An MRI Study on the Articulatory Properties of Italian Consonants

(Article begins on next page)

22 May 2024

AN *MRI* STUDY ON THE ARTICULATORY PROPERTIES
OF ITALIAN CONSONANTS

ANTONIO ROMANO
Dept Scienze del Linguaggio, LFSAG - Università degli Studî di Torino, Italy
antonio.romano@unito.it

PIERRE BADIN
Dept Parole & Cognition, GIPSA-lab, INPG, Grenoble, France
pierre.badin@gipsa-lab.inpg.fr

ABSTRACT

MRI (*Magnetic Resonance Imaging*) data have been collected for three male speakers of Italian producing sustained consonants in VCV-context (with V=/{i, a, u}/). For one speaker the resulting database consists of a midsagittal set of 42 Italian articulations (/p/, /f/, /t͡s/, /t/, /s/, /ʃ/, /t͡ʃ/, /k/, /ʎ/, /l/, /r/, /m/, /n/ and /ɲ/) plus 9 dialectal sound configurations and 12 scans for specific nasal allophones. It is associated with vowels, jaw and teeth references and dental casts. A subset of images is also available, however, for a limited selection of articulations produced by two other control speakers. It has been collected in view of dialect studies and it includes midsagittal and coronal scans. As in previous partial publications of these data, the aim of the present paper is to discuss only place and manner of articulation in a descriptive framework.

Keywords: *articulatory phonetics, Italian consonants, Magnetic Resonance Imaging.*

RESUMEN

Los datos de *MRI* (*Magnetic Resonance Imaging*) han sido reunidos a partir de consonantes sostenidas, emitidas por tres hablantes de italiano masculinos, en un contexto de VCV (donde V = /{i, a, u}/). Para cada emisor la base de datos resultante consta de un conjunto midsagital de 42 articulaciones en italiano (/p/, /f/, /t͡s/, /t/, /s/, /ʃ/, /t͡ʃ/, /k/, /ʎ/, /l/, /r/, /m/, /n/ y /ɲ/) junto con 9 configuraciones fónicas dialectales y 12 registros sobre ciertos alófonos nasales. Se asocia con vocales, tomando referencias dentales, maxilares y dentales. Dispone asimismo de un grupo de imágenes para una selección limitada de articulaciones producidas por dos hablantes controles. Todo ello está orientado hacia los estudios dialectales y comprende muestras midsagitales y coronales. Al igual que en publicaciones monográficas anteriores, la finalidad de este artículo consiste en analizar solo el lugar y el modo de articulación en un marco descriptivo.

Palabras clave: *fonética articulatoria, consonantes del italiano, Magnetic Resonance Imaging.*

1. INTRODUCTION

The articulatory properties of Italian consonants (both the peculiar ones and those common to other languages) are well known thanks to various organic instrumental surveys (such as the radiographic one by L. Croatto in Tagliavini, 1965) and to scattered studies which appeared in different places and at different times (cp. Magno Caldognetto, 1988; we shall quote among the numerous palatographic works Farnetani, 1986).

At present, however, among the various techniques used for the investigation of the articulatory characteristics of Italian sounds, *MRI* (*Magnetic Resonance Imaging*) seems not to have raised much interest, yet.

Despite the initial scepticism towards its use in the phonetic field (see the list of advantages and drawbacks compiled by Stone, 1997), nowadays this technique for the acquisition of articulatory information, which uses magnetic fields and radio waves for the representation of anatomical and physiological characteristics either of the vocal tract or of the brain during the production of sounds, is gaining international consensus as for its use in descriptive or experimental linguistic research.

The distrust of this type of surveys in the field of articulatory phonetics – partly related to their being extremely difficult to get and to their questioned representativeness – seems to be over at last thanks to recent progress which has allowed to overcome certain procedural *impasses* and to improve the acquisition and presentation of the results[1].

These improvements have enhanced the use of *MRI*, which has been common since the early 90's for the measurement of the volumes and the forms of the resonators and for the observation of the position of the mobile articulators in the static configurations they assume during the articulation of a sound or for the reconstruction of the movements from one configuration to another, in experimental linguistics.

---

[1] Among the most quoted drawbacks, some are related to the temporal and spatial integration with which images are obtained. In the first applications, the acquisition time for a good resolution ranged from 3 to 20 seconds, but the most recent *MRI scanners* only take 50 ms per scan. As for the width of the section (which initially was in the order of 1 cm), scans (at present 5 mm wide) can now be obtained by integrating 1.5 mm wide sections.

One of the drawbacks still at issue, though often overcome with the use of mobile magnets, is the fact that the informer has to assume either a prone or a supine position. However, a satisfactory answer has been given to these objections in dynamic terms or in terms of compensation (cp. for these aspects Tiede *et alii*, 2000, Kitamura *et alii*, 2005, Kedrova *et alii*, 2006)[2]. Considering that *MRI* is the only technique capable of providing images in transversal (coronal) section, and keeping in mind the improving acquisition speed and the wider spatial resolution achieved in the last years, it is easy to understand the growing use of *MRI* for the acquisition of information on the articulatory characteristics during the production of sounds in replacement for the traditional radiographic images[3].

Even though in the international literature this technique is often used to investigate specific phenomena – by applying sophisticated computational and/or volumetric techniques, e.g. for research in coarticulation, vocal tract estimation and inversion[4] –, in this contribution, situated outside advanced experimental approaches, the observation of available scans has the aim of suggesting an objective evaluation of the modes and places of articulation merely for descriptive purposes.

2. DATA COLLECTION

The Magnetic Resonance (*MR*) images dealt with in this study have been acquired at the Regional University Hospital (CHRU) of Grenoble, France, in 2001 and at the Radiodiagnostic Service of the "Molinette" Hospital of Turin, Italy, in 2004 and in 2008.

Even though data have been collected for three speakers, the main database discussed here is based on productions by the author AR. The reference to a

[2] In conditions of immobility, because of the force of gravity, the tongue shows a tendency to concentrate in the rear region of the vocal tract while the uvula shows a tendency to lower its position. Also the setting of the larynx is modified in the supine position because of the lack of gravity on the abdominal organs of the expiratory system which generally keep the larynx in a lower position. Finally, the orientation of the head limits the movement possibilities of the jaw because of the reduced space (v. Kitamura *et alii*, 2005).

[3] The fact that studies of this type have increasingly gained ground is witnessed by the works by Narayanan *et alii* (1997, 1999, 2004).

[4] For further references see e.g. Story *et alii* (1996), Kröger *et alii* (2000), Birkholz & Kröger (2006), Proctor *et alii* (2006).

limited number of subjects usually allows to simplify the data collection and analysis, but may result in the description of non-representative, speaker-specific conditions. The advantages are, however, relevant since the corpus size can be increased significantly and data quality better verified. Moreover, when using a known speaker easily available, the possibility to obtain natural speech also provides a reference that can be used when verifying vocal tract shapes. Larger possibilities of successfully combining data from different acquisition methods are also offered in these conditions.

2.1. Acquisition procedures

The scans were realised in three one-two hour sessions with a 1 Tesla *MRI* Scanner "Philips GyroScan T10-NT"[5]. The midsagittal images had a size of 256x256 pixels and a final resolution of 1 mm/pixel. They were acquired to cover the largest vocal tract length possible, with maximally lowered larynx and maximally protruded lips[6].

In all three cases, the subjects lay in supine position in the *MR* machine with their head inside a Radio Frequency (*RF*) coil. A padded crane support was used in the *RF* coil to minimise head movements. Therefore, the subjects' heads were not fixed, but movements were limited by the crane support and the coil provided a reference frame for keeping the head in position (see Engwall & Badin, 2000).

The acquisition time was about 10 seconds for the first set and approximately 4 seconds for the subsets. During this time, the subjects held the articulation in full apnoea or breathing out very slowly (fricatives). For stops and affricates (but also other full contact sounds) the scans were carried out during the contact phase[7].

---

[5] The type of equipment used is similar to that illustrated in Kitamura *et alii* (2005) and which has been object of evaluation by Tiede *et alii* (2000). In particular, a detailed description of the technical specifications has been discussed by Engwall & Badin (2000) who obtained their scans with the same equipment and in the same conditions as the data illustrated here.

[6] In the third subset, the images were taken as slices, 3.6 mm thick, and sampled every 4.0 mm. As first proposed by Demolin *et alii* (1996) and then systematically used by Apostol *et alii* (1999) and Engwall & Badin (2000), the acquisition was carried out using stacks of parallel slices.

[7] The long acquisition time is of course a drawback, as the subject could not phonate naturally during image acquisition. It is however lower than in earlier studies (cp. Stone,

2.2. Corpus

As stated above, we analysed three static *MRI* sets of native Italian male speakers: a full set of 70 midsagittal scans for AR (33 years), a speaker coming from a South-Eastern region (even though without any specific diatopic marks); a partial corpus of 6 scans (midsagittal and coronal) for GM (65 years), a speaker coming from a Northern region (also investigated for his dialectal palatal articulation); and a partial corpus of 39 scans (midsagittal and coronal) for FG (24 years), a speaker coming from Southern Italy (also investigated for his dialectal cacuminal articulation). All the corpus contains images of sustained sounds.

The first one is intended to allow for a complete and fine-grained description of place and manner of /p/, /f/, /t͡s/, /t/, /s/, /ʃ/, /t͡ʃ/, /k/, /ʎ/, /l/, /r/, /m/, /n/ and /ɲ/, but it includes vowels, jaw and teeth references[8]; it also includes 9 scans related to cacuminal and prepalatal specific dialectal sounds and 12 scans related to nasal combinatory variants[9].

The second corpus contains one midsagittal and one coronal *MRI* slices scanned for three articulatory places (postalveolar, palatal, velar)[10]. The third corpus contains 7 sagittal and 6 coronal *MRI* slices scanned during the articulation of three affricates (dental, cacuminal, postalveolar). Whenever it was possible, voiceless articulations were observed for each articulatory place. Furthermore, all the scans for each consonant (C) were acquired during its production in the utterances /ˈaCː#ˈCa/, /ˈiCː#ˈCi/ and /ˈuCː#ˈCu/.

Original scans for the /ˈaCː#ˈCa/ sequence for speaker AR (with /a/ > [a]) are shown in figures 1-6[11].

1997). It was also considered better to scan the entire vocal tract during one sustained production than having the subject repeat the articulation several times as done with other techniques (see for instance Narayanan *et alii*, 1995, 1997; Kim *et alii*, 2005).

[8] These scans were used to obtain the reference configurations with upper and lower incisors aligned and touching.

[9] Scans about postalveolar and prepalatal articulations have already been used for partial publication in Romano (2002) and Romano (2007).

[10] These scans have been discussed in Molino & Romano (2004) and Romano (2007).

[11] Except for the expected velar stop and palatal lateral approximant (see §3). Fairly different contoids realised /k/ in the three contexts. The articulation of the /ˈaCː#ˈCa/

2.3. Outline tracing and final verification

The midsagittal airway boundaries of all the scans were hand-traced on the computer. Unfortunately the manual tracing was carried on without using any edge detector. This required careful evaluations and had as a consequence a longer processing time. The different profiles for the same consonant were superimposed as shown in figures 7-12.

Plaster casts of the upper and lower jaws and teeth of speaker AR were also taken but they have not been used at this preliminary stage. Outside a careful verification on the basis of proprioceptive contrasts (even though objectivity reserves could be raised), interspeaker comparisons were carried out for common articulations and lead to the description sketched below.

3. PLACES OF ARTICULATION

In the selection of images presented here, we observed a particularly systematic use of the following places of articulation (see Figures 1-6)[12]:

1. the *bilabial*, in the realisations of /p/[13];

---

utterance for /ʎ/, as well as a few other captures for /ˈiCː#ˈCi/ and /ˈuCː#ˈCu/ sequences, was not well kept during scanning and produced soft-focus shapes.

[12] It should be noted that consonants with different mode of articulation are possible for every place of articulation, sometimes with voice distinctions. Where possible, images have been obtained during the articulation of voiceless consonants. However, when a voiceless sound in a certain place of articulation was not present in the Italian phonological system, the original image refers to the articulation of a voiced consonant. In the description of the places of articulation observed, which might at times be exceedingly detailed, we have considered the data illustrated in these images. Aiming at a generalisation, it is nevertheless necessary to take into account a certain degree of individual variability.

[13] Italian /b/, /m/ and, partly, /w/ (voiced labial–velar approximant) are articulated in this region as well. It should be noted that a component of (bi-)labiality is present in the labialised allophones of many other sounds, among which [sʷ] and [zʷ] in the realisations of /s/ and /z/ are particularly evident (see a greater labiality in /VCV/ context when V= /u/ in Figure 9-11). Besides, the protrusion and lip rounding are also generally significant for the articulation of /ʃ/, /t͡ʃ/ and /d͡ʒ/ (cp. Proctor *et alii*, 2006; see also Fig. 3 – right image).

2. the *labio-dental*, in the realisations of /f/[14];

3. the *alveodental*: with low tip, as in the realisations of /t͡s/[15] or more apico-dental, in the common realisations of /t/[16]; as well as alveolar (with high tip), in the realisations of /n/ and /l/; lamino-alveolar, in the realisations of /s/[17]; apico-alveolar (more or less advanced), in the realisations of /r/[18];

4. the *(post)alveolar*, in the realisations of /ʃ/ and /t͡ʃ/[19];

5. the *palatal*, in the realisations of /ɲ/ and /ʎ/[20];

6. the *velar*, in the realisations of /k/[21].

If we refer to superimposed outline tracings of the realisations in the three contexts (Figures 7-12), beside the position of the jaw (occasionally different), we can observe evident coarticulatory effects in the sporadic greater raising of the tongue dorsum towards the velum if V=/u/ (e.g. for /p/ or /s/) or in the greater predorsal convexity present for /t/, /t͡ʃ/ or /l/ if V=/i/.

---

[14] /v/ and the allophone [ʋ] of /w/ are articulated in this place as well. It is also possibile to remark a greater (bi-)labiality for /f/ in /VCV/ when V= /u/ (in Fig. 7 – right image).

[15] /d͡z/ and the allophone [n̪] of /n/ and [l̪] of /l/ are articulated in this place as well.

[16] /d/ is also articulated in this place.

[17] /z/ is articulated in this place as well. Even though this cannot be verified by the data presented here, a concave longitudinal configuration of the tongue appears in the articulation of these fricatives and produces a more evident transversal groove in the case of /ʃ/.

[18] Tap realisations are articulated in a similar manner.

[19] /d͡ʒ/ and the allophone [n̠] of /n/ are articulated in this place as well.

[20] /j/ is also articulated here. It can be noted that a prepalatal place is not uncommon in Italian in the articulation of *palatalised* realisations of velar stops /k/ and /g/ (and of the preceding nasal, [ŋ], if any) which commonly, before /i/ or /j/, extend ([kʲ], [gʲ] and [ŋʲ]), advance ([k̟], [g̟] and [ŋ̟]), or totally change their articulation place to a prepalatal one ([ç], [ʝ] and [ɲ̟]). This coarticulation phenomenon is the most evident in our data (see Fig. 12).

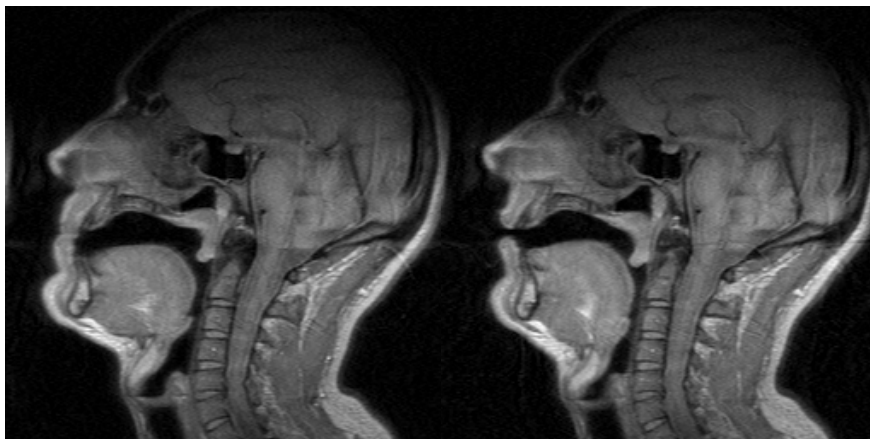[21] /g/, the allophone [ŋ] of /n/ and partly /w/ are articulated in this place too.

Figure 1. *MRI scans for C in an utterance of* /ˈaCːˈCa/ *with C articulated in the* bilabial *place for a realisation of* /p/ *(left) and in the* labio-dental *place for a realisation of* /f/ *(right).*
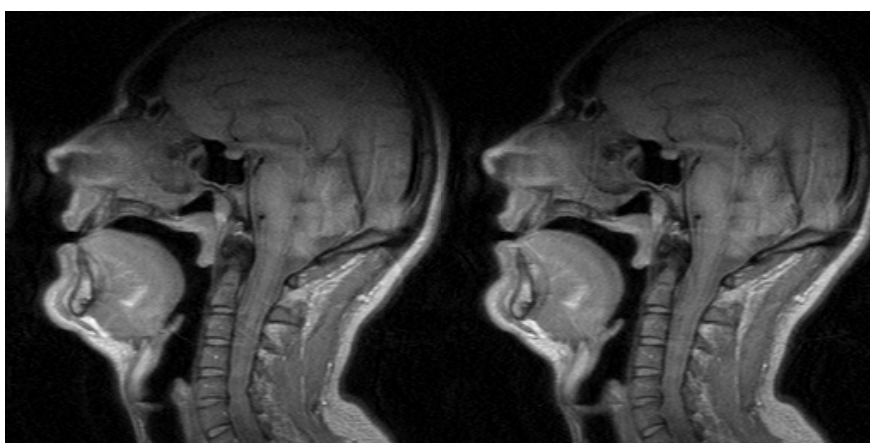

Figure 2. *MRI scans for C in an utterance of* /ˈaCːˈCa/ *with C articulated in the* alveodental *place for a realisation (with low tip) of* /t͡s/ *(left) and for a similar (perhaps more apical) realisation of* /t/ *(right).*
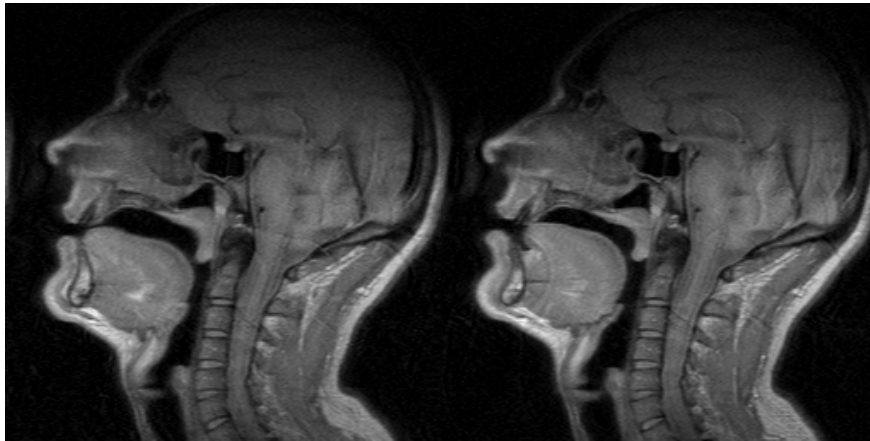
Figure 3. *MRI scans for C in an utterance of* /ˈaCː#ˈCa/ *with C articulated in a* alveodental *place for a lamino-alveolar realisation of* /s/ *(left) and in a* (post)alveolar *place for a labialised realisation of* /t͡ʃ/ *(right).*
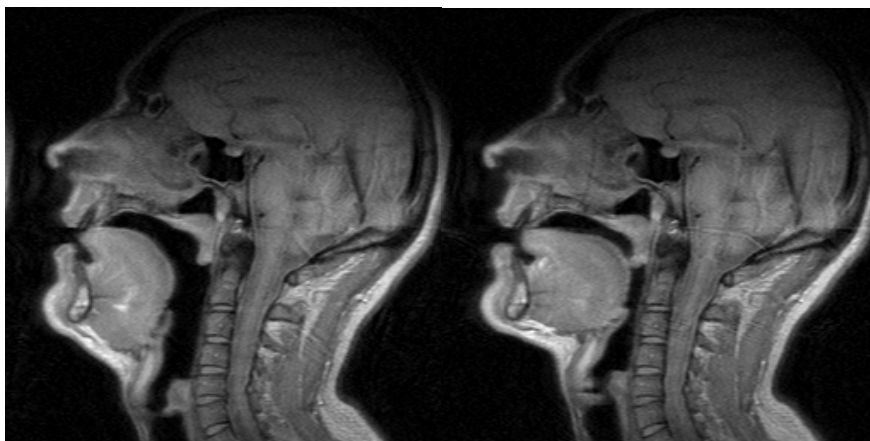

Figure 4. *MRI scans for C in an utterance of* /ˈaCː#ˈCa/ *with C articulated in a* alveodental *place for a lateral lamino-alveolar realisation of* /l/ *(left) and for an* apico-alveolar *(intermittent) realisation of* /ɾ/ *(right).*
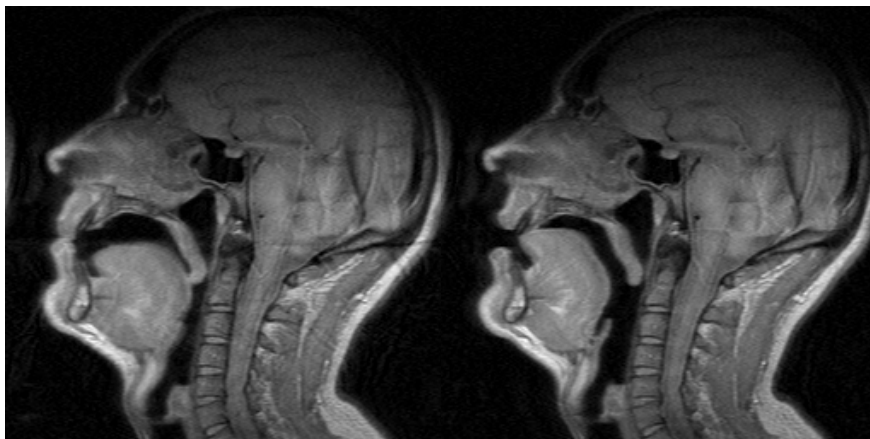
Figure 5. *MRI scans for C in an utterance of* /ˈaCː#ˈCa/ *with C articulated in the nasal manner (see the lowering of the velum) for the* bilabial *realisation of* /m/ *(left) and the* alveodental *realisation of* /n/ *(right).*
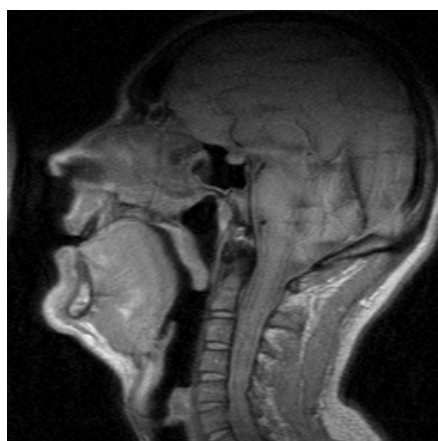

Figure 6. *MRI scans for C in an utterance of* /ˈaCː#ˈCa/ *with C articulated in the nasal manner (see the lowering of the velum) for the* palatal *realisation of* /ɲ/.

The general labialisation of some articulations (e.g. for /ʃ/ and /t͡ʃ/) is definitely more relevant, as well as the evident protrusion when V=/u/ (see above all the articulations in Figures 9-12). Nevertheless, the conditions of greater variation are the ones for /k/, with a place definitely varying from velar to prepalatal if V=/i/.

These descriptions mostly confirm those of other studies (first of all Tagliavini, 1965), but in many cases – if verified on a wider sample of speakers – they may call for a better cure of the terminology used for some articulations (above all those of /ʃ/, /t͡ʃ/ e /d͡ʒ/, for which the data of other speakers are also congruent).
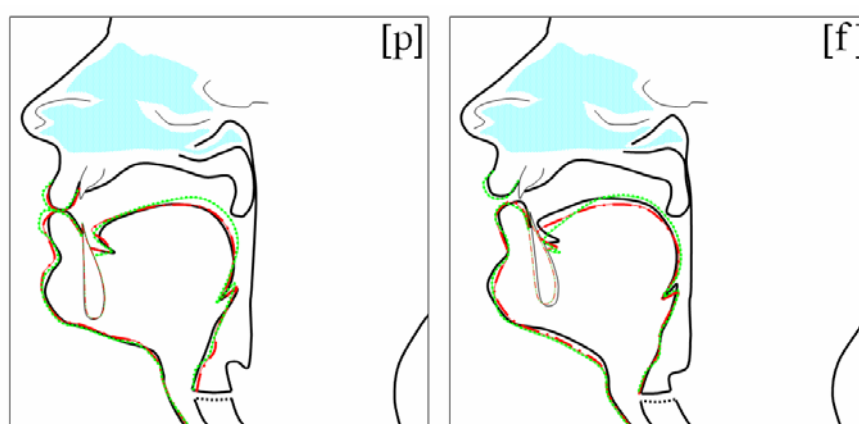
Figure 7. *Outline hand-tracings of the articulatory configurations detected in the* bilabial *and* labio-dental *places for the realisations of* /p/ *(left) and* /f/ *(right) in* /VCV/ *utterances with V=*/a/ *(plain line), V=*/i/ *(dashed line) V=*/u/ *(dotted line).*
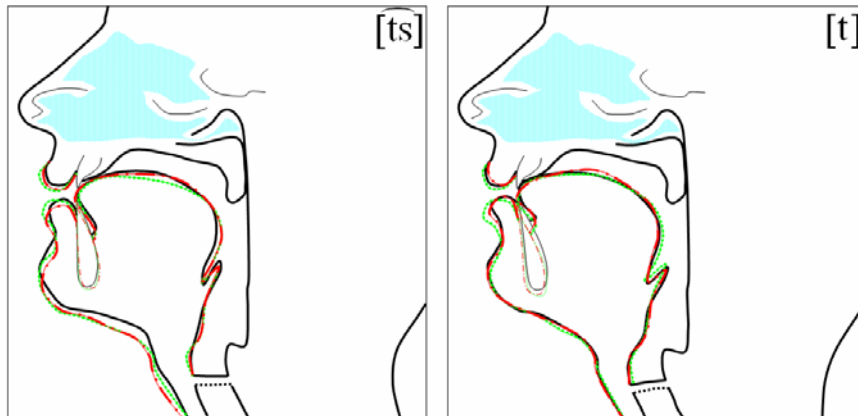
Figure 8. *Outline hand-tracings of the articulatory configurations detected in the* alveodental *place for the realisations of* /t͡s/ *(left) and* /t/ *(right) in* /VCV/ *utterances with V=*/a/ *(plain line), V=*/i/ *(dashed line) V=*/ɯ/ *(dotted line).*

Figure 9. *Outline hand-tracings of the articulatory configurations detected in the* alveodental *place for the realisations of* /s/ *(left) and in the* (post)alveolar *place for the realisations (more or less labialized) of* /t͡ʃ/ *(right) in* /VCV/ *utterances with V=*/a/ *(plain line), V=*/i/ *(dashed line) V=*/ɯ/ *(dotted line).*

Figure 10. *Outline hand-tracings of the articulatory configurations detected in the* (post)alveolar *place for the realisatios (more or less labialized) of /ʃ/ (left) and in the* palatal *place for the lateral realisations of /ʎ/ in /VCV/ utterances with V=/a/ (plain line), V=/i/ (dashed line) V=/ʊ/ (dotted line).*

Figure 11. *Outline hand-tracings of the articulatory configurations detected in the* alveodental *place for the lateral realisations of /l/ (left) and for the trilled realisations of /r/ (right) in /VCV/ utterances with V=/a/ (plain line), V=/i/ (dashed line) V=/ʊ/ (dotted line).*

Figure 12. *Outline hand-tracings of the articulatory configurations detected in the* velar (or prepalatal) *place for the lateral realisation of* /k/ *in* /VCV/ *utterances with V=*/a/ *(plain line), V=*/i/ *(dashed line) V=*/ʉ/ *(dotted line).*

4. REFERENCES

APOSTOL L., PERRIER P., RAYBAUDI M. and SEGEBARTH CHR. (1999). «3D geometry of the vocal tract and interspeaker variability». *Proceedings of the 14th International Congress of Phonetic Sciences* (*ICPhS99,* San Francisco), pp. 443-446.

BIRKHOLZ P. and KRÖGER B.J. (2006). «Vocal Tract Model Adaptation Using Magnetic Resonance Imaging». *Proceedings of the 7th International Seminar on Speech Production* (*ISSP06*, Ubatuba, Brazil, 2006), pp. 493-500.

DEMOLIN D., METENS T. and SOQUET A. (2000). «Real time MRI and articulatory coordinations in vowels". *Proceedings of the 5th Seminar on Speech Production* (Kloster Seeon, Germany, 1999), pp. 93-96.

ENGWALL O. and BADIN P. (1999). «Collecting and analysing two- and three-dimensional MRI data for Swedish». *TMH-QPSR, KTH*, 3-4, pp. 11-38.

ENGWALL O. and BADIN P. (2000). «An MRI Study of Swedish Fricatives: Coarticulatory Effects». *Proceedings of the 5th Seminar on Speech Production:* (Kloster Seeon, Germany, 1999), pp. 297-300.

FARNETANI E. (1986). «Il palatogramma: metodologia e applicazioni cliniche». *Quaderni del CSRF*, Padova, 5, pp. 309-350.

KEDROVA G.Y., ZAKHAROV L.M., PIROGOV YU.A. and ANISIMOV N.V. (2006). «Contrastive study of the MRI representation of Russian vowel articulation (against French, German and Korean analogues)». *Proceedings of the 33rd International Acoustical Conference - EAA Symposium ACOUSTICS* (Štrbské Pleso, Slovakia, 2006), pp. 33-36.

KIM H., HONDA K. and MAEDA SH. (2005). «Stroboscopic-cine MRI study of the phasing between the tongue and the larynx in the Korean three-way phonation contrast». *Journal of Phonetics*, 33, pp. 1-26.

KITAMURA T., TAKEMOTO H., HONDA K., SHIMADA Y., FUJIMOTO I., SYAKUDO Y., MASAKI SH., KURODA K., OKU-UCHI N. and SENDA M. (2005). «Difference in vocal tract shape between upright and supine postures: Observations by an open-type MRI scanner». *Acoustic Science & Technology*, 26/5, pp. 465-468.

KRÖGER B.J., WINKLER R., MOOSHAMMER CHR. and POMPINO-MARSCHALL B. (2000). «Estimation of Vocal Tract Area Function from Magnetic Resonance Imaging: Preliminary Results». *Proceedings of the 5th Seminar on Speech Production* (Kloster Seeon, Germany, 1999), pp. 333-336.

MAGNO CALDOGNETTO E. (1988). «L'apporto delle tecniche sperimentali alla descrizione fonetica: alcune esemplificazioni». In M. Cortelazzo (ed.), *Guida ai dialetti veneti*, X, Padova: Cleup, pp. 61-83.

MOLINO G. and ROMANO A. (2004). «Analisi acustica e articolatoria di alcuni contoidi palatali in un dialetto della Valsesia». *Bollettino dell'Atlante Linguistico Italiano*, 27, Torino (2003), pp. 203-221.

NARAYANAN SHR., ALWAN A. and HAKER K. (1995). «An articulatory study of fricative consonants using magnetic resonance imaging». *JASA*, 98, pp. 1325-1347.

NARAYANAN SHR., BYRD D. and KAUN A. (1999). «Geometry, kinematics, and acoustics of Tamil liquid consonants». *JASA*, 106/4, pp. 1993-2007.

NARAYANAN SHR., KRISHNA N., LEE S., SETHY A. and BYRD D. (2004). «An approach to real-time magnetic resonance imaging for speech production». *JASA*, 115/4, pp. 1771-1776.

PROCTOR M., SHADLE CHR. and ISKAROUS KH. (2006). «An MRI study of vocalic context effects and lip rounding in the production of English sibilants». *Proceedings of the 11th Australasian International Conference on Speech Science and Technology* (Auckland, New Zealand, 2006), pp. 307-312.

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 (ed.), *La fonetica acustica come strumento di analisi della variazione linguistica in Italia*, Roma: Il Calamo, pp. 7-14.

ROMANO A. (2007). «La fonetica sperimentale e gli atlanti linguistici: la sintesi romanza di 'pidocchio' e lo studio degli esiti palatali». In J. Dorta (ed.), *Temas de Dialectología*, Tenerife: Instituto de Estudios Canarios, pp. 178-204.

STONE M. (1997). «Laboratory Techniques for Investigating Speech Articulation». In W.J. Hardcastle & J. Laver (eds.), *The Handbook of Phonetic Sciences*, Oxford: Blackwell, pp. 11-32.

STORY B., TITZE I.R. and HOFFMAN E. (1996). "Vocal tract area functions from magnetic resonance imaging». *JASA*, 100, pp. 537-554.

TAGLIAVINI C. (1965). *La corretta pronuncia italiana: corso discografico di fonetica e ortoepia*. Bologna: Capitol (with 26 records).

TIEDE M., MASAKI S. and VATIKIOTIS-BATESON E. (2000), «Contrasts in speech articulation observed in sitting and supine conditions». *Proceedings of the 5th Seminar on Speech Production* (Kloster Seeon, Germany, 1999), pp. 25-28.