

AperTO - Archivio Istituzionale Open Access dell'Università di Torino

**Nasali eterosillabiche in italiano e spagnolo: l'energia di banda come parametro discriminante nella classificazione dei nessi NC**

**This is a pre print version of the following article:**

*Original Citation:*

*Availability:*

This version is available <http://hdl.handle.net/2318/50691> since

*Publisher:*

ISTC/EDK

*Terms of use:*

Open Access

Anyone can freely access the full text of works made available as "Open Access". Works made available under a Creative Commons license can be used according to the terms and conditions of said license. Use of all other works requires consent of the right holder (author or publisher) if not exempted from copyright protection by the applicable law.

(Article begins on next page)

Antonio Romano, Franca Mancini, Enrico Zovato - In P. Cosi (a cura di),  
"MISURA DEI PARAMETRI - Aspetti tecnologici ed implicazioni nei modelli linguistici"  
(Atti del I convegno naz. AISV, Padova 2004), Padova, EDK, 2005 (CD-ROM), 101-133.

## **NASALI ETEROSILLABICHE IN ITALIANO E SPAGNOLO: L'ENERGIA DI BANDA COME PARAMETRO DISCRIMINANTE NELLA CLASSIFICAZIONE DEI NESSI NC**

Antonio Romano<sup>1</sup>, Franca Mancini<sup>2</sup>, Enrico Zovato<sup>2</sup>  
<sup>1</sup>Dip. di Scienze del Linguaggio - Università di Torino - <sup>2</sup>Loquendo S.p.A.  
*antonio.romano@unito.it, {franca.mancini, enrico.zovato}@loquendo.com*

### **SOMMARIO**

Questo contributo riassume una parte dei risultati ottenuti nell'ambito di un programma di ricerca più ampio in corso sulle caratteristiche acustiche di foni nasali e nasalizzati presenti nella base di dati del sistema di sintesi vocale *LoquendoTTS*<sup>®</sup>. In particolare, i materiali qui presentati riguardano l'analisi acustica degli incontri consonante nasale + C tenendo conto dei processi di assimilazione presenti in italiano e spagnolo.

Anche se la ricerca non è ancora conclusa, l'opportunità di riassumerne i principali risultati già ottenuti si concilia con il tema del primo convegno dell'*AISV* e con la ricorrenza del centenario della Tesi "De la nasalité en italien" presentata, da uno dei pionieri della fonetica sperimentale italiana, Giulio Panconcelli-Calzia, proprio nel 1904.

Sulla base di elaborazioni statistiche relative a durata e configurazione spettrale, la nostra ricognizione, condotta sulle "voci" della base di dati, ci ha portato a riconoscere delle interessanti caratteristiche che contraddistinguono i vari locutori, nonché alcune lievi differenze tra le due lingue in questione.

Oltre alla presenza di *pattern* formantici blandamente differenziati (cfr. Recasens, 1983; Ohde, 1994), abbiamo osservato delle specifiche sovrapposizioni che si manifestano tra la nasale e l'ostruente seguente. In particolare, ci siamo soffermati sulle transizioni tra nasale e occlusive, fricative e affricate e sul noto fenomeno di epentesi di elementi occlusivi all'interno di nessi nasale + fricativa.

Se sul piano articolatorio la comparsa di questi elementi epentetici potrebbe essere giustificata dall'esistenza di fasi di sovrapposizione di configurazioni e/o gesti articolatori (argomento discusso in Farnetani & Busà, 1994), negli studi di fonetica acustica (cfr. Turchi & Gili Fivela 2004), essa è stata associata alle caratteristiche di affermazione più o meno graduale del rumore di frizione e/o alle sue caratteristiche energetiche e temporali.

Partendo dall'osservazione di un gran numero di incontri (circa 2000 per "voce") di *N* + consonante fricativa presenti nelle basi vocali *Loquendo S.p.A.* relative a due locutori italiani e due spagnoli, abbiamo osservato una diversa disposizione degli *speaker* a produrre in determinati contesti delle transizioni tali da dare l'impressione della presenza di elementi occlusivi spurî.

Dal confronto informale con dei nessi *N* + affricata, in cui l'elemento occlusivo avrebbe dovuto invece essere presente e ben analizzabile (anche perché fonologicamente richiesto), è emerso che il principale elemento di differenziazione è in molti casi la diversa distribuzione spettrale dell'energia dei due foni a contatto e, soprattutto, nelle modalità di ridistribuzione di quest'ultima durante la transizione. Una classificazione è stata quindi tentata, in base al ricorso all'energia di banda come parametro discriminante (ad es. di [ns] vs. [nts] e simili), ed è tuttora in fase di validazione su altre voci. La sperimentazione fin qui condotta si è comunque mostrata utile, se non altro, per l'approfondimento delle configurazioni spettrografiche di questi nessi.

Antonio Romano, Franca Mancini, Enrico Zovato - In P. Cosi (a cura di),  
"MISURA DEI PARAMETRI - Aspetti tecnologici ed implicazioni nei modelli linguistici"  
(Atti del I convegno naz. *AISV*, Padova 2004), Padova, EDK, 2005 (CD-ROM), 101-133.

## 1. PORTATA DEL LAVORO

Numerosi sono i fenomeni linguistici legati alla nasalità. Anche solo una rapida rassegna dei contributi sul tema presenti in letteratura porta a collezionare un numero impressionante di titoli<sup>1</sup>.

Trattando della nasalità, e del suo statuto nei vari sistemi linguistici, occorre distinguere a seconda che essa interessi la presenza di:

- consonanti nasali (che sono, in minore o maggiore misura, comuni a tutte le lingue);
- vocali nasali (in quei sistemi linguistici in cui vocali di articolazione orale si oppongono fonologicamente a serie di vocali di simile articolazione ottenute però con l'aggiunta di risonanze delle cavità nasali)<sup>2</sup>;
- vocali nasalizzate in contesto nasale (in varietà linguistiche in cui, pur essendo assenti vocali nasali con statuto fonologico, si trovano vocali nasalizzate in conseguenza di fenomeni di diffusione di tratti fonetici)<sup>3</sup>;
- vocali e altri foni nasalizzati in contesto non nasale (in quelle varietà in cui si può trovare una nasalizzazione come conseguenza dell'affermarsi di caratteristiche sociofonetiche, idiofonetiche o disfonetiche)<sup>4</sup>.

All'interno degli studi che trattano di questi diversi casi, occorrerebbe poi distinguere in prima battuta i diversi approcci di analisi in cui si incorre che, come accade in molti casi, vanno distinti per motivazione, finalità e metodologia. Non essendo questo l'obiettivo principale del presente contributo, ci limitiamo a riassumere in poche righe l'ambito in cui esso si inquadra, inserendolo tra i lavori che si concentrano sulle consonanti nasali e sulla loro variabilità distribuzionale, con obiettivi di descrizione fonetica e acustica di questi suoni in due lingue romanze: italiano e spagnolo peninsulare<sup>5</sup>.

Il lavoro qui riassunto, parte da una rassegna dei fenomeni fonetici prevedibili (descritti dalla letteratura specialistica) per le varietà in osservazione per poi orientarsi verso una caratterizzazione acustica dei dati a disposizione osservati, assoggettata alle necessità pratiche di un'automatizzazione, senza alcuna pretesa di esaustività. Allo stato attuale della ricerca, ci siamo interessati esclusivamente degli aspetti acustici senza entrare nel merito della verifica di aspetti articolatori, a cui sarà invece sicuramente riservata molta più attenzione negli sviluppi futuri di questi lavori.

---

<sup>1</sup> Per una rassegna di contributi riguardanti la nasalità in diverse lingue anche in prospettiva evolutiva si veda Hajek (1997). Un contributo di fonetica storica, per le nasali romanze, è in Straka (1955).

<sup>2</sup> Vocali di questo tipo caratterizzano ad es. il francese, il portoghese e il polacco. Per una rassegna sulle lingue del mondo, si veda Schourup (1973).

<sup>3</sup> Simili casi sono presenti, in misura minore o maggiore, in tutte le lingue (v. Kurowski & Blumstein, 1984; Repp & Svastikula, 1986; Pettorino, 1988; Solé, 1992). La totale mancanza di nasalizzazione delle vocali in contesti nasali può infatti addirittura condizionare una percezione non nasale della stessa consonante nasale (per l'italiano v. Pettorino, 1988; Giannini *et alii*, 1990; Maturi, 1991; Albano Leoni & Maturi, 1995).

<sup>4</sup> Come sottolineavamo accennando a Panconcelli-Calzia (1904), la nasalità come fenomeno diffuso è stato l'oggetto di numerosi studi sin dai primordi della fonetica sperimentale in Italia. Rappresentativa delle difficoltà di stabilire una separazione netta tra la rilevazione sperimentale di caratteristiche accidentali e di tratti sistematici, nella fonetica all'inizio del XX secolo, è l'attribuzione da parte di F.M. Josselyn di tratti di nasalità a elementi del diasistema sonoro italiano che solo accidentalmente possono esibire simili proprietà, in un numero limitato di soggetti e/o di varietà (v. Romano, 2002).

<sup>5</sup> Diversamente da Busà (2003) che si concentra su dati molto precisi (articolatori e acustici) riguardanti soprattutto la nasalizzazione vocalica in contesti VNCV delle nostre osservazioni sono soprattutto le caratteristiche energetiche e temporali di *N* (desumibili da dati acustici) e le fasi di transizione verso *C* (cfr. Farnetani & Busà, 1994).

Antonio Romano, Franca Mancini, Enrico Zovato - In P. Cosi (a cura di),  
“MISURA DEI PARAMETRI - Aspetti tecnologici ed implicazioni nei modelli linguistici”  
(Atti del I convegno naz. *AISV*, Padova 2004), Padova, EDK, 2005 (CD-ROM), 101-133.

Le esigenze iniziali che hanno determinato la ricerca sono legate a difficoltà procedurali comparse nelle fasi di verifica delle segmentazioni dei segnali delle basi di dati etichettate usate dal sistema di sintesi vocale *LoquendoTTS*<sup>®</sup> (cfr. Balestri *et alii*, 1999). Tali difficoltà si possono riassumere nei seguenti punti:

- classificazione dei tassofoni;
- segmentazione del *continuum* acustico;
- classificazione dei nessi.

In particolare, si è visto che le descrizioni teoriche dei sistemi fonetici delle due lingue in questione, utilizzate dal sistema *LoquendoTTS*<sup>®</sup>, possono presentare alcune incoerenze e criticità soprattutto in chiave contrastiva. Una volta superato il problema descrittivo e teorico restano però quello procedurale e classificatorio in fase di etichettatura.

## 2. ASPETTI TEORICI

### 2.1 *Nasali intervocaliche*

Negli studi descrittivi di carattere più generale, dominano di solito gli approfondimenti sui fonemi nasali in posizione intervocalica. È in quest'ambito che troviamo molte somiglianze (ma anche alcune significative differenze) tra italiano e spagnolo.

Mentre per [m] e [n] intervocalici si ha una notevole analogia di comportamenti nelle due lingue (se si escludono le differenze legate ai diversi vincoli che lo statuto della geminazione consonantica induce nel sistema fonologico italiano, che ha quattro fonemi contro i due dello spagnolo), è nell'articolazione palatale di [ɲ] che si possono trovare delle interessanti divergenze. Innanzitutto in italiano (escluse le varietà settentrionali) il fonema si presenta sempre con realizzazioni intrinsecamente lunghe (mentre in spagnolo questo tratto è soggetto a una maggiore variabilità accidentale)<sup>6</sup>. In secondo luogo le sue realizzazioni in italiano presentano di solito un rilascio prolungato (che non era sfuggito a Jones, 1950, e che assume le caratteristiche di una lunga fase di transizione sulla vocale seguente)<sup>7</sup>.

---

<sup>6</sup> Secondo diversi autori, la /ɲ/ dello spagnolo, ortograficamente notata con <ñ>, è più breve rispetto a quella dell'italiano standard (cfr. Mioni, 1973; Canepari, 1979). In italiano la possibilità di esistenza per /ɲ/, con realizzazioni non lunghe, non opponibili a quelle lunghe che sono invece diffusissime in posizione intervocalica (e che fanno preferire la rappresentazione del fonema come /ɲɲ/), sono limitate al contesto iniziale (*gnocco*, *gnomo*, *gnu* etc.) e ai nessi interni. Di questi ultimi sono esempi le parole *corgnolo*, *fargna* e *lorgnette*, rispettivamente due varianti fonetiche e grafiche (settentrionali) di parole diffuse altrove con il nesso /ɲj/ (sintomo appunto della neutralizzazione dell'opposizione /ɲj/ ~ /j/ tipica nelle varietà del Nord della penisola) e un prestito dal francese. Accettabilità e diffusione dello stesso nesso /ɲj/ al Nord sono visibili anche in toponimi e antroponimi piemontesi: *Cuornè*, *Quargnento*, *Vergnano*.

<sup>7</sup> Approfittiamo per ricordare che, sul piano fonetico-articolatorio, le consonanti nasali sono, nell'ambito della loro articolazione orale, delle vere e proprie occlusive, caratterizzate da una fase di tenuta (durante la quale le vibrazioni delle corde vocali mettono in risonanza le cavità nasali producendo un suono che viene irradiato dalle narici) e una fase di rilascio che, in funzione della coordinazione con i movimenti del velo palatino -chiamato a escludere simultaneamente le vie nasali se il suono seguente programmato ha un'articolazione esclusivamente orale-, dovrebbe convogliare improvvisamente l'aria in vibrazione verso l'esterno, attraverso la bocca.

Antonio Romano, Franca Mancini, Enrico Zovato - In P. Cosi (a cura di),  
 “MISURA DEI PARAMETRI - Aspetti tecnologici ed implicazioni nei modelli linguistici”  
 (Atti del I convegno naz. *AISV*, Padova 2004), Padova, EDK, 2005 (CD-ROM), 101-133.

## 2.2 Assimilazione regressiva e nasali preconsonantiche

È noto che in numerose lingue comprendenti numerose varietà iberoromanze e italo-romanze, tra cui castigliano e italiano, un processo di assimilazione regressiva di luogo d'articolazione interessa tutti i nessi eterosillabici il cui primo elemento sia una consonante nasale<sup>8</sup>.

Il fenomeno può essere facilmente riassunto con degli esempi come quelli riportati in Tabella I:

Italiano		Spagnolo cast.		
<i>campo</i>	/ˈkampo/	[ˈkampo]	<i>amparo</i>	/amˈparo/ [amˈparo]
<i>canfora</i>	/ˈkanfora/	[ˈkan̩fora]	<i>enfermo</i>	/enˈfermo/ [en̩ˈfermo]
<i>avanzo</i>	/aˈvantsɔ/	[aˈvantsɔ]	<i>lanzar</i>	/lanˈθar/ [lan̩ˈθar]
<i>canto</i>	/ˈkanto/	[ˈkanto]	<i>cantar</i>	/kanˈtar/ [kan̩ˈtar]
<i>espanso</i>	/esˈpansɔ/	[esˈpansɔ]	<i>cansar</i>	/kanˈsar/ [kan̩ˈsar]
<i>lancio</i>	/ˈlantʃo/	[ˈlantʃo]	<i>ensanchar</i>	/ensanˈtʃar/ [en̩san̩ˈtʃar]
<i>granchio</i>	/ˈgrankjo/	[ˈgrank̩k̩jo]	<i>conyugal</i>	/konjuˈɣal/ [kon̩ˈjuˈɣal]
<i>stanco</i>	/sˈtanko/	[sˈtank̩ko]	<i>naranjada</i>	/naraŋˈxaða/ [naraŋ̩ˈxaða]

Tabella I: Schema riassuntivo dei contesti in cui compaiono i principali tassofoni nasali descritti per italiano e spagnolo e ispirati a quelli di Mioni (1973), Canepari (1979), Quilis (1981), Miotti (1998).

Dagli esempi si osserva che anche in spagnolo dunque la consonante nasale assimila solitamente il suo luogo di articolazione a quello della consonante seguente. Come in italiano, ciò accade anche in fonetica sintattica (*un peso, en frente, en casa*).

È noto che il fenomeno qui descritto si presenta invece con minore evidenza nell'italiano settentrionale dove la resa più frequente di questi nessi si fonda prevalentemente su una velarizzazione incondizionata della nasale (associata a una più consistente nasalizzazione vocalica e a una latente cancellazione della consonante nasale)<sup>9</sup>.

<sup>8</sup> Sul piano evolutivo e con particolari riflessi sulle scelte grafiche nelle diverse lingue di cultura europee, si può considerare che il fenomeno interessi già, forse con una minore portata, anche latino e greco. Si pensi ad esempio all'alternanza di grafemi nasali in parole come lat. TĒMPUS e TRIŪMPHUM (o gr. τύμπανον e ἔμφασις) vs. lat. CANTĀRE (o gr. ἄκωνθος) e lat. CANCER, TRŪNCUM, VINCŪLUM etc. (o gr. ἐλεγχος, λάρυγξ/-γγος etc.). Altri esempi possono essere lat. AMPHITHEĀTRUM e gr. ἀμφιθέατρον, lat. NŪMPHAM e gr. νύμφη. Una diversa disposizione si nota invece nelle soluzioni presentate a confine di morfema; si confronti ad esempio lat. CONFĒRRE e CONFĪNE (ma lat. med. CONCLĀVE < CUM + CLAVEM che denuncia ancora una preferenza per N davanti a C, oppure solo l'estensione analogica di CON-) con gr. συμφωνία, comp. di σὺν- e -φωνία e σύγχρονος, comp. di σὺν- e -χρόνος.

<sup>9</sup> Questo può favorire l'insorgere di altri interessanti fenomeni, come la tendenza alla generalizzazione della sonorizzazione di /s/ (sistematica in posizione intervocalica) anche nella realizzazione di contesti /VnsV/. A Torino (e nell'area piemontese), il fenomeno (che appare molto più generalizzato ad esempio in parlanti di Trento o di Trieste) è esteso a tutte le formazioni col prefisso *trans-* che sono inevitabilmente pronunciate con [z] (*transizione, transitorio, transito, transatlantico, transenna*,

Antonio Romano, Franca Mancini, Enrico Zovato - In P. Cosi (a cura di),  
"MISURA DEI PARAMETRI - Aspetti tecnologici ed implicazioni nei modelli linguistici"  
(Atti del I convegno naz. AISV, Padova 2004), Padova, EDK, 2005 (CD-ROM), 101-133.

In spagnolo, nel prefisso *trans-* e nei prefissi *in-* e *con-*, soprattutto quando seguiti da *s*, *n* è molto debole e tende a ridursi, soprattutto nel parlato veloce (es. *instante*, *transformar*). Il fenomeno non è sconosciuto all'italiano anche se reso più raro dalla minor frequenza di questi nessi (comuni prevalentemente in derivati per i quali sia stata preferita una grafia conservativa)<sup>10</sup>.

A sottolineare le differenze tra spagnolo e italiano, ricordiamo che in spagnolo in posizione finale si trova solo /n/ (realizzato come velare [ŋ] in alcune varietà, o addirittura non realizzato) anche quando graficamente notato *-m*<sup>11</sup>. La velarizzazione - in alcune varietà in cui è particolarmente presente - può avvenire anche a fine sillaba (es. *canción*, realizzata spesso come *ca[ŋ]ció[ŋ]*)<sup>12</sup>.

### 2.3 Consonanti nasali e strutture formantiche

Per inquadrare correttamente il problema sul piano acustico sono necessarie alcune considerazioni preliminari in ordine alla modellizzazione acustica del tratto vocale (v. Hattori *et alii*, 1958; Fujimura, 1962; Calliope, 1989).

Nel caso di suoni orali, l'attivazione supplementare (da parte del velo palatino) delle cavità nasali conduce a condizioni di accoppiamento acustico, con l'apparizione - nello spettro d'energia dei suoni relativi a queste articolazioni- delle formanti relative alla risonanza di queste cavità, sovrapposte -in misura più o meno consistente e più o meno variabile - con quelle delle cavità orali. A questo proposito si vedano i numerosi studi articolatori disponibili (citati per es. in Farnetani, 1979a, 1979b), le teorie acustiche di Fujimura (1962) e Feng (1987), le osservazioni spettrografiche di Vaggies *et alii* (1975) e Ferrero *et alii* (1979) e le interessanti verifiche percettive di Pettorino (1988), Giannini *et alii* (1990), Maturi (1991)<sup>13</sup>.

---

altri esempi in Canepari, 1999: 88). Il fenomeno è maggiormente significativo in altre aree del Nord Italia (in particolare, è segnalato -con minori o maggiori oscillazioni- proprio in Trentino, v. Canepari, 1999: 396, e Venezia-Giulia, v. Canepari, 1999: 407).

<sup>10</sup> Per esempio: *constare*, *constatare*, *insperato*, *inspirare*, *instabile*, *instaurare*, *instradare* (ma si pensi anche ad *avanspettacolo*, *sanscrito*). Si veda Panconcelli-Calzia (1904) e il più recente lavoro di Turchi & Gili Fivela (2004).

<sup>11</sup> Oltre a esempi come *canción*, i casi descritti solitamente sono quelli di *álbum*, *máximo* etc. (notare però che in fonosintassi, davanti a vocale si può recuperare una pronuncia di tipo [m], es. *álbum italiano*).

<sup>12</sup> Nel Nord della Spagna, in Galizia e nelle Asturie, per esempio. Si veda a questo proposito D'Introno *et alii* (1995). V. anche, più avanti, §5 e Fig. 8.

<sup>13</sup> Sul piano acustico, in accordo con Delattre *et alii* (1955) e Fujimura (1962), Ferrero *et alii* (1979: 77) osservano - come effetti dell'inserimento delle cavità nasali - delle modifiche nella curva di risposta della vocale orale corrispondente ("spostamento delle frequenze di risonanza (F<sub>1</sub>, F<sub>3</sub>, ecc.)"; "comparsa di nuove risonanze (FN<sub>1</sub>, FN<sub>2</sub>, ecc.) e anche di anti-risonanze, cioè di zone dello spettro in cui l'intensità delle armoniche è fortemente attenuata"). Mettendo a confronto l'involuppo spettrale di una vocale nasale con quello della vocale orale corrispondente, gli stessi autori (Ferrero *et alii*, 1979: 79) rilevano dei fatti che non possono però essere descritti come generali. Una sicura conferma può venire riguardo all'assenza di cambiamenti significativi di F<sub>0</sub>; all'apparizione di una prima formante nasale con frequenza di circa 250 Hz; alla forte attenuazione di L<sub>1</sub>; all'apparizione di una seconda formante nasale verso i 1000 Hz e agli spostamenti di F<sub>3</sub> e F<sub>4</sub>. Non si possono invece generalizzare le altre osservazioni, che sono valide solo in alcuni casi: la tendenza di F<sub>1</sub> ad assumere un valore medio di 500 Hz dipende infatti dal luogo di articolazione; la posizione di F<sub>2</sub> dipende, oltre che da questo, anche dalla posizione degli zeri di nasalità (v. Nakata, 1959); l'apparizione di una terza formante nasale verso i 2000 Hz dipende invece, anche secondo le nostre osservazioni, oltre che dalla qualità vocalica, anche da caratteristiche individuali.

Nel paragrafo che Ladefoged & Maddieson (1996) dedicano alla struttura acustica delle nasali, troviamo la cavità orale descritta come una camera di risonanza laterale rispetto al percorso ideale - del flusso d'aria proveniente dai polmoni - lungo il tubo acustico che si crea dalla laringe alle narici (cfr. anche Ferrero *et alii*, 1979: 138, che parlano di "inserzione delle cavità nasali in parallelo alla cavità orale chiusa"). Il contributo acustico di questa camera si manifesterebbe con l'aggiunta di un'antirisonanza, cioè uno zero spettrale, nella composizione formantica dei suoni che così si generano. La frequenza di questo zero (individuato come minimo picco negativo tra  $F_1$  e  $F_2$ ) aumenterebbe con l'arretramento dell'articolazione orale nell'area uvulare, ma come effetto di questo arretramento viene anche segnalato un aumento della prima risonanza nasale (Ladefoged & Maddieson, 1996: 116). Le formanti si presentano con un'energia rapidamente decrescente all'aumentare della frequenza, al punto che a volte solo la prima ha un'energia sufficiente per emergere dalla sequenza di armoniche<sup>14</sup>.

Per diverse articolazioni, oltre alla variazione di frequenza di alcune formanti orali, grazie all'accoppiamento tra cavità orali e cavità nasali, anche i valori di frequenza delle risonanze nasali possono subire delle variazioni (*formant shifting*)<sup>15</sup>, spostandosi in base al loro luogo di articolazione orale. In realtà il risultato di quest'accoppiamento è la presenza di poli e zeri nella funzione di trasferimento complessiva e la variazione della loro posizione legata a questo grado di libertà può causare degli effetti di cancellazione di alcune frequenze, come pure degli effetti di scissione o divisione formantica (*formant splitting*)<sup>16</sup>.

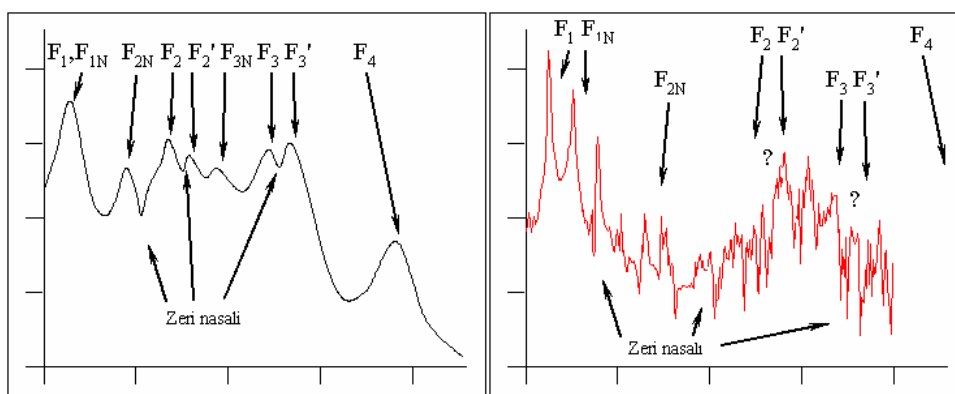


Figura 1: Schema di tutte le potenziali risonanze che si possono manifestare in un suono nasale o orale nasalizzato (a sinistra), affianco a un esempio di spettro FFT (a destra).

<sup>14</sup> Molti autori, proponendosi di dettagliare la struttura formantica dei suoni nasali, si limitano a osservare, dalle frequenze più basse a quelle più alte, le maggiori concentrazioni energetiche incontrate, senza discutere la loro affiliazione a cavità orali o nasali.

<sup>15</sup> "On devrait distinguer les pics de nasalité des pics des formants oraux déplacés (*shifted formants*)" (Feng, 1987: 6)

<sup>16</sup> Degli esempi di divisione delle formanti orali per fraposizione di uno zero di nasalità sono commentati da Feng (1987: 34), mentre per la cancellazione si veda Feng (1987: 68). Anche gli schemi sistematici di Vaggés *et alii* (1975), riportati poi in Ferrero *et alii* (1979), consideravano già formanti "splitted".

Antonio Romano, Franca Mancini, Enrico Zovato - In P. Cosi (a cura di),  
"MISURA DEI PARAMETRI - Aspetti tecnologici ed implicazioni nei modelli linguistici"  
(Atti del I convegno naz. *AISV*, Padova 2004), Padova, EDK, 2005 (CD-ROM), 101-133.

L'assunzione di partenza è che lo schema generale della composizione spettrale di suoni nasali o orali nasalizzati sia affidato a una sequenza di valori formantici ordinati come in Fig. 1 (a sinistra, schema generale valido nel caso di un'articolazione vocalica anteriore piuttosto bassa e con labializzazione contenuta; a destra, esempio delle difficoltà di riconoscere i vari contributi sullo spettro *FFT* di una /i/ nasalizzata)<sup>17</sup>. Come anticipato sopra, in realtà, a ogni  $F_N$  sembra corrispondere una concentrazione di energia non sempre tale da concorrere in maniera significativa, con i suoi valori di picco, a una sostanziale modifica degli involucri spettrali. In genere si può anzi presentare solo come una debole deviazione di questi, cui possono essere associate fluttuazioni negative, testimoni di fenomeni di antirisonanza.

Ovviamente per consonanti nasali lo spettro si riduce essenzialmente ai contributi di  $F_{1N}$ ,  $F_{2N}$  e  $F_{3N}$ , anche se le formanti orali possono dominare nella prima fase di transizione (con i fenomeni di *shifting* visti sopra, associati a frequenti allargamenti di banda e in presenza di indici di divisione formantica che interessano soprattutto  $F_3$ , ma si manifestano visibilmente, per certe regioni articolatorie, anche per  $F_2$ ).

Un'attenta valutazione, in termini statici e dinamici, delle formanti delle diverse realizzazioni nasali presenti nel *database* (in seguito *DB*) è in corso, a parziale conferma delle osservazioni in Romano (1997). Un confronto con le misure medie rese disponibili da altri lavori più generali (in molti casi su varietà vicine) sta rivelando diffuse inconsistenze, probabilmente legate proprio all'ambiguità dei diversi contributi energetici presenti nelle diverse voci (i valori considerati provengono da Fant, 1960; Fujimura, 1962; Ferrero *et alii*, 1979; Quilis, 1981; Recasens, 1983; Fernández *et alii*, 1988; Giannini & Pettorino, 1992; Albano Leoni & Maturi, 1995; Machuca, 1991a).

### 3. INDICI ACUSTICI PER LA SEGMENTAZIONE E LA CLASSIFICAZIONE

Nell'ottica di una segmentazione automatica, l'esistenza dei numerosi fenomeni di cui sopra, unita ai noti effetti di smorzamento (già descritti da Delattre *et alii*, 1955, e Fujimura, 1962), costituisce un'ottima base per la scansione in segmenti distinti. Oltre alla segmentazione basata sulla presenza di indici energetici (Mermelstein, 1977), esistono infatti algoritmi multiparametrici già collaudati per l'individuazione, in un parlato sufficientemente controllato, delle regioni maggiormente candidate a stabilire delle frontiere tra la vocale precedente e la nasale seguente (con tutti i *caveat* necessari quando si ha a che fare con un parlato meno controllato, v. Machuca, 1991b; Busà, 2003)<sup>18</sup>.

#### 3.1 Transizioni formantiche sulla vocale

Accantonati per il momento i risvolti pratici del riconoscimento spettrografico dei vari contributi spettrali e della segmentazione, si pone inoltre il problema della classificazione su base acustica della consonante nasale in base a quanto visto nel §2.2.

Se la consonante nasale, per effetto dell'assimilazione anticipatoria, sposta il suo luogo di articolazione orale in corrispondenza di quello della consonante che segue, ci aspetteremmo delle configurazioni spettrali differenziate per ciascuna delle varianti combinatorie in *Tabella I*. Tuttavia, sulla base di quanto esposto nel §2.3 gli schemi

<sup>17</sup> Rammentiamo che in un contesto di vocale posteriore labializzata una  $F_2$  potrebbe trovarsi su valori più bassi di quelli della seconda risonanza nasale. È questa una delle situazioni in cui forse sono più evidenti gli effetti di *splitting* su  $F_2$  causati dall'allineamento con questa del primo zero nasale.

<sup>18</sup> In realtà, in termini di affidabilità, si è andati decisamente più avanti con le tecniche statistiche e con i metodi *HMM* (v. Cosi *et alii*, 1991). Resta però il problema di una corrispondenza tra il risultato ottenuto da tali algoritmi e le indicazioni spettrografiche, in base alle esigenze dell'utente fonetista.



Antonio Romano, Franca Mancini, Enrico Zovato - In P. Cosi (a cura di),  
“MISURA DEI PARAMETRI - Aspetti tecnologici ed implicazioni nei modelli linguistici”  
(Atti del I convegno naz. *AISV*, Padova 2004), Padova, EDK, 2005 (CD-ROM), 101-133.

formantici dei segmenti nasali si presentano essenzialmente dominati dalle risonanze nasali e da un affollarsi di complessi fenomeni di accoppiamento e da interazioni tra componenti che non ne facilitano il discernimento. Visto che però le diverse varianti combinatorie hanno un diverso luogo d'articolazione, e conoscendo gli effetti di coarticolazione che intervengono al passaggio VC, è facile constatare come i diversi tassofoni si rivelino già, almeno in parte, grazie a un diverso schema di transizioni delle formanti della vocale precedente (cfr. Malécot, 1956; Nakata, 1959; Mermelstein, 1977; Recasens, 1983; Kurowski & Blumstein, 1987; Harrington, 1994; Harding & Meyer, 2003).

Passando in rassegna numerosi nessi di nasale + C presenti nei dati a nostra disposizione, per l'italiano e per lo spagnolo, abbiamo cercato di mettere ordine tra questo tipo di indici acustici presenti nelle configurazioni spettrali delle vocali.

L'ispezione informale, seguita da una ancora in corso con procedure di rilievo dei valori più rigorose, ha confermato la presenza di *pattern* formantici blandamente differenziati per i vari tassofoni (cfr. Recasens, 1983; Ohde, 1994). Inoltre ha messo in evidenza come le diffuse approssimazioni che portano a trascurare alcune varianti - di più difficile individuazione sulla base di sommarie ispezioni articolatorie - sono ribaltate sul piano acustico, rendendo più evidente ad esempio la caratterizzazione di una [ɲ] (una variante che compare davanti a /tʃ/ e /dʒ/<sup>19</sup>) di quanto non lo siano una [ɲ] (spesso considerata anche in testi divulgativi) o una [ɲ] (che si è presentata sempre estremamente mutevole nei nostri dati, cfr. Recasens, 1983: 1347). Di questo vogliamo dare una dimostrazione nelle Figg. 2, 3 e 4 che permettono di constatare la variabilità della nasale in alcune parole pronunciate dalla voce Carmen del sistema *LoquendoTTS*<sup>®</sup> (spagnolo castigliano). Al di là delle influenze derivanti dalla presenza di una diversa consonante prevocalica, sulle vocali accentate di queste parole ci saremmo aspettati di vedere una deviazione di F<sub>2</sub> verso i 1700 Hz per /ant/ e verso i 2500 Hz per /ank/. Nei casi riportati in figura, tenendo conto dei fenomeni descritti al §2.3, ciò sembra effettivamente accadere, ma in una misura assai limitata (valori più bassi<sup>20</sup>). È invece evidente, anche se disturbata da uno *splitting*, la deviazione verso i 1900 Hz tipica delle [a] di questa locutrice in contesti di tipo /antʃ/ (Fig. 4) e confermata da molti altri esempi (v. Fig. 5).

<sup>19</sup> Di un arretramento del punto di articolazione della nasale in questo contesto testimoniano Farnetani & Busà (1994: 228-229) che parlano di assimilazione categorica solo per /nk/ e /ntʃ/ (p. 241).

<sup>20</sup> Anche questo è in accordo con i rilievi articolatori di Farnetani & Busà (1994) che osservano come, nel caso di /nt/, per i locutori da loro considerati, più che di assimilazione del luogo di articolazione della nasale a quello della consonante seguente, si debba tener conto invece di una coarticolazione. Il risultato (p. 227) si manifesta in un contatto più esteso (di quelli presenti nelle articolazioni di /n/ e /t/ intervocaliche da parte degli stessi locutori) e più arretrato ([nt] invece di un ipotetico [n̠t]).

Antonio Romano, Franca Mancini, Enrico Zovato - In P. Cosi (a cura di),  
“MISURA DEI PARAMETRI - Aspetti tecnologici ed implicazioni nei modelli linguistici”  
(Atti del I convegno naz. *AISV*, Padova 2004), Padova, EDK, 2005 (CD-ROM), 101-133.

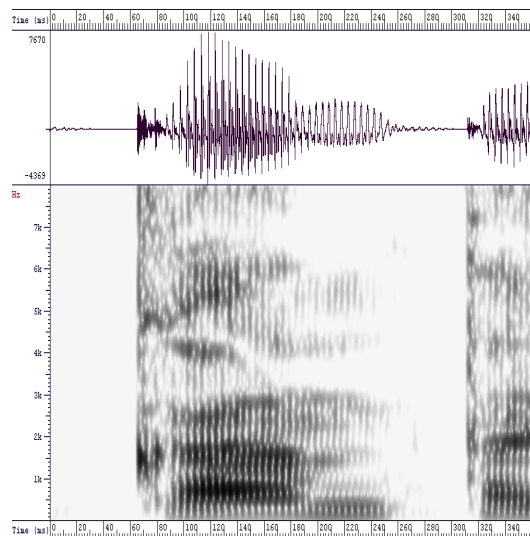


Figura 2: Esempio di transizione in [ant] della parola *canta* pronunciata dalla voce Carmen.

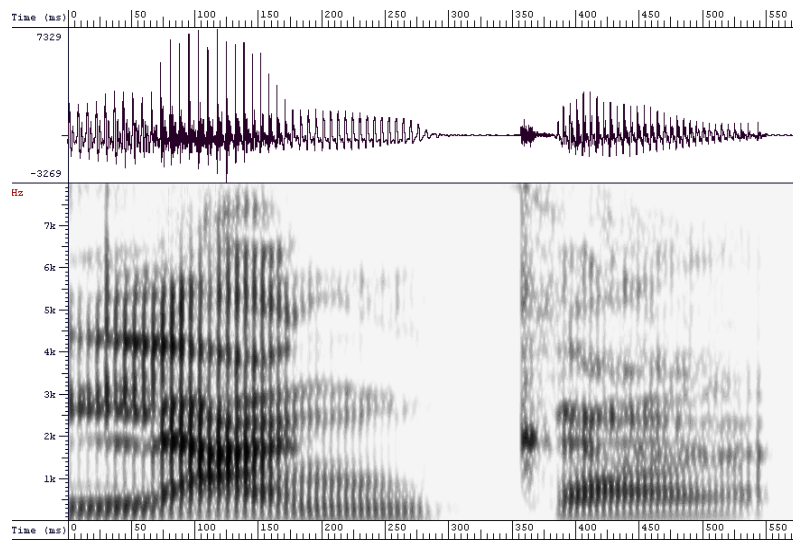


Figura 3: Esempio di transizione in [anj] della parola *Lanka* pronunciata dalla voce Carmen.

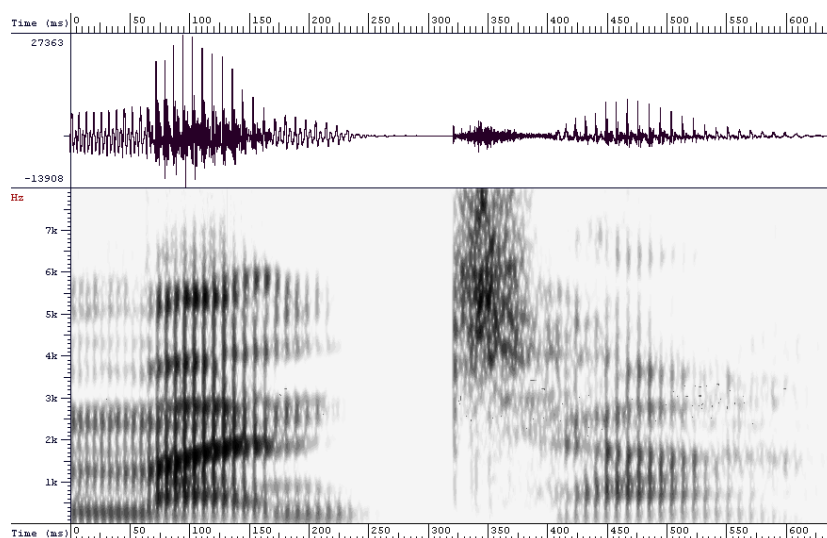


Figura 4: Esempio di transizione in [antʃ] della parola *mancha* pronunciata dalla voce Carmen.

In Fig. 5 è possibile constatare una buona convergenza (anche se non sono assenti occasionali deviazioni non standard) tra le varie strutture formantiche delle diverse realizzazioni della stessa vocale in tre contesti distinti: /'anta/, /'anka/, /'antʃa/<sup>21</sup>. In particolare si noterà che le regioni di convergenza di F<sub>2</sub> per entrambe le *speaker* e in contesti velari o postalveolari non differiscono di molto<sup>22</sup>. Mentre le misure delle frequenze di tali regioni, nel caso di /ant/ vanno dai 1600 ai 1800 Hz, per /ank/ gli ultimi valori registrati si aggirano più variabilmente tra i 1500 e i 2000 Hz<sup>23</sup>. Maggiormente concentrate sembrano invece le misure effettuate nel caso di /antʃ/, che si localizzano tra i 1800 e i 2000 Hz. La differenza principale tra le [a] di /ank/ e le [a] di /antʃ/ è nei tempi con cui avviene di solito la deviazione, che nel caso di /antʃ/ parte visibilmente con un certo anticipo (più sistematico per la voce Carmen, la quale presenta anche la migliore separazione tra [ŋ] e [ɲ])<sup>24</sup>.

<sup>21</sup> Entrambe le locutrici hanno una seconda formante di /a/ piuttosto alta.

<sup>22</sup> Le misure sono effettuate nelle ultime 10 finestre (da 120 campioni) di vocale, prima del confine convenzionale con la nasale seguente. Il demarcatore è posto automaticamente dal segmentatore utilizzato da *Loquendo* (cfr. Brugnara *et alii*, 1993) e validato da un operatore esperto.

<sup>23</sup> Notare che, la voce Paola si caratterizza anche per un sistematico fenomeno di cambiamento di affiliazione di cavità di risonanza per F<sub>3</sub> e F<sub>4</sub> nel passaggio da [a] a [ŋ].

<sup>24</sup> Trovandoci in posizione implosiva, la salienza di questi indici è ovviamente ridotta (cfr. Farnetani & Busà, 1994: 242). Riguardo allo scarso rilievo che hanno ricevuto in letteratura le differenze tra [ŋ] e [ɲ] si veda però Recasens (1983: 1347).

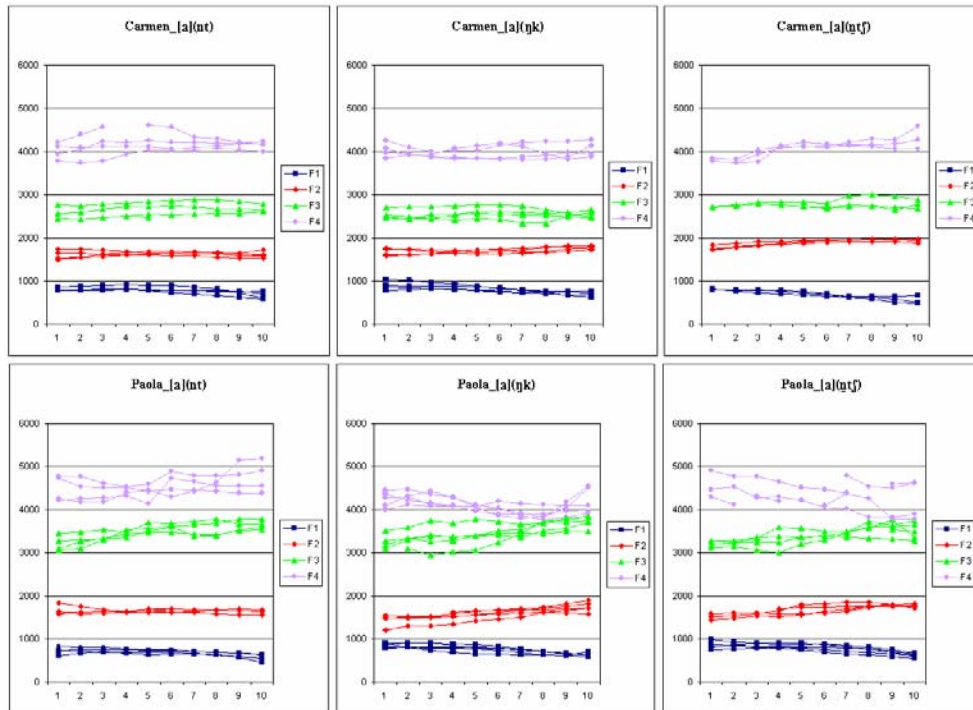


Figura 5: Esempi di transizioni delle prime quattro formanti orali in realizzazioni di /a/ (nei diversi contesti indicati in didascalia). Le 30 parole considerate sono pronunciate dalle voci Carmen e Paola dei sistemi *LoquendoTTS*<sup>®</sup> per lo spagnolo castigliano e per l'italiano.

### 3.2 Fenomeni di contatto tra nasali e consonanti seguenti

Il passaggio dalla nasale alla consonante seguente nei nessi /NC/, almeno per certi tipi di C, sembrerebbe non dover porre problemi di segmentazione. E in effetti, salvo i casi in cui la C sia un'occlusiva sonora, la delimitazione dei foni è di solito abbastanza agevole<sup>25</sup>.

Resta però il problema dell'etichettatura delle parti. Infatti spesso il segmento delimitato come C potrebbe includere elementi spurî - dall'apporto percettivo più o meno importante - che si generano al contatto con la nasale precedente per sovrapposizione di *pattern* acustici o forse anche articolatori. Nel caso di fricative ad esempio, è ben noto il processo descritto per l'inglese (ad esempio in Wells, 1997: 22-23) sulla base della latente omofonia di coppie come *sense-cents* e *prince-prints* e di altri frequenti casi di epentesi (*emphasis* con [mpf] o *instance* con [nst]). Sul piano dell'osservazione strumentale, al fenomeno è stata dedicata un'ampia trattazione in Ohala (1997), che approfondisce il tema delle "ostruenti emergenti" in diverse lingue e discute l'affermazione di questi elementi epentetici sulla base di indagini articolatorie e percettive.

<sup>25</sup> Come vedremo, nel caso di occlusive o affricate sorde, però l'esatta delimitazione della fase di tenuta dell'occlusione, resta inficiata dalle condizioni e dalle apparecchiature di registrazione: strascichi delle risonanze di bassa frequenza della nasale invadono infatti spesso i tempi in cui avviene la fase d'impostazione dell'occlusiva. Il fenomeno è molto limitato nei materiali *Loquendo*, registrati in studi e con apparecchiature professionali.

Antonio Romano, Franca Mancini, Enrico Zovato - In P. Cosi (a cura di), "MISURA DEI PARAMETRI - Aspetti tecnologici ed implicazioni nei modelli linguistici" (Atti del I convegno naz. *AISV*, Padova 2004), Padova, EDK, 2005 (CD-ROM), 101-133.

Per le parlate italo-romanze e per l'italiano parlato in alcune regioni d'Italia si veda Canepari (1999: 20, 1999: 88): toglie l'area fiorentina, il processo sembra interessare tutta l'Italia sotto la linea La Spezia-Rimini. Secondo Canepari (1999: 380, 483), il fenomeno sarebbe ugualmente attestato (anche se in forme meno evidenti) nel Ticino e nell'area lombarda nord-orientale<sup>26</sup>. Un contributo di impianto acustico è stato proposto di recente per la sola area toscana, da Turchi & Gili Fivela (2004).

Nessuna attestazione in letteratura sembra interessarsi dell'epentesi di occlusive in contesto di nasale + fricativa (inter-)dentale per lo spagnolo<sup>27</sup>.

Della diversa tendenza mostrata da italiani settentrionali e meridionali testimoniano i dati sperimentali dello stesso Panconcelli-Calzia (1904), quando commenta i chimogrammi raccolti con parlanti di Atina (Caserta, p. 94), Ancona (ma non sempre, p. 103) e le sue stesse produzioni (simili a quelle di qualunque altro parlante romano che pronunci questi nessi senza sorvegliarsi, p. 104). Il fenomeno in questione è descritto non come epentesi, ma come realizzazione semi-occlusiva di /s/<sup>28</sup>, in contrasto con i fenomeni di nasalizzazione della vocale (con conseguente tendenza alla perdita della consonante nasale) tipica dei soggetti del Nord (p. 105)<sup>29</sup>.

Nel nostro caso, avendo constatato una certa varietà di realizzazioni di questo tipo, ci siamo interessati anche noi ai casi di nasale + fricativa, con o senza epentesi, e al confronto di questi con casi di nasale + affricata primaria, quando esistenti nella lingua in questione.

Dall'osservazione informale di un certo numero di spettrogrammi relativi a materiali del *DB Loquendo* a nostra disposizione per italiano e spagnolo (v. §4), abbiamo avuto la netta impressione che si inseriscano elementi occlusivi in condizioni e con modalità spesso non prevedibili sulla base degli indici spettrografici descritti in letteratura (Dorman *et alii*, 1979, Ali *et alii*, 1979). In particolare, in molti casi non era la presenza di passaggi silenziosi a favorire la percezione di un'esplosione all'attacco della fricativa postnasale. Al contrario, i contributi energetici delle due consonanti, a contatto, sembravano proprio confondersi - sovrapponendosi - in una fase di transizione che, in base ai tempi di affermazione del rumore di frizione, poteva lasciar percepire (o no) un'affricazione<sup>30</sup>.

Inoltre, mentre nelle produzioni degli *speaker* italiani la percezione di un'esplosione resta limitata ai casi di affricata, per gli *speaker* spagnoli, un elemento esplosivo di tipo

---

<sup>26</sup> Nelle aree partenopea e cosentina, sarebbe concomitante con la latente sonorizzazione delle consonanti postnasali.

<sup>27</sup> In base alle informazioni raccolte da Ana Ma. Fernández Plana e da Elisa Fernández Rei (che qui approfittiamo per ringraziare), il fenomeno, seppur evidente in alcune pronunce iberiche, non è stato ancora indagato con tecniche strumentali, ma neanche descritto come fatto fonetico accidentale (cfr. anche Massone, 1988). Per l'inglese, oltre ai riferimenti citati sopra, si vedano anche Barnitz (1974), col suo ammiccante titolo, Ali *et alii* (1979), Fourakis & Port (1986) e Warner & Weber (2001).

<sup>28</sup> Molto interessante è, a questo proposito, la testimonianza che l'autore dà quando dice che, nei suoi tentativi di evitare questa pronuncia affricata (cercando di preservare una *s* "pura"), tende a perdere la *n*, nasalizzando inevitabilmente la vocale precedente (Panconcelli-Calzia, 1904: 104).

<sup>29</sup> Di questa nasalizzazione, tengono conto, in un lavoro di palatografia dinamica, Farnetani & Busà (1994) e ancora, recentemente, Busà (2003), in un contributo sperimentale in cui si fa ricorso a misure di flusso nasale per studiare realizzazioni di tipo [ṽ<sup>n</sup>s]. Numerosi sono i lavori su fenomeni simili in altre lingue, come il francese o il portoghese: si può partire dalla rassegna di Lacerda & Companys (1958), fino ai più recenti lavori di Maeda (1993) e Teixeira *et alii* (1999a, 1999b, 2001, 2003).

<sup>30</sup> Anche Turchi & Gili Fivela (2004: 5) riportano simili osservazioni, attribuendo la percezione di un'affricata al *rise time* del rumore di frizione.

Antonio Romano, Franca Mancini, Enrico Zovato - In P. Cosi (a cura di),  
"MISURA DEI PARAMETRI - Aspetti tecnologici ed implicazioni nei modelli linguistici"  
(Atti del I convegno naz. *AISV*, Padova 2004), Padova, EDK, 2005 (CD-ROM), 101-133.

dentale può apparire in alcuni casi, marcando le produzioni di /nθ/, e un altro di tipo velare caratterizza la voce Carmen nelle realizzazioni di /ns/<sup>31</sup>.

Per valutare più attentamente questi casi, abbiamo fatto ricorso a una rappresentazione a cascata delle componenti energetiche da 0 a 8 kHz (grafico di diversi profili *LPC* in successione temporale ogni 20 punti, con finestre da 160 punti; si vedano gli esempi riportati nelle Figg. 4 e 5). Tale rappresentazione permette di osservare meglio la variabilità temporale, nei diversi contesti, delle diverse concentrazioni energetiche e dell'andamento generale dell'energia.

In Fig. 6, riportiamo l'esempio della transizione presente in /nθ/ nella realizzazione della parola *atención* della voce Carmen. La realizzazione in questo caso è del tipo [n<sup>h</sup>θ] come si può verificare anche grazie al diagramma "waterfall" (LPC a 20 poli su finestre di 10 ms) in Fig. 7. Una traccia di *spike* è presente al passaggio dalla nasale alla fricativa anche se le formanti nasali persistono (anche quelle più alte) fino all'inizio (piuttosto brusco) della frizione. L'impressione uditiva di una leggera occlusione al passaggio dalla nasale alla fricativa è in parte confermata dall'improvvisa comparsa di energia nella zona alta di frequenze (1900-2900, poi 4200-5500-6900 Hz).

Anche l'esempio in Fig. 8 (/ns/ della parola *tensión* pronunciata dalla stessa voce) riguarda una sequenza (contrariamente alle aspettative, di tipo [n<sup>h</sup>s]) che presenta la comparsa di un elemento ostruente spurio di tipo piuttosto velare all'ascolto (quasi fosse [n<sup>h</sup>s<sup>h</sup>])<sup>32</sup>. Sono probabilmente i rapidi rilasci delle costrizioni nasali che causano delle transizioni di tipo esplosivo, le quali rappresentano degli importanti indici per la percezione di occlusive (Ali *et alii*, 1979: 86).

---

<sup>31</sup> Il fenomeno è simile a quello che si può percepire nelle realizzazioni di simili nessi dell'italiano da parte di quei parlanti settentrionali che generalizzano una nasale velare in tutte le posizioni di coda sillabica.

<sup>32</sup> La transizione di F<sub>2</sub> sulla vocale lascerebbe però pensare a un'articolazione più anteriore. Deve trattarsi di uno di questi casi, descritti anche da Recasens (1983: 1347), in cui una [ŋ] può manifestarsi con un abbassamento della seconda risonanza (orale).

Antonio Romano, Franca Mancini, Enrico Zovato - In P. Cosi (a cura di),  
 "MISURA DEI PARAMETRI - Aspetti tecnologici ed implicazioni nei modelli linguistici"  
 (Atti del I convegno naz. AISV, Padova 2004), Padova, EDK, 2005 (CD-ROM), 101-133.

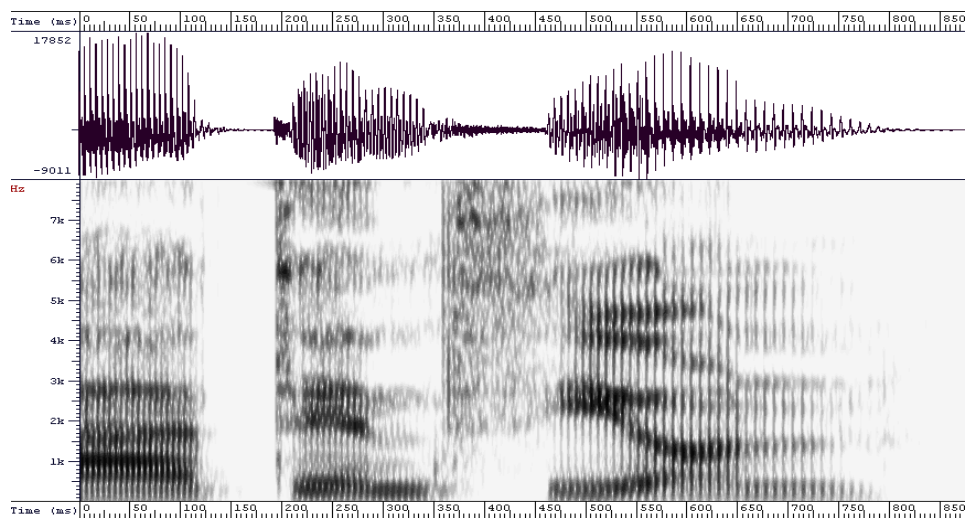


Figura 6: Esempio di transizione in /nθ/ della parola atención pronunciata dalla voce Carmen; la realizzazione in questo caso è del tipo [n<sup>h</sup>θ].

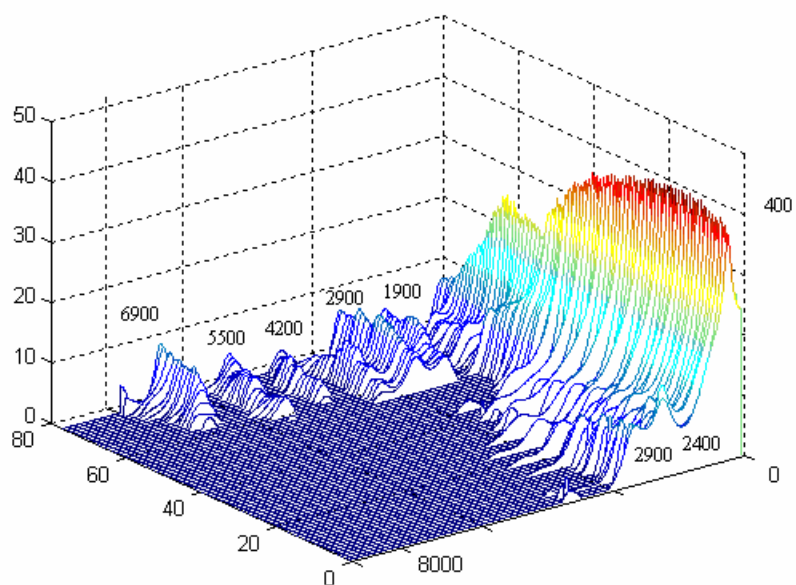


Figura 7: Diagramma "waterfall" (LPC a 20 poli su finestre di 10 ms) della transizione /nθ/ della parola in Fig. 6; una traccia di *spike* è presente al passaggio dalla nasale alla fricativa (in questa rappresentazione disposto "obliquamente" lungo l'asse delle frequenze da destra a sinistra; il confine tra i fonici in questo tipo di grafici è posto per convenzione in corrispondenza della 38<sup>a</sup> sezione).

Antonio Romano, Franca Mancini, Enrico Zovato - In P. Cosi (a cura di),  
 "MISURA DEI PARAMETRI - Aspetti tecnologici ed implicazioni nei modelli linguistici"  
 (Atti del I convegno naz. *AISV*, Padova 2004), Padova, EDK, 2005 (CD-ROM), 101-133.

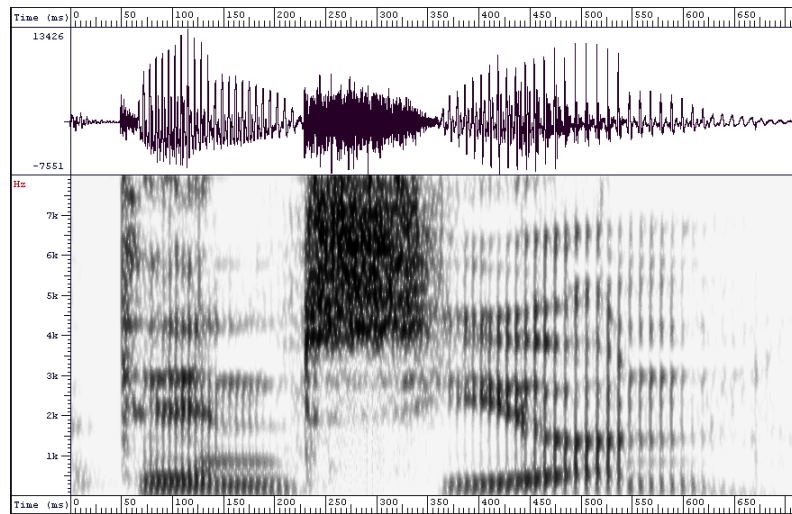


Figura 8: Esempio di transizione in /ns/ della parola *tensión* pronunciata dalla voce Carmen; contrariamente alle aspettative, la realizzazione in questo caso è del tipo [ɲʂ] con la comparsa di un elemento ostruente spurio di tipo piuttosto velare all'ascolto.

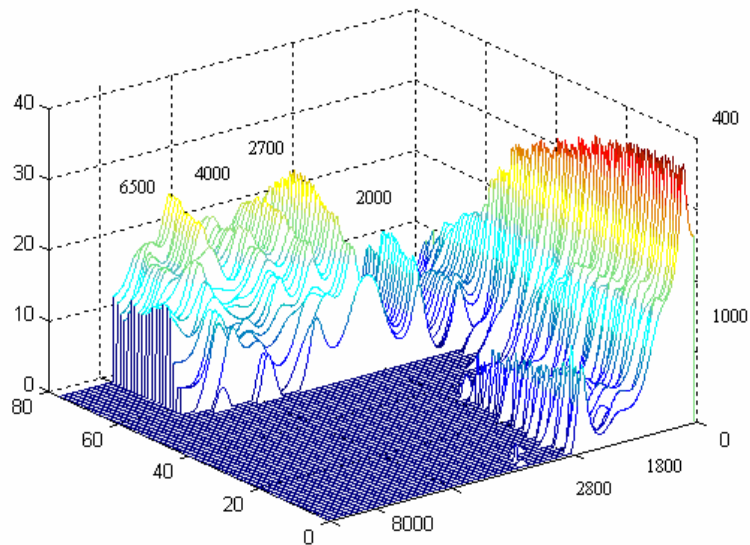


Figura 9: Diagramma "waterfall" (LPC a 20 poli su finestre di 10 ms) della transizione /ns/ della parola in Fig. 8; l'impressione uditiva di una netta occlusione al passaggio dalla nasale alla fricativa è confermata dall'improvvisa comparsa di energia nella zona alta dello spettro frequenze con un evidente *spike* che segue una fase di notevole riduzione energetica.



Antonio Romano, Franca Mancini, Enrico Zovato - In P. Cosi (a cura di),  
"MISURA DEI PARAMETRI - Aspetti tecnologici ed implicazioni nei modelli linguistici"  
(Atti del I convegno naz. *AISV*, Padova 2004), Padova, EDK, 2005 (CD-ROM), 101-133.

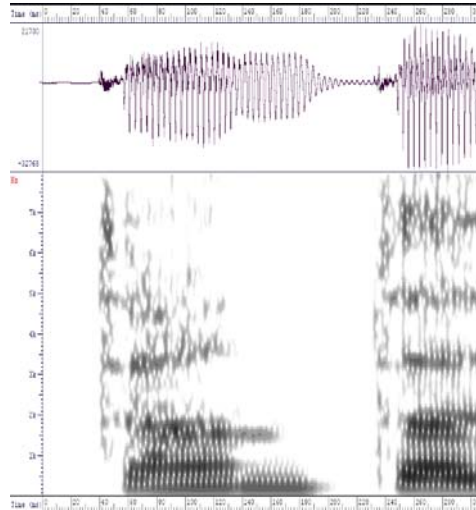


Figura 10: Esempio di transizione in /nt/ della parola [tanta](#) pronunciata dalla voce Paola.

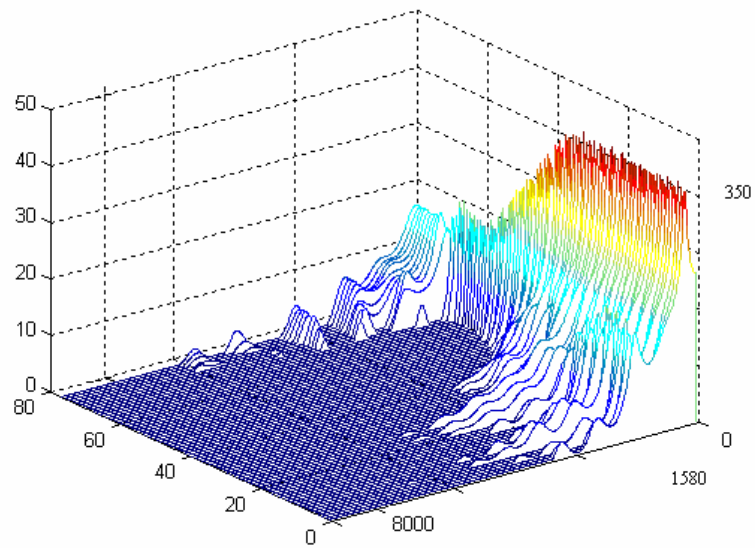


Figura 11: Diagramma "waterfall" (LPC a 20 poli su finestre di 10 ms) della transizione /nt/ della parola in Fig. 10.

Antonio Romano, Franca Mancini, Enrico Zovato - In P. Cosi (a cura di),  
"MISURA DEI PARAMETRI - Aspetti tecnologici ed implicazioni nei modelli linguistici"  
(Atti del I convegno naz. AISV, Padova 2004), Padova, EDK, 2005 (CD-ROM), 101-133.

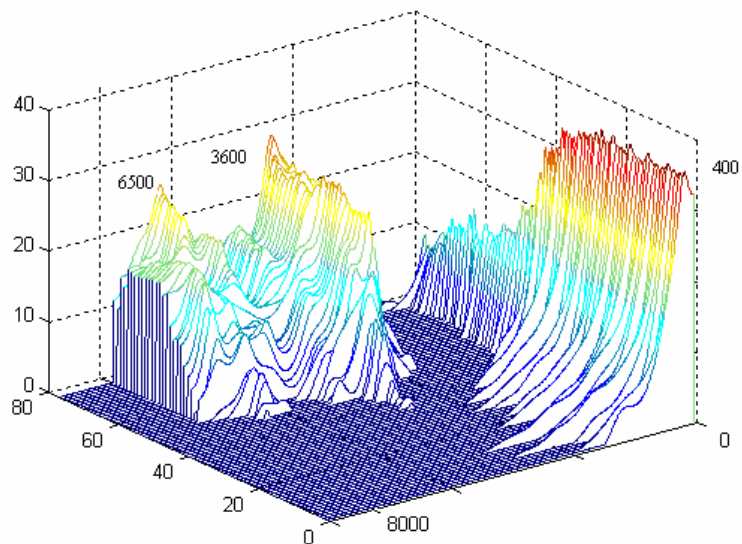


Figura 12: Diagramma "waterfall" (LPC a 20 poli su finestre di 10 ms) della transizione /nt/ della parola lancio pronunciata dalla voce Paola.

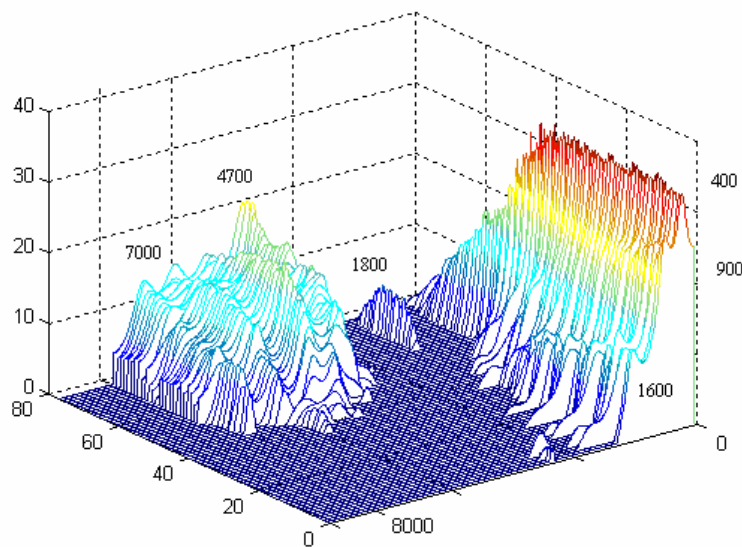


Figura 13: Diagramma "waterfall" (LPC a 20 poli su finestre di 10 ms) della transizione /ns/ della parola mansione pronunciata dalla voce Paola (cfr. con la realizzazione di *tensi3n* della voce Carmen nelle Figg. 8 e 9).

Antonio Romano, Franca Mancini, Enrico Zovato - In P. Cosi (a cura di),  
"MISURA DEI PARAMETRI - Aspetti tecnologici ed implicazioni nei modelli linguistici"  
(Atti del I convegno naz. *AISV*, Padova 2004), Padova, EDK, 2005 (CD-ROM), 101-133.

Per permettere una migliore valutazione dell'utilità dei diagrammi a cascata, riportiamo nelle Figg. 10 e 11 un esempio di realizzazione del nesso /nt/ da parte della voce italiana Paola. I grafici consentono di apprezzare la netta separazione tra la parte conclusiva della nasalità e l'apparizione del *burst*. Solo uno strascico di bassa frequenza prosegue, con un importante smorzamento, durante la fase di tenuta dell'occlusiva.

Ci saremmo aspettati un comportamento simile nel caso del nesso /ntʃ/ nella parola *lancio*; invece dobbiamo constatare che la manifestazione dell'affricata (perfettamente percepibile), la cui fase di tenuta continua quella della nasale - che è pur sempre un'occlusiva orale -, è legata alla gradualità con cui si presentano i contributi di alta frequenza del rilascio in concomitanza con la brusca riduzione dell'energia di banda bassa (v. anche Fig. 20).

Dal confronto del nesso /ns/ della Fig. 13 con quello /ntʃ/ della Fig. 12, in cui l'elemento occlusivo avrebbe dovuto invece essere presente e ben analizzabile (anche perché fonologicamente richiesto), emerge che il termine di differenziazione deve essere cercato nella diversa distribuzione spettrale dell'energia dei due foni a contatto e, soprattutto, nelle modalità di redistribuzione di quest'ultima durante la transizione.

#### 4. MATERIALI E PROCEDURE

Partendo da tutte le osservazioni riassunte nei paragrafi precedenti, abbiamo condotto uno spoglio sistematico di un gran numero di nasali intervocaliche e di nessi di nasale + C presenti nel *DB* di *LoquendoTTS*<sup>®</sup>. Le misure attualmente verificate e disponibili per la pubblicazione sono quelle relative alle durate e alle variazioni energetiche presenti nelle transizioni *NC*.

##### 4.1 Materiali

I nostri dati fanno riferimento ai modelli di pronuncia presenti nel *DB* del sistema di sintesi vocale *LoquendoTTS*<sup>®</sup> (<http://actor.loquendo.com>). Le voci considerate, in parte già presentate negli esempi dei paragrafi precedenti, sono quelle di:

Carmen e Jorge, per lo spagnolo, e  
Paola e Matteo, per l'italiano.

Si tratta di un numero finito di produzioni (frasi e elementi isolati) di *speaker* professionisti utilizzate per costituire la base di dati sonori a cui attinge il sistema di sintesi<sup>33</sup>. I materiali originari sono stati acquisiti presso studi di registrazione professionali, dove si sono svolte le fasi di *editing* e digitalizzazione (16 bit,  $F_c=16000$  Hz). I segnali non sono stati sottoposti ad alcun tipo di filtraggio ed equalizzazione. A livello di intensità sonora si è cercato di mantenere sempre lo stesso guadagno.

Il corpus è diviso in due sezioni:

- 1) *VNV*: 13853 nasali intervocaliche;
- 2) *VNCV*: 17396 nessi (8372 sono quelli che abbiamo finora analizzato per nasali+occlusive, nasali+fricative, nasali+affricate).

---

<sup>33</sup> I "donatori" italiani sono entrambi dicitori di professione originari dell'area torinese. Nei brani registrati, selezionati dagli specialisti di *LoquendoS.p.A.*, esibiscono delle pronunce controllate per caratteristiche timbriche e ortoepiche che solo raramente tradiscono la loro provenienza settentrionale. Le due voci spagnole invece, anch'esse giudicate ortoepiche, sono state selezionate da esperti fonetisti della *U.A.B. (Universitat Autònoma de Barcelona)* e sono prodotte da due *speaker* professionisti originari rispettivamente dell'area catalana e dell'area madrilenà.

Antonio Romano, Franca Mancini, Enrico Zovato - In P. Cosi (a cura di), "MISURA DEI PARAMETRI - Aspetti tecnologici ed implicazioni nei modelli linguistici" (Atti del I convegno naz. *AISV*, Padova 2004), Padova, EDK, 2005 (CD-ROM), 101-133.

#### 4.2 Trattamenti

Tutti i campioni audio sono stati analizzati con gli strumenti grafici interattivi *X-Waves*, *Wavesurfer*, *Praat* e *Matlab*.

Il corpus *VNV* si basa su misure di durata estratte automaticamente dal *DB* etichettato e sottoposte a trattamento statistico in fogli di *Microsoft Excel*.

Oltre all'analisi delle configurazioni spettrali, i nessi *NC* del corpus *VNCV* sono invece stati sottoposti a una finestrazione avente come scopo quello di calcolare delle "energie di banda a breve termine" (con  $W = 10$  ms e passo = 5 ms) i cui risultati sono stati inseriti anche in questo caso in fogli di *Microsoft Excel* (v. §5).

Le procedure da noi seguite possono sembrare piuttosto approssimative, ma sono quelle dettate dalle necessità pratiche della segmentazione. Riportiamo qui di seguito alcune nostre riflessioni preliminari che ci hanno consentito di definire un metodo, che fosse il meno arbitrario possibile, di delimitazione delle consonanti nasali rispetto ai foni adiacenti.

#### 4.3 Caratteristiche temporali

L'obiettivo di questo paragrafo è di illustrare brevemente le differenze osservate nelle durate delle realizzazioni nasali delle diverse voci analizzate.

Riguardo alle caratteristiche temporali di queste consonanti nelle due lingue è stato condotto uno studio statistico che ha portato a confermare alcune differenze di durata esistenti tra le diverse articolazioni. Le medie e le deviazioni standard delle durate di nasali in posizione intervocalica sono riassunte nella *Tabella II* (che dettaglia anche la consistenza di ciascuna consonante per ciascuna voce nel *DB*), mentre l'istogramma di Fig. 13 consente un colpo d'occhio globale sulla variazione osservata.

I risultati non permettono di generalizzare alcuna differenza sistematica tra le quattro voci raggruppate per lingua. È invece confermata, anche se in alcuni casi con debole significatività, la maggiore lunghezza intrinseca delle articolazioni palatali che, in tre casi su quattro, presentano una durata quasi doppia rispetto alla media delle altre nasali<sup>34</sup>.

Nella *Tabella III* sono riportate le medie e le deviazioni standard delle durate di nasali in posizione preconsonantica. La Fig. 14 permette di evidenziare come in entrambe le lingue si possa riscontrare una minor durata delle nasali prefricative, rispetto a quella delle preocclusive<sup>35</sup>.

Questo fatto può essere legato alle condizioni di anticipazione della frizione (visibile nell'energia di alta frequenza) rispetto ai tempi di chiusura velofaringale (come anticipato al §3.2) che può indurre una segmentazione che attribuisce una maggiore estensione alla fricativa. Questa anticipazione, oltre che sulla durata della consonante nasale stessa, sembra influire anche sulla nasalizzazione della vocale precedente (e questo, tanto più, quanto più la fricativa presenta energie e durate intrinseche maggiori; v. anche Fig. 19)<sup>36</sup>. Nel caso delle occlusive e delle affricate, come già anticipato in nota 26, potrebbero essere le nasali preocclusive a essere leggermente più lunghe per via della loro possibilità di estendersi, con

<sup>34</sup> Proprio a proposito di questo suono, si registra una curiosa divergenza significativa tra i due *speaker* iberici:

<sup>35</sup> Questo dato è in accordo con quanto osservato da Busà (2003). Le durate in quel caso riguardano però soprattutto i tempi di nasalizzazione vocalica o trans-segmentale in *VN*.

<sup>36</sup> Molto interessanti a questo proposito le riflessioni riportate in Farnetani & Busà (1994: 243). Nel caso di /nʃ/, all'ipotesi della salienza percettiva della fricativa che, in questa posizione, porta in italiano (e non solo) alla perdita storica della nasale, si contrappone quell'impossibilità dei due gesti necessari per l'articolazione di *N* e *C*, di coarticolarsi senza ledere la distintività della *C*.

Antonio Romano, Franca Mancini, Enrico Zovato - In P. Cosi (a cura di),  
 "MISURA DEI PARAMETRI - Aspetti tecnologici ed implicazioni nei modelli linguistici"  
 (Atti del I convegno naz. *AISV*, Padova 2004), Padova, EDK, 2005 (CD-ROM), 101-133.

alcuni strascichi, nella fase di tenuta della consonante seguente (inclusi dal segmentatore nel segmento nasale).

È inoltre interessante osservare la tendenziale maggior durata delle nasali spagnole in posizione preconsonantica (preocclusiva o prefricativa sorda), rispetto alla corrispondente durata delle nasali italiane (in parziale contraddizione con le ipotesi che in italiano vorrebbero più lunghe le consonanti eterosillabiche di coda in sillabe accentate).

Carmen (nasali misurate: 2665)			
	$\mu$	$\sigma$	no.
m	<b>82</b>	<b>14</b>	1007
n	<b>64</b>	<b>14</b>	1507
ɲ	<b>140</b>	<b>24</b>	151

Jorge (nasali misurate: 2855)			
	$\mu$	$\sigma$	no.
m	<b>65</b>	<b>11</b>	1044
n	<b>58</b>	<b>12</b>	1654
ɲ	<b>95</b>	<b>16</b>	157

Paola (nasali misurate: 4082)			
	$\mu$	$\sigma$	no.
m	<b>61</b>	<b>11</b>	1294
n	<b>47</b>	<b>11</b>	2404
ɲ:	<b>103</b>	<b>21</b>	192
mm	<b>104</b>	<b>18</b>	74
nn	<b>107</b>	<b>25</b>	118

Matteo (nasali misurate: 4251)			
	$\mu$	$\sigma$	no.
m	<b>66</b>	<b>13</b>	1403
n	<b>52</b>	<b>12</b>	2512
ɲ:	<b>121</b>	<b>20</b>	195
mm	<b>112</b>	<b>15</b>	57
nn	<b>105</b>	<b>27</b>	85

*Tabella II:* Schema riassuntivo delle durate di consonanti nasali in posizione intervocalica. Vengono indicate la media, la deviazione standard e la consistenza di ciascuna consonante per ciascuna voce nel *DB*.

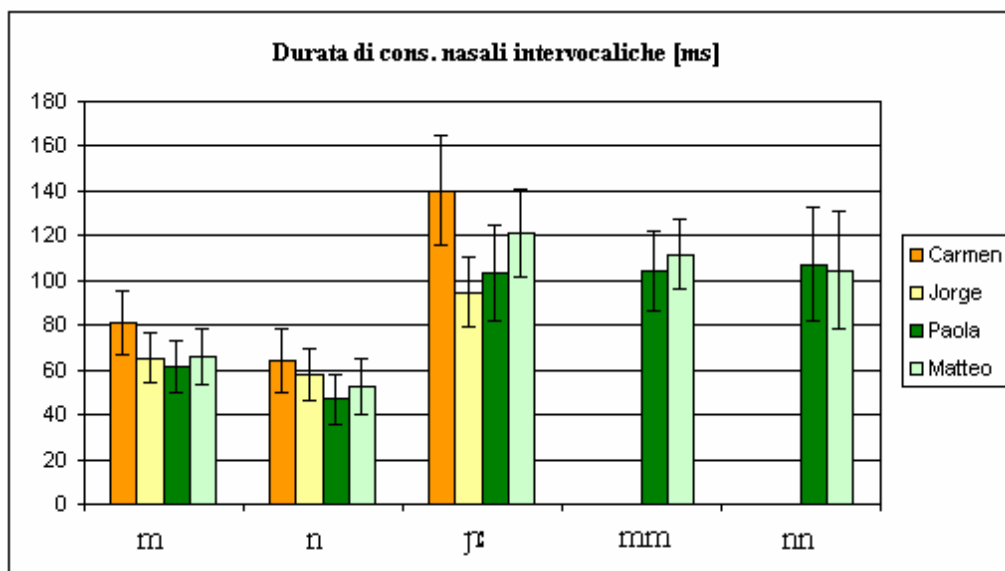


Figura 13. Durate medie di consonanti nasali intervocaliche ( $\pm$ dev.st.), calcolate sull'intera base di dati (13853 occorrenze) per le quattro voci.

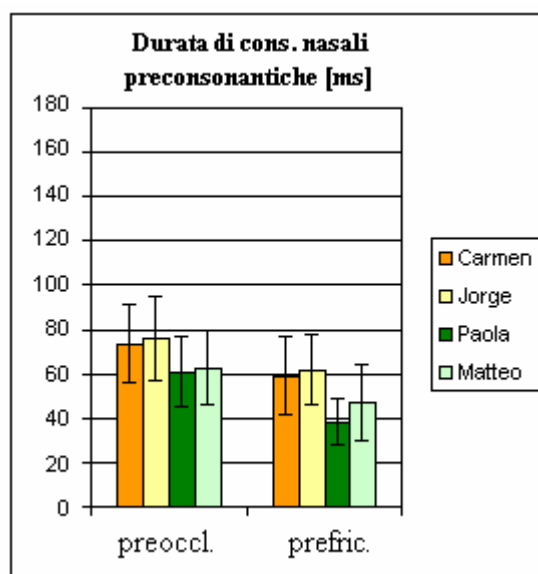


Figura 14. Durate medie di consonanti nasali preconsonantiche ( $\pm$ dev.st.), calcolate sull'intera base di dati (8372 occorrenze) e distinte, per le quattro voci, nelle due classi preocclusive e prefricative.

Antonio Romano, Franca Mancini, Enrico Zovato - In P. Cosi (a cura di),  
 “MISURA DEI PARAMETRI - Aspetti tecnologici ed implicazioni nei modelli linguistici”  
 (Atti del I convegno naz. *AISV*, Padova 2004), Padova, EDK, 2005 (CD-ROM), 101-133.

Carmen (nasali misurate: 2107)						
	preocclusiva			prefricativa		
	$\mu$	$\sigma$	no.	$\mu$	$\sigma$	no.
m	<b>72</b>	<b>15</b>	263	/	/	/
$\eta$	/	/	/	<b>61</b>	<b>21</b>	93
$\underline{\eta}/n$	<b>78</b>	<b>16</b>	1006	<b>69</b>	<b>18</b>	396
$\underline{n}/n^j$	<b>60</b>	<b>22</b>	29	/	/	/
$\eta$	<b>83</b>	<b>20</b>	284	<b>46</b>	<b>16</b>	36

Jorge (nasali misurate: 2026)						
	preocclusiva			prefricativa		
	$\mu$	$\sigma$	no.	$\mu$	$\sigma$	no.
m	<b>74</b>	<b>14</b>	236	/	/	/
$\eta$	/	/	/	<b>64</b>	<b>15</b>	94
$\underline{\eta}/n$	<b>74</b>	<b>16</b>	961	<b>60</b>	<b>15</b>	381
$\underline{n}/n^j$	<b>74</b>	<b>28</b>	29	/	/	/
$\eta$	<b>81</b>	<b>17</b>	290	<b>62</b>	<b>19</b>	35

Paola (nasali misurate: 2009)						
	preocclusiva			prefricativa		
	$\mu$	$\sigma$	no.	$\mu$	$\sigma$	no.
m	<b>61</b>	<b>14</b>	257	/	/	/
$\eta$	/	/	/	<b>39</b>	<b>11</b>	91
n	<b>61</b>	<b>15</b>	1022	<b>38</b>	<b>10</b>	186
$\underline{n}$	<b>60</b>	<b>18</b>	114	/	/	/
$n^j/\eta$	<b>61</b>	<b>16</b>	339	/	/	/

Matteo (nasali misurate: 2230)						
	preocclusiva			prefricativa		
	$\mu$	$\sigma$	no.	$\mu$	$\sigma$	no.
m	<b>62</b>	<b>16</b>	274	/	/	/
$\eta$	/	/	/	<b>48</b>	<b>15</b>	195
n	<b>58</b>	<b>17</b>	1102	<b>46</b>	<b>19</b>	99
$\underline{n}$	<b>66</b>	<b>19</b>	177	/	/	/
$n^j/\eta$	<b>65</b>	<b>17</b>	383	/	/	/

*Tabella III:* Schema riassuntivo delle durate di consonanti nasali in posizione preconsonantica (preocclusiva e prefricativa). Vengono indicate la media, la deviazione standard e la consistenza di ciascuna consonante per ciascuna voce nel *DB*.

## 5. CLASSIFICAZIONE DI NESSI *N*+FRICATIVA VS. *N*+AFFRICATA

### 5.1 L'energia di banda come parametro di discriminazione

Un altro aspetto indagato nella caratterizzazione acustica di consonanti nasali è quello della loro individuazione in termini di cambiamento di pendenza spettrale e di ridistribuzione della concentrazione energetica al passaggio da e verso foni adiacenti.

Studi pionieristici sulle caratteristiche spettrali dei suoni nasali hanno mostrato una notevole riduzione in ampiezza di tutte le componenti spettrali al di sopra di  $F_{1N}$  (Fant, 1960, descrive una pendenza spettrale di 1,6 dB/ottava). Tuttavia l'andamento complessivo dello spettro dipende fortemente dall'interazione tra cavità orali e cavità nasali e dall'apparizione di antiformanti che, in funzione del loro allineamento con i picchi di formante (a cui sono pur sempre associate da relazioni di accoppiamento, si veda per questo Hattori *et alii*, 1958; Fujimura, 1962) possono portare a depressioni significative molto localizzate come pure al rinforzo di concentrazioni energetiche in determinate aree di frequenza (si veda anche Ladefoged & Maddieson, 1996: 116). Ad ogni modo, queste configurazioni sono globalmente caratterizzate da una concentrazione in bassa frequenza che contrasta nettamente con le configurazioni degli altri suoni consonantici.

Per individuare la frontiera tra i foni, il principio è allora molto semplice e si basa sull'osservazione del cambiamento di distribuzione di energia spettrale al passaggio da una configurazione tipica di un suono nasale (Fig. 15, a sinistra) a quella tipica di uno fricativo (Fig. 15, a destra). Il principio non è nuovo: si tratta anzi di un metodo collaudato già a partire da Mermelstein (1977).

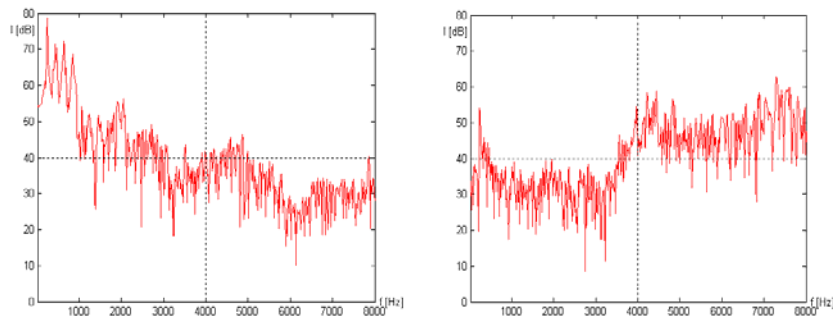


Figura 15: Distribuzioni di energia spettrale di una nasale e di una fricativa calcolate per *FFT* (finestre di 10 ms) nella porzione centrale dell'evoluzione temporale dei due suoni.

In Fig. 15, si nota facilmente come l'energia si distribuisca quasi specularmente per le due consonanti: per la nasale la quasi totalità dei contributi al di sopra dei 40 dB si situa al di sotto dei 4 kHz, mentre per la fricativa, le maggiori concentrazioni di energia si trovano nella banda frequenziale superiore.

In particolare, nel nostro caso, abbiamo fatto riferimento a due bande frequenziali 0÷4 kHz e 4÷8 kHz nello spettro dei suoni nasali e dei suoni del loro *entourage* e ne abbiamo studiato le variazioni temporali.



### 5.2 Calcolo dell'energia di banda

Per il calcolo dell'energia nelle due bande, in contesti fonetici particolari è stata utilizzata la procedura descritta nei punti seguenti:

1) Selezione dell'intervallo di segnale da analizzare. Sono stati considerati intervalli di 10 ms di segnale (160 campioni, con frequenza di campionamento pari a 16000 Hz), traslati tra di loro di 5 ms (80 campioni). Questa operazione di finestrazione è stata applicata ai primi 100 ms successivi al confine fonetico del contesto da analizzare e ai 100 ms precedenti lo stesso confine. In tutto, per ogni contesto, si sono avuti pertanto 40 valori di energia di banda. Nel caso di contesti di durata inferiore a 200 ms, sono stati impostati a zero tutti i valori di energia in eccesso presenti agli estremi del contesto.

2) Moltiplicazione del segnale selezionato per una finestra di Hamming di lunghezza  $WinLen$  pari a 160 campioni:

$$w(n) = 0.54 - 0.46 \cdot \cos(2\pi n / (WinLen - 1)) \quad n = 0 \dots WinLen - 1$$

3) Calcolo dell'energia nel dominio della frequenza mediante trasformata di Fourier, facendo riferimento al Teorema di Parseval. Per la trasformata, si è fatto uso di un algoritmo di *FFT* con decimazione nel tempo, radice di 2, su 256 punti. Nel caso di un segnale a tempo discreto il calcolo dell'energia è dato da:

$$E = \sum_{n=-\infty}^{\infty} |x[n]|^2 = \frac{1}{2\pi} \int_{-\pi}^{\pi} |X(\omega)|^2 d\omega$$

4) Pesatura dei punti in frequenza ovvero filtraggio in frequenza, mediante utilizzo di coefficienti moltiplicativi (impostati a 0 o 1). In questo caso specifico, posto  $WinLen = 160$  ed essendo  $X(k)$  la *FFT* di  $x(n)$  calcolata con  $FFTLen = 256$  si ha:

$$E_w = \sum_{n=1}^{WinLen} x_w[n]^2 = \frac{1}{FFTLen} \cdot \sum_{k=1}^{FFTLen} p_k \cdot |X(k)|^2$$

5) Calcolo dei valori di energia espressi in dB:

$$E_w' = 10 \cdot \log_{10}(E_w)$$

I risultati sono stati riassunti, per le quattro voci, e per tutti i contesti fonetici del *DB*, in diversi fogli di un *file* di *Excel*, all'interno dei quali è stato possibile selezionare e raggruppare i contesti, sottoponendo i dati a valutazioni grafiche sulla base di rappresentazioni come quelle nelle Figg. 16 e 17.

Si noterà, dal confronto delle Figg. 16 e 17, che la /s/ delle voci spagnole è generalmente più intensa di quella delle voci italiane<sup>37</sup>. Per inciso si noti come il rilascio

<sup>37</sup> Trasversalmente, presenta maggiori contributi in banda bassa per le voci maschili. Il fatto che l'articolazione sia prevalentemente apico-alveolare per gli *speaker* spagnoli non trova una corrispondenza sistematica in termini di concentrazioni nelle due bande considerate.

fricativo di /ts/, oltre a delle peculiarità dinamiche, presenti una minor energia complessiva rispetto a /s/. Un altro dato interessante riguarda la conferma di una minor energia complessiva delle fricative interdentali a cui è associato generalmente uno spettro assai piatto (cfr. Ladefoged & Maddieson, 1996: 175).

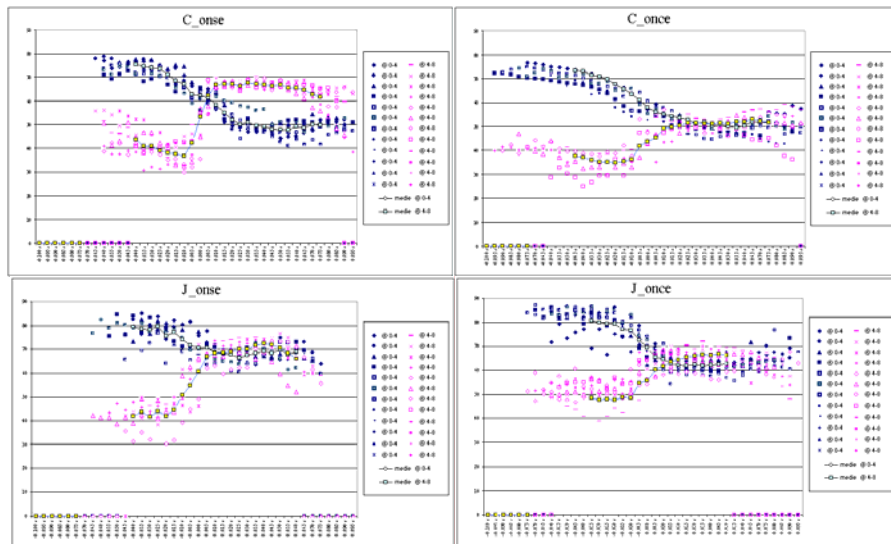


Figura 16: Andamenti delle energie di banda per gli *speaker* spagnoli (48 realizzazioni). Opposizioni /onse/ (a destra) ~ /onθe/ (a sinistra). Per la voce Carmen l'andamento segnala la presenza di un'ostruente epentetica (inattesa) solo nel primo caso (in alto a sinistra).

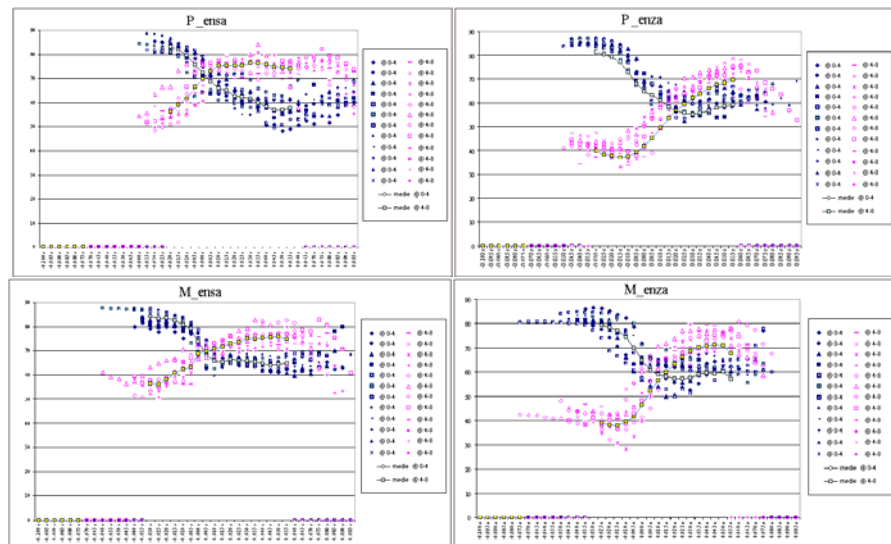


Figura 17: Andamenti delle energie di banda per gli *speaker* italiani (48 realizzazioni). Opposizioni /ensa/ (a destra) ~ /entsa/ (a sinistra). Indici che segnalano la presenza di un'affricazione (attesa) compaiono per le due voci solo nel secondo caso (grafici a destra).

In Fig. 16, osservando per la voce Carmen (in alto a sinistra) la caduta di energia di banda bassa, si possono notare degli indici che segnalano la presenza di un'ostruente epentetica (inattesa). Il dato è tanto più evidente se si confronta questo andamento con quello delle affricate in italiano. In Fig. 17, si può infatti verificare che gli indici che segnalano la presenza di un'affricazione (attesa) compaiono per le due voci solo nel secondo caso (grafici a destra).

Gli stessi dati possono essere visualizzati in un altro modo che ne rende forse più evidente l'utilizzabilità. I grafici in Fig. 18 rappresentano infatti gli andamenti di uno dei due parametri derivati (“Media di banda” e “Delta di banda”) a cui abbiamo fatto altresì ricorso. Si può constatare come un andamento concavo sia indicatore di una depressione solitamente associata a un'affricazione (in italiano) o all'apparizione di una maggiore ostruzione (eventualmente seguita da un salto energetico associato a condizioni esplosive) che possono creare le condizioni per la percezione di un'occlusiva epentetica (in spagnolo).

Se, per entrambi gli speaker, questo può dar luogo a sporadiche realizzazioni del tipo [ɲtθ], nel caso di Carmen, si conferma la presenza di un rilascio brusco che genera l'impressione dell'epentesi di un'occlusiva velare (del tipo [ɲ<sup>q</sup>ʂ]). Quest'esempio ci permette di osservare come siano anche gli indici specifici di luogo di articolazione della nasale che, interagendo con le proprietà acustiche della fricativa seguente, stabiliscono le caratteristiche dell'elemento esplosivo<sup>38</sup>.

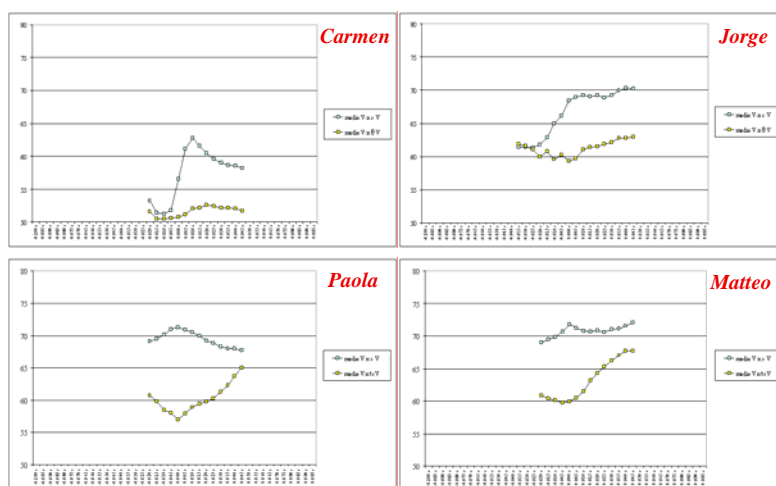


Figura 18: Grafici dell'andamento del parametro “Media di banda” che permettono di individuare il confine tra la nasale e la consonante seguente, distinguendo i casi di fricativa da quelli di affricata. In alto gli andamenti per /ns/ e /nθ/ dei locutori spagnoli. In basso quelli per /ns/ e /nts/ dei locutori italiani.

<sup>38</sup> La presenza di questi indici può essere sfruttata per una classificazione delle fricative dei vari nessi fonetici, come fricative pure o fricative preoccluse (una corretta etichettatura è però rilevante solo per un sistema di sintesi basato su *DB* insufficienti). D'altra parte, le conseguenze teoriche di questa considerazione si riflettono anche sulle valutazioni riguardanti quelle varietà di italiano caratterizzate dall'affricazione di /ns/ i cui casi potrebbero essere inseriti tra quelli più generali di epentesi (v. Ohala, 1997): l'elemento inserito sarebbe in quel caso di tipo apicale o predorsale perché quei parlanti avrebbero articolazioni non-velari per /n/ e predorso-alveolari per /s/.

Antonio Romano, Franca Mancini, Enrico Zovato - In P. Cosi (a cura di),  
“MISURA DEI PARAMETRI - Aspetti tecnologici ed implicazioni nei modelli linguistici”  
(Atti del I convegno naz. *AISV*, Padova 2004), Padova, EDK, 2005 (CD-ROM), 101-133.

Un'ultima considerazione che facciamo riguarda la minore o maggiore facilità con cui avviene la nasalizzazione della vocale precedente: oltre che alla posizione rispetto alle varie prominenze, il fenomeno ci sembra possa essere ricollegato al peso prosodico degli elementi post-nasali e/o alla risillabificazione di segmenti flottanti. Questi aspetti suggeriscono un'ulteriore ipotesi di lavoro per spiegare il motivo per cui la /i/ di *insperato* (v. Panconcelli-Calzia, 1904: 94-104) o di *inscenerà* (Fig. 19) presenti una maggior predisposizione a nasalizzarsi di quanto avvenga ad esempio per la /i/ di *vincerà* (Fig. 20).

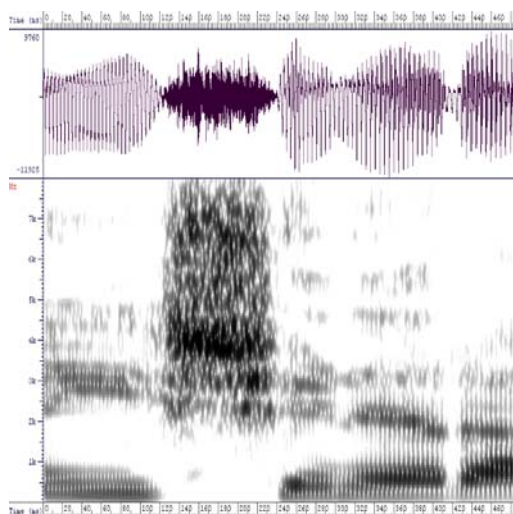


Figura 19: Esempio di perdita della nasale (con totale nasalizzazione della vocale precedente) nel caso di /nʃ/ della parola *inscenerà* pronunciata dalla voce Paola.

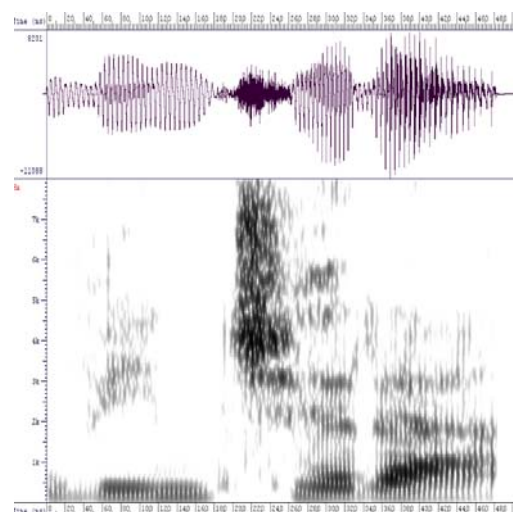


Figura 20: Esempio di mantenimento della nasale nel caso di /ntʃ/ della parola *vincerà* pronunciata dalla voce Paola.

Antonio Romano, Franca Mancini, Enrico Zovato - In P. Cosi (a cura di), "MISURA DEI PARAMETRI - Aspetti tecnologici ed implicazioni nei modelli linguistici" (Atti del I convegno naz. *AISV*, Padova 2004), Padova, EDK, 2005 (CD-ROM), 101-133.

## 6. CONCLUSIONI E PROSPETTIVE

In quest'articolo abbiamo cercato di riassumere alcuni dei risultati ottenuti nell'ambito di un nostro progetto di ricerca sulle consonanti nasali. Il lavoro è ancora in corso e si propone di approfondire le caratteristiche acustiche delle nasali presenti nella base di dati *Loquendo S.p.A.*.

D'altra parte, avendo definito una procedura operativa per la segmentazione e la classificazione dei nessi, e avendone verificato la validità su un numero limitato di casi, sono ancora da valutare le sue prestazioni in automatico e l'incidenza dei fenomeni descritti all'interno della base di dati considerata (inclusendo altre voci di *LoquendoTTS*<sup>®</sup>).

Nonostante la grande variabilità osservata, le nostre conclusioni sono in favore di una possibilità discriminativa di alcuni tassofoni nasali su base acustica, a condizione di procedere con un'attenta valutazione dei contributi energetici che si manifestano per le diverse voci nei diversi contesti.

Un altro aspetto indagato riguarda le realizzazioni fonetiche di nessi *N*+occlusiva, *N*+fricativa e *N*+affricata, riflettendo sulla necessaria cautela nella loro descrizione e categorizzazione in base alle reali caratteristiche acustiche che essi presentano. In particolare, ci siamo soffermati sulla discussione dei problemi di epentesi o di percezione di elementi occlusivi o esplosivi spurî.

Per trovare delle conferme strumentali nell'ambito di ricerche articolatorie, le nostre osservazioni avrebbero bisogno di essere estese su dati acustici raccolti simultaneamente a misure di parametri articolatori. In vista di una simile estensione, è attualmente in fase di sperimentazione un metodo di stima della nasalità vocalica in italiano basato sull'osservazione dell'*OQ* (*opening quotient*) dell'onda glottidale (per applicazione relative al portoghese si vedano i diversi contributi di Teixeira *et alii* in bibliografia). Una prima ricerca è stata condotta su dati registrati con il ricorso a un elettroglottografo (Kay Elemetrics, *EGG 6103*) collegato a un *CSL 4400*. I dati raccolti - parte di un mini-corpus di produzioni di due parlanti italo-foni - sono attualmente in fase di spoglio<sup>39</sup>.

## BIBLIOGRAFIA

- Albano Leoni, F. & Maturi, P. (1995), *Manuale di Fonetica*, Roma: La Nuova Italia Scientifica.
- Ali L., Daniloff R. & Hammarberg, R. (1979), Intrusive Stops in Nasal-Fricative Clusters: An Aerodynamic and Acoustic Investigation, *Phonetica*, 36, 85-97.
- Anderson, S.R. (1975), The Description of Nasal Consonants and the Internal Structure of Segments, in C. Ferguson, L. Hyman & J. Ohala (eds.), *Nasalfest: Papers from a symposium on nasal and nasalization*, Stanford, Language Universals Project, 1-26.
- Balestri, M., Pacchiotti, A., Quazza, S., Salza, P.L. & Sandri, S. (1999), Choose the best to modify the least: A new generation concatenative synthesis system, in *Proceedings of Eurospeech'99* (Budapest, 1991), 2291-2294.
- Barnitz, J.G. (1974), Bloom-p-field, Chom-p-sky, and phonetic open-t-thesis, *Studies in the Linguistic Sciences*, Vol. 4, no. 2, 1-13.

---

<sup>39</sup> Uno studio di varie configurazioni della forma d'onda glottidale è in Romano (1994).

- Antonio Romano, Franca Mancini, Enrico Zovato - In P. Cosi (a cura di), "MISURA DEI PARAMETRI - Aspetti tecnologici ed implicazioni nei modelli linguistici" (Atti del I convegno naz. *AISV*, Padova 2004), Padova, EDK, 2005 (CD-ROM), 101-133.
- Brugnara, F., Falavigna, D. & Omologo, M. (1993), Automatic Segmentation and Labeling of Speech based on Hidden Markov Models. *Speech Communication*, 12, 357-370.
- Busà, M.G. (2003), Vowel Nasalization and Nasal Loss in Italian, in *Proceedings of the 15<sup>th</sup> ICPHS* (Barcelona, 2003), paper 725.
- Calliope* (Tubach, J.P. ed.) (1989). *La parole et son traitement automatique*. Paris: Masson.
- Canepari, L. (1979), *Introduzione alla Fonetica*, Torino: Einaudi.
- Canepari, L. (1999), *Manuale di Pronuncia Italiana*, Bologna: Zanichelli.
- Chen, M.Y. (1997), Acoustic Correlates of English and French Nasalized Vowels, *Journal of the Acoustical Society of America*, 102, 2360-2370.
- Cosi, P., Falavigna, D. & Omologo, M. (1991), A preliminary statistical evaluation of manual and automatic segmentation discrepancies, in *Proceedings of Eurospeech'91* (Genova, 1991), 2, 693-696.
- Delattre, P., Liberman, A.M. & Cooper, F.S. (1955). Acoustic loci and transitional cues for consonants. *Journal of the Acoustical Society of America*, Vol. 27, no. 4, 769-773.
- D'Introno, F., Del Teso, E., & Weston, R. (1995), *Fonética y fonología actual del español*, Madrid: Cátedra.
- Dorman, M.F., Raphael, L. & Liberman, A.M. (1979), Some experiments on the sound of silence in phonetic perception, *Journal of the Acoustical Society of America*, 65, 6, 1518-1532.
- Fant, G. (1960), *Acoustic Theory of Speech Production*, The Hague: Mouton.
- Farnetani, E. (1979a), Aerodinamica della nasalizzazione, *Rivista Italiana di Acustica*, Vol. III, no. 1, 5-22.
- Farnetani, E. (1979b), Foni nasali e nasalizzazione, *Acta Phoniatica Latina*, 1, 30-57.
- Farnetani, E. (1986), A Pilot-Study of the Articulation of /n/ in Italian Using Electropalatography and Airflow Measurements, in *Actes des 15<sup>èmes</sup> Journées d'Étude sur la Parole du GALF* (Aix-en-Provence, 1986), 23-26.
- Farnetani, E. & Busà, M.G. (1994), Consonant-to-consonant interactions in clusters: Categorical or continuous processes? *Quaderni del Centro di Studi per le Ricerche di Fonetica del CNR*, XIII, 219-245.
- Feng, G. (1987), Étude articulatoire-acoustique des voyelles nasales du français, *Bulletin de l'Institut de Phonétique de Grenoble* (Institut de la Communication Parlée), 16, 1-102.
- Feng, G. & Castelli, E. (1996), Some acoustic features of nasal and nasalized vowels: a target for vowel nasalization, *Journal of the Acoustical Society of America*, 99, 3694-3706.
- Fernández, H., Fernández, N., Garrido, J.M. & Prieto, P. (1988), Análisis acústico de las consonantes nasales castellanas, Universidad Autónoma de Barcelona (inédito, cit. in Machuca).
- Fernández, H., Fernández, N., Garrido, J.M. & Prieto, P. (1989), Análisis y síntesis de las consonantes nasales en español, *Revista Española de Lingüística*, Vol. 19, no. 1, 145.

Antonio Romano, Franca Mancini, Enrico Zovato - In P. Cosi (a cura di), "MISURA DEI PARAMETRI - Aspetti tecnologici ed implicazioni nei modelli linguistici" (Atti del I convegno naz. *AISV*, Padova 2004), Padova, EDK, 2005 (CD-ROM), 101-133.

- Ferrero, F. & Magno Caldognetto, E. (1986), "Elementi di Fonetica Acustica", in L. Croatto (a cura di), *Trattato di foniatria e logopedia. Aspetti fonetici della comunicazione*, Vol. III, Padova: La Garangola, 155-196.
- Ferrero, F., Genre, A., Boë, L.J. & Contini, M. (1979), *Nozioni di Fonetica Acustica*, Torino: Ed. Omega.
- Flanagan, J. L. (1965), *Speech Analysis Synthesis and Perception*. Berlin: Springer.
- Fourakis, M. & Port, R. (1986), Stop epenthesis in English, *Journal of Phonetics*, 14, 197-221.
- Fujimura, O. (1962), Analysis of nasal consonants, *Journal of the Acoustical Society of America*, 34, 1865-1975.
- Giannini, A. & Pettorino, M. (1992), *La Fonetica Sperimentale*. Napoli: Edizioni Scientifiche Italiane.
- Giannini, A., Maturi, P. & Pettorino, M. (1990), Il ruolo della nasalità nella fonologia dell'Italiano, in *Atti del XVIII Conv. Naz. dell'AIA* (L'Aquila, 1990), 191-196.
- Hajek, J. (1997), *Universals of Sound Change in Nasalization*, Oxford: Blackwell.
- Harding, S. & Meyer, G. (2003), Changes in the perception of synthetic nasal consonants as a result of vowel formant manipulations, *Speech Communication*, 39, 173-189.
- Harrington, J. (1994), The contribution of the murmur and vowel to the place of articulation distinction in nasal consonants, *Journal of the Acoustical Society of America*, Vol. 96, no. 1, 19-32.
- Hattori, S., Yamamoto, K. & Fujimura, O. (1958), Nasalization of Vowels in Relation to Nasals, *Journal of the Acoustical Society of America*, 30, 267-274.
- House, A.S. (1957), Analog Studies of Nasal Consonants, *Journal of Speech and Hearing Disorders*, 2, 190-204.
- Jones, D. (1950), *The Phoneme, its nature and use*, Cambridge: Heffer.
- Kurowski, K. & Blumstein, S. (1984), Perceptual integration of the murmur and formant transitions for place of articulation in nasal consonants, *Journal of the Acoustical Society of America*, Vol. 76, no. 2, 383-390.
- Kurowski, K. & Blumstein, S.E. (1987), Acoustic properties for place of articulation in nasal consonants, *Journal of the Acoustical Society of America*, Vol. 81, no. 6, 1917-1927.
- Lacerda, A., de & Companys, M. (1958), Notes pour l'étude de l'activité nasale en français, *Revista do Laboratório de Fonetica Experimental*, Univ. de Coimbra, IV, 183-205.
- Ladefoged, P. & Maddieson, I. (1996), *The Sounds of the World's Languages*, Oxford: Blackwell.
- Liberman, A.M., Delattre, P.C., Cooper, F.S. & Gerstman, L.J. (1954), The role of consonant-vowel transition in the perception of the stop and nasal consonants. *Psychol. Monograph*, 8/379, 1-14 (anche in D.B. Fry (a cura di) (1976), *Acoustic Phonetics: a course of basic readings*, Cambridge: CUP).

Antonio Romano, Franca Mancini, Enrico Zovato - In P. Cosi (a cura di),  
"MISURA DEI PARAMETRI - Aspetti tecnologici ed implicazioni nei modelli linguistici"  
(Atti del I convegno naz. *AISV*, Padova 2004), Padova, EDK, 2005 (CD-ROM), 101-133.

- Machuca, M.J. (1991a), *Estudio de consonantes nasales del español en habla espontánea y en habla de laboratorio*, Proyecto de Investigación de Tercer Ciclo, Dep. de Filología Española, Facultad de Letras, Universidad Autónoma de Barcelona.
- Machuca, M.J. (1991b), Acoustic description of the Spanish nasal consonants in continuous speech, in *Proceedings of the 12<sup>th</sup> ICPHS* (Aix-en-Provence, 1991), 2, 414-417.
- Maeda, S. (1993), Acoustics of Vowel Nasalization and Articulatory Shifts in French Nasal Vowels, in M. K. Huffman & R. Krakow (eds.), *Nasals, Nasalization, and the Velum*, San Diego: Academic Press, 147-167.
- Malécot, A. (1956), Acoustic Cues for Nasal Consonants: an Experimental Study Involving Tape-splicing Technique, *Language*, 32, 274-284.
- Malécot, A. (1960), Vowel Nasality as a Distinctive Feature in American English, *Language*, 36, 222-229.
- Massone, M.I. (1988), Estudio acústico y perceptivo de las consonantes nasales y líquidas en español, *Estudios de Fonética Experimental*, III, Barcelona, 13-34.
- Maturi, P. (1991), The perception of consonant nasality in Italian: conditioning factors, in *Proceedings of the 12<sup>th</sup> ICPHS* (Aix-en-Provence, 1991), 50-53.
- Mermelstein, P. (1977), On Detecting Nasals in Continuous Speech, *Journal of the Acoustical Society of America*, Vol. 61, no. 2, 581-587.
- Mioni, A.M. (1973), *Fonematica contrastiva*, Bologna: Patron.
- Miotti, R. (1998), Descrizione fono-tonetica delle varietà regionali dello spagnolo d'America e di Spagna, *Annali di Ca' Foscari (Rivista della Fac. di Lingue e Lett. Str. dell'Univ. di Venezia)*, Vol. XXXVII, nn. 1-2, 393-440.
- Nakata, K. (1959), Synthesis and Perception of Nasal Consonants, *Journal of the Acoustical Society of America*, Vol. 31, no. 6, 661-666.
- Nord, N. (1976), Perceptual Experiments with Nasals, *Quarterly Progress Status Report*, Stockholm: Royal Institute of Technology, Speech Transmission Laboratory, 2-3.
- Ohala, J.J. (1975), Phonetic explanation for nasal sound patterns, in C. Ferguson, L. Hyman & J. Ohala (eds.), *Nasalfest: Papers from a symposium on nasal and nasalization*, Stanford, Language Universals Project, 289-316.
- Ohala, J.J. (1997), Emergent stops, in *Proceedings of the Seoul International Conference on Linguistics*, Seoul: Linguistic Society of Korea, 84-91.
- Ohala, J.J. & Busà, M.G. (1995), Nasal Loss Before Voiceless Fricatives: A Perceptually-based Sound Change, *Rivista di Linguistica*, Vol. 7, no. 1, 125-144.
- Ohala, J.J. & Ohala, M. (1993), The Phonetics of Nasal Phonology: Theorems and Data, in M.K. Huffman & R.A. Krakow (eds.) *Phonetics and Phonology, vol. 5 - Nasals, Nasalization, and the Velum*, San Diego: Academic Press, 225-249.
- Ohde, R.N. (1994), The development of the perception of cues to the [m]-[n] distinction in CV syllables, *Journal of the Acoustical Society of America*, Vol. 96, no. 2, 675-686.



- Antonio Romano, Franca Mancini, Enrico Zovato - In P. Cosi (a cura di), "MISURA DEI PARAMETRI - Aspetti tecnologici ed implicazioni nei modelli linguistici" (Atti del I convegno naz. *AISV*, Padova 2004), Padova, EDK, 2005 (CD-ROM), 101-133.
- Panconcelli-Calzia, G. (1904), De la nasalité en italien, *Thèse présentée pour le doctorat de l'Univ. de Paris* (Lettres), Inst. de Laryngologie et Orthophonie.
- Pettorino, M. (1988), Nasality: segmental or suprasegmental feature?, in *Speech'88, Proceedings of the 7<sup>th</sup> FASE Symposium* (Edinburgh, 1988), 105-112.
- Quilis, A. (1981), *Fonética Acústica de la Lengua Española*, Madrid: Gredos.
- Recasens, D. (1983), Place Cues for Nasal Consonants with special reference to Catalan, *Journal of the Acoustical Society of America*, 73, 1346-1353.
- Repp, B.H. & Svastikula, K. (1988), Perception of the [m]- [n] distinction in VC syllables. *Journal of Acoustical Society of America*, Vol. 83, no. 1, 237-247.
- Romano, A. (1994), Analisi della funzione di trasferimento del tratto vocale negli intervalli di chiusura della glottide ed estrazione dell'eccitazione glottidale, Tesi di Laurea in Ing. Eln.- Politecnico di Torino - *Institut de la Communication Parlée* di Grenoble.
- Romano, A. (1997), Misure acustiche di consonanti nasali in due locutori salentini, (Ms. inedito, basato sul materiale raccolto nell'ambito della Tesi di Dottorato in Romano 2001).
- Romano, A. (2001), *Analyse des structures prosodiques des dialectes et de l'italien régional parlés dans le Salento: approche linguistique et instrumentale*, Lille: Presses Univ. du Septentrion.
- Romano, A. (2002), La fonetica strumentale applicata ai dialetti d'Italia a un secolo dall'"Etude sur la phonétique italienne" di F.M. Josselyn, in A. Regnicoli (a cura di), *La fonetica acustica come strumento di analisi della variazione linguistica in Italia, Atti delle XII Giornate di Studio del GFS* (Macerata, 2001), Roma: Il Calamo, 7-14.
- Rossato, S., Badin, P. & Feng, G. (2000), Détermination de la position du voile du palais à partir du signal de parole pour les nasales du français, in *Actes des XXIII<sup>èmes</sup> Journées d'Étude sur la Parole* (Aussois, 2000), 137-140.
- Schourup, L. (1973), A Cross-Language Study of Vowel Nasalisation, *Working Papers in Linguistics*, 15, 190-221.
- Solé, M.J. (1992), Phonetic and Phonological Processes: the case of nasalization, *Language and Speech*, 35, 29-43.
- Straka, G. (1955), Remarques sur les voyelles nasales, leur origine et leur évolution en français, *Revue de Linguistique Romane*, 19, 245-274.
- Teixeira, A., Castro Moutinho, L. & Coimbra, R.L. (2001), Acerca das vogais nasais do Português Europeu, *Revista da Universidade de Aveiro - Letras*, 18, 241-274.
- Teixeira, A., Castro Moutinho, L. & Coimbra, R.L. (2003), Production, Acoustic and Perceptual Studies on European Portuguese Nasal Vowels Height, in *Proceedings of the 15<sup>th</sup> ICPhS* (Barcelona, 2003).
- Teixeira, A., Vaz, F. & Príncipe, J.C. (1999a), Effects of Source-Tract Interaction in Perception of Nasality, in *Proceedings of Eurospeech'99* (Budapest, 1999), 161-164.
- Teixeira, A., Vaz F. & Príncipe J.C. (1999b), Influence of Dynamics in the Perceived Naturalness of Portuguese Nasal Vowels, in *Proceedings of the 14<sup>th</sup> ICPhS'99* (San Francisco, 1999).

Antonio Romano, Franca Mancini, Enrico Zovato - In P. Cosi (a cura di),  
"MISURA DEI PARAMETRI - Aspetti tecnologici ed implicazioni nei modelli linguistici"  
(Atti del I convegno naz. *AISV*, Padova 2004), Padova, EDK, 2005 (CD-ROM), 101-133.

Turchi, L. & Gili Fivela, B. (2004), L'affricazione di /s/ postconsonantico nella varietà  
pisana di italiano, in F. Albano-Leoni, F. Cutugno, M. Pettorino & R. Savy (a  
cura di), *Il Parlato Italiano* (Atti del Conv. Naz. di Napoli, 2003), Napoli:  
D'Auria (CD-ROM, art. A06).

Vaggies, K., Ferrero, F.E., Magno Caldognetto, E. & Lavagnoli, C. (1978), Some Acoustic  
Characteristics of Italian Consonants, *Journal of Italian Linguistics*, 3, 69-85  
(articolo presentato allo 8<sup>th</sup> *ICPhS*, Leeds, 1975, preprint 23 pp.).

Warner, N. & Weber, A. (2001), Perception of epenthetic stops, *Journal of Phonetics*, 29,  
53-87.

Wells, J.C. (1997), Whatever happened to Received Pronunciation?, in *Actas de las II  
Jornadas de Estudios Ingleses* (Jaén, 1997), 19-28.