and Phonology: interactions and interrelations*, Amsterdam: John Benjamins.

Voghera, M. (1990), “Gruppi tonali e strutture sintattiche nell'italiano parlato spontaneo”, *The Italianist*, 10: 175-203.

Voghera, M. (1992), *Sintassi e intonazione nell'italiano parlato*, Bologna: il Mulino.

Whalen, D.H., Coopert, A.M. & Fowler, C.A. (1991), "P-Center judgments are generally insensitive to the instructions given", *Phonetica*, 46: 197-203.

White, L., Payne, E., & Mattys, S.L. (2009), "Rhythmic and prosodic contrast in Venetan and Sicilian Italian", in M. Vigario, S. Frota & M.J. Freitas (eds.), *Phonetics and Phonology: Interactions and Interrelations*, Amsterdam: John Benjamins, pp. 137-158.

## Sitografia

Albano Leoni, F. (2006), Il corpus CLIPS: presentazione del profetto [on line]. Disponibile da [www.clips.unina.it](http://www.clips.unina.it) [ultimo accesso lì8 gennaio 2015].

Bonomi A., Falcone M. & Barone A. (2007), Definizione e caratterizzazione di un database vocale ortofonico realizzato da parlanti professionisti in camera anecoica [on line]. Disponibile da [http://www.clips.unina.it/it/documenti/8\\_definizione\\_database\\_ortofonico.pdf](http://www.clips.unina.it/it/documenti/8_definizione_database_ortofonico.pdf) [ultimo accesso il 18 marzo 2015].

Cerrato, L. (2007), Sulle tecniche di elicitazione di dialoghi di parlato semi-spontaneo [on line]. Disponibile da [http://www.clips.unina.it/it/documenti2\\_tecniche\\_di\\_elicitazione\\_dialogica.pdf](http://www.clips.unina.it/it/documenti2_tecniche_di_elicitazione_dialogica.pdf) [ultimo accesso il 10 gennaio 2015].

CLIPS, [www.clips.unina.it](http://www.clips.unina.it) [ultimo accesso il 18 febbraio 2015].

CORRELATORE, <http://www.lfsag.unito.it/correlatore/index.html> [ultimo accesso il 10 ottobre 2014].

Savy, R. (2007), Specifiche per l'etichettatura dei livelli segmentali [on line]. Disponibile da [http://www.clips.unina.it/itdocumenti/12\\_specifiche\\_di\\_etichettatura.pdf](http://www.clips.unina.it/itdocumenti/12_specifiche_di_etichettatura.pdf) [ultimo accesso il 10 gennaio 2015].

Sobrero, A. & Tempesta, I. (2007), Definizione delle caratteristiche generali del corpus: informatori, località [on line]. Disponibile da [http://www.clips.unina.it/it/documenti/1\\_scelta\\_informatori\\_e\\_località.pdf](http://www.clips.unina.it/it/documenti/1_scelta_informatori_e_località.pdf) [ultimo accesso il 15 gennaio 2015].