THE EXTENT OF COARTICULATION OF ENGLISH LIQUIDS: AN ACOUSTIC AND ARTICULATORY STUDY

Paula West
Phonetics Laboratory, University of Oxford, U.K.

ABSTRACT

This study explores the production of long-domain coarticulatory patterns associated with English /l/ and /r/ by investigating the extent and nature of differences in articulation and acoustics. A speaker of Southern British English was recorded using simultaneous EPG and EMA. Six l/r minimal word pairs were recorded in a frame sentence. Strong local coarticulatory effects were found in the vowels adjacent to the liquid. Non-local anticipatory differences in F_3, lip and tongue position were found in vowels not adjacent to the liquid, with lower F_3, more lip rounding and backer and higher tongue position preceding an /r/. Non-local perseverative differences proved more elusive. EPG data showed significant differences in contact patterns for consonants up to two syllables before the liquid.

1. INTRODUCTION

English /l/ and /r/ have secondary articulations, whose coarticulatory effects have been claimed to have temporal extent longer than the phonological foot [5]. These patterns of long-domain coarticulation, largely manifested in F_2 and F_3, are perceptually available to listeners [13]. Whilst claims about the articulatory settings of long-domain patterns have been made, little articulatory study has been undertaken. Studies of long-distance coarticulation have been primarily acoustic [6, 8], although some articulatory results have been reported [10]. This study explores the production of long-domain coarticulatory patterns associated with English /l/ and /r/ by investigating the extent and nature of differences in articulation and acoustics, using electromagnetic articulography (EMA) [9] and electropalatography (EPG) [1]. EMA is an effective tool for investigating coarticulation, in particular lingual coarticulation [2]. Most studies to date have concentrated on local effects, although longer domain anticipatory lip-rounding at least a syllable before a rounded vowel has been reported [4]. Some studies report the use of EPG in conjunction with EMA to investigate coarticulation [*e.g.* 3], and the reliability of such data acquired simultaneously has been investigated [11]. The experiment reported here demonstrates that simultaneous EMA and EPG recordings provide complementary data helpful in building up a fuller picture of the nature and extent of coarticulation.

2. PROCEDURE

A male speaker of standard Southern British English was recorded producing six l/r pairs (leap/reap, lip/rip, lap/wrap, lope/rope, lobe/robe, lob/rob) in the frame sentence ‘Have you uttered a __ at home?’. Simultaneous acoustic, EMA and EPG recordings were made at Queen Margaret College, Edinburgh (QMC), using the Carstens AG100 Articulograph (sampling rate 500 Hz) and Reading EPG system (sampling rate 200 Hz). The audio signal was recorded using a Shure 849 electret microphone directly connected to a SoundBlaster analogue-to-digital converter in a PC and digitised with 16 bit resolution at 16kHz. The audio, EPG and EMA signals are synchronised to within the 5 ms accuracy imposed by the EPG sampling rate. A second audio recording was made with an Audio Technica AT803B electret lapel microphone attached to the helmet worn by the subject, an Alice Mic-Amp-Pack 2 preamplifier and a Sony DTC690 DAT player. This recording was used for the subsequent acoustic analysis; all acoustic interpretation of articulatory data is based on the synchronised recordings made using the PC, including the first audio recording, as the quality of that recording was more than adequate to identify acoustic landmarks.

Three EMA coils were placed on the subject’s tongue (approximately 1 cm, 3.5 cm and 5.5 cm back from the tip) and one each on the upper lip, lower lip and the gum beneath the lower incisors. Reference coils were placed on the bridge of the nose and the gum above the upper incisors. A plastic T-bar was used to approximate the subject’s occlusal plane, and to determine a co-ordinate system in which to situate the EMA data [14]. 7 repeats of each utterance were obtained and analysed.

Immediate post-processing of the EMA data was done at QMC by Alan Wrench. The reference coil positions were smoothed by filtering, and the data was corrected for head movement. The post-processed data was rotated so that the x axis was parallel to the subject’s occlusal plane, with the origin of the system at the junction of the central-maxillary incisor diastema and the incisors’ exposed tips (estimated using the T-bar). Coil positions should be interpreted as follows: x values increase with fronting, and y values increase with raising.

3. RESULTS

The EMA and EPG data were processed in *Matlab*, using modifications of a set of Matlab macros written by Noel Nguyen [7], and routines written by the author. EMA and EPG data were extracted from the relevant data matrices at points of interest identified from the acoustic signal. The first three formant frequencies and (x,y) co-ordinates of all the coils were measured at the midpoint of the vowels adjacent to the liquid (local effects) and the schwas of ‘uttered’ and ‘at’ (non-local). EPG contact data was examined for the alveolar consonants in ‘uttered’. Acoustic measurements, unless otherwise stated, were made in *Waves* using 18 pole Burg spectra with a 50 ms Hanning window, checked manually against wide-band spectrograms and DFT spectra.

Most of the data is modelled using multivariate general linear models (GLMs) in *SPSS* as the variables are inter-correlated. Measurements are assumed to be independent, i.e. the error term is assumed to be a vector of independent errors. Equivalence of the variance-covariance matrices was checked where possible, using Box's *M* test, supplemented by Levene's test. Multivariate normality of distribution is difficult to check and the test is robust provided the data exhibits symmetry, so each variable was tested separately for the normal distribution, with the one-sample Kolmogorov-Smirnov test. The data modelled by multivariate GLMs in this paper met the required assumptions as far as they could be tested, and the results of the tests will not be reported here. The *F* statistics reported are estimated by Pillai's test.

3.1. Local coarticulation

Strong coarticulatory effects were found in vowels adjacent to the liquid, for both the acoustic and EMA data. Consistent effects are lowered F_3, rounded lips and retracted tongue in /r/ contexts, reflecting the speaker's production of /r/ with labialization and a strongly retracted tongue position. EPG data was not examined at these points, due to minimal tongue-palate contact.

3.1.1 Anticipatory coarticulation. A multivariate GLM was constructed for the acoustic and EMA data from the midpoint of the schwa preceding the liquid, with word pair and liquid as factors. Both factors were highly significant ($F(15,51) \approx 45.557$, $p < 0.001$ and $F(75,275) \approx 2.445$, $p < 0.001$ for liquid and word pair respectively), and there was no interaction between them. Univariate tests show that the factor liquid was significant ($\alpha = 0.01$) for all of the acoustic and articulatory variables except tongue back y and F_2. The upper and lower lip, jaw and tongue positions of the schwa before the liquid differ significantly between /l/ and /r/ contexts, with differences in means ranging from 1.1 mm for jaw x and upper lip y to 3.6 mm for tongue tip x position. The tongue is backed and raised, the jaw and lower lip raised and protruded and the upper lip lowered and protruded (evidence of lip rounding) before the /r/ relative to the /l/. The acoustic manifestation of this is a difference in F_1 and F_3, with both F_1 and F_3 lower before /r/, with differences in means of 18 and 195 Hz respectively.

Spearman correlations were calculated between the EMA coil positions and formant frequencies. Significant correlations (2-tailed) are shown in table 1. F_1 is negatively correlated with upper lip x, lower lip and lower incisor x and y, and positively correlated with upper lip y. Thus the decrease in F_1 is correlated with lip rounding. F_2 is positively correlated with jaw, lower lip and tongue raising (y data). F_3 has significant positive correlations with tongue tip, mid and back x, and upper lip y. It is negatively correlated with upper lip x and lower lip y. Thus, F_3 lowers with tongue retraction and lip rounding (protrusion and lowering of the upper lip, raising of the lower lip). This corresponds exactly to the articulatory behaviour expected for the /r/ context in which F_3 is lower: F_3 lowering is an expected outcome of lip rounding.

Partial correlations controlling for variation due to liquid and word pair show only one significant result: F_2 is positively correlated with lower incisor y ($r = 0.3786$, $p < 0.001$), and thus increases with jaw raising. This suggests that most of the Spearman correlations can be partly attributed to liquid context.

	F_1		F_2		F_3	
	ρ	$p <$	ρ	$p <$	ρ	$p <$
LIX	-.302	0.005	n.s.	n.s.	n.s.	n.s.
ULX	-.279	0.010	n.s.	n.s.	-.466	0.001
LLX	-.298	0.006	n.s.	n.s.	n.s.	n.s.
TTX	n.s.	n.s.	n.s.	n.s.	.515	0.001
TMX	n.s.	n.s.	n.s.	n.s.	.491	0.001
TBX	n.s.	n.s.	n.s.	n.s.	.510	0.001
LIY	-.288	0.008	.427	0.001	n.s.	n.s.
ULY	.341	0.001	n.s.	n.s.	.477	0.001
LLY	-.293	0.007	.349	0.001	-.324	0.004
TTY	n.s.	n.s.	.430	0.001	n.s.	n.s.
TMY	n.s.	n.s.	.476	0.001	n.s.	n.s.
TBY	n.s.	n.s.	.545	0.001	n.s.	n.s.

Table 1. Significant Spearman correlations ($\alpha = 0.01$) for EMA coil positions and formant frequencies for the vowel before the liquid. N = 84 for F_1 and F_2, N = 77 for F_3.

3.1.2. Perseverative coarticulation. Perseverative effects were examined in the monophthongs /ɪ/, /æ/ and /ɒ/ (/i/ was excluded due to extensive devoicing). Acoustic measurements proved more difficult as many of the vowels were nasalized, displaying prominent nasal formants, so 22 pole Burg spectra with 25 ms windows were used in an attempt to pick out F_1, F_2 and F_3. F_3 could not be consistently measured for /ɒ/, and these measurements were omitted. A multivariate GLM constructed for the EMA variables, with liquid and word pair as factors, showed a significant interaction between liquid and word pair ($F(24,52) \approx 5.941$, $p < 0.001$), with both factors liquid and word pair highly significant ($F(12,25) \approx 16.834$, $p < 0.001$ and $F(24,52) \approx 88.224$, $p < 0.001$ respectively). Tamhane's T2 post-hoc tests showed significant differences ($\alpha = 0.01$) between all three vowels in all but a few variables. For lower incisor x and tongue tip y, /æ/ and /ɒ/ are not significantly different; /ɪ/ and /ɒ/ are not significantly different in lower lip y.

To investigate the interaction between word pair and liquid, independent samples *t*-tests were conducted separately for each vowel (variables did not differ significantly from the normal distribution, and homogeneity of variance was tested for). There were several significant differences between the liquid contexts at the 1% level (2-tailed test). These results are shown in table 2. Both vowels for which F_3 was measured, /ɪ/ and /æ/, showed a significantly lower third formant in the /r/ context, with differences in means of 128 and 443 Hz respectively. Lower incisor and lip positions were significantly different for /ɪ/ (fronter/more protruded by about 2.9 and 2.8 mm respectively) and /ɒ/ (higher by about 2.3 and 2.1 mm respectively) in the /r/ context. The speaker produces /r/ with a strongly retracted tongue, and the tongue middle is significantly backer after /r/ for /æ/ (by about 2.4 mm) and for /ɒ/ (3.3 mm). The tip is also significantly retracted for /ɒ/ after /r/ (by about 3.2 mm). For /ɪ/ and /æ/ the tongue back was significantly lowered in the /r/ context relative to the /l/ context (by approximately 1.9 mm and 1.3 mm), whereas for /ɒ/ the tongue back was significantly lowered (by approximately 3.5 mm) in the /l/ context. This difference is perhaps attributable to strong coarticulation of the /l/ with the following vowel and suggests that /r/, whose articulatory configuration has lower tongue back height than /i/ and /æ/, but higher tongue back height than the low vowel /ɒ/, is less strongly coarticulated.

	/ɪ/		/æ/		/ɒ/	
	t(12)	*p* <	*t*(12)	*p* <	*t*(12)	*p* <
F3	3.76	0.003	9.50	0.001	n/a	n/a
LIX	-6.71	0.001	n.s.	n.s.	n.s.	n.s.
LIY	n.s.	n.s.	n.s.	n.s.	-5.06	0.001
LLX	-4.26	0.001	n.s.	n.s.	n.s.	n.s.
LLY	n.s.	n.s.	n.s.	n.s.	-3.44	0.005
TTX	n.s.	n.s.	n.s.	n.s.	3.90	0.002
TMX	n.s.	n.s.	3.57	0.004	6.42	0.001
TBY	5.52	0.001	4.96	0.001	-5.08	0.001

Table 2. Significant results of independent samples *t*-test ($\alpha = 0.01$) on EMA and acoustic data divided by liquid context for three of the vowels after the liquid.

Spearman correlations for acoustic and EMA data were calculated for each vowel, with some significant results, mainly correlations between F$_3$ and articulatory variables. A partial correlation analysis controlling for vowel and liquid showed no significant correlations between the formants and any EMA coil positions. Most of the variability is accounted for by the combination of vowel and liquid. Within these categories there is no correlation between formant frequency variation and articulatory variation as measured by EMA.

3.2. Long-domain coarticulation

Distant coarticulatory effects were found in the /təd/ of ‘uttered’, which differs systematically depending on the following liquid, exhibiting lip rounding and a higher and backer tongue position before an /r/.

3.2.1. Vowels. Anticipatory coarticulatory effects were found in the schwa of ‘uttered’, a vowel not adjacent to the liquid. A multivariate GLM with liquid and word pair as factors was constructed for the EMA and acoustic data. Both factors were significant ($F(15,53) \approx 5.638$, $p < 0.001$, $F(75,285) \approx 1.742$, $p < 0.001$) with no interaction between them. Univariate tests show that the significance of the factor liquid is due to differences in F$_3$, tongue mid y, upper lip x and y. Tongue mid y is higher in /r/ than in /l/ contexts and upper lip is fronter, i.e. more protruded, and lower in /r/ contexts, again evidence of lip rounding. F$_3$ is significantly lower in /r/ than in /l/ contexts, with a difference in means of 51 Hz. The articulatory differences are less than 1 mm in magnitude. Spearman correlations calculated for the EMA and acoustic data were not significant for $\alpha = 0.001$. Partial correlations controlling for liquid and word pair showed a significant correlation between F$_1$ and tongue tip y ($r = -0.3613$, $p < 0.001$).

There were no significant acoustic differences in the schwa in the perseverative domain. The non-parametric Kolmogorov-Smirnov *Z* test (used because of unequal variances) showed no significant differences due to liquid context in the data set as a whole. Tests carried out for each word pair separately showed a few significant results in some EMA variables, but nothing systematic. Spearman’s ρ shows significant positive correlations (2-tailed test, $\alpha = 0.01$) between all three formants and many of the EMA variables; formants increase with fronting and raising of the articulators. The results are presented in Table 3.

	F_1		F_2		F_3	
	ρ	*p* <	ρ	*p* <	ρ	*p* <
TTX	.465	0.001	.488	0.001	.414	0.001
TMX	.322	0.003	.554	0.001	.371	0.001
TBX	n.s.	n.s.	.594	0.001	.352	0.001
LIY	.394	0.001	.576	0.001	n.s.	n.s.
LLY	.538	0.001	.587	0.001	.349	0.002
ULY	.411	0.001	.601	0.001	n.s.	n.s.
TTY	n.s.	n.s.	.593	0.001	.452	0.001

Table 3. Significant Spearman correlations ($\alpha = 0.01$) for EMA coil positions and formant frequencies for the vowel of ‘at’. N = 84 for F$_1$ and F$_2$, N = 79 for F$_3$.

Partial correlations controlling for variation due to vowel and liquid showed F$_2$ positively correlated with tongue tip y and tongue middle and tongue back x ($r = .363$, $.411$ and $.529$, $p < 0.001$) and F$_3$ positively correlated with tongue tip x ($r = .394$, $p < 0.001$). Thus F$_2$ increases with tongue fronting and tip raising, and F$_3$ increases with fronting of the tongue tip.

3.2.2. Consonants. EPG data, extracted at the point of maximal contact, was examined for two consonants: the /t/ and /d/ of ‘uttered’. Data is missing from the subject's leftmost contact five rows from the front, as the contact on the artificial palate was broken. Figure 1 summarises the contact patterns for each consonant in the two contexts, asterisks mark areas of difference.

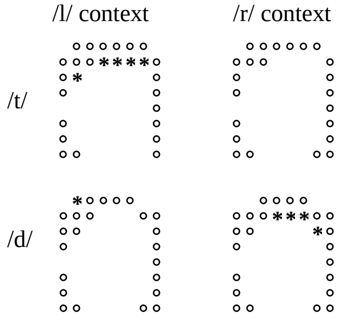

Figure 1. EPG contact patterns for the /t/ and /d/ of ‘uttered’ in /l/ and /r/ contexts. Open circles show contacts made consistently in each context, asterisks mark contacts different across contexts.

Non-parametric two-sample Kolmogorov-Smirnov *Z* tests conducted for various summary measures showed significant differences in the different liquid contexts. /t/ exhibited more contact in the front region (first three rows) in an /l/ than an /r/ context ($Z = 1.855$, $p < 0.002$), whereas /d/ showed less contact in the /l/ context for row 2 ($Z = 1.637$, $p < 0.009$).

Kolmogorov-Smirnov *Z* tests conducted on EMA data at these points did not show significant differences in tongue tip position. This is unsurprising, as the ‘tongue tip’ coil was 1 cm behind the tip of the tongue, and not measuring movement of the part of the tongue which made contact with the front of the EPG palate. For /t/ the tongue back was significantly higher in the /r/ context ($Z = 1.530$, $p < 0.018$) with a difference in mean position of 0.7 mm. For /d/ upper lip x and y placement differed significantly between the two contexts ($Z = 1.789$, $p < 0.003$ and $Z = 1.541$, $p < 0.017$), being more protruded and lower in the /r/ context by 0.5 mm in each direction.

3.3. The extent of coarticulation

Dynamic time warping (DTW) is a method of non-linearly aligning two signals [12]. The technique was used to align the multiple repetitions of EMA data, so that comparisons of articulatory events could be made across the naturally produced repeats which varied slightly in timing. The analysis is supplementary to the pointwise analyses previously described, and confirms those results. For each word, the 7 repetitions were aligned using an algorithm which warped all the EMA coil traces of a token simultaneously based on the best overall fit to another (reference) token. For each pair, the /l/ and /r/ tokens were then time aligned using a similar procedure, and compared. Because of the difficulty of performing repeated statistical tests on correlated data, 95% confidence intervals were calculated from the aligned signals for each word. Where confidence intervals for any pair do not overlap, data can safely be assumed to differ, although this is probably a conservative estimate of the extent of differences. An example is given in Figure 2, which shows the extent of differences in articulation for the tongue tip x and upper lip x and y of the minimal pair ‘leap/reap’ (from the schwa 2 syllables before the liquid to the vowel after the liquid). The waveform for the ‘leap’ utterance is given as a reference, with segmental labelling.

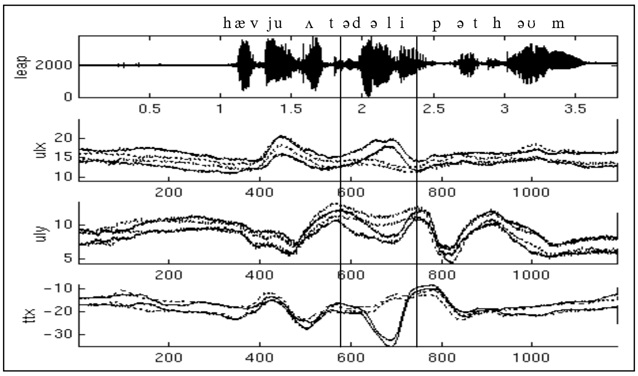

Figure 2. Waveform from a ‘leap’ utterance and 95% confidence intervals (solid line ‘reap’, dashed line ‘leap’ stimuli) from time warped data for tongue tip x and upper lip x and y coils. Vertical bars mark the extent of clear coarticulatory difference.

4. CONCLUSIONS

Long-distance coarticulatory effects associated with English liquids extend up to two syllables before the liquid for this speaker. EMA and EPG data show that there are significant differences in tongue and lip placement for consonants and vowels up to two syllables before a liquid, depending on the liquid they precede. The acoustic differences are manifested mainly in F_3, which is lower in /r/ contexts. Correlations between EMA articulatory data and acoustic data are significant in some cases, largely due to variation associated with liquid and word pair. Partial correlations, controlling for variation due to the liquid, produced fewer significant results, showing that the relationship between EMA and acoustic data is complex. The lack of correlations may be due to the magnitude of the differences being measured (often under 1 mm for the EMA data), the limited nature of the EMA data and to inherent measurement error. Those articulatory-acoustic correlations found tend to confirm the predictions of acoustic theory about correlations between vocal tract configurations and formant frequencies. The experiment confirms that phonological distinctions are not just made locally, and that both the tongue and lips can play a role in long-distance coarticulation.

ACKNOWLEDGEMENTS

I would like to thank John Coleman, Burton Rosner, Andrew Slater, Mario Cortina, Alan Wrench and the subject for their help.

REFERENCES

[1] Hardcastle, W.J. and F. Gibbon. 1997. Electropalatography and its Clinical Applications. In Ball, M and C. Code (eds.), *Instrumental Clinical Phonetics.* London: Whurr.

[2] Hoole, P. and N. Nguyen. 1997. Electromagnetic articulography in coarticulation research. *Forschungsberichte des Instituts für Phonetik und Sprachliche Kommunikation der Universität München,* 35, 177-184.

[3] Hoole, P., N. Nguyen-Trong and W. Hardcastle. 1993. A comparative investigation of coarticulation in fricatives: electropalatographic, electromagnetic and acoustic data. *Language and Speech,* 36 (2-3), 235-260

[4] Katz, W., J. Machetanz, U. Orth and P. Schönle. 1990. A Kinematic Analysis of Anticipatory Coarticulation in the Speech of Anterior Aphasic Subjects Using Electromagnetic Articulography. *Brain and Language,* 38 (4), 555-575.

[5] Kelly, J. and J. Local. 1986. Long-domain resonance patterns in English. *International Conference on Speech Input/Output; Techniques and Applications.* Conference publication no. 258, 56-59 London: Institution of Electrical Engineers.

[6] Magen, H.S. 1997. The extent of vowel-to-vowel coarticulation in English. *Journal of Phonetics,* 25, 187-205.

[7] Nguyen, N. 1996. A Matlab package for the analysis of EMA/EPG data. *QMC Working Papers in Speech and Language Sciences,* 1, 5-12.

[8] Öhman, S.E.G. 1966. Coarticulation in VCV utterances: spectrographic measurements. *Journal of the Acoustical Society of America,* 24, 979-998.

[9] Perkell, J. S., M. H. Cohen, M. A. Svirsky, M. L. Matthies, I. Garabieta and M. T. T. Jackson. 1992. Electromagnetic midsaggital articulometer systems for transducing speech articulatory movements. *Journal of the Acoustical Society of America,* 92 (6), 3078-3096.

[10] Recasens, D. 1989. Long-range coarticulation effects for tongue-dorsum contact in VCVCV sequences *Speech Communication,* 8 (4), 293-307.

[11] Rouco, A. and D. Recasens. 1996. Reliability of electromagnetic midsaggital articulometry and electropalatography data acquired simultaneously. *Journal of the Acoustical Society of America,* 100 (5), 3384-3389.

[12] Strik, H. and L. Boves. 1990. A dynamic programming algorithm for time-aligning and averaging physiological signals related to speech. *Journal of Phonetics,* 19, 367-378.

[13] West, P. 1997. *Perception of /l/ and /r/ in English dialects.* M.Phil. thesis, University of Oxford.

[14] Westbury, J. R. 1994. On coordinate systems and the representation of articulatory movements. *Journal of the Acoustical Society of America,* 95 (4), 2271-3.