AN ARTICULATORY AND ACOUSTIC STUDY OF THE EUROPEAN PORTUGUESE /l/

Catarina Oliveira[a], *Paula Martins*[a], *António Teixeira*[b], *Isabelle Marques*[a] *& Pedro Sá-Couto*[c]

[a]Escola Superior de Saúde / IEETA, Universidade de Aveiro, Portugal;
[b]Dep. de Eletrónica, Telecomunicações e Informática / IEETA, Universidade de Aveiro, Portugal;
[c]Dep. de Matemática / CIDMA, Universidade de Aveiro, Portugal
coliveira@ua.pt; pmartins@ua.pt; ajst@ua.pt;
p.sa.couto@ua.pt; tf_isabelle_marques@hotmail.com

ABSTRACT

In this paper, MRI images and acoustic data for several speakers of European Portuguese (EP) were utilized in order to investigate articulatory and acoustic characteristics of /l/. Syllabic position differences and effects of the surrounding vowels were further evaluated. In general, the EP /l/ showed no substantial articulatory or acoustic differences at word edges, with low F2 frequencies and small pharyngeal/velar areas due to tongue-root retraction and/or tongue dorsum raising. The amount of vocalic effect on both the articulatory and acoustic data was speaker-dependent.

Keywords: European Portuguese, /l/ velarization, MRI, articulatory data, acoustic data

1. INTRODUCTION

There are conflicting positions in the literature concerning the allophonic variation of the EP /l/ conditioned by its affiliation to distinct syllable positions. Most authors (e.g. [4]) consider that EP /l/ is categorically associated with a non-velarized allophone ("light") in syllable onset, and a velarized one ("dark") in coda position. However, according to recent articulatory and acoustic data [1, 7], this syllabic light-dark allophony is quite subtle, and /l/ may be consistently dark even in onset position. The main articulatory difference between the two allophones lies in the gestural composition: in the dark /l/, the tongue tip constriction is accompanied by a post-dorsal velar or pharyngeal gesture [6, 9, 11]. Acoustically, retraction of the posterior tongue body in the dark /l/ causes F2 to lower and F1 to raise [10]. Furthermore, the contact between tongue tip and alveolar ridge is weaker in coda position than it is in onset position. Dark /l/ may even lack a lingual contact [3, 5, 6]. In terms of interarticulator timing, in syllable-onset position, tongue tip constriction of /l/ tends to occur earlier than, or in simultaneous with, tongue dorsum gesture, while in final syllable position roughly the reverse situation occurs [3, 5, 11]. Previous studies [10] show that, in comparison with the light /l/, the dark /l/ is more resistant to vowel coarticulation at tongue dorsum level.

While in many languages (especially in English) the articulatory and acoustic properties of /l/ have been relatively well-studied [6, 9, 11], most of the earlier studies on EP described /l/ in impressionistic terms [12], and the few recent studies that have provided acoustic or articulatory descriptions [1, 7] had several limitations (e.g. data collected from a small number of speakers and contexts; only acoustic and articulatory data).

MRI is well suited to study 3-D tongue and vocal tract shapes during sound articulation. However, it also features some well-known drawbacks (e.g. artificially sustained and hyperarticulated productions). Complementarily to the MRI data, acoustic measurements were also performed separately. As stated above, formant frequencies (mainly F2) are sensitive to modifications in tongue shape and constriction place, and therefore the degree of /l/ velarization can be inferred from the acoustic data. The major goal of this paper is to improve the knowledge of the articulatory and acoustic properties of the EP /l/ through the study of differences in its phonetic implementation as function of the syllable position and the vowel context.

2. METHOD

2.1. MRI set-up

The subjects were 7 native EP speakers: 3 female (CO, MC, ER) and 4 male (JPM, LCR, JH, AS) speakers, ages ranging from 21 to 39, with no known hearing or speech disorders. An MRI screening form and informed consent were obtained before their participation in the study. MRI corpus included the EP /l/ next to the vowels

/i/, /a/, /u/ in different word positions: word-initial (e.g *litro* “liter” ([ˈlitɾu]), intervocalic (e.g *b ílis* “bile” ([ˈbiliʃ]) and word-final (e.g *til* “tilde”[til]).

The MRI experiment was carried out at IBILI-Coimbra, using a 3.0 T scanner (Magneton Tim Trio, Siemens, 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 were used in all data acquisition sessions. The subjects were positioned comfortably in supine position using headphones. After acquiring reference images, a T1 W 5 mm thickness midsagittal MRI slice of the vocal tract was obtained in 6 s (0.938 x 0.938mm pixel size). A volume covering the entire vocal tract was also 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 (1x1x2mm). The speakers sustained the sound during the period of acquisition; the sequence was launched when the /l/ was produced. Finally, a 3-D high-resolution sequence was obtained for each of the speakers, without phonation, so as to allow the extraction and co-registration of dental casts (see fig. 1). A detailed description of the image processing and analysis procedures (e.g. measurements of the vocal tract and lateral passages areas) is provided in [8]. Segmentation of both the vocal tract and tongue was based on the general deformable models framework. Image processing, data analysis and visualizations were performed in MevisLab (MeVis Medical Solutions), ITK-SNAP toolkit [13] and Matlab.

2.2. Acoustic set-up

The subjects engaged in the acoustic experiment were 9 female and 11 male EP native speakers (including the MRI speakers) with ages ranging from 18 to 39. Although more contexts were recorded, we will only report those that were included in the MRI experiment. Each stimulus was repeated three times in a total of 27 tokens *per* speaker. The recordings were performed after the MRI scans in a sound-treated room with a condenser microphone connected to an external 24-bit sound system. The speech signal was recorded at a sampling rate of 22 kHz. Each lateral consonant was manually segmented and labeled in the digitized sound wave by using the program Praat [2]. Burg algorithm as built into the Praat program was used to compile values for F1 and F2, in Hertz, at the mid-point of the liquid.

Two acoustic measures were taken as indicative of the degree of /l/ velarization: F2 frequency, and distance between F2 and F1 (F2-F1).

The statistical investigations were conducted with SPSS 15.0. A three-way mixed analysis of variance (ANOVA) was carried out, with word position and vowel context as within-subject factors and gender as a between-subject factor. In general, the ANOVA assumptions of residual normality, homogeneity of variance and sphericity were validated. Bonferroni multiple comparisons were applied to significant results. Two speakers were excluded from the statistical analysis as they were identified as severe outliers. In all statistical analysis, the level of significance was $p<0.05$.

3. RESULTS

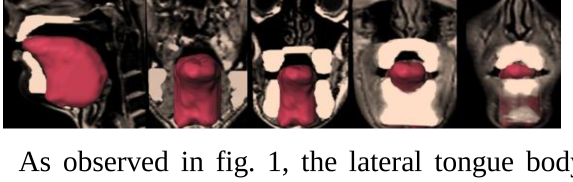
3.1. MRI results

Although the analyses of the 2-D MRI images have revealed that the tongue configurations were different among subjects, several common characteristics (in 3-D) were identified (section 3.1.1). The effects of vocalic context and word position on the articulation of /l/ are addressed in section 3.1.2 and 3.1.3, respectively.

3.1.1. (General) articulatory characteristics

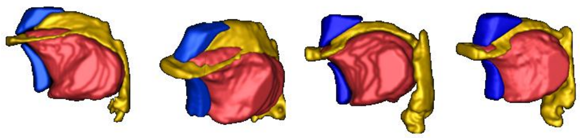
An analysis of the 2-D midsagittal tongue body contours revealed that the /l/ was produced with linguo-alveolar contact (see midsagittal view in fig. 1). The /l/ articulations were either laminal or apical, but the former were prevalent. The extension of the alveolar contact varies between 0.2-1.2 cm (for all speakers, contexts and positions acquired). A tendency for a slightly short lingual contact in word-final position was verified. Complete loss of alveolar contact was observed in 2 tokens (speakers JH, MC), in word-final position.

Figure 1: 3-D volume rendering (speaker LCR) for /l/ in the word [ˈlakɐ]. From left to right: midsagittal view; coronal views from velum to the lips.



As observed in fig. 1, the lateral tongue body compression enables airflow along both sides of the tongue. In this speaker, separate lateral channels along the sides of the tongue appeared in the alveolar contact area and continued behind the constriction. The extent of these lateral channels seems to be speaker-dependent. In general, for the female speakers, these channels are short, and can be observed mostly along the linguo-alveolar contact (see fig. 2), without any major differences between the word-positions analyzed. The middle and posterior tongue body exhibits a convex shape. The presence of grooving along the midsagittal line is not a common feature in /l/ production, for most speakers.

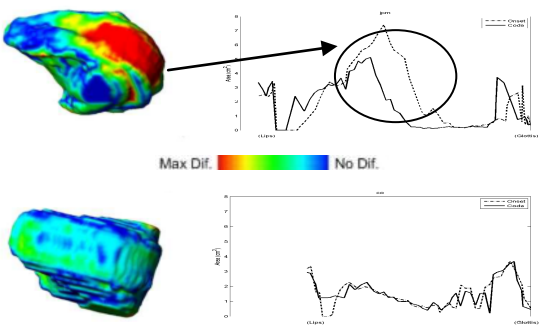
Figure 2: 3-D vocal tract and tongue meshes (lateral and oblique views) for speakers JPM (left) and CO (right), during /l/ production.



3.1.2. Positional effects

Five speakers revealed similar patterns of vocal tract areas for word-initial, intervocalic and word-final /l/, with decreased areas at velar and/or pharyngeal region, depending on the part of the tongue body involved in the formation of the "constriction." When dissimilarities occurred between /l/s motivated by word position, these could be observed either in the palatal area, in the region behind the constriction, where word-final /l/ exhibited larger areas, or in the velar and upper pharyngeal regions. Superimposed meshes of the tongue and corresponding area functions for two subjects, with different behaviours, are shown in fig. 3.

Figure 3: Mesh comparison (left) and corresponding area functions (right) for /l/ as produced in [ˈlakɐ] and [sal] for the speakers JPM (top) and CO (bottom).

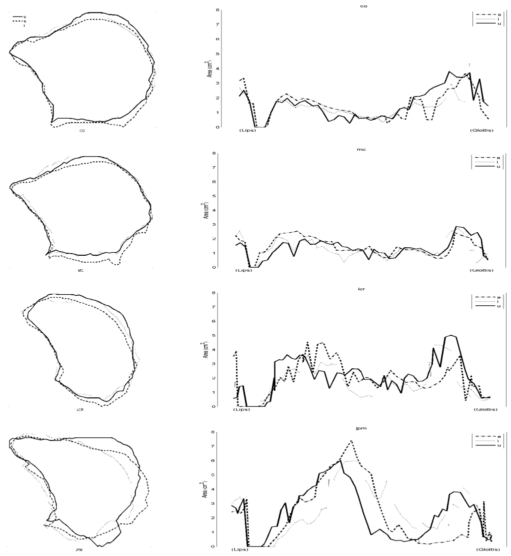


3.1.3. Vowel effects (Coarticulation)

According to 2-D (tongue body configuration) and 3-D data (vocal tract areas), the vowel effects were more visible behind the linguo-alveolar contact, although in general they were not very prominent for most speakers (speaker JPM presented the highest degree of variability, as shown in fig. 4). Nevertheless, word-initial /l/ was more affected by vowel context than intervocalic and word-final /l/s.

Also, male speakers showed more variation with vocalic context than female speakers, as shown in the example of the fig. 4.

Figure 4: 2-D contour superimposition of the tongue (left) and vocal tract area functions (right), obtained from 3-D data, for onset /l/ in context of /i, a, u/, for female (on the top) and male (on the bottom) speakers.



3.2. Acoustic results

The results in terms of means (M) and standard deviations (SD) for F2 and F2-F1 (as function of word position, vowel context and gender) are shown in table 1.

Regarding F2, the vowel context (F(1.6,26.1)=4.39), word position (F(2,32)=5.84), gender (F(1,16)=7.40) and interaction between word position and vowel context (F(3.6,57.4)=3.15) were statistically significant ($p<0.05$). As revealed by table 1, mean F2 frequencies were found to vary in the progression /i/ > /u, a/. A statistical difference was found between /i/ and /a/ ($p=0.003$). Furthermore, F2 values were higher for intervocalic /l/s, followed by word-initial and word-final /l/s. A statistical difference was found between intervocalic and word-final position ($p=0.013$). Moreover, female speakers exhibited higher F2 frequencies than male speakers (see table 1).

ANOVAs on the F2-F1 data revealed that vowel context (F(1.6,25.3)=8.33) and interaction between position and vowel context (F(3.4,55.1)=4.64) were statistically significant ($p<0.05$). This parameter presented higher values when the contextual vowel was /i/, followed by /u/ and /a/ (see table 1). A statistical difference was found between /i/ and /a/ ($p<0.001$).

Table 1: Mean F2 and F2-F1 values as function of word position, vowel context, and gender (N=18; 9 males and 9 females). The results are in M±SD.

		F2 (Hz)	F2-F1 (Hz)
Word position	Initial	955.7±135.3	532.4±115.0
	Intervocalic	1012.3±154.9	702.4±136.3
	Final	869.3±149.6	588.9±153.7
Vowel context	/a/	910.8±111.6	532.4±115.0
	/i/	1016.8±135.5	702.4±136.3
	/u/	910.6±179.6	638.6±184.0
Gender	Male	879.6±103.8	575.0±106.1
	Female	1012.6±103.8	674.0±106.1

4. DISCUSSION AND CONCLUSIONS

In this paper, the articulatory and acoustic properties of the EP /l/, as well as the influence of word-position and vocalic context in its production, were studied using MRI data and acoustic recordings of several native speakers of EP.

MRI images for /l/ in different word-positions revealed some invariant characteristics, which were hypothesized by Narayanan [9] to be the primary features for laterals: linguo-alveolar contact; convex tongue body shape; inward-lateral compression that enables the creation of lateral channels. Most of the speakers revealed a similar tongue body configuration in the three positions analyzed, with a narrowing of the vocal tract at the velar and/or pharyngeal regions, due to raising (velarization) and/or retraction (pharyngealization) of the posterior tongue body. These results were consistent with the low F2 frequencies obtained in the acoustic study, and point to a strongly dark realization of EP /l/ in all word-positions and vowel contexts [1, 10].

Differences in articulation of /l/ due to vocalic context were observed (namely in tongue posture, behind the constriction), although these coarticulatory effects were more noticeable for male speakers and for word-initial position. This articulatory variation corresponded to significant acoustic variation between /l/s produced in different vocalic contexts.

The results obtained provided important insights into the EP /l/ articulation and revealed that there are important sources of variation (speaker, gender, vowel context, and word position) that must be considered. This information is crucial for the improvement of our articulatory synthesizer.

5. AKNOWLEDGEMENTS

This work was supported by FCT grant (SFRH/BD/65183/2009) and FCT project HERON II (PTDC/EEA-PLP/098298/2008). Authors thank IBILI director, technical staff of the MRI unit and our speakers.

6. REFERENCES

[1] Andrade, A. 1999. On /l/ velarization in European Portuguese. *Proc. ICPhS 99* San Francisco, 543-546.

[2] Boersma, P., Weenink, D. 2011. Praat:doing phonetics by Computer [Computer Program]. Version 5.2.16. 2011. *http://www.praat.org/*.

[3] Browman, C.P., Goldstein, L. 1995. Gestural syllable position effects in American English. In Bell-Berti, F., Raphael, L.J. (eds.), *Producing Speech: Contemporary Issues, for Katherine Safford Harris*. New York: AIP-Press, 19-33.

[4] Cunha, C., Cintra, L. (eds.). 1984. *Nova Gramática do Português Contemporâneo*. Lisboa: Edições João Sá da Costa.

[5] Gick, B. 2003. Articulatory correlates of ambisyllabicity in English glides and liquids. In Local, J., Ogden, R., Temple, R. (eds.), *Papers in Laboratory Phonology VI* Cambridge: Cambridge University Press, 222-236.

[6] Giles, S.B., Moll, K.L. 1975. Cinefluorographic study of selected allophones of English /l. *Phonetica* 31, 206-227.

[7] Martins, P., et al. 2008. European Portuguese MRI based speech production studies. *Speech Communication* 50 (11-12), 925-952.

[8] Martins, P., et al. 2010. Articulatory characteristics of European Portuguese laterals: a 2D & 3D MRI study. *Proc. FALA* Vigo, Spain, 33-36.

[9] Narayanan, S., Alwan, A., Haker, K. 1997. Toward articulatory-acoustic models for liquid approximants based on MRI and EPG data. Part I. The laterals. *JASA* 101(2), 1064-1077.

[10] Recasens, D., Espinosa, A. 2005. Articulatory, positional and coarticulatory characteristics for clear /l/ and dark /l/: evidence from two Catalan dialects. *JIPA* 35(1), 1-25.

[11] Sproat, R., Fujimura, O. 1993. Allophonic variation in English /l/ and its implications for phonetic implementation. *Journal of Phonetics* 21, 291-311.

[12] Viana, A.R. 1883. Essai de phonétique et phonologie de la langue portuguaise d'áprés le dialecte actuel de Lisbonne. Lisboa.

[13] Yushkevich, P.A., et al. 2006. User-guided 3D active contour segmentation of anatomical structures: Significantly improved efficiency and reliability. *Neuroimage* 31(3), 1116-1128.