Gestural Reduction, Phonetic Context, and Word Frequency: A Case Study of Word-Final Singleton /l/ in British English[1]

Mitsuhiro NAKAMURA

1. **Introduction**

This study reports on the continuing investigation into the articulatory realisation of the voiced alveolar lateral approximant /l/ in English and contributes to an understanding of phonetic variation and articulatory organisation. We describe an exploratory study examining the effects of phonetic context and word frequency on gestural reduction in word-final singleton /l/s in British English. Both phonetic context and word frequency influence the phonetic shape of the consonant in question, but the word frequency effects on gestural reduction are not fully understood, particularly at the level of articulatory organisation. This issue is considered by Lin *et al.* (2011, 2014), an articulatory study of gestural reduction in the /lC/ sequence in American English. The current study aims to further investigate the patterns of gestural reduction and to relate them to the frequency effects and progressive sound changes of /l/ vocalisation.

In section 2 of this paper, the pattern of allophonic variation in /l/ is outlined and phonetic conditioning factors for clear, dark, and vocalised /l/ are summarised. Section 3 reviews two potential causes of phonetic variation, namely phonetic context and word frequency. We focus on one landmark study on gestural reduction of the English /l/, Lin *et al.* (2011, 2014) and discuss their hypotheses, what we call, 'place-conditioned' gestural reduction and 'frequency-conditioned' gestural reduction, which I believe are relevant to the understanding of phonetic variation and articulatory organisation. Section 4 describes how the spatiotemporal characteristics of the /l/ articulation are identified and analysed. The articulatory data presented in this paper was drawn from the multichannel articulatory (MOCHA) database (Wrench 1999, 2000). Section 5 presents the results of an articulatory analysis of the /l/s and evaluates the above hypotheses. We discuss our findings, in a preliminary way, in relation to the results reported by Lin *et al.* (2011, 2014). Finally, section 6 describes a summary of

[1] This paper is a revised and enlarged version of my poster presentation 'Effects of phonetic context and word frequency on gestural reduction' at the 14th Australasian International Conference on Speech Science and Technology, Macquarie University, Sydney 3-6 December 2012. This research is supported in part by a grant from the Japan Society for the Promotion of Science, Grant-in-Aid for Scientific Research (C) 2022, Grant no. 21K00487.

our observations and analyses and a prospect for future research.

2. Clear/dark allophony and /l/ vocalisation[2]

It is generally acknowledged that many varieties of English exhibit allophonic variation in /l/ (e.g. Schneider *et al*. 2004; Wells 1982). According to Wells (1982: 45), 'in RP clear [l] and dark [ɬ] are in complementary distribution, the rule being that the former occurs before a vowel or /j/, the latter elsewhere.' The RP allophonic rules for dark [ɬ] and /l/ vocalisation are formulated as follows (Wells, 1982: 258f):

(1a)　l → ɬ　/　_____ { ‖ / #$_0$ C }

(1b)　ɬ → o　/　_____ { ‖ / #$_0$ C }

The rule in (1a) sets a clear allophone of /l/ as the basic and changes it into the dark one before a consonant or a pause (The diacritic '‖' indicates a major boundary). Thus, we find the dark allophone in words such as *feel*‖ [fiːɬ], *milk* [mɪɬk], and *small corner* [smɔːɬ kɔːnə]. The rule in (1b) converts the dark [ɬ] into a vowel of the [o] or [ʊ] type: this phenomenon is called /l/ vocalisation. The pronunciation of *feel*‖, *milk*, and *small corner* becomes [fiːo], [mɪok], and [smɔːo kɔːnə] respectively. Note in passing that the development of the dark [ɬ], or the distinction between clear and dark /l/, is arguably the essential condition for the actuation of /l/ vocalisation (Johnson & Britain 2007).

From the point of view of parametric phonetics (e.g. Abercrombie 1965), the lateral approximant in English is produced with two lingual gestures, the tongue tip and the tongue dorsum (e.g. Browman & Goldstein 1995; Jones 1966). The formation of the dark [ɬ] involves a primary articulation by the tongue tip closing at the alveolar ridge and a secondary articulation by the tongue dorsum arching towards the velum (velarisation) or backing towards the pharynx (pharyngealisation). In the production of the vocalised /l/, the dark [ɬ] can lose its tongue tip constriction, while a tongue dorsum configuration

[2] The description and some ideas in this section are taken from my earlier works, namely Nakamura (2009, 2011)

similar to a high back vowel remains[3]. Thus, /l/ vocalisation can be considered as gestural reduction, a coda weakening of the tongue tip gesture. The current study focuses on the magnitude of the tongue tip raising gesture, not the dorsum gesture, and investigates the effects of phonetic context and word frequency on the production of the English /l/.

3. Effects of phonetic context and word frequency on gestural reduction[4]

Articulatory studies conducted by Lin *et al.* (2011, 2014) explicitly raise the question of whether word frequency effects can be seen when examining the magnitude of the tongue tip gesture in the production of the English /l/. Using the ultrasound imaging technique and acoustic analysis, they examined whether the magnitude of the raising gesture of the tongue tip in coda /lC/ clusters is conditioned not only by place of articulation of the following consonant (C) but also by the frequency of the word containing the /lC/ cluster. Previous articulatory studies have shown the place-dependent tendency that the complete closure of the /l/ tends to be reduced in the /lC/ cluster where the final consonant is a labial (e.g. *help*, *golf*) or a velar (e.g. *milk*) (e.g. Borowsky 2001; Giles & Mill 1975; Nakamura 2009; Scobbie, Pouplier, & Wrench 2007; Stuart-Smith *et al*. 2006; Wrench & Scobbie 2003). Also, studies of the word frequency effects on sound change have demonstrated that reduction phenomenon is observed in high-frequency words more often than low-frequency words (e.g. Bybee 2002). Based on these observations, Lin *et al*. (2014: 12) formulated the hypothesis: 'ultrasound image of the tongue tip and blade region will show weaker alveolar constrictions (i.e., greater apertures) in labial (*help*) and velar (*milk*) context than in alveolar (*built*) contexts, and that reduction will be greater in high-frequency (*help, milk*) than in low-frequency (*whelp, ilk*) words.' This hypothesis is supported by the main novel finding that the constriction by the tongue tip is weaker in high-frequency words in the coda /lC/ cluster with a labial or a velar consonant. This result serves as evidence for their proposal that such weakening of the tongue tip gesture in the /lC/ cluster would be further reduced to /l/ vocalisation since high-frequency words involve the well-practiced articulatory routines (Lin *et al*. 2014). The mechanism of gestural reduction and word frequency is appealing but has not been substantiated fully by the empirical data.

The aim of the current study is to further investigate the effects of phonetic context and word

[3] This interpretation may be supported by Gick *et al*. (2002), an MRI study that demonstrates the similarity between the pharyngeal configuration of /l/ and /ɔ:/ spoken by three native speakers of American English.

[4] Sociolinguistic factors that influence the use of /l/ vocalisation have been investigated by various researchers: for instance, Tolfree (1999), Horvath & Horvath (2001), and Stuart-Smith *et al*. (2006).

frequency on the production of the /l/. The study conducted by Lin *et al.* (2011, 2014) focuses on the production of the words with the /lC/ cluster in a 'non-vocalising' variety of American English. The current study examines the articulation of the word final singleton /l/s in British English, namely a 'vocalising' variety of English. The two hypotheses Lin *et al*. have formulated are recapitulated here. The first one is characterised as a 'place-conditioned' gestural reduction: the raising gesture of the tongue tip would be spatially reduced when the /l/ is followed by a non-coronal consonant (i.e. labial or dorsal). The second one is characterised as a 'frequency-conditioned' gestural reduction: reduction including /l/ vocalisation would be greater for high frequency words than for low frequency words. In order to test these two hypotheses and to further investigate the different roles of phonetic context and word frequency on gestural reduction, we examine the /l/s in word-final position (e.g. *all, bowl, compile, deal, fill, male, roll, tadpole, will*) and how the effects occur across word boundaries.

4. Methodology

4.1 Articulatory database used

The data analysed in this paper came from a larger and continuing articulatory-acoustic study on phonetic variation in English[5]. The speech materials used were drawn from the multichannel articulatory database MOCHA-TMIT described by Wrench (1999, 2000). This database comprises the articulatory and acoustic data of 460 phonetically balanced sentences read by native speakers of English. Data acquisition was made with Carstens Electromagnetic Articulograph (EMA, 500Hz sample rate), Reading Electropalatograph (EPG, 200Hz sample rate), Electrolaryngograph (Lx), and acoustic recordings. In the EMA data recording, the tongue tip (TT) coil was placed at about 5-10mm and the tongue dorsum (TD) coil was placed at about 50-70mm from the tongue tip, as measured with the tongue extended. In this study, the EMA data and the EPG contact patterns were analysed for the selected utterances spoken by three speakers of Southern British English referred to as SE (female), SA (male), and AP (male).

4.2 Analysis

The analysis of the MOCHA data employs MATLAB analysis tools. The configurational characteristics were studied for EPG contact patterns in an alveolar zone[6]. This zone, as shown in Figure 1(a),

[5] The description of the MOCHA-TIMIT database and /l/-zone (next section) in this paper is based on my previous work, Nakamura (2009).

[6] The Reading EPG provides tongue-palate contact data at 62 normalised positions on the hard palate.

covers the front three rows of 18 contacts. Following the index used by Scobbie *et al.* (2007), alveolar contact in the production of /l/ was defined by reference to an '/l/-zone' in the three front rows on the palate as shown in Figure 1(b). If any contact occurred in this zone during the /l/ production, the token was classified as 'consonantal.' The EPG frame for the analysis was the one specified by the tangential velocity minimum of the EMA tongue tip trajectory. For some vocalised tokens, the minimum velocity of the dorsum trajectory was used as a reference point.

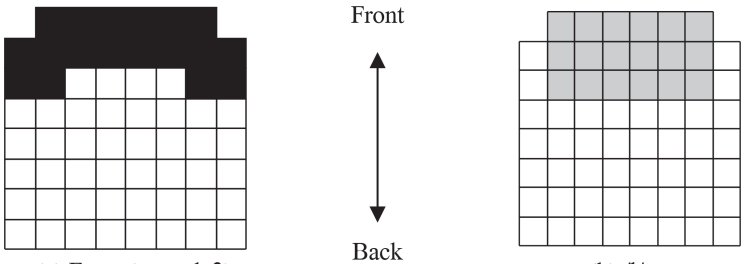


Figure 1: EPG prototypical palate and /l/ zone

A dataset of 241 tokens was used for analysis: 84 word-final /l/s × 3 speakers−11 unanalysable cases. Table 1 below shows the rate of vocalisation coded by the following context (consonant, vowel, and pause). The /l/ vocalisation is typical when followed by a consonant (e.g. *small boy, football team, clamshell near, deal with, creole cooking, all colleges*), but less predictable when followed by a pause (e.g. *compile‖, bowl‖, doll‖*).

Table 1. Rate of vocalisation by the following context

	Consonant	Vowel	Pause	Total
Vocalised	110	5	18	133
Consonantal	26	55	27	108

The EMA data was used for measuring the vertical height of the tongue tip (TTy). Figure 2 and 3 show the EMA recording of the tongue tip vertical movement (TTy), the tongue dorsum horizontal movement (TDx), the tongue dorsum vertical movement (TDy), and the lower lip vertical movement (LLy). Figure 2 illustrates the EMA data of the vocalised /l/ in the production of *small boy*. The TTy height in this particular example is measured at the point indicated by the thin vertical line (on the TTy panel): it might be helpful to follow the phonetic transcription and segmentation at the top. Figure 3 demonstrates the EMA trajectories of the consonantal /l/ in the production of *wall with*. The measurement point of the TTy height is indicated by the vertical line. In this study, TTy height was

standardised across all speakers by dividing individual values by average values for a given speaker.

Figure 2: EMA data of the vocalised /l/ in the production of *small boy* (SE)

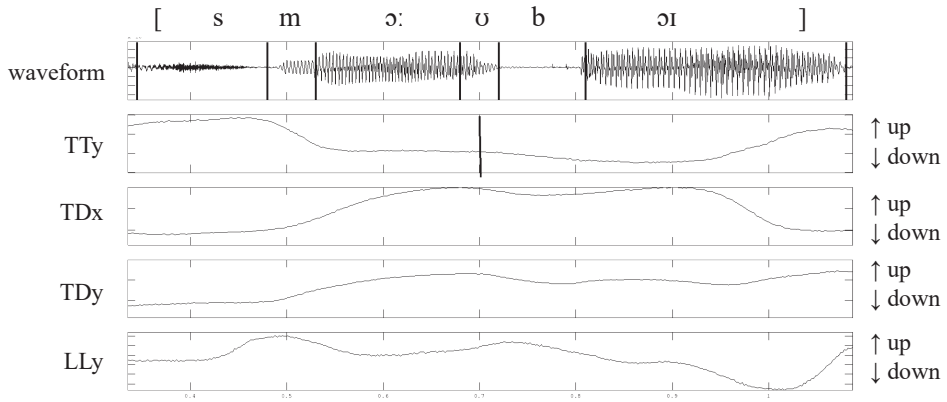
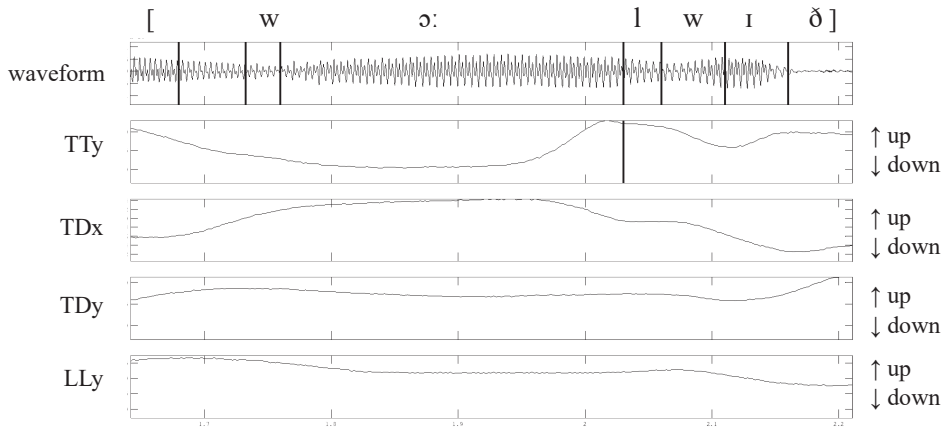


Figure 3: EMA data of the consonantal /l/ in the production of *wall with* (SA)



To test the hypothesis of place-conditioned gestural reduction (see section 3), the /l/ tokens followed by a consonant were examined (e.g. *small corner*). 136 tokens were coded by places of articulation for the following consonant: labial, coronal, and dorsal (e.g. *full moon*, *all times*, *all colleges*). To examine the hypothesis of frequency-conditioned gestural reduction, word frequency (log10 transformed) was calculated based on Kilgarriff (1998) which records the raw number of occurrences. Using a cut-off point (4.04), all 241 tokens were divided into two groups: a high frequency (e.g. *all, tell*) and low frequency (e.g. *yell, pill*) group. For the examination of the two hypotheses above, statistical analyses (one-way ANOVA, t-test) were performed for vocalised and consonantal tokens separately.

5. Results and discussion

5.1 Effects of the following consonant

Table 2 summarises mean values of standardised TTy height by place of articulation for the following consonants, labial, coronal, and dorsal. The results of the vocalised /l/ tokens pooled across three speakers demonstrate that TTy height is significantly higher in the coronal context than in the labial and dorsal contexts [$F(2,107)=6.42$, $p=0.002$]. Bonferroni post-hoc test shows significant coronal-labial and coronal-dorsal differences ($p<0.01$). In contrast, for the consonantal /l/ tokens, the distinction leveled out [$F(2,23)=1.55$, $p=0.23$]: LSD post-hoc exact test indicates non-significant labial-coronal ($p=0.29$), coronal-dorsal ($p= 0.27$), and labial-dorsal ($p=0.11$) differences.

Table 2. Mean standardised TTy height by places of articulation (mm)

(a)Labial

	n	mean (SD)	min	max
Vocalised	28	-0.86 (0.73)	-2.63	0.33
Consonantal	2	0.58 (0.16)	0.46	0.73

(b)Coronal

	n	mean (SD)	min	max
Vocalised	50	-0.13 (0.61)	-1.72	1.48
Consonantal	15	1.01 (0.45)	0.17	1.74

(c)Dorsal

	n	mean (SD)	min	max
Vocalised	32	-0.81 (1.59)	-2.94	6.57
Consonantal	9	1.25 (0.63)	-0.44	2.75

*N=110 (Vocalised); N=26 (Consonantal)

For the case of the vocalised /l/, it was found that the location of the tongue tip (TTy height) is significantly influenced by the place of articulation of the following consonant. This implies that the coarticulatory effects extend beyond a word boundary. Hardcastle & Barry (1989), Lin *et al*. (2011, 2014), Scobbie & Pouplier (2010), and Wrench & Scobbie (2003) have suggested that the origin of /l/ vocalisation is the coda reduction of the tongue tip and/or dorsum gesture. This characterisation is further substantiated by our findings.

In contrast, for the case of the consonantal /l/, our results do not support the hypothesis of place-

conditioned gestural reduction. It is predicted that the following non-coronal consonant (i.e. labial or dorsal) might influence the degree of complete closure by the tongue tip: the raising gesture of the tip is weakened, although there is still a complete closure. Our results indicate that the magnitude of the tip gesture does not significantly vary with the place of articulation of the following consonant. The TTy height is actually highest when followed by a dorsal consonant. This might be related to the combination of words: four out of nine tokens are *Will you* and show the higher mean TTy height value (1.6) than the other five (e.g. *small corner*): speakers do not use the vocalised /l/ even if the phonetic/phonological conditions are met. These results are inconsistent with those reported by Lin *et al*. (2014) on the case of the /lC/ sequence. One possible reason might be the choice of the accent of English, non-vocalising variety vs. vocalising variety. Another reason might be the presence/absence of a word boundary, /lC/ sequence vs. /l/+word boundary+consonant. Also, the nature of the boundary type might be related to the extent of coarticulatory effects of the following consonant.

5.2 Effects of word frequency

In Table 3 below, the mean values of standardised TTy height are summarised by word frequency. The statistical test shows that TTy height difference is non-significant for the vocalised /l/ tokens [t=1.03, df=131, p= 0.304] but is significant for the consonantal /l/ tokens [t=-2.17, df=106, p=0.032].

Table 3. Mean standardised TTy height by word frequency (mm)

(a)Vocalised

	n	mean (SD)	min	max
Low	66	-0.48 (1.12)	-1.99	1.12
High	67	-0.66 (0.91)	-2.94	1.48

(b)Consonantal

	n	mean (SD)	min	max
Low	57	0.82 (0.38)	-0.06	1.75
High	51	0.99 (0.42)	0.17	2.75

*N=133 (Vocalised); N=108 (Consonantal)

The hypothesis of frequency-conditioned gestural reduction predicts that reduction including /l/ vocalisation would be greater for high frequency words than for low frequency words. Our result indicates that, when the /l/ is consonantal, TTy height for low frequency words was significantly higher than that for high frequency words. This is compatible with the results reported by Lin *et al.* (2014). Thus, the consonantal cases in the current study are suggestive of the frequency effect on gradual articulatory changes.

In contrast, the vocalised cases do not show the effects of word frequency: the TTy height of the /l/ in high frequency words is not significantly different from that in low frequency words. As we have seen in the previous section, the TTy height of the vocalised /l/ varies significantly with the place of articulation of the following consonant. It appears that phonetic context, rather than word frequency, might impact on the degree of gestural reduction in the vocalised /l/ which has already undergone change in a vocalising variety of English[7]. This view assumes the cognitive process of abstraction and categorisation of the new phonetic form (e.g. MacKenzie 2019; Pierrehumbert 2006) and would be elaborated in further research.

6. Summary and prospects

The current study has explored the effects of phonetic context and word frequency on gestural reduction in the production of the word final singleton /l/s in British English. The preliminary results revealed that (i) labial and dorsal onsets of the following word contribute to the tip gesture reduction in the vocalised /l/s and that (ii) word frequency influences the reduction in the consonantal /l/s. While these findings are not totally consistent with previous studies on articulatory organisation and language use, they encourage reconsideration of the vocalised /l/, a phonetic form which has already undergone change in a vocalising variety of English. And our findings support the view that vocalisation is not simply a coda weakening of the tongue tip/blade gesture. Further research is needed to investigate whether and how the role of phonetic context and word frequency is different between a non-vocalising variety and a vocalising variety of English.

One aspect which has not been considered in this study is the effects of individual speakers. It is reasonable to assume that speakers use idiosyncratically preferred forms of the word final /l/s, the vocalised one or not. Stevens & Harrington (2014: 6) mention that '[s]peech production is idiosyncratic and differences between speakers can be attributed to learned behaviour as well as to physiology.'

[7] Labov (2006) illustrates the dominant effects of phonetic context on creating categorical allophones.

This suggests that the 'cognitive' effects of frequency[8], both contextual and lexical, could be related to aspects of speech motor control in variable realisations of the /l/ in English. Within the usage-based framework, Hay & Foulkes (2016: 300) mention that '...linguistic representations are, at some level, constantly updated with experience and are rich in phonetic and indexical detail.' Given this, it could be assumed that the phonological/gestural representation of a given word would be updated or reorganised to the form individual speakers select. The interaction between phonetic context and word frequency in the reorganisation of articulatory parameters will be substantiated by subsequent research.

Finally, the current study has employed the data collected from the multichannel articulatory database. Controlled experiments are also needed to validate the articulatory patterns we have observed. Much remains to be done. It is hoped, however, a parametric approach explored in this study provides a step forward in understanding the relationship between gestural reduction, phonetic context, and word frequency.

Bibliography

Abercrombie, David. (1965) 'Parameters and phonemes.' In his *Studies in Phonetics and Linguistics.* Oxford: Oxford University Press, pp. 120-124.

Borowsky, Tony. (2001) 'The vocalization of dark l in Australian English.' In D. Blair and P. Collins. (Eds.) *English in Australia*. Amsterdam: John Benjamins Publishing Co. 69-87.

Browman, Catherine and Louis Goldstein. (1995) 'Gestural Syllable Position Effects in American English.' F. Bell-Berti and L.J. Raphael (Eds.) *Producing Speech: Contemporary Issues.* New York: AIP Press, pp. 19-33.

Bybee, Joan. (2002) 'Word frequency and context of use in the lexical diffusion of phonetically conditioned sound change.' *Language Variation and Change*, 14, pp. 261-290.

Gick, Brian, A.Min Kang, and Douglas H. Whalen. (2002) 'MRI evidence for commonality in the post-oral articulations of English vowels and liquids.' *Journal of Phonetics*, 30, pp. 357-371.

Giles, Stephen B. and Kenneth L. Moll. (1975) 'Cinefluorographic Study of Selected Allophones of English /l/.' *Phonetica*, 31, pp. 206-227.

Hardcastle, William and William Barry. (1989) 'Articulatory and perceptual factors in /l/ vocalisations in English.' *Journal of the International Phonetic Association*, 15(2), pp. 3-17.

Hay, Jennifer and Paul Foulkes. (2016) 'The Evolution of medial /t/ over real and remembered time.' *Language*, Vol. 92(2) pp. 298-330.

Horvath, Barbara M. and Ronald J. Horvath. (2001) 'A multilocality study of a sound change in progress: The case of /l/ vocalisation in New Zealand and Australian English.' *Language Variation and Change*, 13, pp. 37-57.

[8] See Bybee (2002) and Pierrehumbert (2006) for discussion.

Johnson, Wyn and David Britain. (2007) 'L-vocalisation as a natural phenomenon: explorations in sociophonology.' *Language Sciences*, 29, pp. 294-315.

Jones, Daniel. (1966) *The Pronunciation of English*. Fourth Edition. Cambridge: Cambridge University Press.

Kilgarriff, Adam. (1998) 'BNC database and word frequency lists.' Online: http://www.itri.bton.ac.uk/~Adam.Kilgarriff/bnc-readme.html, assessed on 4 Sept, 2011. Currently, the page is moved to https://www.kilgarriff.co.uk/bnc-readme.html.

Labov, William. (2006) 'A sociolinguistic perspective on sociophonetic research.' *Journal of Phonetics*, 34, pp. 500-515.

Lin, Susan S., Patrice S. Bedder, and Andries W. Coetzee. (2011) 'Gestural Reduction and Sound Change: An Ultrasound study.' *Proceedings of the XVII International Congress on Phonetic Science, Hong Kong, 2011*, pp. 1250-1253.

Lin, Susan, Patrice S. Bedder, and Andries W. Coetzee. (2014) 'Gestural reduction, lexical frequency, and sound change: A study of post-vocalic /l/.' *Laboratory Phonology*, 5(1), pp. 9-36.

MacKenzie, Laurel. (2019) 'Perturbing the community grammar: Individual differences and community-level constraints on sociolinguistic variation.' *Glossa: a journal of general linguistics*, 4(1), pp. 1-23.

Nakamura, Mitsuhiro. (2009) 'Exploring Vocalization of /l/ in English: an EPG and EMA study.' *Proceedings of Interspeech 2009, 6-10 September, Brighton, U.K.*, pp. 2311-2314.

Nakamura, Mitsuhiro. (2011) 'A Parametric Approach to L Vocalisation in English: Some Findings and Hypotheses.' *Festschrift Commemorating the Retirement and the 70th Birthday of Professor John C. Wells* (*English Phonetics* No. 14 & No. 15, The English Phonetic Society of Japan) , pp. 99-108.

Pierrehumbert, Janet B. (2006) 'The next toolkit.' *Journal of Phonetics*, 24, pp. 516-530.

Schneider, Edgar W., Kate Burridge, Bernd Kortmann, Rajend Mesthrie, and Clive Upton. (2004) (Eds.) *A Handbook of Varieties of English*. Volume 1: Phonology. Berlin Mouton de Gruyter.

Scobbie, James M., Maria Puplier, and Alan Wrench. (2007) 'Conditioning factors in external sandhi.' *Proceedings of the XVI International Congress of Phonetic Science, Saarbrücken, Germany*, pp. 441-444.

Scobbie, James M. and Maria Pouplier. (2010) 'The role of syllable structure in external sandhi: An EPG study of vocalisation and retraction in word-final English /l/.' *Journal of Phonetics*, 38 (2), pp. 240-259.

Stevens, Mary and Jonathan Harrington. (2014) 'The individual and the actuation of sound change.' *Loquens* 1(1), pp. 1-10.

Stuart-Smith, Jane, Clare Timmins, and Fiona Tweedie. (2006) 'Conservation and innovation in a traditional dialect: L-vocalization in Glaswegian.' *English World Wide*, 27(1), pp. 71-87.

Tolfree, L. (1999) 'South East London English: Discrete vs. continuous modelling of consonant reduction.' In P. Foulkes & G. Docherty. (Eds.) *Urban voices: Accent studies in the British Isles.* London: Arnold, pp. 163-184.

Wells, John. (1982) *Accents of English: An Introduction*. Volume 1. Cambridge: Cambridge University Press.

Wrench, Alan. (1999) "The MOCHA-TIMIT Articulatory Database." University of Edinburgh, available at http://www.cstr.ed.ac.uk/research/projects/artic/ mocha.html., November, 1999.

Wrench, Alan. (2000) 'A multi-channel/multi-speaker articulatory database for speech recognition research.' *Phonus* 5, 1-13.

Wrench, Alan and James Scobbie. (2003) 'Categorising Vocalisation of English /l/ Using EPG, EMA, and

ULTRASOUND.' *Proceedings of the Sixth International Seminar on Speech Production, Sydney, Australia, December 7 to 10, 2003*, pp. 314-319.