Dialect variation in formant dynamics: The acoustics of lateral and vowel sequences in Manchester and Liverpool English

Sam Kirkham,[1, a)] Claire Nance,[1] Bethany Littlewood,[1] Kate Lightfoot,[1] and Eve Groarke[1]

*Department of Linguistics and English Language, Lancaster University,
County South, Lancaster LA1 4YL, United Kingdom*

This study analyses the time-varying acoustics of laterals and their adjacent vowels in Manchester and Liverpool English. We use Generalized Additive Mixed-Models (GAMMs) for quantifying time-varying formant data, which allows us to model nonlinearities in acoustic time series while simultaneously modelling speaker and word level variability in the data. We compare these models to single time-point analyses of lateral and vowel targets in order to determine what analysing formant dynamics can tell us about dialect variation in speech acoustics. The results show that lateral targets exhibit robust differences between some positional contexts and also between dialects, with smaller differences present in vowel targets. The time-varying analysis shows that dialect differences frequently occur globally across the lateral and adjacent vowels. These results suggest a complex relationship between lateral and vowel targets and their coarticulatory dynamics, which problematizes straightforward claims about the realization of laterals and their adjacent vowels. We further discuss these findings in terms of hypotheses about positional and sociophonetic variation. In doing so, we demonstrate the utility of GAMMs for analysing time-varying multisegmental acoustic signals, and highlight the significance of our results for accounts of English lateral typology.

---

[a] s.kirkham@lancaster.ac.uk

I. INTRODUCTION

A. Variation in English laterals

The present study aims to quantify time-varying acoustic patterns in lateral and vowel sequences and, secondarily, to determine the nature of dialect differences and positional contrast in the lateral systems of two varieties of British English (Manchester and Liverpool). The allophony of English lateral production is most commonly framed in terms of 'clear' versus 'dark' allophones of /l/ (Recasens, 2012), and the presence or absence of positional variants (Sproat and Fujimura, 1993). The terms 'clear' and 'dark' represent abstractions on different ends of a continuum (Recasens and Espinosa, 2005). Articulatorily, clear /l/s involve raising and fronting of the tongue body, while dark /l/s involve tongue dorsum lowering and retraction (Narayanan *et al.*, 1997; Recasens and Espinosa, 2005). Clear /l/s also typically involve the tongue tip gesture occurring simultaneous with (or prior to) the tongue dorsum gesture, whereas dark /l/s typically show tongue dorsum retraction prior to the tongue tip gesture (Sproat and Fujimura, 1993). Lateral clearness-darkness has also been conceptualised as a single gesture in terms of amounts of predorsum lowering and postdorsum retraction (Recasens and Espinosa, 2005). These complex articulatory and timing relations and how they interact with the surrounding vowels make the time-varying nature of lateral production highly significant (see Section I B).

In terms of acoustic consequences, clear laterals typically have high F2 and low F1, while dark laterals have low F2 and high F1 (Carter and Local, 2007; Ladefoged and Maddieson, 1996; Lehiste, 1964; Recasens, 2012). Accordingly, many studies have used the F2 minus

F1 measure (F2−F1) to quantify lateral quality, with higher values indicating clearer /l/s (Carter, 2002; Kirkham, 2017; Lehiste, 1964; Nance, 2014; Sproat and Fujimura, 1993; Turton, 2014). F3−F2 is also typically higher for darker /l/ than for clearer /l/, due to a low F2 and high F3 (Recasens and Espinosa, 2005).

In the context of British English dialect typology, Southern British English is described as having clear /l/ in syllable-onsets and dark /l/ in syllable-rimes (Wells, 1982, 370), resulting in positional contrast between word-initial and word-final productions. However, many British English varieties do not show such strong positional effects and may display dark /l/s in all positions, such as Leeds, while others show clearer /l/s in all positions, such as Newcastle (Carter and Local, 2007). Within dark /l/ varieties, there is also a distinction between those that show positional differences between initial and final /l/ (e.g. Leeds) and those that do not (e.g. Sheffield) (Kirkham, 2017). There are other dialects that occupy a more contested status on the clear-dark continuum, as will be discussed below.

The dialects in this study are Liverpool English and Manchester English. Liverpool and Manchester are both located in the north west of England and are only 35 miles apart by road. However, these two dialects are reported to be extremely different, with Liverpool in particularly being one of the most distinctive accents in England (Baranowski and Turton, 2015; Nance *et al.*, 2015; Watson, 2007). In terms of laterals, Manchester English is widely described as having dark /l/s in all positions (Carter, 2002; Kelly and Local, 1986; Turton, 2014). Turton (2014) reports that middle-class speakers produce an acoustic and articulatory contrast between initial and final /l/, whereas working-class speakers do not.

The realization of Liverpool /l/ is less documented and its status is contested in the literature. Jones (1966, 92) speculates that Liverpool /l/ may be clear in all positions, stating that 'its existence there is probably due to Irish influence', with many varieties of Irish English having very clear /l/s. Knowles (1973, 256) claims that /l/ in Liverpool is frequently 'velarised' and produced in similar ways across positions. One of the few sources of instrumental data on Liverpool /l/ comes from Turton (2014), who reports acoustic and ultrasound data on a single male speaker. She finds that he produces the initial∼final contrast in /l/, but that he also produces word-final /l/ with distinct velarisation, as opposed to the more pharyngealised articulations documented for other British English varieties. This also suggests a potentially 'intermediate' realization for Liverpool /l/, which may lie towards the middle of a continuum between clear and dark.

In this study, we address the relationship between time-varying lateral and vowel formant dynamics. Accordingly, we briefly overview previous research on vowels in each variety. Manchester English shows features typical of many northern Englishes, such as the lack of a FOOT-STRUT or TRAP-BATH split and monophthongal productions of canonical diphthongs (Baranowski and Turton, 2015). Liverpool English typically merges the NURSE and SQUARE vowels (Knowles, 1973; Watson, 2007) and has complex patterns of raising in PRICE and MOUTH before nasal-obstruent clusters (Cardoso, 2015).

A concrete difference between dialects that we predict will have an effect on our results is the final vowel in words such as *belly* (Wells 1982 calls this the HAPPY vowel). Manchester is reported to produce very low and back variants of HAPPY (Baranowski and Turton, 2015), which we do not expect to see in Liverpool. Finally, we discuss pre-lateral vowels, which are particularly significant for our study. Fronting of /u/ is typically inhibited before coda /l/ in some varieties of English (Kleber *et al.*, 2011), although the articulatory interpretation of this is not straightforward (Strycharczuk and Scobbie, 2017). However, Baranowski (2017) finds a clear social class effect on pre-lateral /u/ fronting in Manchester, with a strong negative correlation between social class and fronting in this context. While we are not aware of any studies of pre-lateral /u/ in Liverpool, our own impressions suggest that fronting of /u/ before coda /l/ is widespread in this dialect.

B. Time-varying spectral analysis

The significance of the time-varying properties of sonorant sounds has been comprehensively documented in the literature (Elvin *et al.*, 2016; Fox and Jacewicz, 2009; Strycharczuk and Scobbie, 2017; Watson and Harrington, 1999; Williams and Escudero, 2014). This is particularly pertinent to a study of laterals, which are inherently non-static due to the timing relations outlined in Section I A, as well as the existence of strong interactions between laterals and the surrounding vowels. This interaction also makes it challenging to place reliable segmental boundaries between a lateral and any adjacent vowels. This is even more pronounced when comparing clear and dark laterals, which vary in terms of of their acoustic structure (Recasens and Espinosa, 2005), transitions into and out of the steady-state of the lateral phase, and duration of the steady-state phase (Carter, 2002).

The above findings have theoretical and methodological implications for how to treat adjacent lateral and vowel targets. Many studies have isolated the lateral target by identifying an F2 steady-state and then more holistically analysed syllable-level formant transitions across the lateral and surrounding vowels (Carter and Local 2007; Kirkham 2017; Nance 2014; Stuart-Smith *et al.* 2015). However, the relationship between lateral targets and adjacent vowel targets is not necessarily straightforward, as we expect a strong coarticulatory relationship between them, especially for clearer initial /l/s (Recasens and Espinosa, 2005). Therefore, a primary aim of this study is to analyse lateral and vowel sequences in terms of (i) steady-state targets for adjacent laterals and vowels; (ii) time-varying formant dynamics across the sequence of both segments. This allows us to establish whether patterns of dialect variation can be captured by targets alone, or whether time-varying information further contributes to dialect differences.

Previous research on lateral formant trajectories has quantified non-linear differences using methods such as Smoothing-Spline ANOVA (SS-ANOVA) (Kirkham, 2017; Nance, 2014; Simonet *et al.*, 2008). Such methods fit smooth functions to the data using a computationally-derived smoothing penalty that aims to avoid under-/over-fitting. This has an advantage over, for example, polynomial regression, as the analyst only needs to set an upper bound on non-linearity, rather than specifically determine the degree of non-linearity in advance. However, these methods are unable to incorporate a random effects structure into the model, which leads to anti-conservative estimates due to the fact that, for example, repeated productions from an individual speaker do not represent independent observations. One alternative is to use linear mixed-effects models with random intercepts and slopes (Stuart-Smith *et al.*, 2015). These models adequately account for the kinds of variability mentioned previously, but can only model linear trends in the data and are therefore inappropriate for modelling non-linearities.

Generalized Additive Mixed-Models (GAMMs) are an ideal solution to the above problems (Wood 2017; see Sóskuthy 2017; Wieling 2018 for excellent tutorials applying GAMMs to phonetic data). Similar to SS-ANOVA or Generalized Additive Modelling, GAMMs provide a data-driven method for quantifying non-linear trends, but they also allow for the inclusion of *random smooths*, which can capture group or individual variation in non-linear effects. This is similar to the use of random intercepts and slopes in a linear mixed-effects model, but instead of only the height and slope being allowed to vary, random smooths permit modelling of non-linearities in the relationship between predictor and outcome variables. This has the benefit of more comprehensively capturing dependencies between adjacent data points and allows us to better model variance in the data.

C. Hypotheses

In this study we compare the production of laterals and their surrounding vowels in Liverpool and Manchester English, focusing on (i) lateral and vowel targets; (ii) time-varying formant dynamics across the lateral and adjacent vowels. In light of the research reviewed above, we make the following predictions with respect to our study:

H1. Initial laterals will have higher F2−F1 and lower F3−F2 than final laterals.

H2. Liverpool non-final laterals will have higher F2−F1 and lower F3−F2 than Manchester non-final laterals.

H3. Liverpool will have higher F2−F1 in medial trochaic V2 than Manchester.

H4. Liverpool and Manchester will differ in a non-linear fashion across non-final time-varying lateral and vowel intervals, due to the prediction that there will be bigger dialect differences in the laterals (H2) than in the surrounding vowels.

We do not predict specific dialect differences in any other surrounding vowels except for those specified in H3. We have no reason to predict sociophonetic gender differences, but we anticipate that female speakers may produce higher formant values across the board. As a consequence, we do not predict significant interactions between gender and either position or dialect.

II. METHODS

A. Sampling and data collection

Data were collected from 46 speakers. 24 speakers were from Liverpool (12 female, 12 male) and 22 speakers were from Manchester (13 female, 9 male). All speakers were aged between 19–27 years old, were born in their respective cities, and had lived there until at least the age of 18.

All recordings were carried out in a sound attenuated booth in Lancaster University Phonetics Lab using a Beyerdynamic Opus 55 headset microphone, preamplified and digitized using a Sound Devices USBPre2 audio interface, and recorded to a desktop computer at 44.1 kHz with 16-bit quantization. Stimuli were presented using PsychoPy in standard English orthography. Thirteen target words were elicited in the carrier phrase ‘she said X’, where X was a word with a lateral in one of four positional contexts: word-initial (*lead, lad, Lord,* *lute, like*); word-medial trochaic (monomorphemic) (*belly, Bally*); word-medial morpheme boundary (*filing, stalling*); word-final (*peel, pal, Paul, pool*). Each word was produced once by each speaker, except for *like*, which was produced twice by each speaker due to this word being elicited for an additional planned analysis. There were 93 non-lateral words in the same test block, which served as distractors and were the subject of another experiment. 18 tokens were discarded due to recording errors or mispronunciations, leaving a total of 626 tokens for analysis.

B. Data processing and acoustic analysis

The audio recordings were downsampled to 22.05 kHz and low-pass filtered at 11 kHz. Two acoustic intervals were then labelled using Praat: (1) a steady-state period of the lateral; (2) the entire lateral-vowel (initial tokens), vowel-lateral-vowel (medial tokens) or vowel-lateral (final tokens) interval. The steady-state period of the lateral was defined as a period during the lateral at which the F2 trajectory was steady or as close to steady as could be achieved, representing an unambiguously lateral phase (Carter and Local, 2007; Kirkham, 2017; Nance, 2014). Praat TextGrids were converted to EMU annotation files for use with the EMU Speech Database Management System (Winkelmann *et al.*, 2017).

We carried out formant estimation via Linear Predictive Coding using a 22-order autocorrelation method (Markel and Gray, 1976). Resonance frequencies were obtained by root solving of the filter polynomial and formants were classified using the Split Levinson Algorithm (Delsarte and Genin, 1986). This procedure was implemented using the wrassp::forest R function (Bombien *et al.*, 2016) in order to interface with the EMU-webApp. LPC analysis was based on a 20 ms Hamming window with 5 ms window shift, which was applied across the entire signal file. Visual inspection of formant trajectories for every token was carried out using the EMU-webApp (Winkelmann and Raess, 2014) and formant trajectories were hand-corrected when the values visibly diverged from the formants on the wideband spectrogram.

We report measurements of F2−F1 as a proxy for clearness/darkness in laterals, with lower values suggesting darker laterals (Sproat and Fujimura, 1993). In addition to this, we report analyses of F3−F2 because darker laterals are more likely to have low F2 and high F3 (Recasens and Espinosa, 2005), which means that we expect this measure to further discriminate between positional variants and also potentially between dialects.

We anticipate that the acoustics of lateral and vowel targets will interact due to coarticulation. Accordingly, in order to compare lateral and vowel targets, we also report F2−F1 and F3−F2 from an adjacent vowel. In the case of word-medial contexts, we specifically analyse V1 in morpheme boundary words (e.g. *stalling*) and V2 in medial trochaic words (e.g. *belly*), because this is where we expect dialect differences to be largest in each context (see Section I A). We note that our use of formant ratios, such as F2−F1, provide some degree of speaker normalization, but no further normalization such as z-scoring was applied to the data. This is because we are not only interested in the relationship between positional variants within each variety, but also in the absolute clearness/darkness of laterals between varieties.

For the time-varying analysis, we extracted measurements at 11 equidistant points from the onset to the offset of the interval containing the lateral and surrounding vowels in each word. Time normalization assumes that phonetically similar events occur at proportionally similar times across tokens with different durations, which may not always be the case. This is magnified when normalizing across different contexts, such as lateral-vowel versus vowel-lateral-vowel. The latter issue is not relevant here as our GAMMs focus only on within-context dialect differences. In order to resolve the former issue, we fitted linear mixed-effects models to the duration of the interval encompassing the lateral and its adjacent vowels. The null model had interval duration as the outcome variable, with speaker and word random intercepts and by-speaker random slopes for position. The test model added a position*dialect interaction to the null model. We found no significant difference between these two models ($\chi^2(7) = 12.57$, $p = .083$). As a consequence, we discount the role of interval duration differences as a potential explanation for our findings.

C. Statistical analysis

Data and code for all analyses reported in this article are publicly available at: `https://osf.io/5u6ez/`.

For the lateral and vowel targets analysis, linear mixed-effects models were fitted to the F2−F1 and F3−F2 values extracted from the the midpoint of (i) the lateral steady-state interval; and (ii) the vowel adjacent to the lateral. Models were fitted to the data using the lme4 package in R (Bates *et al.*, 2015). The models had either F2−F1 or F3−F2 as the outcome variable, with fixed effects of dialect, gender and position, and interactions between dialect*gender, position*gender and position*dialect. We included random intercepts for speaker and word, as well as by-speaker random slopes for the effect of position.

Significance testing was conducted using likelihood ratio tests to compare a full model to a nested model that excluded the term being tested for significance. When interaction terms are significant, we do not report p-values for the main effects that are part of the relevant interaction, but refer the reader to accompanying figures and model summaries. In cases where all interactions in a given model are non-significant at $p > 0.3$, we test the significance of main effects by comparing a model containing only main effects against a series of nested models that each exclude the main effect of interest (Aikin and West, 1991; Harrell, 2015).

The time-varying analysis uses Generalized Additive Mixed-Models (Wood, 2017). Formant values were sampled at eleven equidistant points between the beginning and end of the entire lateral and vowel sequence and separate GAMMs were fitted to the time-varying F2$-$F1 and F3$-$F2 data at each position using the mgcv::bam function in R (Wood, 2017). Predictor variables included a parametric term of dialect and smooth terms of normalised time and a normalised time-by-dialect interaction. In order to improve statistical power and model simplicity, the GAMMs exclude gender as a predictor, so all model estimates are derived from collapsing over gender groups. We also fitted random smooths of time-by-speaker and time-by-word. We tested the significance of dialect and the time-by-dialect smooth by conducting model comparison as follows (Sóskuthy, 2017; Sóskuthy *et al.*, 2018):

1. We compare a full model (containing the dialect parametric term and time-by-dialect smooth term) to a nested model excluding those terms, which allows us to test overall effects of dialect and time-by-dialect on the trajectory.

2. If there is a significant difference in (1) then we specifically test for differences in the shape of the trajectory by comparing the full model to a nested model excluding the

time-by-dialect smooth term. If this comparison is significant then we conclude that there is a difference in shape of the two dialect's trajectories. If not, then we conclude that there is a difference only in the height of the two dialect's trajectories.

All model comparison was conducted using the itsadug::compareML function (van Rij *et al.*, 2017). Autocorrelation in trajectories was corrected using a first-order autoregressive (AR1) model. We initially set the AR1 correlation parameter (ρ) as the autocorrelation value at lag 1 for each model, but changing this value to $\rho = 0.3$ decreased autocorrelation in the residuals to a greater degree for all models.

III. RESULTS

In this section we focus on positional, dialect and gender differences in lateral steady-state and vowel midpoint formant values. The statistical analysis reports significance testing of predictor variables via model comparison, followed by a more holistic interpretation of the patterns via data visualization. Full summaries for all models in this section can be found in Appendix A.

A. Lateral steady-state

A linear mixed-effects regression model fitted to the lateral steady-state F2−F1 values shows significant interactions between position*dialect ($\chi^2(3) = 9.06$, $p = .028$) and dialect*gender ($\chi^2(1) = 5.40$, $p = .020$), but no significant position*gender interaction ($\chi^2(3) = 3.46$, $p = .327$). As all main effects are also included as part of higher-level interactions, we do not report their significance as they are not straightforwardly interpretable in the presence of interactions. Figure 1 shows that there is robust contrast between initial and final tokens for all groups, and that Liverpool typically has higher values than Manchester. However, the significant position*dialect and dialect*gender interactions can be clearly seen in the plots. For instance, Liverpool and Manchester females produce very similar final /l/s, with Manchester females having slightly higher values (and thus a smaller initial∼final contrast). In contrast, Manchester males produce final /l/ with lower values than Liverpool males.

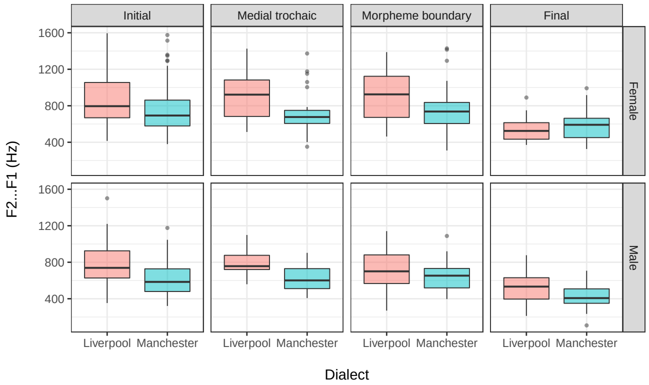

FIG. 1. F2−F1 values for /l/ steady-state. (Colour online).

The F3−F2 lateral steady-state model shows significant effects of position ($\chi^2(3) = 14.07$, $p = .003$), dialect ($\chi^2(1) = 10.36$, $p = .001$) and gender ($\chi^2(1) = 11.29$, $p < .001$), with no significant interactions between any of these variables ($p > .35$ for all interactions). Figure 2 shows that final tokens have higher values than non-final tokens, Manchester speakers have higher values than Liverpool speakers, and female speakers have higher values than male speakers. While the F3−F2 measurements largely mirror the F2−F1 values, there are some differences, such as the existence of dialect differences in final /l/ for both female and male speakers.

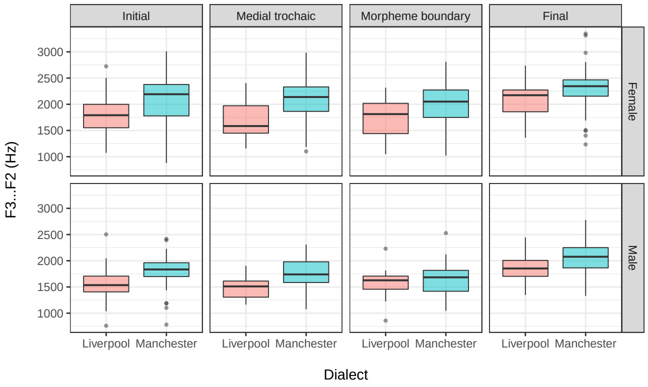

FIG. 2. F3−F2 values for /l/ steady-state. (Colour online).

B. Vowel midpoint

A linear mixed-effects regression model fitted to the vowel midpoint F2−F1 values shows significant interactions between position*gender ($\chi^2(3) = 17.59$, $p < .001$) and position*dialect ($\chi^2(3) = 31.54$, $p < .001$), but not dialect*gender ($\chi^2(1) = 0.01$, $p = .924$). As all main effects are also included as part of higher-level interactions, we do not report their significance as they are not easily interpretable in the presence of interactions. Figure 3 shows that final tokens typically have lower values than non-initial tokens. Liverpool typically has slightly higher values across all positions, except for morpheme boundary position where dialect differences are very minor. It also appears that the magnitude of dialect differences is greatest in the medial trochaic context, where Liverpool has higher values than Manchester. Note that these vowel results are largely in the same direction as for the lateral target analysis, but the difference between dialects is typically smaller in magnitude. There are also instances in which the vowel distributions heavily overlap between dialects, such as morpheme boundary and final contexts.

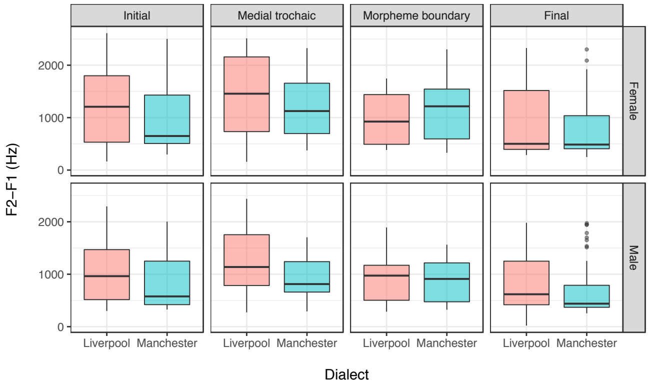

FIG. 3. F2−F1 values for vowel midpoint. (Colour online).

The F3−F2 model shows a significant interaction between position*dialect ($\chi^2(3) = 20.71$, $p < .001$), but no significant dialect*gender ($\chi^2(1) = 0.42$, $p = .518$) or position*gender ($\chi^2(3) = 4.18$, $p = .243$) interactions. Due to the significant position*dialect interaction, we do not report the significance of any main effects, but note that the very low t-value for the gender main effect ($\beta = -41.58$, $SE = 49.77$, $t = -0.84$) means that there is unlikely to be meaningful gender differences in vowel F3−F2. Figure 4 shows that final tokens have higher values than non-final tokens and Manchester has higher values than Liverpool in all contexts except morpheme boundary position. Again, these results are largely similar to the lateral target analysis, but the vowel dialect differences are consistently smaller in magnitude.

In summary, we observe relatively similar patterns across the lateral and vowel targets analyses, with Liverpool generally showing higher F2−F1 and lower F3−F2 than Manch-

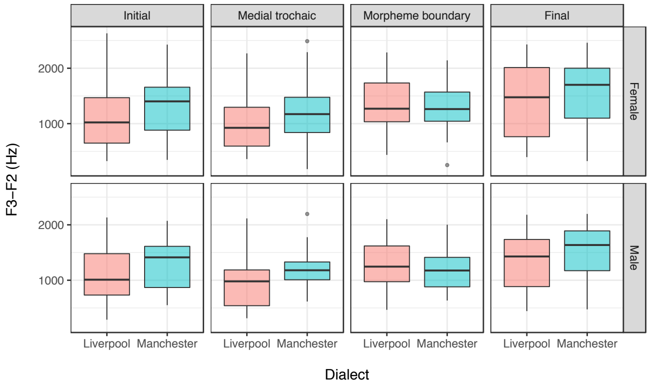

FIG. 4. F3−F2 values for vowel midpoint. (Colour online).

ester. However, while we see dialect differences across all positional contexts (except for word-final /l/ amongst females), these differences are typically of a smaller magnitude in the vowels. In some cases, such as morpheme boundary position, the dialects produce near-identical vowel realizations. Overall, this suggests that there exists positional and dialect variation in laterals, accompanied by a smaller degree of positional and dialect variation in the surrounding vowels.

C. Time-varying analysis

In this section we report the GAMM analysis that models the effects of time and dialect on formant values across the entire lateral and vowel(s) sequence at each position. We fit separate models to each positional context and focus on dialect differences within contexts. This is because (i) time-varying formants between positional contexts are unsurprisingly different due to a different sequencing of the lateral and vowel phases between contexts; (ii) time normalization across non-equivalent intervals (e.g. initial lateral-vowel versus medial vowel-lateral-vowel) renders direct comparison of different positions somewhat problematic.

However, while we do not statistically test comparisons across positional contexts, they can still be observed in the graphical model fits.

Table I shows the model comparisons used to test the significance of dialect and time-by-dialect on F2−F1. For the initial tokens we find no overall effect of dialect. Medial trochaic /l/ shows an overall effect of dialect, but further testing shows no significant effect of shape, suggesting that the two dialects only differ in the height of the F2−F1 trajectory. For morpheme boundary and final contexts we find no overall significant effect of dialect on F2−F1.

The model fits for F2−F1 are visualized in Figure 5. In line with the model comparisons, word-medial trochaic tokens show a difference only in the height of the trajectory, with Liverpool speakers showing higher F2−F1 across the lateral and vowel(s). Morpheme boundary and final contexts also show an absence of non-linear differences, in addition to no significant differences in the height of the trajectory. Word-final tokens in particular show almost complete overlap between dialects, while word-initial tokens show only very small dialect differences.. All trajectories only show a slight degree of non-linearity, so the data also do not confirm our prediction of significant non-linear differences between dialects.

Table II shows the model comparisons used to test the significance of dialect and time-by-dialect on F3−F2. For the initial and medial trochaic tokens we find an overall effect of dialect, but further testing shows no significant effect of shape. This suggests that the two dialects only differ in the height of the F3−F2 trajectory in these contexts. Morpheme boundary context shows an overall effect of dialect, while specific testing of the time-by-

TABLE I. Model comparisons for F2−F1 GAMMs.

Comparison	χ^2	*df*	$p(\chi^2)$
Initial			
Overall: dialect	2.70	3	.145
Shape: dialect	—	—	—
Medial trochaic			
Overall: dialect	4.62	3	.026
Shape: dialect	0.77	2	.463
Morpheme boundary			
Overall: dialect	2.92	3	.120
Shape: dialect	—	—	—
Final			
Overall: dialect	2.15	3	.231
Shape: dialect	—	—	—

dialect smooth term also shows a significant effect, suggesting significant dialect differences in the shape of the trajectory. For the word-final tokens we find no overall effect of dialect.

The model fits for F3−F2 are visualized in Figure 6. The patterns for initial and medial trochaic tokens show differences only in height rather than shape, with little-to-no overlap

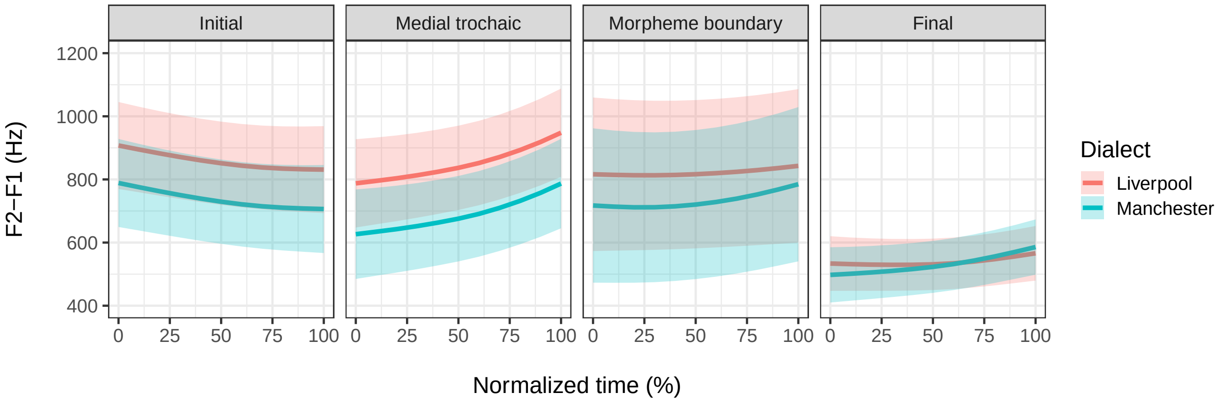

FIG. 5. Generalized Additive Mixed Model fits of the effects of normalized time-by-dialect on F2−F1 (Hz) at each positional context. Each panel shows the full model fit for that positional context with a mean smooth and 95% confidence interval for each dialect. (Colour online).

in confidence intervals. Word-final position shows a small difference in height, but this difference was not significant according to the model comparison. The morpheme boundary context is the only example of a non-linear significant difference between dialects in our time-varying data. While the differences in the overall height of the trajectory are smaller than the other contexts, the Manchester group shows a more non-linear trajectory for these tokens, with F3−F2 showing the biggest dialect differences around the interval midpoint and becoming most similar over the latter 50% of the V1-lateral-V2 interval. Our lateral and vowel targets analysis found no significant dialect differences in the morpheme boundary V1, while the GAMMs here show even fewer differences in V2 for the same context. Note that, despite the lack of overall non-linear *differences* between dialects, there is a visibly greater degree of non-linearity in the F3−F2 trajectories when compared with F2−F1.

TABLE II. Model comparisons for F3−F2 GAMMs.

Comparison	χ^2	*df*	$p(\chi^2)$
Initial			
Overall: dialect	5.62	3	.011
Shape: dialect	0.16	2	.849
Medial trochaic			
Overall: dialect	6.68	3	.004
Shape: dialect	0.93	2	.395
Morpheme boundary			
Overall: dialect	6.80	3	.004
Shape: dialect	4.52	2	.011
Final			
Overall: dialect	2.93	3	.119
Shape: dialect	—	—	—

D. Summary of results

In summary, Liverpool speakers generally produce higher F2−F1 and lower F3−F2 than Manchester speakers in non-final /l/ contexts and in the adjacent vowels. In final /l/, Manchester males produce darker /l/s than Liverpool males, whereas female speakers pro-

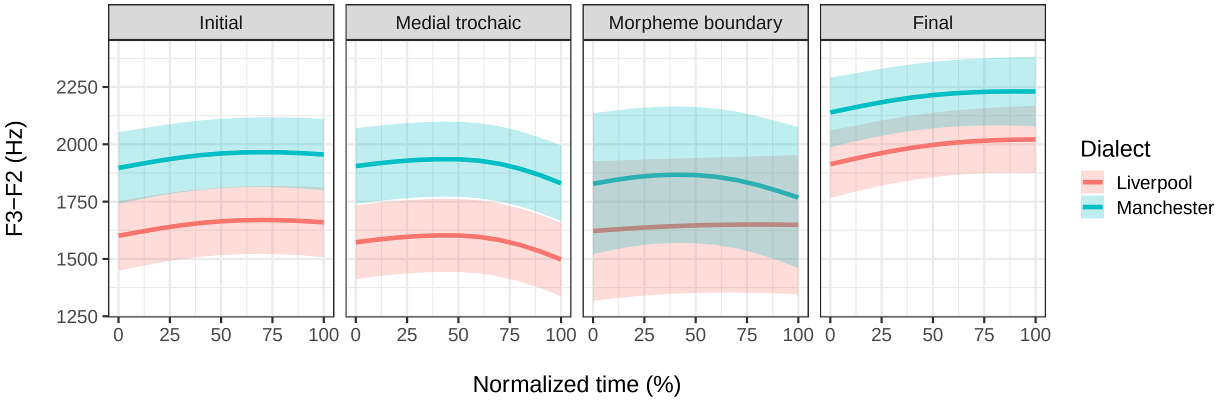

FIG. 6. Generalized Additive Mixed Model fits of the effects of normalized time-by-dialect on F3−F2 (Hz) at each positional context. Each panel shows the full model fit for that positional context with a mean smooth and 95% confidence interval for each dialect. (Colour online).

duce roughly similar F2−F1 values in this context. All groups produce contrast between initial and final /l/ to some extent, although this is largest in Liverpool speakers and smallest in Manchester females. The time-varying results collapsed the data across gender groups, so we only observed dialect differences in this analysis. Accordingly, the GAMMs show global differences in the height of the trajectory in F2−F1 for medial trochaic /l/, and in F3−F2 for all non-final contexts. However, the morpheme boundary F3−F2 model shows significant non-linear differences, which are largest in the first 50% of the interval (roughly equivalent to V1 plus lateral) and smallest during V2. In the following section, we discuss these results with respect to our hypotheses and illuminate their broader significance.

IV. DISCUSSION

A. Time-varying formant patterns

One of the major aims of our study was to offer a conceptual comparison between an analysis of the lateral/vowel targets and an analysis of the time-varying lateral and vowel formants. We find evidence of global F2−F1 and F3−F2 differences across the lateral and vowel in medial trochaic contexts, and for F3−F2 in all non-final contexts. Surprisingly, the only non-linear difference between dialects is in F3−F2 for morpheme boundary sequences. Here we see the biggest difference in the middle of sequence (roughly representing the /l/) and the smallest at the end of the sequence (roughly representing V2). This was not predicted; in fact, we actually predicted that we would find non-linear differences in all contexts (H4), with the magnitude of non-linearity largest in medial trochaic context (H3).

The non-linear difference in morpheme boundary context potentially represents the fact that the two dialects differ in the lateral but not V2. This stands in contrast to medial trochaic tokens, where we predicted and found differences in V2 (H3). A potential explanation for this could lie in the morphological conditioning of /l/ and its subsequent influence on the adjacent vowel. Medial trochaic contexts potentially allow for clearer realizations (Hayes, 2000; Lee-Kim *et al.*, 2013; Sproat and Fujimura, 1993) and, therefore, arguably greater potential for dialect variation. This may explain why we also see larger dialect differences in medial trochaic vowels, while Figures 3 and 4 show little-to-no dialect differences in morpheme boundary V1. Under this view, the medial trochaic vowel differences would be a coarticulatory consequence of dialect differences in /l/, while the lack of such differences in morpheme boundary vowels are due to the smaller dialect differences in /l/ in this context.

We believe that a more convincing explanation for these patterns is the likelihood of robust dialect variation in medial trochaic vowels. Medial trochaic V2 was always what Wells (1982) terms the HAPPY vowel, which is well-known to vary between dialects of British English. In the south of England, this vowel is undergoing change from [ɪ] to [i] (Fabricius, 2002; Harrington, 2006), whereas in many northern varieties there are a range of backed and centralized realizations, including [ɛ̈] (Hughes *et al.*, 2005; Kirkham, 2015). Manchester English in particular is stereotyped for its centralised production of this vowel, which is prevalent in working-class speakers (Baranowski and Turton, 2015). There is little prior data on this vowel in Liverpool English, but the acoustic evidence in this study suggests that it is produced with higher F2−F1 values, which would place it closer to [ɪ] and [i]. It is likely that the coarticulatory relationship between clearer /l/s and higher-fronter vowels, and between darker /l/s and lower-backer vowels, is magnified when both segments co-occur. Indeed, this explanation has been pursued in previous work in which there are known differences in the quality of this vowel between dialects (Kirkham, 2017) and this proposal may explain why dialect differences tend to be largest for both the lateral and the adjacent vowel in medial trochaic context.

Unsurprisingly, it is difficult to conclude whether the dialect differences we see here are a consequence of the lateral (which then exerts coarticulatory influence on the vowel) or the vowel (which then exerts coarticulatory influence on the lateral). In practice, the lateral and its adjacent vowels appear to vary in tandem in some instances, although the targets analysis suggests that the magnitude of the dialect difference at the lateral steady-state is larger than at the vowel midpoint. Despite this, we did not find the predicted non-linear time-varying differences at medial trochaic position in our GAMM analysis. One reason for this could be that speaker and word level variance in the time-varying patterns is too large to support significant non-linear differences between dialects. Indeed, this could suggest that there is greater within-dialect variability in cross-segmental formant dynamics than in segmental targets, which could lend support to the view that segmental targets are a more prominent goal than cross-segmental dynamics. A more comprehensive investigation into the relationship between segmental targets and time-varying acoustics is required in order to address this issue further.

B. English lateral typology

Our static and time-varying analyses both find that Liverpool non-final /l/s typically have higher F2−F1 and lower F3−F2 values than Manchester /l/s, which supports our predictions (H2). Based on previous work on the relationship between acoustic measures and impressionistic description (Kelly and Local, 1989; Recasens, 2012), this suggests that Liverpool has clearer realizations of /l/. However, while these results are predicted by the literature and clearly evident in the data, it is important to note that these differences may be comparably small when placed in context with other British English dialects. For example, Kirkham (2017) reports data from Sheffield Asian (Punjabi-influenced) speakers producing the same or very similar words as in the present study and reports mean F2−F1 values in initial /l/ of 1679 Hz for male speakers and 1599 Hz for female speakers. The comparison is somewhat hindered by the age difference between samples (13–14 in Kirkham 2017; 19–27 in the present study). However, with this caveat in mind, the highest F2−F1 value for a Liverpool female speaker producing initial /l/ is 1595 Hz, with most tokens around or below 1000 Hz. Therefore, in comparison to Sheffield Asian English – a variety with unusually clear /l/s – Liverpool is probably best considered an 'intermediate' /l/ variety. This is in line with previous impressionistic reports (Knowles, 1973; Wells, 1982), as well as the instrumental evidence available on Liverpool English (Turton, 2014).

Another salient aspect of /l/ variation is the implementation of positional contrast. Unsurprisingly, initial laterals have higher F2−F1 and lower F3−F2 than final laterals (thus confirming H1), with little evidence that word-medial laterals are significantly different from initial laterals. However, we find that the initial∼final contrast appears to be larger in Liverpool than in Manchester. This may reflect larger differences in the production of initial /l/s in the two dialects, which we see in the absence of significant time-varying F2−F1 differences across the entire vowel-lateral interval in final position. We note that while positional contrast in dark /l/ varieties of English, such as Manchester, is widely attested (Carter and Local, 2007; Turton, 2014), the production of initial∼final contrast is not inevitable. For example, previous work finds that some dark /l/ dialects of Catalan (Recasens and Espinosa, 2005) and English (Kirkham, 2017) do not show such positional variants.

To this end, one unexpected difference is gender variation in the initial∼final contrast. While Manchester males show lower F2−F1 than Liverpool males in initial and final position, Manchester females have similar or slightly higher values than Liverpool females in final position. Individual-level data show that Manchester females are more variable in the implementation of the initial∼final contrast, with some speakers producing a small or no difference between positions. The size of these effects is relatively small and we did not predict their existence, so we do not wish to attach too much weight to them. However, in terms of possible explanations, Turton (2014) finds that working-class Manchester speakers may not produce an initial∼final contrast in laterals, whereas middle-class speakers do. We did not collect information on the socioeconomic background of our participants, but it could be the case that some of the Manchester female speakers in our study are from more working-class backgrounds, which may interact with variation in the production of the initial∼final contrast. Establishing the robustness of such effects motivates a need for tighter control over social stratification in experimental phonetic studies.

V. CONCLUSIONS

In this article we reported acoustic data on laterals, vowels and their time-varying formant dynamics in two major dialects of British English. We find that Liverpool generally has clearer non-final /l/s than Manchester. However, we propose that Liverpool English is best considered an 'intermediate' variety that lies towards the middle of the clear-dark continuum in English dialects. Our comparison of steady-state and time-varying results shows that the two analyses generally agree with each other, but the time-varying analysis further highlights the strong coarticulatory interactions between laterals and vowels in each dialect. This analysis also demonstrates that GAMMs are a versatile tool for modelling formant dynamics across multi-segmental sequences. In conclusion, analysing formant dynamics reveals that making strong claims about independent lateral and vowel targets should be approached with caution, and future research into segmental targets and time-varying spectral information should seek to further address the specific nature of their relationship.

ACKNOWLEDGMENTS

This research was supported by a Lancaster University Summer Project Research Internship grant and a Lancaster University FASS Research Fund grant, both awarded to Sam Kirkham and Claire Nance. We would like to thank Ewa Jacewicz, Márton Sóskuthy and an anonymous reviewer for their valuable feedback on previous versions of this article. Data and code for all analyses are available at: `https://osf.io/5u6ez/`.

APPENDIX A: LMER MODEL SUMMARIES

For all models, baseline variables are Dialect = Liverpool, Position = Initial, Gender = Female. Random effects in each model include word and speaker random intercepts and by-speaker random slopes for the effect of position.

TABLE III. Lateral steady-state: F2−F1.

Variable	β	SE	t	$p(\chi^2)$
Intercept	883.21	64.38	13.72	—
Dialect				—
Manchester	−76.60	53.34	−1.44	
Position				—
Medial trochaic	−19.20	96.02	−0.20	
Morpheme boundary	−7.93	95.13	−0.08	
Final	−341.53	81.30	−4.20	
Gender				—
Male	−68.98	53.82	−1.28	
Dialect x gender				.020
Manchester:Male	−120.13	48.16	−2.50	
Position x gender				.327
Medial trochaic:Male	27.74	40.85	0.68	
Morph. boundary:Male	−35.09	38.38	−0.91	
Final:Male	39.27	44.80	0.88	
Position x dialect				.028
Medial trochaic:Manchester	−45.31	40.66	−1.11	
Morph. boundary:Manchester	22.21	38.20	0.58	

TABLE IV. Lateral steady-state: F3−F2.

Variable	β	SE	t	$p(\chi^2)$
Intercept	1793.50	85.25	21.04	—
Dialect				.001
Manchester	261.67	99.20	2.64	
Position				.003
Medial trochaic	−56.67	96.52	−0.59	
Morpheme boundary	−40.14	95.91	−0.42	
Final	335.54	94.17	3.56	
Gender				$< .001$
Male	−254.55	100.92	−2.52	
Dialect x gender				.872
Manchester:Male	22.42	137.28	0.16	
Position x gender				.881
Medial trochaic:Male	−23.99	62.18	−0.39	
Morph. boundary:Male	34.79	61.89	0.56	
Final:Male	−6.00	80.27	−0.08	
Dialect x position				.354
Medial trochaic:Manchester	37.72	61.89	0.61	
Morph. boundary:Manchester	−77.26	61.57	−1.26	

TABLE V. Vowel midpoint: F2−F1.

Variable	β	SE	t	$p(\chi^2)$
Intercept	1253.78	216.56	5.79	—
Dialect				—
Manchester	−227.71	42.02	−5.42	
Position				—
Medial trochaic	926.46	407.66	2.27	
Morpheme boundary	−618.89	404.45	−1.53	
Final	−357.80	324.24	−1.10	
Gender				—
Male	−175.76	42.50	−4.14	
Dialect x gender				.924
Manchester:Male	−4.62	47.25	−0.10	
Position x gender				$< .001$
Medial trochaic:Male	−247.55	90.23	−2.74	
Morph. boundary:Male	143.49	68.19	2.10	
Final:Male	104.43	53.20	1.96	
Position x dialect				$< .001$
Medial trochaic:Manchester	−298.23	89.85	−3.32	
Morph. boundary:Manchester	268.38	67.86	3.96	

TABLE VI. Vowel midpoint: F3−F2.

Variable	β	SE	t	$p(\chi^2)$
Intercept	1146.84	197.90	5.80	—
Dialect				—
Manchester	166.07	49.03	3.39	
Position				—
Medial trochaic	−577.74	368.40	−1.57	
Morpheme boundary	455.20	367.33	1.24	
Final	300.99	294.62	1.02	
Gender				—
Male	−41.58	49.77	−0.84	
Dialect x gender				.518
Manchester:Male	41.97	63.70	0.66	
Position x gender				.243
Medial trochaic:Male	63.67	69.54	0.92	
Morph. boundary:Male	−98.14	62.07	−1.58	
Final:Male	−79.15	49.48	−1.60	
Position x dialect				< .001
Medial trochaic:Manchester	159.38	69.22	2.30	
Morph. boundary:Manchester	−266.24	61.77	−4.31	

Aikin, L. S., and West, S. G. (**1991**). *Multiple Regression: Testing and Interpreting Interactions* (Sage, London).

Baranowski, M. (**2017**). "Class matters: The sociolinguistics of GOOSE and GOAT in Manchester English," Language Variation and Change **29**(3), 301–339.

Baranowski, M., and Turton, D. (**2015**). "Manchester English," in *Researching Northern Englishes*, edited by R. Hickey (John Benjamins, Amsterdam), pp. 293–316.

Bates, D., Maechler, M., Bolker, B., and Walker, S. (**2015**). "Fitting linear mixed-effects models using lme4," Journal of Statistical Software **67**(1), 1–48.

Bombien, L., Winkelmann, R., and Scheffers, M. (**2016**). "wrassp: An R wrapper to the ASSP Library" R package version 0.1.4.

Cardoso, A. (**2015**). "Variation in nasal–obstruent clusters and its influence on PRICE and MOUTH in Scouse," English Language and Linguistics **19**(3), 505–532.

Carter, P. (**2002**). "Structured variation in British English liquids," Ph.D. thesis, University of York, York.

Carter, P., and Local, J. (**2007**). "F2 variation in Newcastle and Leeds English liquid systems," Journal of the International Phonetic Association **37**(2), 183–199.

Delsarte, P., and Genin, Y. (**1986**). "The split Levinson algorithm," IEEE Transactions on Acoustics, Speech, and Signal Processing **34**(3), 470–478.

Elvin, J., Williams, D., and Escudero, P. (**2016**). "Dynamic acoustic properties of monophthongs and diphthongs in Western Sydney Australian English," Journal of the Acoustical Society of America **140**(1), 576–581.

Fabricius, A. H. (**2002**). "Weak vowels in modern RP: an acoustic study of happY-tensing and KIT/schwa shift," Language Variation and Change **14**(2), 211–237.

Fox, R. A., and Jacewicz, E. (**2009**). "Cross-dialectal variation in formant dynamics of American English vowels," Journal of the Acoustical Society of America **126**(5), 2603–2618.

Harrell, F. E. (**2015**). *Regression Modeling Strategies: With Applications to Linear Models, Logistic and Ordinal Regression, and Survival Analysis*, second ed. (Springer, Switzerland).

Harrington, J. (**2006**). "An acoustic analysis of 'happy-tensing' in the Queen's Christmas broadcasts," Journal of Phonetics **34**(4), 439–457.

Hayes, B. (**2000**). "Gradient well-formedness in Optimality Theory," in *Optimality Theory: Phonology, Syntax, and Acquisition*, edited by J. Dekkers, F. van der Leeuw, and J. van de Weijer (Oxford University Press, Oxford), pp. 88–120.

Hughes, A., Trudgill, P., and Watt, D. (**2005**). *English Accents and Dialects: An Introduction to Social and Regional Varieties of English in the British Isles* (Hodder, London).

Jones, D. (**1966**). *The Pronunciation of English* (Cambridge University Press, Cambridge).

Kelly, J., and Local, J. (**1986**). "Long domain resonance patterns in English," Proceedings of IEEE Conference on Speech Input/Output: Techniques and Applications 304–309.

Kelly, J., and Local, J. (**1989**). *Doing Phonology: Observing, Recording, Interpreting* (Manchester University Press, Manchester).

Kirkham, S. (**2015**). "Intersectionality and the social meanings of variation: Class, ethnicity and social practice," Language in Society **44**(5), 629–652.

Kirkham, S. (**2017**). "Ethnicity and phonetic variation in Sheffield English liquids," Journal of the International Phonetic Association **47**(1), 17–35.

Kleber, F., Harrington, J., and Reubold, U. (**2011**). "The relationship between perception and production of coarticulation during a sound change in progress," Language and Speech **55**(3), 383–405.

Knowles, G. (**1973**). "Scouse: The urban dialect of Liverpool," Ph.D. thesis, University of Leeds, Leeds.

Ladefoged, P., and Maddieson, I. (**1996**). *The Sounds of the World's Languages* (Blackwell, Oxford).

Lee-Kim, S.-I., Davidson, L., and Hwang, S. (**2013**). "Morphological effects on the darkness of English intervocalic /l/," Laboratory Phonology **4**(2), 475–511.

Lehiste, I. (**1964**). *Acoustical Characteristics of Selected English Consonants* (Mouton, The Hague).

Markel, J., and Gray, A. (**1976**). *Linear Prediction of Speech* (Springer-Verlag, Berlin).

Nance, C. (**2014**). "Phonetic variation in Scottish Gaelic laterals," Journal of Phonetics **47**(1), 1–17.

Nance, C., Kirkham, S., and Groarke, E. (**2015**). "Intonational variation in Liverpool English," in *Proceedings of the XVIII International Congress of Phonetic Sciences*, paper ICPHS0633, University of Glasgow, Glasgow, pp. 1–5.

Narayanan, S. S., Alwan, A. A., and Haker, K. (**1997**). "Toward articulatory-acoustic models for liquid approximents based on MRI and EPG data: Part I the laterals," Journal of the Acoustical Society of America **101**(2), 1064–1077.

Recasens, D. (**2012**). “A cross-language acoustic study of initial and final allophones of /l/,” Speech Communication **54**(3), 368–383.

Recasens, D., and Espinosa, A. (**2005**). “Articulatory, positional and coarticulatory characteristics for clear /l/ and dark /l/: evidence from two Catalan dialects,” Journal of the International Phonetic Association **35**(1), 1–25.

Simonet, M., Rohena-Madrazo, M., and Paz, M. (**2008**). “Preliminary evidence for incomplete neutralization of coda liquids in Puerto Rican Spanish,” in *Selected proceedings of the 3rd Conference on Laboratory Approaches to Spanish Phonology* (Cascadilla Proceedings Project, Somerville, MA), pp. 72–86.

Sóskuthy, M. (**2017**). “Generalised additive mixed models for dynamic analysis in linguistics: A practical introduction” arXiv:1703.05339 [stat:AP].

Sóskuthy, M., Foulkes, P., Hughes, V., and Haddican, B. (**2018**). “Changing words and sounds: The roles of different cognitive units in sound change,” Topics in Cognitive Science **10**(4), 787–802.

Sproat, R., and Fujimura, O. (**1993**). “Allophonic variation in English /l/ and its implications for phonetic implementation,” Journal of Phonetics **21**(2), 291–311.

Strycharczuk, P., and Scobbie, J. M. (**2017**). “Fronting of Southern British English high-back vowels in articulation and acoustics,” Journal of the Acoustical Society of America **142**(1), 322–331.

Stuart-Smith, J., Lennon, R., Macdonald, R., Robertson, D., Sóskuthy, M., José, B., and Evers, L. (**2015**). “A dynamic acoustic view of real-time change in word-final liquids in spontaneous Glaswegian,” in *Proceedings of the XVIII International Congress of Phonetic*

Sciences, paper ICPHS1028, University of Glasgow, Glasgow, pp. 1–5.

Turton, D. (**2014**). “Variation in English /l/: Synchronic reflections of the life cycle of phonological processes,” Ph.D. thesis, University of Manchester, Manchester.

van Rij, J., Wieling, M., Baayen, R. H., and van Rijn, H. (**2017**). “itsadug: Interpreting time series and autocorrelated data using GAMMs” R package version 2.3.

Watson, C. I., and Harrington, J. (**1999**). “Acoustic evidence for dynamic formant trajectories in Australian English vowels,” Journal of the Acoustical Society of America **106**(1), 458–468.

Watson, K. (**2007**). “Liverpool English,” Journal of the International Phonetic Association **37**(3), 351–360.

Wells, J. C. (**1982**). *Accents of English: Volumes 1–3* (Cambridge University Press, Cambridge).

Wieling, M. (**2018**). “Generalized additive modeling to analyze dynamic phonetic data: A tutorial focusing on articulatory differences between L1 and L2 speakers of English,” Journal of Phonetics **70**, 86–116.

Williams, D., and Escudero, P. (**2014**). “A cross-dialectal acoustic comparison of vowels in Northern and Southern British English,” Journal of the Acoustical Society of America **136**(5), 2751–2761.

Winkelmann, R., Harrington, J., and Jänsch, K. (**2017**). “EMU-SDMS: Advanced speech database management and analysis in R,” Computer Speech & Language **45**, 392–410.

Winkelmann, R., and Raess, G. (**2014**). “Introducing a web application for labeling, visualizing speech and correcting derived speech signals,” in *Proceedings of the Ninth International*

Conference on Language Resources and Evaluation (LREC'14), edited by N. Calzolari, K. Choukri, T. Declerck, H. Loftsson, B. Maegaard, J. Mariani, A. Moreno, J. Odijk, and S. Piperidis (European Language Resources Association (ELRA), Nicoletta), pp. 4129–4133.

Wood, S. N. (**2017**). *Generalized Additive Models: An Introduction with R*, second edition ed. (CRC Press, Boca Raton, FL).