Interspeech 2022
18-22 September 2022, Incheon, Korea

Acquisition of allophonic variation in second language speech: An acoustic and articulatory study of English laterals by Japanese speakers

Takayuki Nagamine[1]

[1]Lancaster University
t.nagamine@lancaster.ac.uk

Abstract

Acquisition of positional allophonic variation is seen as the foundation of a successful L2 speech learning. However, previous research has mostly focused on the phonemic contrast between English /l/ and /r/, providing little evidence in the acquisition of positional allophones, such as those in English /l/.

The current study investigates the acoustics and articulation of allophonic variations in English laterals produced by Japanese speakers, focusing on the effects of syllabic positions and flanking vowels. Acoustic and articulatory data were obtained from five Japanese speakers in a simultaneous audio and high-speed ultrasound tongue imaging recording set-up while they read sentences containing syllable-initial and -final tokens of English /l/ in four different vowel contexts. Acoustic analysis was conducted on 500 tokens using linear-mixed effects modelling and the articulatory data were analysed using generalised additive mixed modelling.

Syllable position and vowel context had significant effects on acoustics, while midsagittal tongue shape was more influenced by vowel context, with fewer positional effects. The results demonstrate that differences in acoustics not always be mirrored exactly by midsagittal tongue shape, suggesting multidimensionality of articulation in second language speech.

Index Terms: articulation, acoustics, English laterals, second language speech, ultrasound tongue imaging

1. Introduction

In second language speech learning, it is well-established that the phonetics and phonology of a learner's first language (L1) interferes with acquisition of nonnative sounds in a native-like manner [1], [2]. The Speech Learning Model (SLM) posits that acquisition of position-sensitive allophones is seen as the fundamental mechanisms of second language (L2) speech learning [3]. However, while acquisition of the English liquid contrast between /l r/ by Japanese speakers has been well studied, there have been relatively fewer studies that investigate acquisition of allophonic variation in English /l/.

The two canonical allophonic variants of English /l/, typically referred to as *clear-L* and *dark-L*, are acoustically distinguished by F2 and the distance between F2 and F1 [4]. *Clear-L*, typically occurring in a syllable-initial position or before a consonant, shows higher F2 and greater F2-F1 distance than the darker, syllable-final variant [5]. The clear and dark /l/s exhibit different sequence and magnitude concerning the tongue tip (TT) and tongue dorsum (TD) gestures [6], [7]; Lowering of F2 for the *dark-L* and resultant smaller F2-F1 value could correspond to the retraction and raising of TD [6], [8].

Previous articulatory studies show that the TD gesture is the key in understanding allophonic variations in laterals. However, due to lack of articulatory data in L2 speech research, we do not know what articulatory properties correspond to lateral allophony, given that L2 speakers could plausibly use different strategies from L1 speakers. In the case of Japanese, previous research demonstrated that Japanese /r/ lacks the specific TD gestural target, based on the extent of variability across vowel contexts, which could also affect the way Japanese speakers articulate syllable-initial and final /l/ [9]. In addition, a great degree of individual variation can be expected in the articulatory domain; for instance, at least seven patterns of tongue postures were discovered in production of English /r/ by native Japanese learners of English with varying proficiency in English, which made it difficult to establish a clear relationship between acoustics and articulation (e.g., lower F3) [10].

The current study examines two predictions: (1) if the articulatory properties of nonnative sounds are influenced by those of the closest native category, given the variability in the TD gesture for Japanese /r/, the posterior tongue is also likely to be highly variable depending on the vowel context in English /l/; (2) changes in articulatory properties may not always be manifested in acoustics, making it difficult to generalise what articulatory strategies would contribute to the changes in F2 and F2-F1 in the case of English /l/. In the following section, I will report a study that investigates Japanese speakers' production of the allophonic variations in English laterals. Specifically, we compare the effects of: 1) syllable position and 2) vowel contexts on acoustics and articulation.

2. Methods

2.1. Participants

Five L1 speakers of Japanese (two female: "JP01F" and "JP04F", and three male: "JP02M", "JP03M", and "JP05M") participated in this study. They were aged between 23 and 30 years (M = 24.6 years) and enrolled in a university in the UK for a postgraduate study. All of them defined themselves as native speakers of Tokyo Japanese, and their English proficiency was considered to be high given their enrolment status at the UK university as well as their self-reported assessment. None of them reported speech and hearing impairment at the time of the recording.

2.2. Materials

The stimuli were developed to elicit /l/s occurring in syllable-initial and -final positions in a controlled manner. The list of target words is shown in Table 1. The material development conceptually followed previous studies [11], [12]. First, target words were determined such that words embedding both onset and coda tokens of /l/ share same sets of phonemes, differing

10.21437/Interspeech.2022-11020

only in their sequences. For example, we used *leave* /liːv/ for a syllable-initial lateral whereas *veal* /viːl/ for a syllable-final lateral, with flanking vowels being either /i/ or /a/.

Second, each token was flanked by another word before or after the target token depending on the syllable position of /l/s to control the vowel environment surrounding /l/. These words were always in /hVp/ structure, with the vowels either /i/ or /a/. The consonants /h/ and /p/ were chosen to minimise consonantal effects on tongue movement while enabling us maximally to discern the lateral tokens from the neighbouring sounds in the acoustic signal.

As a result, we developed a list of 32 phrases to elicit initial and final /l/s (16 tokens x 2 vowel conditions). These phrases were then embedded in a carrier phrase '*(Someone) said "X Y" to (someone's) boss.*', in which *X Y* corresponds to the two-word phrase that contains a target word. The participants recorded each sentence at least three times, yielding at least 96 tokens per speaker.

Table 1: *List of stimuli and example phrases. (# indicates syllable boundary.)*

Vowel	Initial	Final	Example phrase
high	leap	peal	heap leap (i#li), peal heap (il#i)
	lead	deal	heap lead (i#li), deal heap (il#i)
	lean	kneel	hap lean (a#li), kneel hap (il#a)
	leave	veal	hap leave (a#li), veal hap (il#a)
low	lap	pal	heap lap (i#la), pal heap (al#i)
	lag	gal	heap lag (i#la), gal heap (al#i)
	lab	bal	hap lab (a#la), bal hap (al#a)
	lack	Cal	hap lack (a#la), Cal hap (al#a)

2.3. Data collection

The recording took place at the Phonetics Lab at Lancaster University, UK. Ultrasound data were obtained using a Telemed MicrUs system, with a 64 element probe of 20 mm radius. Midsagittal tongue views were imaged with a 2MHz probe frequency, 80 mm depth, 74.5% field of view and 52 scan lines, resulting in a framerate of ca. 100 per second. Simultaneous acoustic signals were also collected with the signal pre-amplified and digitized using a TASCAM US 4x4 audio interface, and then recorded onto a laptop computer at 44.1 kHz with 16-bit quantisation. The research project has been approved by the Lancaster University ethics committee.

2.4. Data analysis

Audio recording and tongue spline data were exported from the Articulate Assistant Advanced software [13] for analysis. Segmentation was carried out by the author with Praat [14] onto the accompanying TextGrid files for each audio recording. The audio files were low-pass filtered to 11,025 Hz, down-sampled to 22,050 Hz, and scaled for its peak intensity of 0.8. Following previous research, the lateral portion was labelled based on a steady-state of F2, which therefore excluded formant transitions in and out of the flanking vowels [15]. Specifically, the lateral was identified based on the following multiple cues: 1) an abrupt change in F2, 2) decrease in resonance energy on spectrograph and 3) decrease in waveform amplitude [15], [16].

For the acoustic analysis, F1 to F3 were estimated and extracted at the 11 equal intervals throughout the lateral portion using Fast Track, an optimized formant analysis Praat plug-in [17]. Fast Track automatically adjusts optimal ceiling formant values within a range of 4,700 and 7,550 Hz and returns a 'winning' analysis for each token based on 24 regression analyses. In this paper, I report the analysis of F2-F1 measure in 500 tokens (248 initial and 252 final). F2-F1 is often used in the previous research to give an approximate estimation of lateral quality [4], [6], [15], [16]; A higher F2-F1 corresponds to a clearer variant of English lateral. The F2-F1 values were then normalised by speaker so that it better captures within-speaker contrast in realisations of English /l/. Data were visualised through *ggplot* functions in the tidyverse suite [18].

For articulatory analysis, based on the TextGrid file that delimited the lateral portion on the acoustic signal, tongue splines were automatically fitted to the ultrasound tongue images through the best-fit batch processing function in the AAA software. In the current study, I focus on the tongue spline data at the lateral midpoint, given that the tongue shape at midpoint is shown to well-represent the articulation of /l/ in British English [7]. The analysis focused on the 497 tokens (252 initial and 245 final), as a result of some tokens having to be excluded due to errors in audio-ultrasound synchronization.

2.5. Statistical analysis

2.5.1. Acoustic data

The F2-F1 values were entered to Linear Mixed-Effect modelling (LME) to evaluate the effects syllabic position and vowel context using *lmer* package [18] on R [20]. For significance testing, the full model was tested against models that did not contain factors of interest through likelihood ratio testing [21]. This enables us to formally assess whether presence or absence of factors could improve the model fit. As a result, a maximal model was constructed for the acoustic data with z-scored F2-F1 values ('*z.F2F1*') as the outcome variable and the syllabic position (i.e., initial vs final: '*position*') and the vowel environment (i.e., /a_a/, /a_i/, /i_a/, and /i_i/ '*vowel*') as the fixed effects. Also, the by-speaker random intercepts were included for *position* and *vowel*, whereas the random intercept for words was not included as it did not improve the model fit. Finally, the interaction between position and vowel context was included as it improved the model fit significantly. The final model specification was determined as: *z.F2F1 ~ position + vowel + (1 + position + vowel | speaker) + position:vowel.*

2.5.2. Articulatory data

For the articulatory analysis, tongue splines were fitted with generalised additive mixed models (GAMMs) using the *mgcv* and *itsadug* packages on R [22], [23]. The exported tongue spline data from the AAA software consisted of X and Y Cartesian coordinates based on the pixel intensity differences in the ultrasound tongue images. The X and Y values were then converted into polar coordinates using the *rticulate* package [24]. Note that data visualisation is based on unrotated tongue spline data, meaning that x-axis is not parallel to the occlusal plane. Therefore, I only focus on within-subject differences in the shape of tongue splines.

3. Results

3.1. Acoustic analysis

In Table 2, mean and standard deviation for F1, F2, and F2-F1 frequency values by position (initial vs final) are summarised across speakers. Overall, F1 was higher syllable-finally than syllable-initially, except for the speaker JP04F. F2 was consistently lower for the syllable-final laterals than for the syllable-initial ones. Consequently, higher F2-F1 values were seen for syllable-initial /l/ than for syllable-final /l/. These results suggest that the Japanese speakers in this study mostly conform with the native-like allophonic variation pattern between syllable-initial and syllable-final /l/s acoustically.

Table 2: *Mean and SD (in bracket) of F1, F2 and F2-F1 values (Hz) at the lateral midpoint.*

Speaker	Position	F1	F2	F2-F1
JP01F	Initial	360.86	1850.17	1489.32
		(68.56)	(190.81)	(226.87)
	Final	594.48	1134.23	539.75
		(86.25)	(50.70)	(73.37)
JP02M	Initial	447.84	1290.76	842.92
		(95.75)	(255.35)	(332.02)
	Final	504.60	882.45	377.85
		(55.32)	(78.98)	(109.74)
JP03M	Initial	377.52	1717.50	1339.98
		(48.30)	(240.25)	(267.79)
	Final	472.61	1685.55	1212.94
		(85.98)	(301.24)	(362.94)
JP04F	Initial	471.87	1361.04	889.18
		(75.15)	(193.07)	(217.92)
	Final	461.89	1059.48	597.60
		(97.85)	(118.93)	(88.71)
JP05M	Initial	383.49	1394.87	1011.38
		(57.25)	(129.18)	(143.97)
	Final	497.24	1161.98	664.74
		(58.98)	(124.55)	(163.64)

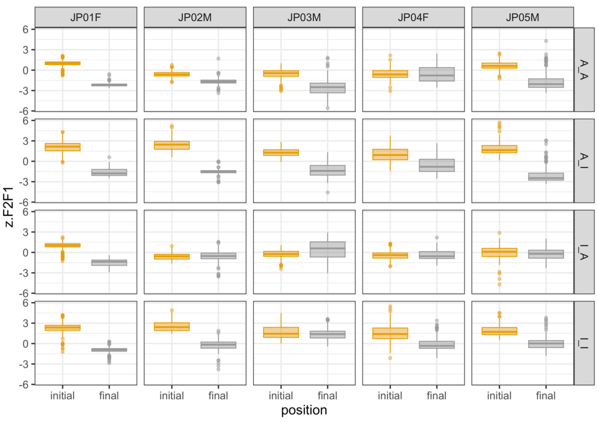

Figure 1 *Comparisons of F2-F1 values (z-scored) for initial (in orange) and final (in grey) for different vowel positions across five speakers.*

Figure 1 summarises the F2-F1 values extracted from the lateral midpoint by position (initial vs final) across four vowel environments across five speakers. Speakers JP01F, JP02M and JP05M made a contrast between initial and final /l/s in most of the vowel environments; While JP01F was consistent in making a clear-dark contrast in all vowel environments, such contrast was levelled in the I_A context for JP02M and JP05M. JP04F differentiated the two position-dependent variants in the A_I and I_A conditions but less so in the A_A and I_I contexts. JP03M differentiated the two allophones in the A_A and A_I contexts but to a lesser extent in the I_A and I_I environments. Finally, comparisons of several linear mixed-effect models demonstrated that there were significant effects of *position* ($\chi^2(1) = 8.801$, $p = .003$), *vowel* ($\chi^2(3) = 10.727$, $p = .013$), and the interaction between *position* and *vowel* ($\chi^2(3) = 171.800$, $p < .001$).

In summary, the acoustic analysis based on the F2-F1 values suggests that Japanese speakers distinguish syllable-initial and syllable-final laterals, but the magnitude of this varies by vowel context. Particularly, visual inspections of the boxplots suggest that there is difference in the initial-final contrast depending on the preceding vowel (/i/ vs /a/). We will then turn to the articulatory analysis to examine whether similar trends can be observed in the articulatory data.

3.2. Articulatory analysis

The participants' tongue spline data are visualised by speaker in Figure 2. Based on the results from the acoustic analysis, the four vowel conditions have been collapsed into two categories for the preceding vowels /i/ vs /a/. GAMM models based on the polar coordinate were constructed using the *polar_gam* function in the *rticulate* package separately for each speaker [25] with the tongue height (Y) as the outcome variable and separate smoothing spline terms of the X coordinate by position and vowel. Residual autocorrelations were reduced by specifying AR1 models for each speaker as the amount of autocorrelation at lag 1 [26]. Significance testing was conducted through model comparisons using the *compareML* function between the full model (*Y ~ position + vowel + s(X, by = position) + s(X, by = vowel)*) and a model that did not contain a parametric term and a smooth term with the variable in by-parameter (Position model: *Y ~ position + s(X, by = position)*; Vowel-context model: *Y ~ vowel + s(X, by = vowel)*).

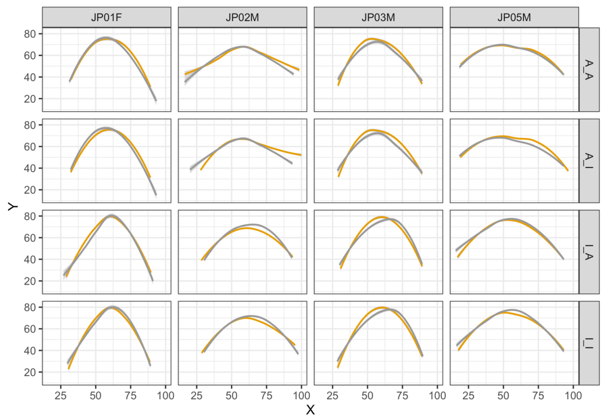

Figure 2 *Comparisons of tongue splines for initial (in orange) and final (in grey) for different vowel positions across four speakers with tongue tip pointing rightward (based on the Cartesian coordinates). A speaker JP04F was excluded due to the poor imaging quality.*

Overall, the full models significantly improved the model-fit against the position-only models (i.e., without terms associated with vowel context) across speakers (JP01F; $\chi^2(5.00) = 3589.510$, $p < .001$, JP02M; χ^2 $(5.00) = 500.325$, $p < .001$, JP03M; χ^2 $(5.00) = 522.015$, $p < .001$, JP05M; χ^2 $(5.00) = 231.851$, $p < .001$) suggesting that vowel context significantly improved the model fit. On the other hand, there were no significant differences between the full models and the vowel-context models (i.e., without terms associated with position), with only a slight decrease in the AIC values, suggesting that the positional effect did not improve the model fit. In summary, the GAMM models show that the vowel context had a greater effect on the participants' tongue shapes for English /l/ than did syllable position.

In addition, the difference smooths plots show that the way the speakers differentiated the initial-final laterals differed from one another; Whereas JP05M differentiated them mainly at the middle of the tongue, JP03M did so on wider parts spanning tongue anterior and posterior when the vowel /i/ preceded the lateral /l/ (see Figure 3).

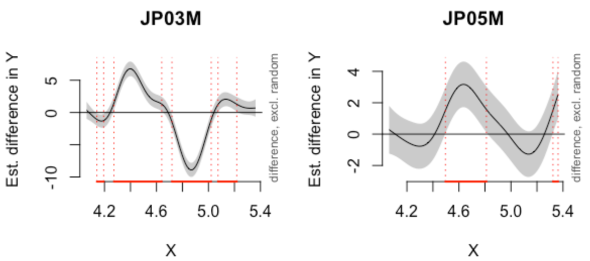

Figure 3 *Difference smooths between syllable-initial and -final /l/s that follow the vowel /i/ produced by JP03M (left) and JP05M (right).*

4. Discussion

The current study investigated the production of English laterals by Japanese speakers. The main finding of this study is that Japanese speakers contrast initial and final English /l/ in acoustics, but there is not strong support for this in midsagittal tongue shape. Instead, tongue shape appears to be more strongly affected by vowel context.

The acoustic analysis has shown that the participants were generally able to distinguish syllable-initial and -final /l/ in an expected native-like manner. The participants produced syllable-initial lateral with lower F1 and higher F2 as shown in Table 2, which generally agrees with documentation in the previous literature [5], [8]. While the interaction between *position* and *vowel* is not surprising, there were some between-speaker differences in the degree of vowel coarticulation. It is known that clear and dark /l/s differ in the degree of vowel coarticulation, in which dark /l/ is coarticulated with the neighbouring vowels to a lesser degree than clear /l/ [5]. This tendency was true of one of the speakers, JP01F, who produced syllable-initial /l/s with higher F2-F1 than syllable-final /l/ consistently. In contrast, the other two speakers (JP03M and JP04F) showed little contrast between the initial and final tokens across the four vowel contexts, with the case of the /l/ preceded by /i/ showing lack of the initial-final distinction compared to when /l/ followed /a/.

While the distinction between the initial-final contrast was seen relatively clearly in the acoustic results, such tendency was not observed in the articulatory results, suggesting that static midsagittal tongue shape may not capture the most articulatory salient dimension of this contrast [27]. In the articulatory domain, the preceding vowels instead have a greater effect on the tongue shape for English laterals. Despite smaller degrees of initial-final contrasts in the acoustic data when /i/ preceded /l/, JP03M's tongue shapes were different throughout the anterior-posterior dimensions of the tongue, which was also somewhat evident from the visual inspection of Figure 2. This could suggest that it was not the TD gesture that JP03M uses to differentiate initial and final /l/. As well as in the case of English /r/ [10], the articulation of English /l/ in L2 speech may also be diverse across speakers, making it difficult to pin down a single articulatory property (e.g., TD gesture) that is crucial in the initial-final contrast in English /l/.

One possible explanation for this mismatch between acoustics and articulation may be the dynamic nature of acoustics and articulatory realisations of English liquids over time [28], [29]. While the mid-point analysis can adequately represent the broad phonetic quality of English /l/, dynamic analyses could provide a greater amount of information, particularly regarding co-articulatory influence of the flanking vowels [28]. Previous research finds that the articulatory gesture is often present before changes in acoustic signals were observed [29]. It could also be true that the participants might utilise several articulatory strategies to achieve a certain acoustic target (e.g., along the F2-F1 dimension), which was implicated in an earlier study of English /r/ [10]. In future research, auditory analysis could be incorporated to investigate what acoustic/articulatory properties are important for English laterals, which may not necessarily be manifested in the acoustics or articulatory data at the temporal mid-point.

The current focus on the midsagittal tongue shapes may have reduced the dimensionality of the complex articulatory properties involved in English /l/, particularly regarding tongue lateralisation that could only be captured on the tongue's coronal plane. Tongue lateralisation has been proposed to be an active articulatory gesture in English /l/ [30]. However, the transfer of the lateral gesture does not always happen from L1 to L2, such that a Japanese speaker who lateralises the tongue in L1 does not always succeed in producing English /l/ in a native-like manner [31]. In investigating tongue lateralization, obtaining coronal tongue data would be helpful.

Finally, an anonymous reviewer pointed that it would be helpful to include data collected from L1 English speakers. Whereas the current study makes reference to the English variety that exhibits the clear-dark allophony, such as Standard Southern British English (SSBE), the degree of the clear-dark contrast and the degree of 'darkness' are known to vary from one accent to another [7], [15]. The current results would be more enhanced if L1 English data were to be included.

5. Conclusions

This study investigated the effects of the syllabic position and vowel environment on the acoustics and articulation of English laterals produced by Japanese speakers. The results demonstrated that the picture may be more complicated than a mere syllable-initial and -final distinction, particularly in the articulatory dimension. The current study also adds novel evidence into the articulatory properties in L2 speech, in which L2 speakers' articulatory strategies may be diverse.

6. Acknowledgements

I thank all the participants who gave up their time for the data collection. I thank Dr Sam Kirkham and Dr Claire Nance for helpful comments and support throughout the research project. The work has improved thanks to comments from three anonymous reviewers. This research is financially supported by Graduate Scholarship for Degree Seeking Students, Japan Student Services Organization (JASSO) awarded to the author.

7. References

[1] J. E. Flege, 'Second Language Speech Learning Theory, Findingsand Problems', in *Speech Perception and Linguistic Experience: Issues in Cross-Language Research*, W. Strange, Ed. York Press, 1995.

[2] C. T. Best and M. D. Tyler, 'Nonnative and second-language speech perception: Commonalities and complementarities', in *Language experience in second language speech learning: In honor of James Emil Flege*, O.-S. Bohn and M. J. Munro, Eds. Amsterdam: John Benjamins Publishing Company, 2007, pp. 13–34. doi: 10.1075/lllt.17.07bes.

[3] J. E. Flege and O.-S. Bohn, 'The Revised Speech Learning Model (SLM-r)', in *Second Language Speech Learning: Theoretical and Empirical Progress*, 1st ed., R. Wayland, Ed. Cambridge University Press, 2021, pp. 3–83. doi: 10.1017/9781108886901.002.

[4] P. Carter and J. Local, 'F2 variation in Newcastle and Leeds English liquid systems', *Journal of the International Phonetic Association*, vol. 37, no. 2, pp. 183–199, Aug. 2007, doi: 10.1017/S0025100307002939.

[5] D. Recasens, 'A cross-language acoustic study of initial and final allophones of /l/', *Speech Communication*, vol. 54, no. 3, pp. 368–383, Mar. 2012, doi: 10.1016/j.specom.2011.10.001.

[6] R. Sproat and O. Fujimura, 'Allophonic variation in English /l/ and its implications for phonetic implementation', *Journal of Phonetics*, vol. 21, no. 3, pp. 291–311, Jul. 1993, doi: 10.1016/S0095-4470(19)31340-3.

[7] D. Turton, 'Categorical or gradient? An ultrasound investigation of /l/-darkening and vocalization in varieties of English', *Laboratory Phonology: Journal of the Association for Laboratory Phonology*, vol. 8, no. 1, p. 13, May 2017, doi: 10.5334/labphon.35.

[8] S. S. Narayanan, A. A. Alwan, and K. Haker, 'Toward articulatory-acoustic models for liquid approximants based on MRI and EPG data. Part I. The laterals', *The Journal of the Acoustical Society of America*, vol. 101, no. 2, pp. 1064–1077, Feb. 1997, doi: 10.1121/1.418030.

[9] N. Yamane and P. Howson, 'An ultrasound examination of taps in Japanese', in *Proceedings of the 18th International Congress of Phonetic Sciences*, Aug. 2015, p. 5. [Online]. Available: https://www.internationalphoneticassociation.org/icphs-proceedings/ICPhS2015/Papers/ICPHS0815.pdf

[10] J. Moore, J. Shaw, S. Kawahara, and T. Arai, 'Articulation strategies for English liquids used by Japanese speakers', *Acoust. Sci. & Tech.*, vol. 39, no. 2, pp. 75–83, Mar. 2018, doi: 10.1250/ast.39.75.

[11] B. Gick, F. Campbell, S. Oh, and L. Tamburri-Watt, 'Toward universals in the gestural organization of syllables: A cross-linguistic study of liquids', *Journal of Phonetics*, vol. 34, no. 1, pp. 49–72, Jan. 2006, doi: 10.1016/j.wocn.2005.03.005.

[12] F. Campbell, B. Gick, I. Wilson, and E. Vatikiotis-Bateson, 'Spatial and Temporal Properties of Gestures in North American English /r/', *Lang Speech*, vol. 53, no. 1, pp. 49–69, Mar. 2010, doi: 10.1177/0023830909351209.

[13] Articulate Instruments, *Articulate Assistant Advanced version 2.18*. Edinburgh: Articulate Instruments, 2019.

[14] P. Boersma and D. Weenink, *Praat: doing Phonetics by Computer*. 2022. Accessed: Feb. 21, 2022. [Online]. Available: https://www.fon.hum.uva.nl/praat/

[15] S. Kirkham, D. Turton, and A. Leemann, 'A typology of laterals in twelve English dialects', *The Journal of the Acoustical Society of America*, vol. 148, no. 1, pp. EL72–EL76, Jul. 2020, doi: 10.1121/10.0001587.

[16] C. Nance, 'Phonetic variation in Scottish Gaelic laterals', *Journal of Phonetics*, vol. 47, pp. 1–17, Nov. 2014, doi: 10.1016/j.wocn.2014.07.005.

[17] S. Barreda, 'Fast Track: fast (nearly) automatic formant-tracking using Praat', *Linguistics Vanguard*, vol. 7, no. 1, p. 20200051, Jan. 2021, doi: 10.1515/lingvan-2020-0051.

[18] H. Wickham *et al.*, 'Welcome to the Tidyverse', *Journal of Open Source Software*, vol. 4, no. 43, p. 1686, Nov. 2019, doi: 10.21105/joss.01686.

[19] D. Bates, M. Mächler, B. Bolker, and S. Walker, 'Fitting Linear Mixed-Effects Models Using lme4', *Journal of Statistical Software*, vol. 67, pp. 1–48, Oct. 2015, doi: 10.18637/jss.v067.i01.

[20] R Core Team, *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing, 2021. [Online]. Available: https://www.R-project.org/

[21] B. Winter, 'Statistics for Linguists: An Introduction Using R', *Routledge & CRC Press*, 2020. https://www.routledge.com/Statistics-for-Linguists-An-Introduction-Using-R/Winter/p/book/9781138056091 (accessed Mar. 19, 2022).

[22] S. N. Wood, *Generalized Additive Models: An Introduction with R*, 2nd ed. New York: Chapman and Hall/CRC, 2017. doi: 10.1201/9781315370279.

[23] J. van Rij, M. Wieling, R. H. Baayen, and H. van Rijn, *itsadug: Interpreting Time Series and Autocorrelated Data Using GAMMs*. 2020. Accessed: Mar. 20, 2022. [Online]. Available: https://CRAN.R-project.org/package=itsadug

[24] S. Coretta, *rticulate: Ultrasound Tongue Imaging in R*. 2021. Accessed: Mar. 20, 2022. [Online]. Available: https://github.com/stefanocoretta/rticulate

[25] S. Coretta, 'Vowel duration and consonant voicing: A production study', PhD Thesis, University of Manchester, Manchester, 2020. Accessed: Mar. 20, 2022. [Online]. Available: https://stefanocoretta.github.io/phd-thesis/index.html

[26] M. Wieling, 'Analyzing dynamic phonetic data using generalized additive mixed modeling: A tutorial focusing on articulatory differences between L1 and L2 speakers of English', *Journal of Phonetics*, vol. 70, pp. 86–116, Sep. 2018, doi: 10.1016/j.wocn.2018.03.002.

[27] D. Turton, 'Variation in English /l/: Synchronic reflections of the life cycle of phonological processes', PhD Thesis, University of Manchester, Manchester, 2014. Accessed: Mar. 20, 2022. [Online]. Available: https://www.research.manchester.ac.uk/portal/en/theses/variation-in-english-l-synchronic-reflections-of-the-life-cycle-of-phonological-processes(dfa11693-a112-45e2-99f6-2a08cf5f117b).html

[28] S. Kirkham, C. Nance, B. Littlewood, K. Lightfoot, and E. Groarke, 'Dialect variation in formant dynamics: The acoustics of lateral and vowel sequences in Manchester and Liverpool English', *The Journal of the Acoustical Society of America*, vol. 145, no. 2, pp. 784–794, Feb. 2019, doi: 10.1121/1.5089886.

[29] J. Ying, J. A. Shaw, C. Kroos, and C. T. Best, 'Relations Between Acoustic and Articulatory Measurements of /l/', in *Proceedings of the 14th Australasian International Conference on Speech Science and Technology*, Sydney, Dec. 2012, pp. 109–112.

[30] J. Ying, J. A. Shaw, C. Carignan, M. Proctor, D. Derrick, and C. T. Best, 'Evidence for active control of tongue lateralization in Australian English /l/', *Journal of Phonetics*, vol. 86, p. 101039, May 2021, doi: 10.1016/j.wocn.2021.101039.

[31] M. Morimoto, 'Articulatory Preference in Japanese Liquids and F3 in English: A Preliminary Report', *ICU Working Papers in Linguistics: Selected Papers from the 5th Asian Junior Linguists Conference (AJL5)*, vol. 15, pp. 1–6, Mar. 2021.