The effects of syllable and utterance position on tongue shape and gestural magnitude in /l/ and /r/

Eleanor Lawson[a], Gregory Leplatre[b], Jane Stuart-Smith[c], James M. Scobbie[a]

[a]Queen Margaret University, Edinburgh
[b]Edinburgh Napier University
[c]University of Glasgow

ABSTRACT

This paper is an ultrasound-based articulatory study of the impact of syllable-position and utterance position on tongue shape and tongue-gesture magnitude in liquid consonants in American, Irish and Scottish English. Mixed effects modelling was used to analyse variation in normalised tongue-gesture magnitude for /r/ and /l/ in syllable-onset and coda position and in utterance-initial, medial and final position. Variation between onset and coda mean midsagittal tongue surfaces was also quantified using normalised root-mean-square distances, and patterns of articulatory onset-coda allophony were identified. Despite the fact that some speakers in all varieties used tip-up /r/ in syllable-onset position and bunched /r/ in coda position, RMS distance results show greater degrees of similarity between onset and coda /r/ than between onset and coda /l/. Gesture magnitude was significantly reduced for both /l/ and /r/ in coda position. Utterance position had a significant effect on /l/ only.

Keywords: Articulatory Phonetics; Ultrasound Tongue Imaging; allophony.

1. INTRODUCTION

It has long been recognised that consonants sound different depending on their location within the syllable [13]. The notion that there is a phonetic basis for the cross-linguistic tendency of consonants to be phonetically “stronger”, or more consonantal, in syllable-onset position and “weaker”, or more vowel-like, in coda position, has long been of interest to phoneticians, and phonologists. In addition to syllable-conditioned variation in phonetic quality, research suggests that proximity to prosodic boundaries also affects phonetic quality [12], [4], further strengthening or weakening consonants. In fact, it has been claimed that that we cannot understand segmental articulation independently of prosodic structure [2].

Two aspects of articulation that seem to be modified by both syllable and utterance position are (1) tongue configuration and (2) magnitude of gestures. One study providing evidence of the former for /r/ is [11]: in a sample of 100 speakers of English, from multiple articulatory datasets, it was found that around one third of speakers studied used different configurational variants of /r/ in each syllable position: tip-up /r/ in onset and bunched /r/ in coda position (T-B), see example in Fig. 1. Other speakers in the study used the same tongue configuration in both onset and coda position: either tip-up (T-T), or bunched (B-B). Only one (somewhat ambiguous) instance of the pattern (B-T) was identified.

For the latter type of phonetic variation, gesture magnitude, we find evidence for syllable- and utterance-position-conditioned variation in tongue gesture magnitude from studies using a variety of articulatory techniques. Using cineradiography and Electromagnetic Articulography (EMA) respectively [1], [6] and [7] showed a reduction in gesture magnitude for lingual consonants in coda position, compared to onset position. Using electropalatography (EPG), [4] studied the impact of utterance position on tongue-palate contact patterns and found greater degrees of palatal contact in nasals when they were in utterance-initial position, compared to utterance-final position.

The current study considers articulatory strengthening/weakening in the so-called liquid consonants /l/ and /r/; articulatorily-complex consonant sounds that often exhibit extreme patterns of syllable-based allophony [9]. Variation in tongue configuration and gesture magnitude in /l/ and /r/ is quantified and analysed across three broad accent categories in English: American, Scottish and Irish. Our research questions are:

(1) How much do tongue shapes for /l/ and /r/ vary between syllable-onset and coda position?
(2) How do the syllable and utterance positions of a liquid consonant affect tongue gesture magnitude?

2. METHOD

6 Scottish speakers, 10 American speakers and 4 Irish speakers (from the Republic of Ireland) were recorded with ultrasound tongue imaging (UTI – Sonix RP machine, 111fps, 150° fan angle) reading

aloud 80 sentences (c11 syllables long), e.g. *Perm equipment damages dry brittle hair*. Each sentence contained at least one stressed monosyllabic key word, e.g. *perm* and *hair*, with a liquid consonant in onset or coda position, and each key word was positioned at the beginning, middle, or end of the utterance. There were circa 8 instances of /l/ and /r/ produced in each syllabic-utterance position per speaker. Utterance-medial liquids were always phrase medial, i.e. not positioned adjacent to a phrase boundary. All tokens of /r/ in the study are approximants. Lingual consonants were avoided in the two segments adjacent to the key segment in order to mitigate against consonantal coaritculatory effects.

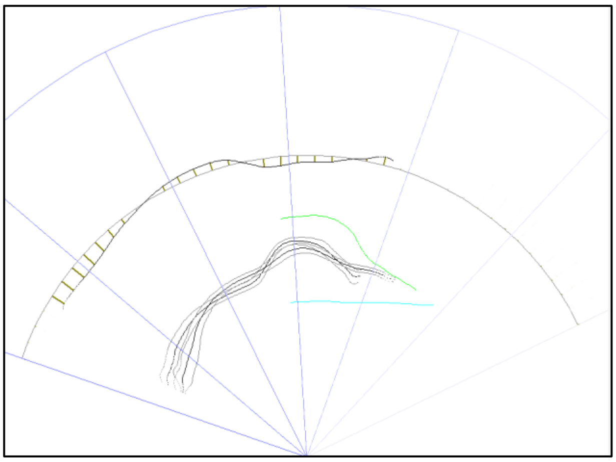
2.1. Articulatory measures

2.1.1 Quantifying variation in tongue shape across syllable positions

Using Articulate Assistant Advanced (AAA) v2.16.12 [14], the maxima of the anterior lingual gestures of /r/ (N=784) and /l/ (N=859) were annotated by the first author. For /l/, the anterior maximum was annotated at the temporal point where the tongue tip reached its highest position, this was also the case for tip-up /r/ variants. 24 vocalised variants of /l/ with no tongue-tip gesture were also annotated at the maximum of the dorsal gesture. For bunched /r/ variants, the maximum was annotated when the tongue dorsum/front reached its highest position. Splines were then automatically fitted to the midsagittal tongue surface and hand corrected where necessary.

As speakers were very consistent in their syllable-based production strategies, variation between liquid consonants in onset and coda position for each speaker was quantified by first creating mean midsagittal tongue-surface splines for onset and coda /l/ and /r/. This mean tongue spline was created by averaging the distance where all individual splines intersected each of 42 radial axes of a superimposed fan grid, whose point of origin was set to match the virtual point of origin of radial ultrasound pulses emitted by the probe, see Fig.1. Thereafter AAA's *distance* function was used to determine the raw distances (in mm) between the mean onset and coda splines along each radial fan axis, see Fig. 1, and their root-mean-square (RMS) was calculated. In order to account for inter-speaker variation in oral-cavity size, the distance between the ultrasound probe surface and the top of each speaker's palatal arch was measured and used to normalise the raw RMS measures. Normalised RMS measures = raw RMS/probe-to-palatal arch distance.

Figure 1: Showing tip-up and bunched /r/ mean tongue-surface splines for speaker ScotF7, with standard deviations. Uppermost arc shows raw distance measures between the splines along radial fan axes. The green spline represents the speaker's hard palate and the blue spline their occlusal plane.



2.1.2 Quantifying gesture magnitude across syllable and utterance positions.

Quantification of gesture magnitude was carried out as follows: all tongue-surface splines were rotated so that the speaker's occlusal-plane trace (obtained using a bite plate) was set to horizontal; all individual tongue-surface splines were then exported as sets of Cartesian coordinates for measurement using R v.3.3.2 [3]. An R script automatically identified the height of the tongue tip for /l/ and tip-up variants of /r/ and the highest point on the bunched tongue dorsum/front for bunched variants of /r/. Tongue surfaces were plotted and their measurement points annotated with an "x" to determine whether the location of automatic measures was correct. One token of /r/ was excluded after eyeballing the plots – a single instance of tip up /r/, produced by a speaker who habitually produced bunched /r/s. Again, raw tongue height measures were normalised by expressing them as a proportion of the probe-to-palatal-arch distance.

2.2. Statistical analysis

ANOVAs were used to identify significant variation in normalised RMS distances between factors (1) liquid consonant: (i) /l/ and (ii) /r/; (2) accent: (i) American, (ii) Irish and (iii) Scottish, (3) syllable-based allophonic patterns of /r/: (i) T-T, (ii) B-B and (iii) T-B (T=tip up, B=bunched). (4) syllable-based allophonic patterns of /l/: (i) P-P, (ii) V-V, (iii) P-V, (P=palatalised, V=velarized).

Variation in gesture magnitude was analysed using mixed-effects modelling in R [3]. Separate models were run for /l/, tip-up /r/ and bunched /r/. The dependent variable was *normalised tongue gesture magnitude*. Fixed factors were: (1) *syllable position*, with levels (i) onset, (ii) coda; (2) *utterance position*, with levels (i) initial, (ii) medial, (iii) final and (3) *speaker accent*, with levels (i) Irish, (ii) American, (iii) Scottish. In all models, we tested for two and three-way interactions between *utterance position*, *syllable position* and *accent*. Random factors included in all models were: *speaker* and *word*. The step( ) function in the LmerTest package [8] was used to find best fit models for the datasets.

3. RESULTS

3.1. Configurational variation in onset and coda: RMS distance

ANOVAs were run for both raw and normalised RMS distances and showed the same patterns of significant variation, therefore we report only on the results for normalised RMS distances here.

A significant difference was found between the two liquid consonants F=7.34 p<0.05, with /l/ showing significantly greater normalised RMS distances between onset and coda tongue shapes (mean= 0.039) than /r/ (mean=0.024), see Fig. 2.

Significant variation was found between accents for /r/ only: F=4.971 p<0.05. Post-hoc Tukey tests showed a significant difference between the Scottish and American accent groups, (Scottish norm. RMS mean=0.037, American norm. RMS mean=0.017). Table 1 and Fig. 2 provide an explanation as to why Scottish speakers have greater normalised RMS distances between mean onset and coda tongue surfaces for /r/, namely that half of the Scottish speakers had the asymmetric allophonic pattern (T-B), which has the greatest RMS distance of the three rhotic allophonic categories. Only 1 American speaker had this pattern.

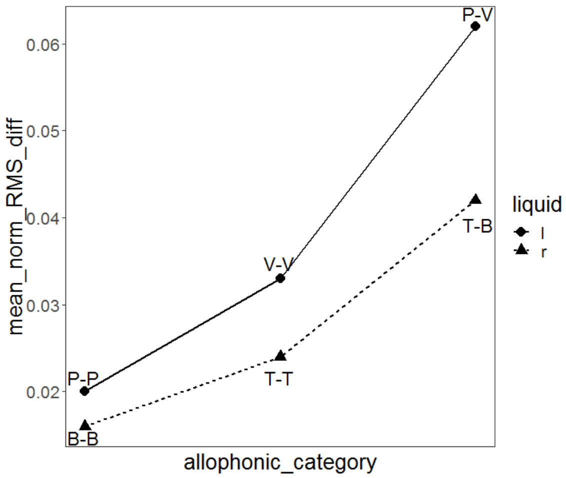
For syllable-based allophonic patterns for /r/, a significant difference was found between the T-B and B-B allophonic categories p<0.01. Unsurprisingly, the asymmetric allophonic pattern (T-B) results in the greatest mean RMS distance, see Fig. 2.

For syllable-based allophonic patterns for /l/, significant differences were found between the P-P and P-V categories p<0.05, and between the V-V and P-V categories p<0.01. Again, unsurprisingly, the asymmetric allophonic pattern (P-V) results in the greatest mean RMS distance, see Fig. 2.

Table 1: Numbers of speakers in each accent group using specific syllable-based allophonic patterns for /r/ and /l/.

	Allophonic pattern	Scottish	Irish	American
/r/	(T-T)	3	1	3
	(T-B)	3	0	1
	(B-B)	0	3	5
/l/	(P-P)	0	2	0
	(V-V)	3	1	9
	(P-V)	3	1	0

Figure 2: Means plot for normalised RMS distance between onset and coda tongue shapes for /l/ and /r/, organised by onset-coda allophonic patterns.



While unsurprisingly non-symmetrical allophonic configurational patterns (P-V for /l/ and T-B for /r/) resulted in the greatest normalised RMS distance measures between onset and coda position, symmetrical allophonic patterns (V-V, P-P for /l/ and T-T, B-B for /r/) showed different degrees of onset-coda difference. It is the palatal configurational variants of /l/ (P-P) and /r/ (B-B) that showed the smallest normalised RMS distance values, in other words, the least variation in tongue location and configuration between onset and coda position.

3.2. Gesture magnitude

3.2. 1 /l/

The best-fit model for normalised tongue-tip height included an interaction between fixed factors: *syllable position*utterance position* F=20.66, p<0.001, illustrated in the interaction plot in Fig. 3. The random factor *speaker* was also significant: χ=653.99, p<0.001.

Figure 3: Interaction plot for factors *syllable position* and *utterance position* for normalised tongue-tip height for /l/.

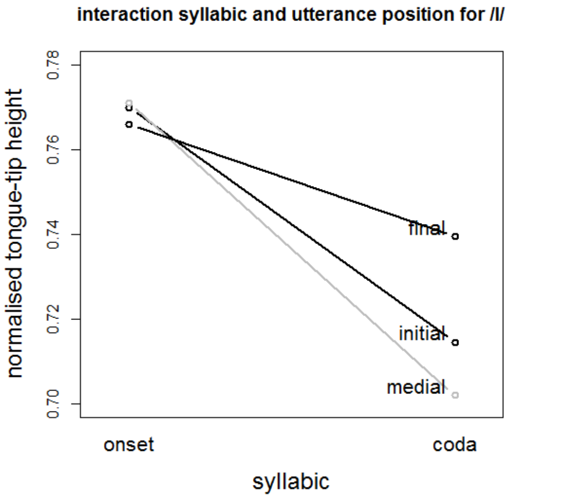
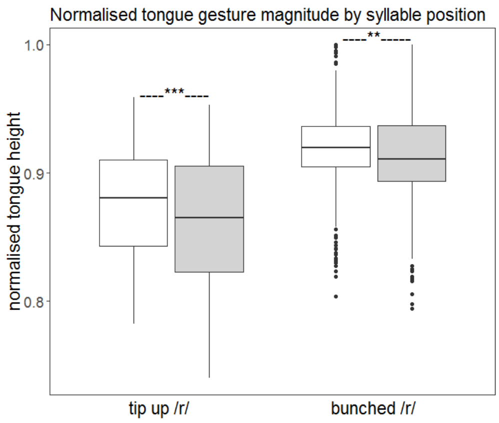


Fig. 3 shows that, for /l/, the magnitude of the tongue-gesture is always reduced in coda position. This finding supports those of earlier consonantal gesture-magnitude studies [12] and [6]; however, we also see that coda /l/ gesture magnitude varies according to utterance position. Perhaps surprisingly given [4]'s finding of greater degrees of palatal contact for consonants in utterance-initial position, tongue-tip gesture magnitude was significantly greater (p<0.001) for coda /r/ in utterance-final position. One possible explanation for this finding is that syllable lengthening in utterance-final position allows sufficient time for the tongue-tip gesture to reach its target, whereas shorter syllable lengths in other utterance positions result in gesture undershoot (see §4). The duration of the vowel + liquid portion of each key word in the study was measured. An ANOVA with post hoc Tukey tests confirmed that its duration was significantly longer (p<0.001) in utterance-final position than in utterance-initial or medial position.

3.2. 2 /r/

Separate models were run for normalised tongue gesture magnitude for tip-up (N=271) and bunched /r/ (N=512) datasets. For both /r/s, best-fit models contained the fixed factor *syllable position* and the random factor *speaker* (p<0.001): tip-up *syllable position* F=11.51, p<0.001; bunched *syllable position* F=9.47, p<0.01. Tongue gesture magnitude was significantly reduced for /r/ in coda position for both types of /r/, see Fig. 4, although greater reduction occurred for tip up /r/ than for bunched /r/.

Figure 4: Boxplots showing the effect of *syllable position* on normalised tongue gesture magnitude for tip-up and bunched /r/. White boxes=onset position, grey=coda position.



4. DISCUSSION

In answer to RQ1: "**How much do tongue shapes for /l/ and /r/ vary between syllable-onset and coda position**?" We have shown that while /r/ has two very distinct articulatory tongue shapes (tip-up and bunched), /l/ shows the greatest amount of positional and configurational variation (measured using normalised RMS distance) between onset and coda position. Palatal variants of both /l/ and /r/ showed the smallest degree of variation between onset and coda position. Previous research has shown that speech sounds with a primary articulatory constriction involving the tongue body are both more coarticulatorily aggressive and more resistant to coarticulation [5] [10]. The findings of the current study show that palatal sounds are also more consistently produced across the syllable.

In answer to RQ2: "**How do the syllable and utterance positions of a liquid consonant affect tongue gesture magnitude?**" In agreement with the findings of earlier studies [12], [1], [6], and [7] , for both /l/ and /r/, tokens in coda position had significantly reduced gesture magnitudes. The picture was less clear for the effects of utterance position. Utterance-position did not have an effect on /r/ magnitude, and for /l/, utterance-final position conditioned significantly greater gesture magnitudes than other positions. This finding is at odds with [4], which identified the strongest contact patterns in utterance-initial nasals. However, it agrees with the findings of [12], which identified a correlation between syllable length and the magnitude of the dorsal gesture for /l/; shorter syllables resulted in gesture undershoot. It would seem that tongue-tip gestures are greater in utterance-final position in our study, because they have time to reach their full target.

5. REFERENCES

[1] Browman, C.P. and Goldstein, L., 1995. Gestural syllable position effects in American English. In: F. Bell-Berti and L.J. Raphael, eds, *Producing Speech: Contemporary Issues*. pp. 19-33. New York: AIP Press

[2] Byrd, D., Kaun, A., Narayanan, S. and Saltzman, E., 2000. Phrasal signatures in articulation. In: M.B. Broe and J.B. Pierrehumbert, eds, *Papers in Laboratory Phonology V*. Cambridge: Cambridge University Press

[3] Core Team, R., 2013. R: A language and environment for statistical computing. R Foundation for Statistical Computing.

[4] Fougeron, C. and Keating, P.A., 1997. Articulatory strengthening at edges of prosodic domains. *J. Acoust. Soc. Am*., 101, 3728-3740.

[5] Fowler, C. & Brancazio, L. (2000). Coarticulation Resistance of American English Consonants and its effects on transconsonantal vowel to vowel coarticulation. *Language and Speech*, 43, 1-41.

[6] Giles, S. and Moll, K., 1975. Cineflourographic study of selected allophones of English /l/. *Phonetica*, 31, pp. 206-227.

[7] Kochetov, A., 2006. Syllable position effects and gestural organization: Evidence from Russian. In: L. Goldstein, D.H. Whalen and C. Best, eds, *Papers in Laboratory Phonology VIII*. New York: Mouton de Gruyter, pp. 565-588.

[8] Kuznetsova, A., Brockhoff, P.B. and Christensen, R.H.B., 2017. lmerTest Package: Tests in Linear Mixed Effects Models. *Journal of Statistical Software*, 82(13), 1-26.

[9] Lawson, E., Stuart-Smith, J., Scobbie, J.M., Yaeger-Dror, M. and Maclagan, M., 2010. Liquids. In: M. Di Paolo and M. Yaeger-Dror, eds, *Sociophonetics: A Student's Guide*. London: Routeledge, pp. 72-86.

[10] Recasens, D. & Espinosa, A. (2009). An articulatory investigation of lingual coarticulatory resistance and aggressiveness for consonants and vowels in Catalan. *J. Acoust. Soc. Am*., 125, 2288-2298.

[11] Scobbie, J.M., Lawson, E., Nakai, S., Cleland, J. and Stuart-Smith, J., 2015. Onset versus coda asymmetry in the articulation of English /r/. Proceedings of the 18th International Congress of Phonetic Sciences. 10th-14th Aug. 2015.

[12] Sproat, R. and Fujimura, O., 1993. Allophonic variation in English /l/ and its implications for phonetic implementation. *J. Phon*., 21, 292-311.

[13] Trager, G.L., 1959. The Systematization of the Whorf Hypothesis. *Anthropological Linguistics*, 1(1), 31-35.

[14] Wrench, A., 2012. Articulate Assistant Advanced User Guide. 2.14 edn. Edinburgh: Articulate Instruments Ltd.