Articulatory Characteristics of European Portuguese Laterals: a 2D & 3D MRI Study

Paula Martins[1], *Catarina Oliveira*[1], *Augusto Silva*[2], *António Teixeira*[2]

[1]Escola Superior de Saúde/IEETA, University of Aveiro, Aveiro, Portugal
[2]Dep. Electrónica Telec. Informática/IEETA, University of Aveiro, Aveiro,Portugal
pmartins@ua.pt, coliveira@ua.pt, augusto.silva@ua.pt, ajst@ua.pt

Abstract

Magnetic Resonance Imaging (2D and 3D) has been acquired in seven EP native speakers during the production of EP lateral sounds /l, L/, with the purpose of obtaining articulatory data and acquiring further knowledge regarding the production of the lateral consonants. To attain these goals the acquired MRI corpus includes the /l/ sound in different syllabic positions (onset, coda and intervocalic) in the context of the three EP cardinal vowels, and the palatal lateral /L/ in intervocalic context. The results indicate the existence of a high degree of inter-subject variability (more evident for /l/). /l/ velarization can take place in all syllabic positions, including in onset, which is probably a speaker-dependent characteristic. A detailed articulatory description of the palatal lateral is also provided.
Index Terms: Speech Production, European Portuguese, Laterals, MRI, Image processing

1. Introduction

1.1. European Portuguese Laterals

Phonologically European Portuguese (EP) has two lateral consonants: /l/ and /L/. The production of /l/ involves a linguo-alveolar contact and one or two lateral channels along the lateral sides of the tongue [1, 2]; the /L/ is produced by one movement of the front of the tongue against the alveolo-palatal zone.

For the last two decades, it has been argued that Portuguese /l/ is categorically associated with a non-velarized ('light') allophone, which typically occurs syllable initially, and a velarized ('dark') one in coda position. However, there are conflicting positions in the literature regarding this light-dark dichotomy.

Unfortunately, the majority of the descriptions was based on gross impressionistic observations, and the only extensive empirical descriptions of Portuguese /l/ come from the acoustic data of Andrade [3]. Her results show that /l/ velarization does take place in onset position, for speakers of the Lisbon variety of EP, although the degree in which it is manifested varies across individual subjects. In agreement with this acoustic data, Recasens and Espinosa [4] stated that EP, together with Russian and Leeds British English, belong to a group of sound systems where /l/ presents essentially the same dark realization word-initially and word-finally.

Previous MRI studies of EP [5] seem to confirm the theory of the existence of a single dark-l for EP. However, the corpus was not exclusively designed for the study of laterals. More recently, in a study based in EMA [6], some syllable position effects were found: coordination patterns for syllable-initial /l/ are distinct from those observed for syllable-final /l/; the tongue dorsum is more retracted in the syllable-final /l/; /l/s in coda revealed, for one of the speakers, a reduction in magnitude of the tongue tip gesture.

There are no extensive and up-to-date data on the articulatory characteristics of /L/. It is traditionally recognized as a palatal lateral, but according to MRI data [5], the articulation appears to happen a little further forward.

1.2. MRI studies with laterals

MRI articulatory data for American English confirmed the existence of some differences between dark and light /l/. Although, both allophones present a similar overall configuration of the tongue body, with alveolar contact, lateral compression and a convex tongue body, there are differences between the two allophones: greater linguo-alveolar contact in light /l/ and a more retracted tongue body in the production of dark /l/ with repercussions in the area functions. The dark /l/ (when compared with the light /l/) presents large areas in the palatal region (behind the constriction), less lateral contacts and small areas in the pharyngeal and velar regions [1].

More recently, [2] also based on MRI data, reported a relatively shorter linguo-alveolar contact for dark /l/(0.8 cm) than for light /l/ (1.7 cm) and shorter lateral channels for the dark /l/. The two lateral channels formed for the light /l/ are asymmetrical (4.9 cm longer on the right and 2.1 cm on the left). A separate supra-lingual space was also found.

Section 1 introduces the problem and reviews related work on laterals. In section 2, the corpus, speakers, and MRI acquisition setup are described. Section 3 describes image segmentation and image processing techniques used to extract information from MRI images. The results, including preliminary 2D and 3D data, are given in section 4. This section is followed by the discussion and the conclusions that can be withdrawn from the present study.

2. Method

2.1. Corpus

MRI corpus included the lateral-alveolar /l/ in different positions: word-initial position (e.g laca "hairspray" [lak6]), intervocalic position (e.g sala "room" [sal6]) and coda position (e.g sal "salt"[sal]) in the context of the three EP cardinal vowels [i,a,u] (in stressed position). For the lateral /L/, due to phonotactic constraints in Portuguese, only intervocalic context was considered (e.g palha "straw" [paL6]).

2.2. Speakers

MRI data was acquired with seven (3 female, 4 male) speakers: six native mono-lingual EP speakers and a bilingual male speaker (EP and Spanish), ages ranging from 21 to 39, with no history of hearing or speech disorders. The speakers were all volunteers, from the midland of the country. Only two of the speakers (CO and JH) had Phonetic or Linguistic knowledge, the other speakers were instructed by the research team and had the opportunity of having a training session before the MRI acquisition session. An MRI screening form and informed consent were obtained before their participation in the study.

2.3. MRI Acquisition

The MRI experiment was carried out in a Magnetic Resonance Imaging Unit at Coimbra (IBILI).The images were acquired using a 3.0 T MR scanner (Magneton Tim Trio, Siemens, Erlanger, Germany) equipped with high performance gradients (Gmax = 45mT/m, rise time= 0.2s, slew Rate= 200 T/m/s; and FOV =50 cm). A standard 12-channel Head and Neck phased-array coils and parallel imaging (GRAPPA) were used in all data acquisition sessions. The Imaging protocol used in the present study was based in a previous MRI study conducted by our research team [5]. The subjects were positioned comfortably in a supine position using headphones. After acquiring reference images, a T1 W 5 mm thickness midsagittal MRI slice of the vocal tract was obtained using a TSE sequence (TR/TE/FA=400 ms/7.8 ms/120), FOV=240x240 mm; matrix (256x256) resulting in a pixel size of (0.938, 0.938). The acquisition time was 6 seconds. After that, a volume covering the entire vocal tract was obtained in the sagittal plane with a T1W 3D Spoiled GE sequence (VIBE), resulting in an acquisition time of 19 seconds; matrix (224x256); voxel size (1.055, 1.055, 2). The speakers sustained the sound during the period of acquisition; the sequence was launched when the /l/ was produced (e.g. sallllllll). Finally, a 3D high resolution sequence (VIBE) in the axial plane was obtained for each of the speakers, without phonation, to allow the extraction and co-registration of the mandible and dental casts.

3. Image Processing

3.1. Image processing techniques

Image segmentation was performed using two open source image processing tools: MevisLab [7] and ITK-Snap toolkit [8]. Some routines and data analysis were performed in Matlab. The codes used were implemented by one of the authors of the paper.

2D contours were extracted from midsagittal MR images using a semi-automatic technique (Live wire routine implemented in Mevislab). The tongue images were segmented using the same technique (e.g. fig. 4) and also using another semi-automatic tool based on a level sets framework (ITK-Snap, e.g.fig. 1).

Segmentation of the vocal tract images was more demanding and time consuming than the tongue. The overall process involves the following steps: 1) Mandible/maxilla segmentation and extraction from an MRI volume acquired for each speaker 2) Mandible and maxilla masks were resampled and co-registered with vocal tract data sets 3) Curved Multiplanar Reconstruction (MPR) of the vocal tract was performed to obtain areas perpendicular to the midline of the tract 4) Segmentation of the re-sliced volume was performed (every 5 mm from the glottis to the velum and every 3 mm throughout the oral cavity). This process was established using different pipelines implemented in MevisLab. Live wire technique was the method used in segmentation. After that, contour lists (CSO) were ex-

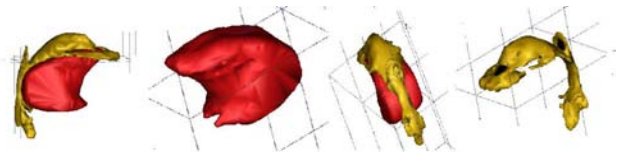

Figure 1: From left to right: ITK-Snap visualizations of the vocal tract and tongue (lateral view), tongue (oblique view), vocal tract and tongue (posterior-oblique view) and vocal tract (anterior-oblique view), during the production of /l/.

ported to Matlab allowing the extraction of area functions and tract visualizations. The inclusion of the teeth is a fundamental step when the goal is the modelation of the small lateral channels established in the production of the lateral sounds.

4. Results

In this section we present 2D MRI data for seven speakers and preliminary 3D information for 2 speakers (CO, female and JPM, male). Place of articulation, tongue configuration (2D and 3D), contextual variability, 3D visualization of the tract and area functions are exploited.

4.1. The /l/

4.1.1. 2D data

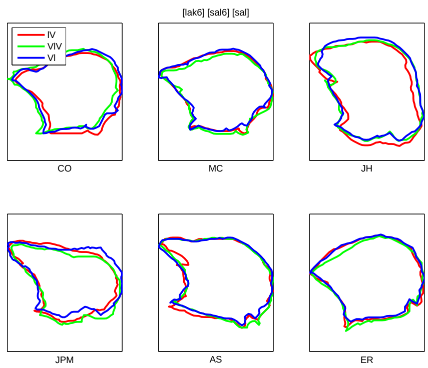

Figure 2: Superimposition of 2D contour for /l/ (6 speakers) in the three syllabic positions (onset, intervocalic and coda) with [a] in stressed position.

Analysis of the midsagittal images with contour superimposition revealed that the [l] was produced with contact of the tongue blade (i.e. laminal articulation) or tongue tip (i.e. apical articulation) against the dento-alveolar region. Laminal articulation was more frequent than apical articulation. Some speakers produced a laminal articulation in onset and a more apical articulation in coda. Some inter-speaker variability was observed regarding tongue body configuration, behind the alveolar contact. The tongue body shape in the three syllabic positions (onset, intervocalic and coda) was quite similar for most of the speakers. Only JPM and JH exhibited more variability (see fig. 2).

When considering the effect of vowel context in tongue configuration for [l] (fig. 3), it was observed that, for some speakers, vowel context has relatively little effect on tongue configuration. The variability appeared to be greater in speakers

JPM, AS and CO.

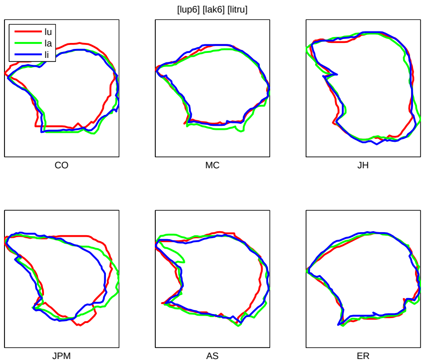

Figure 3: 2D contour superimposition for onset /l/ in vowel context [i, a, u].

4.1.2. 3D data

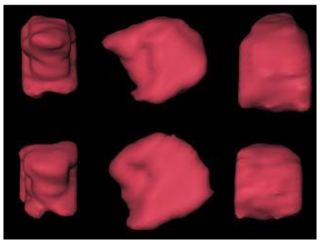

Figure 4: 3D rendering of the tongue (speaker CO) for onset /l/ in the word [lak6] (top) and vowel [a](bottom). From left to right: anterior, lateral and posterior views. Live wire segmentation implemented in MevisLab.

Analysis of 3D information for two of the speakers (JPM and CO) allowed the following considerations: Although the midsagittal contours of the /l/ are, in general, very similar to those of the vowels, 3D data revealed some differences (fig. 4) in tongue body configuration between the lateral and the vowels. The /l/ presents inward-lateral compression and a convex shape of the posterior tongue body; the tongue tip is raised towards the alveolar ridge. On the contrary, the tongue body for the vowel ([a]) has a slightly concave and spread configuration.

Lateral channels (fig. 5) formed around the tongue sides are longer for onset [l] than for coda /l/. For the speaker CO, lateral passages are short and start slightly posterior and through the sides of the alveolar contact, both in onset and coda position. For JPM, lateral channels are longer in onset than in coda: starting back at the velo-palatal area, extending through the oral cavity and continuing along the sides of the alveolar contact. The calculated area for the lateral passages is always less than $0.5cm^2$ for both speakers. The length of the alveolar contact is higher in onset than in coda.

Analysis of JPMs area functions (fig. 6) shows that the main differences between onset and coda /l/ occur at velar and upper pharyngeal regions. At this level the areas for coda /l/ are always lower than those obtained in onset. This information reflects a retracted tongue and tongue back elevation towards the velum, in coda position. For the speaker (CO), the only difference between the two syllabic positions was a slightly lowering of the areas at the pharyngeal region, in coda.

4.2. The /L/

4.2.1. 2D data

In /L/, the contact is made between the tongue blade and predorsum against the alveolo-palatal region for most of the speakers. Only one speaker (ER) articulated the /L/ exclusively at the palatal region. Figure 7 shows a superimposition of the contours for the intervocalic /L/ in the context of vowels /a, i, u/. Mostly, the consonant revealed to be only slightly influenced by the vowel: when observed (speakers JPM, AS, LCR), the variability occurred always behind the contact, being more evident at the tongue root level, an area not involved in the articulation of the lateral.

4.2.2. 3D data

For 3D data, the most relevant findings are related with a quite extensive contact area. Speaker CO presents a higher contact length (2.4 cm) than speaker JPM (2.1cm), which became more evident due to vocal tract length differences (about 3.5 cm shorter for CO than JPM).

Lateral channels are long and can be observed for both speakers. They are relatively symmetric, starting behind the back molars until and throughout the alveolo-palatal contact. Lateral channel areas are higher than those obtained for the lateral-alveolar /l/.

Speaker CO presents large areas at the pharyngeal region (due to a more anterior tongue position), which start to decrease from the velum towards the constriction. The other speaker shows relatively small areas at the low pharyngeal level, reaching the maximum area at the velum, and then following a pattern similar to that observed for CO.

As showed in fig. 8, tongue shape is characterized by an inward compression towards the midline, a convex shape and a lowered tongue tip position, for both speakers.

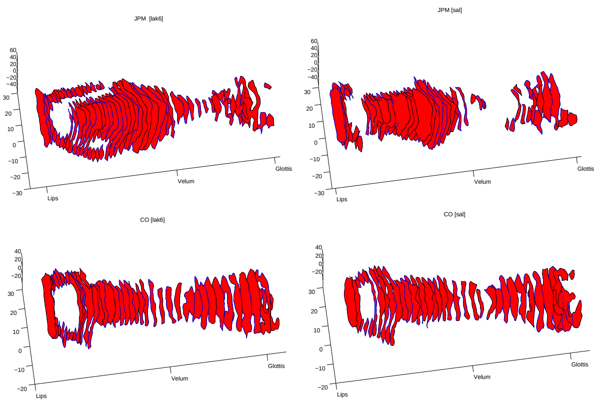

Figure 5: Vocal tract areas and lateral passages (3D data) for speakers JPM (top) and CO (bottom). At left, onset/l/ ([lak6]), and at right coda /l/ ([sal]).

5. Conclusions

This paper investigated articulatory characteristics of EP laterals /l/ an /L/, from a large dataset of MRI images. New 3D data allowed an articulatory description, representing a fundamental step towards the modelation of the lateral sounds.

MRI images revealed the existence of some inter-speaker variability in the production of these consonants, namely in the

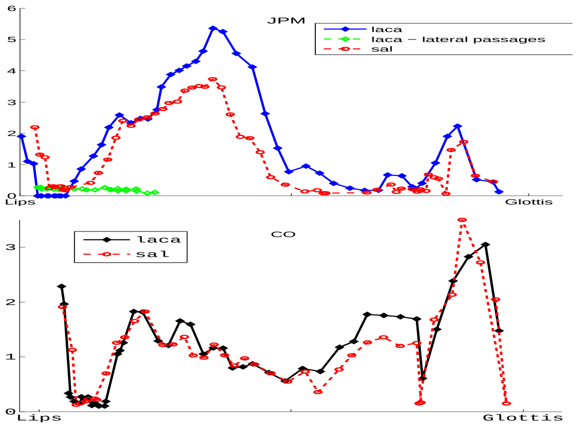

Figure 6: Vocal tract area functions of /l/ in onset (line) and coda (dashed) for JPM (top) and CO (bottom).

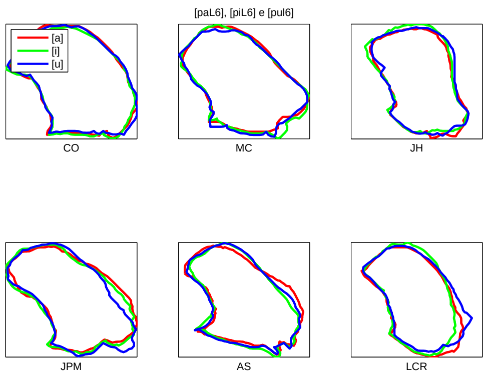

Figure 7: 2D contour superimposition for /L/ in vowel context [i, a, u].

alveolar lateral. Common characteristics for /l/ are: linguo-alveolar contact, either laminal or apical; inward-lateral compression and convex shape of the posterior region of the tongue [1] (though more evident in JPM than in CO), and lateral channels alongside the tongue.

Lateral channels are shorter in coda than in onset, as well as the length of the linguo-alveolar contact.These results are in line with previous studies for American English [1, 2].

Speaker CO presented short lateral channels and small alveolar contact in both syllabic positions. On the other hand, JPM showed extensive lateral channels in onset and short ones in coda. These findings pointed out to a dark realization for CO [3], in both syllabic positions, and two different allophones for JPM.

The palatal lateral is articulated at the alveolo-palatal region and not exclusively at the palatal area. This is in agreement with what has been reported for other Romance languages, as emphasized by [9]. This consonant is also characterized by an extensive contact area, lowered tongue tip, inward compression towards the midline and posterior convex shape resulting in long lateral channels and large pharyngeal and/or velar areas (depending on the speaker).

One of the limitations of this study has to do with difficulties in the co-registration and segmentation of the dental casts. Due to this fact we may assume that the areas of the lateral

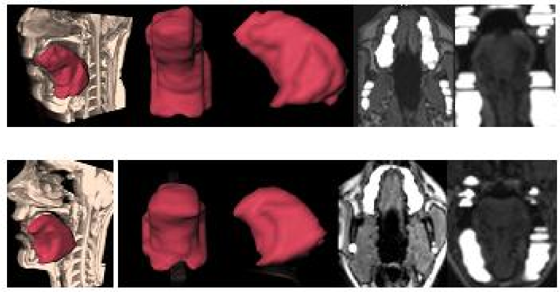

Figure 8: 3D data from JPM (top) and CO (bottom) for /L/ ([paL6]). From left to right: 3D rendering of the tongue (left oblique view) inside the volume, frontal view, lateral view, axial slice (after curved MPR) shows the pattern of the lingual contact and coronal reformatation.

channels were not precisely estimated.

Next steps for complementing this work will include: the segmentation and analysis of 3D data for all the speakers in all contexts, the improvement of segmentation tools (to improve accuracy and efficiency, particularly for vocal tract analysis) and the validation of the methods used to obtain information (e.g mandible and maxilla co-registration on the volumes).

6. Acknowledgements

This research was supported by FCT (Portuguese Research Agency) PhD grant of the first author (SFRH/BD/65183/2009). The authors would like to thank Professor Miguel Castelo-Branco (IBILI Director) and the technical staff of the MRI unit. We are also very grateful to our speakers.

7. References

[1] S. Narayanan, A. Alwan, and K. Haker, “Toward articulatory-acoustic models for liquid approximants based on MRI and EPG data. part I. The laterals,” *Journal of the Acoustical Society of America*, vol. 101, pp. 1064–2007, 1997.

[2] X. Zhou, “An mri-based articulatory and acoustic study of american english liquid souns /r/ and /l/.” Ph.D. dissertation, Faculty of the Graduate School, University of Maryland, Maryland, EUA, 2009.

[3] A. Andrade, “On /l/ velarization in European Portuguese,” in *International Congress of Phonetic Sciences (ICPhS)*, San Francisco, August 1999, pp. 543–546.

[4] D. Recasens and A. Espinosa, “Articulatory, positional and coarticulatory characteristics for clear /l/ and dark /l/: evidence from two Catalan dialects,” *Journal of the International Phonetic Association*, vol. 35, no. 1, pp. 1–25, 2005.

[5] P. Martins, I. Carbone, A. Pinto, A. Silva, and A. Teixeira, “European Portuguese MRI based speech production studies,” *Speech Communication*, vol. 50, no. 11-12, pp. 925–952, 2008.

[6] C. Oliveira, A. Teixeira, and P. Martins, “Towards an articulatory characterization of european portuguese /l/,” in *ISCA Workshop on Experimental Linguistics*, Athens, Greece, 2010.

[7] M. S. MeVis, “Mevislab (medical image processing and visualization) version 2.1,” 2003-2010.

[8] P. A. Yushkevich, J. Piven, H. C. Hazlett, R. G. Smith, S. Ho, J. C. Gee, and G. Gerig, “User-guided 3d active contour segmentation of anatomical structures: Significantly improved efficiency and reliability,” *Neuroimage*, vol. 31, no. 3, pp. 1116–1128, 2006.

[9] D. Recasens and A. Espinosa, “Articulatory, positional and contextual characteristics of palatal consonants: Evidence from majorcan Catalan,” *Journal of Phonetics*, vol. 34, pp. 295–318, 2006.