Laboratory Phonology
Journal of the Association for Laboratory Phonology

Turton, D 2017 Categorical or gradient? An ultrasound investigation of /l/-darkening and vocalization in varieties of English. *Laboratory Phonology: Journal of the Association for Laboratory Phonology* 8(1): 13, pp. 1–31, DOI: https://doi.org/10.5334/labphon.35

JOURNAL ARTICLE

Categorical or gradient? An ultrasound investigation of /l/-darkening and vocalization in varieties of English

Danielle Turton
Newcastle University, GB
danielle.turton@newcastle.ac.uk

This paper presents an empirical analysis of /l/-darkening in English, using ultrasound tongue imaging data from five varieties spoken in the UK. The analysis of near 500 tokens from five participants provides hitherto absent instrumental evidence demonstrating that speakers may display both categorical allophony of light and dark variants, and gradient phonetic effects coexisting in the same grammar. Results are interpreted through the modular architecture of the life cycle of phonological processes, whereby a phonological rule starts its life as a phonetically driven gradient process, over time stabilizing into a phonological process at the phrase level, and advancing through the grammar. Not only does the life cycle make predictions about application at different levels of the grammar, it also predicts that stabilized phonological rules do not replace the phonetic processes from which they emerge, but typically coexist with them, a pattern which is supported in the data. Overall, this paper demonstrates that variation in English /l/ realization has been underestimated in the existing literature, and that we can observe phonetic, phonological, and morphosyntactic conditioning when accounting for a representative range of phonological environments across varieties.

Keywords: ultrasound; dialects of English; life cycle of phonological processes; /l/-darkening; /l/-vocalization; variation and change; diachrony and synchrony; categorical; gradient; modular

1 Introduction

The process of /l/-darkening, whereby /l/ is realized with a delayed and/or reduced tongue-tip gesture, shows a remarkable amount of variation across accents of English. Although it is generally stated that light [l] is found in onsets, e.g., *leap, love*, and dark [ɫ] is found in codas, e.g., *heal, pull*, many investigations into the phonetics of /l/ have claimed there is no allophonic distinction, and that the perceived allophones are merely two ends of a phonetic continuum, conditioned by factors such as duration and neighbouring vowels. Complicating matters of allophony further, previous studies assuming a simple onset/coda distinction have found that morphosyntactic sensitivity to the process may vary between varieties, so that some may show darkening before a vowel in the following word, e.g., word-final prevocalic *heal it*, and other varieties may not.

The current study seeks to investigate claims of phonetic gradience and phonological categoricity in English /l/ realizations using ultrasound tongue imaging. The primary aim is to find articulatory evidence for the claims of allophony, or lack of it, by considering whether the data show categorical light and dark variants of /l/, or whether the patterning seems to be more gradient. Although not the primary focus of this paper, the effects of morphosyntax will also be considered in an attempt to interpret the data with due consideration of all possible linguistic predictors.

In this paper, each variety of English is represented by one speaker. Thus, it is not necessarily being claimed that patterns found in individual speakers are representative of the variety as a whole. It is entirely possible that speakers may show dialectal patterns subject to ongoing community change, variation based on sociolinguistic factors, or even idiolectal differences. Nevertheless, it is interesting to note that, for the most part, speakers' general patterns bear a striking similarity with the descriptions in the existing dialectal literature.

The structure of this paper is in two parts. The first discusses the literature on /l/-darkening, both theoretical and empirical, and summarizes the arguments concerning the grammatical representation of the variation found in English. The second part presents the ultrasound tongue imaging data, providing necessary instrumental evidence to test the claims summarized in the first section.

2 Previous interpretations of /l/-darkening

2.1 Categorical vs. gradient processes

The traditional approach to /l/-darkening posits an allophonic distinction between light [l] and dark [ɫ] (e.g., Chomsky & Halle, 1968; Halle & Mohanan, 1985) which, depending on the interpretation, may or may not map onto onset and coda realizations. Dark [ɫ] is the lenited variant, being articulatorily backer, with a reduced or delayed tongue tip gesture (Gick, 2003; Giles & Moll, 1975; Lee-Kim, Davidson, & Hwang, 2013; Lin, 2011; Sproat & Fujimura, 1993). Acoustically, dark [ɫ] is characterized by a low F2 or, more specifically, a small difference between the first and second formants (Carter, 2003; Ladefoged & Maddieson, 1996; Lehiste, 1964). Although most phonological descriptions of darkening do not usually account for any variation within this observation, Hayes's (2000) analysis of darkening in Optimality Theory is an exception to this. Hayes used a judgement elicitation task in which native speakers of American English rated the acceptability of light and dark /l/s in different phonological contexts. Hayes did not directly control the phonetic quality of /l/ in the stimuli, but rather that of the preceding vowel, justifying this procedure on the grounds that front or high vowels and 'true diphthongs' preceding dark [ɫ] have a schwa off-glide. Therefore, the presence of breaking in the preceding vowel was taken as a proxy for darkening. Hayes argues that acceptability judgements are determined by a variable grammar, and that any variation in /l/-darkening is entirely down to category mixture of discrete allophones.

The results are as we might expect: Participants give low rankings to light [l]s in word final position (e.g., ***bell***) and dark [ɫ]s in initial position (e.g., ***L****ouanne*), the expected variants (i.e., light word initially, dark word finally) are ranked high, and all other contexts in between. Although Hayes acknowledges extra gradient phonetic effects, such as duration, he treats these as low-level, and his analysis reflects a situation where variation is conditioned by probabilistic application of morphosyntactically sensitive, categorical phonological processes. The results are compelling, but the design of the experiment inherently elicits scalar results, meaning that it is difficult (or impossible) to distinguish between variability at the level of the individual, and variability across the speech community as a whole. This results in a potential and probable overestimation of categorical variation in the production of individuals. An additional, albeit peripheral, problem with this approach is that it cannot account for findings such as duration variance being greater in e.g., phrase-final position ('pre-boundary lengthening,' Lehiste, 1980, p. 7) than in other positions.

Hayes's analysis stands in stark contrast to several phonetic studies of /l/-darkening, many of which have argued that /l/-darkening differences reflect a gradient, not categorical, distribution. Under this interpretation, it is argued that the so-called allophonic difference

between light and dark variants is merely two ends of one continuum. The most famous of these claims is made by Sproat and Fujimura (1993) in their X-ray microbeam study of /l/-darkening. In this investigation, /l/s were analyzed in 9 contexts, ranging from word initial position, to pre-consonantal position, and other prosodically conditioned environments in between. Their methodology allowed them to monitor gestural phasing, leading them to uncover a distinction which would regularly be cited as the primary articulatory correlate of light and dark /l/ from then on. They found that in light [l]s the coronal gesture precedes the dorsal gesture, and in dark [ɬ]s the dorsal gesture precedes the coronal (or is timed near-simultaneously), as represented by a subset of environments in **Table 1**.

However, they do not accept that such categories are the best way of characterizing the variation found, and argue that darkening is gradient and is dependent on duration. Although the relative phasing of the coronal and dorsal gestures seems categorical in nature, they argue it is simply a result of their alignment with adjacent segments. For example, in a word such as ***l**eap* the dorsal gesture comes second, as it is aligned with the following vowel. In *pee**l***, the dorsal gestures comes first, as it is aligned with the preceding vowel. They argue that, as the dorsal gesture can take more time to articulate because of its larger displacement, in longer rimes it has more chance to reach its extremum. This is why the darkness is correlated with duration: The longer the rime, the darker the /l/.

Lee-Kim et al. (2013) also conclude that /l/ darkness is on a continuum in their ultrasound tongue imaging study of American English. However, they do not accept the claim that duration is solely responsible for the magnitude of darkness. The authors used ultrasound to study /l/ in three phonological contexts: As a part of a suffix (e.g., *flaw-**l**ess*), before a stem-suffix boundary (e.g., *ta**ll**-est*), and pre-consonantally (e.g., *ta**ll** building*). Their primary articulatory correlate of darkness was tongue-body lowering, with a lower tongue body indicating a darker /l/. Their results indeed show a scalar range of darkness, with the lightest /l/s occurring in suffixal *flaw-**l**ess*-type tokens, the darkest in pre-consonantal *ta**ll** building*-type tokens, with pre-suffix *ta**ll**-est*-type tokens displaying an in between articulation. Working on the fair assumption that it is unlikely that three phonological categories would fall within such a tiny articulatory range, they conclude that darkness must be gradient, not categorical. They discuss an explanation whereby phonetic implementation is conditioned by morphological boundaries, which may also result in an indirect side-effect of duration in some contexts. It is worth noting that this explanation violates the principle of morphology-free phonetics (Bermúdez-Otero, 2007, 2015), the long-standing fundamental assumption of generative linguistics that there is no interface between morphology and phonetics (Kiparsky, 1982; Myers, 2000; see Pierrehumbert, 2002 for an overview of the assumptions of the modular feedforward architecture). In addition, it is not clear whether just three phonological contexts are comprehensive enough to make broad judgements of categoricity and gradience. It is fair to say that we may need data for the entire spectrum of realization possibilities before making such conclusions, i.e., a set of stimuli more like Sproat and Fujimura's than Lee-Kim, et al.'s.

Table 1: Adapted from the results of Sproat and Fujimura (1993).

Environment	Example	Gesture	Realisation
Initial	***L**ikkovsky*	Coronal gesture precedes dorsal	[l]
Intervocalic	*Mr Bee**l**ik*		
Intervocalic pre-boundary	*bee**l**-ing*		
Word-final prevocalic	*Bee**l** equates*	Dorsal gesture precedes coronal	[ɬ]
Final	*Nea**l***		

2.2 Evidence for both categoricity and gradience

Studies which take a more all-encompassing approach to both phonetics and phonology give us further insight into the debate. Lin (2011) considers both categorical effects of onset/coda distinctions alongside gradient factors of gestural timing and duration in her ultrasound production and perception study of American English /l/. Although coda /l/s show an effect of duration for all but 1 speaker, onset /l/s show effects for only 4 out of her 11 speakers. Lin (2011, p. 40–43) demonstrates how this onset effect is due to the prosodic conditions: Utterance initial /l/s are longer than utterance-medial word-initial /l/s, but the lag is generally shorter in utterance initial position, which exhibits a tighter correspondence between the timing of the tongue tip and tongue dorsum gestures. This demonstrates that longer durations are not always associated with darker variants, and that duration can reflect a prosodically stronger position without producing an articulatory darker /l/.

Evidence for both categorical and gradient effects is not restricted to articulatory investigation, but also found in acoustic studies. Yuan and Liberman (2009, 2011) analyzed over 20,000 tokens of American English /l/ in their investigation into /l/-darkening in the SCOTUS corpus (Supreme Court Justice of the United States corpus), which includes 50 years of spoken data from the Supreme Court tapes. Their results show some interesting possibilities for the categoricity against gradience arguments, particularly where duration is concerned. Yuan and Liberman (2009) looked at /l/ in three contexts: Intervocalically before a stressed syllable (e.g., *be**l**ieve*), intervocalically before an unstressed syllable (e.g., *he**l**ix*, *pee**l**-ing*) and word-finally (e.g., *pee**l***).[1] They found a correlation between /l/ darkness and duration, but for the latter two contexts only. In this dataset, those are the phonological contexts which were acoustically realized as dark [ɫ]s. In a follow-up paper, Yuan and Liberman (2011) consider different kinds of word-final /l/ (their 2009 paper grouped pre-consonantal and prevocalic tokens into the same word-final category) as well as initial /l/ (e.g., ***l**eap*). The results reinforce their original findings: Initial and foot-initial tokens (***l**eap* and *be**l**ieve*) reflect lighter /l/s, while the intervocalic and word-final tokens (*he**l**ix, pee**l*** and *pee**l** bananas*) reflect distinctly darker /l/s. Crucially, the correlation between how dark an /l/ was and its duration was found only in the *he**l**ix*, *pee**l*** and *pee**l** bananas*-type tokens, i.e., the dark ones.

Note that, for Yuan and Liberman's dataset, intervocalic *he**l**ix* and *pee**l**-ing*-like tokens are generally dark, although the experiment does not distinguish between monomorphemes and morphologically complex tokens. Intervocalic /l/-darkening is also found in Hayes (2000), whose informants accepted a dark [ɫ] in intervocalic monomorphemes such as *ye**ll**ow* and much more so in intervocalic /l/s preceding a stem-suffix boundary, such as *mai**l**-er*, showing that darkening is rather advanced in these North American varieties, applying to any /l/ which is non-initial in the prosodic foot. Yuan and Liberman's data do not make it possible to distinguish between gradience and variance, but we can see that the darkness scores for the intervocalic tokens are not as high as the pre-consonantal ones (2011, p. 41). However, whether this is phonetic gradience or category mixture is not clear.

The most interesting aspect to Yuan and Liberman's work on /l/-darkening is the finding that duration is only significant for the dark tokens, i.e., only dark variants show the correlation with duration. This is important, as it demonstrates an interaction between the categorical effects of allophony and the gradient effects of duration. However, this is not

[1] Note that Yuan and Liberman (2009) use different terminology to the present study. They refer to *he**l**ix*-type tokens as 'intervocalic coda /l/' and *be**l**ieve*-type tokens as 'intervocalic onset /l/,' using a syllabification of intervocalic consonants in the coda of the preceding syllable unless stressed. This is rephrased in their 2011 follow-up paper.

the first evidence of such effects. In fact, as pointed out by Bermúdez-Otero and Trousdale (2012), when inspecting Sproat and Fujimura's (1993: 303) own plot of darkening vs. duration, the exact same pattern seems to be true of their data also. **Figure 1** is a re-construction of the graph taken from the original 1993 paper, with ellipses and a correlation line added by the author.[2] The plot shows tip delay against duration. In their study, tip delay is taken as the primary articulatory correlate of darkening, due to the relative phasing of the coronal and dorsal gestures, and is plotted on the *y* axis. /l/s which have a negative tip delay, i.e., where the coronal gesture precedes the dorsal one, reflect lighter /l/s and are toward the bottom of the plot. Tokens where the dorsal gesture precedes the coronal reflect darker variants and are at the top of the plot. The *x* axis shows the duration of the rime. Different symbols reflect different phonological contexts, as shown in the legend. The point which Bermúdez-Otero and Trousdale (2012) make is represented

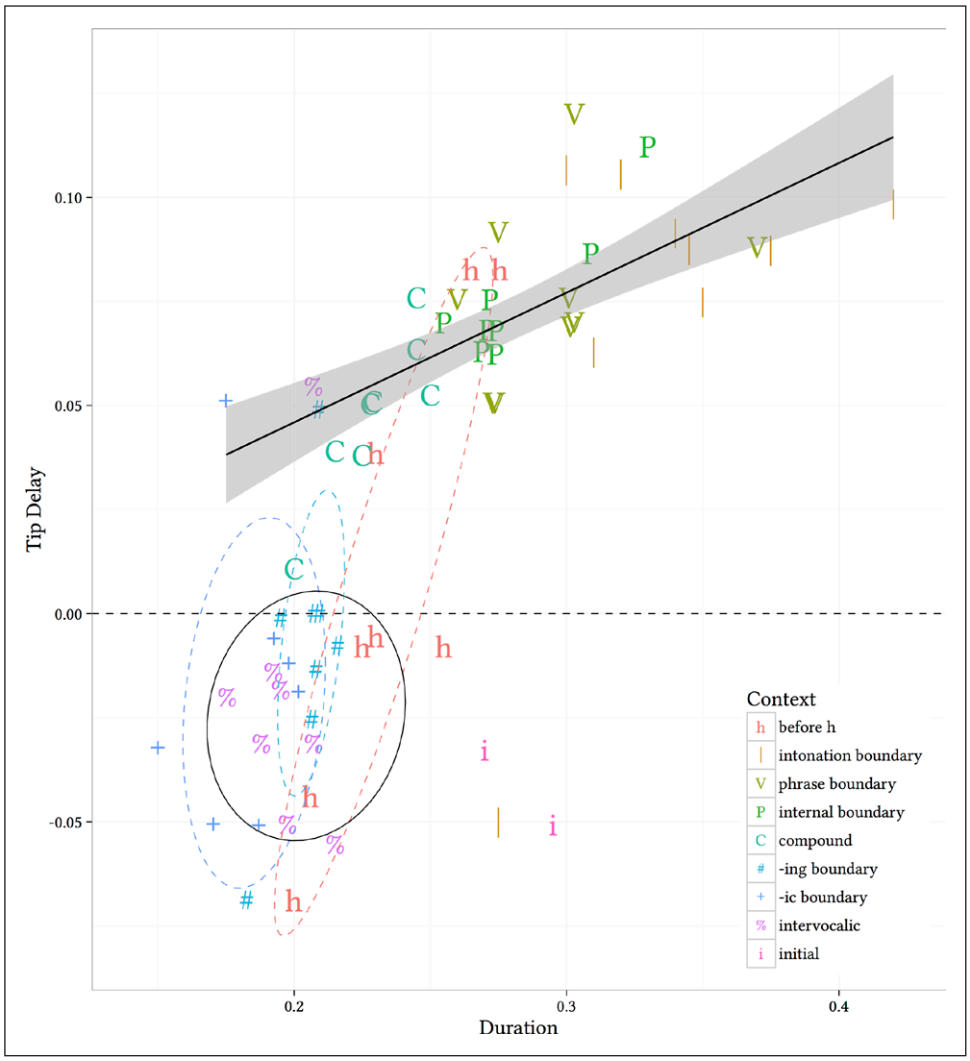

Figure 1: Graph adapted from Sproat and Fujimura (1993: 303) showing tip delay against duration across nine phonological contexts. Different characters represent different phonological contexts. The longer the tip delay, the darker the /l/. Correlation lines and ellipsis not in original plot. Dashed line ellipses represent nearest equivalent phonological environments investigated in Lee-Kim et al. (2013).

[2] The graph was automatically replicated using PlotDigitizer (Huwaldt, 2010). The ellipses were plotted with 68% confidence intervals (i.e., two standard deviations from the mean). The correlation line is fit to all values where the tip delay is positive, and the ellipse is fit to all values where the tip delay is negative.

by the added ellipsis and line, and is that, although the top half of the panel does seem to show a correlation between darkness and duration, the bottom half is more clustered together and shows no such pattern. Once again, this suggests an in-between situation for the debate around categoricity and gradience: Duration does correlate with darkness, but for the dark [ɫ]s only.

Figure 1 also highlights a drawback of Lee-Kim et al.'s investigation, with the nearest phonological environments studied in their 2013 paper shown by the dashed ellipses. By considering three phonological environments only, we have what does indeed look like either three categories in a tiny articulatory space, or gradient darkening, which may lead one to conclude that darkening is gradient. However, when looking at the nine contexts all together, the picture becomes somewhat different, showing the importance of accounting for the whole range of realizations when distinguishing between categoricity and gradience.

Indeed, fitting a linear regression line to the replicated Sproat and Fujimura data gives the suggestion that a model which accounts for both the categorical and gradient aspects of the data gives the best fit. **Table 2** gives the adjusted r^2 for five linear models with the dependent variable of tip delay, which is the articulatory correlate of darkness. The closer to 1 the r^2 value is, the better the fit of the predictors. Model 1 is the Sproat and Fujimura explanation: Tip delay can be explained purely by duration, which has a reasonable fit of 0.51. Model 2 is the Lee-Kim et al. explanation: The morphophonological context is the best way of explaining the variation in tip delay. This provides a slightly better fit than Model 1, with a respectable value of 0.54. However, the best fit is a model which considers the gradient effect of duration alongside an added predictor of category. Model 3 assigns separate categories for tokens with a negative tip delay (i.e., light [l]s) and a positive tip delay (dark [ɫ]) and has an adjusted r^2 of 0.87. Model 4 simply tests whether adding duration to individual context makes a significant improvement (it does not) whilst model 5 considers an interaction between category and duration, so that the model can account for different effects of duration in light [l] than in dark [ɫ]. This model gives the best fit overall, with an an adjusted r^2 of 0.88, which is a significant improvement on Model 3 (likelihood ratio test; $df = 1, F = 6.38, p = 0.01$).

Summarizing so far, within the literature on /l/-darkening in English, we can find analyses which take one of two ways of interpreting the phonological process, and these tend to be strictly gradient or categorical. However, it seems the empirical evidence points towards the existence of both. Purely theoretical approaches tend to overestimate the role of categorical variation, whereas many phonetic analyses seem to underplay it. Nevertheless, there are compelling arguments from both interpretations. This paper will attempt to investigate whether there is an approach in phonological theory which can account for both categorical and gradient effects, retaining the positive aspects of these analyses whilst overcoming their deficiencies.

Table 2: Comparison of adjusted *r*-squared from five linear models. Model 5 is significantly better than model 3 by likelihood ratio test.

Model	adjusted r^2
1. *TipDelay ~ Duration*	0.51
2. *TipDelay ~ Context*	0.54
3. *TipDelay ~ Category + Duration*	0.87
4. *TipDelay ~ Context + Duration*	0.66
5. *TipDelay ~ Category × Duration*	0.88

2.3 A modular approach

Using theoretical and empirical arguments, the present analysis proposes that a modular approach can capture both categorical phonological effects, alongside gradient phonetic effects. In a modular grammar, discrete features are computed in the phonology, whereas the phonetics assigns targets to surface feature realizations. Such approaches date back to early discussions of the architecture of the grammar, furthered by works such as Chomsky and Halle (1968), Kiparsky (1988), Hale and Reiss (2000), and, more recently by Bermúdez-Otero (2011, 2012) and Boersma (2011), amongst others, although the specifics of these theories vary.

Following the ideas behind the life cycle of phonological processes (Bermúdez-Otero, 1999, 2011; Ramsammy, 2015), visualized in **Figure 2**, the phonology itself is stratified with three levels: The stem level, word level, and phrase level. In this theory, phonetically driven innovations enter the grammar as gradient phonetic rules which may eventually stabilize as categorical phonological processes at the phrase level. Domain narrowing occurs in the next stage, resulting from analogical change where the new phonological process climbs to a higher level of the grammar, i.e., a process which formerly applies at the phrase level now applies at the word level.

It is important to point out that the life cycle views the grammar from an amphichronic perspective, where synchronic and diachronic explanation feed each other (Bermúdez-Otero, 2015, p. 374; see also Kiparsky, 2006, p. 222). The diachronic stages of the life cycle are reflected in the current synchronic grammar which often show microtypological variation across different dialects. As Ramsammy (2015, p. 33) explains, there is mutual complementarity between synchrony and diachrony in sound change, and by investigating patterns of interdialectal phonological variation, we can gather a broader picture of possible microtypologies which reflect a series of micro-level sound changes.

Table 3 demonstrates such microtypologies in /l/-darkening, and the effect of morphosyntactic domain narrowing. For example, in Received Pronunciation (henceforth RP), /l/ is light word-finally providing it is prevocalic (Cruttenden, 2008, p. 85). This system varies from American English 1 (Sproat & Fujimura, 1993), where in phrases such as, "Beel equates the actors," the /l/ is dark (the dorsal gesture precedes the coronal). As this /l/ is prevocalic, we may expect it to remain light through resyllabification into the onset, but it does not. Its failure to retain lightness may lead us to the conclusion

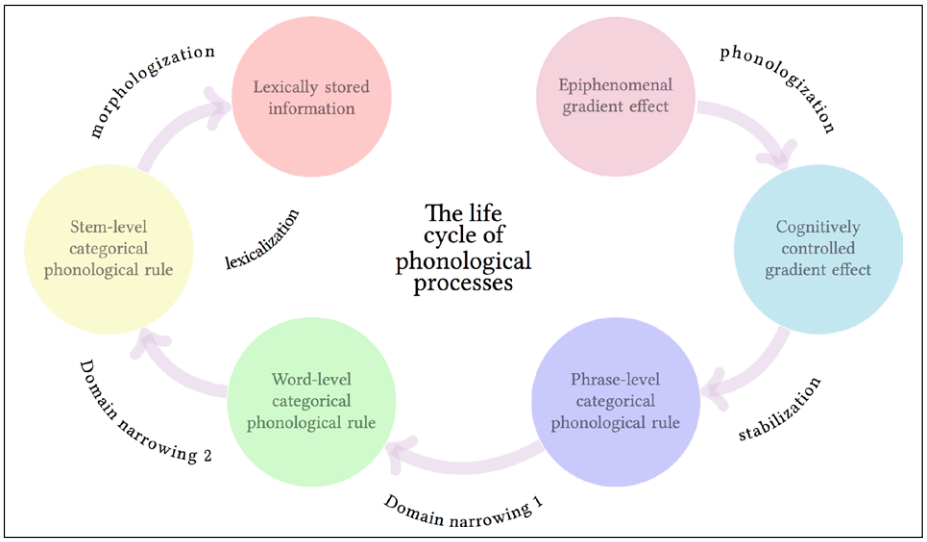

Figure 2: The life cycle of phonological processes (adapted from Ramsammy, 2015, p. 38). Arrows represent the possible diachronic trajectory of a phonological process.

Table 3: /l/-darkening in different morphosyntactic environments. Adapted from Bermúdez-Otero (2007).

	leap	*helix*	*peel-ing*	*heal it*	*peel*	
RP	[l]	[l]	[l]	[l]	[ɫ]	Cruttenden (2008)
Am. Eng. 1	[l]	[l]	[l]	[ɫ]	[ɫ]	Sproat and Fujimura (1993)
Am. Eng. 2	[l]	[l]	[ɫ]	[ɫ]	[ɫ]	Olive et al. (1993)
Am. Eng. 3	[l]	[ɫ]	[ɫ]	[ɫ]	[ɫ]	Hayes (2000)

that darkening shows overapplication in this environment. However, a modular approach under the life cycle can account for this realization by attributing the darkening process to the word level: Darkening occurs when /l/ is in the coda at the word level, and if the following segment is in the next word, it is irrelevant to its application, as a word-level process is only interested in the word itself. For Sproat and Fujimura's American English speakers, it seems as though the process of darkening has moved up from the phrase level to the word level, and /l/ darkens prior to resyllabification. The reanalysis of a process changing from phrase level to word level is easy to imagine. The child hears a word like *heal* in isolation with a dark [ɫ], and learns that word final /l/ is dark regardless of what follows.

The next stage of domain narrowing, from word to stem level, is represented by the third line of **Table 3** (American English 2) and is the pattern reported by Olive et al. (1993). For these speakers, darkening has advanced higher in the grammar: The /l/ darkens in the coda at the stem level, and so cannot see across the stem-suffix boundary. Initially, we may have wanted to group Hayes (2000) informants and Yuan and Liberman's speakers in with American English 2, but in fact they represent a more advanced stage, as they accept dark /l/ not only in *peel-ing*-type tokens but also in *helix*-type tokens. For these speakers, darkening has moved up to the stem level, before advancing through the prosodic hierarchy to non-foot-initial position (contrast *believe*-type tokens, in which the /l/ is stressed and not darkened for these speakers). Of course, the data from these two studies show variation, but nevertheless provide evidence that the next stage of the life cycle has been reached in some capacity. The link between the life cycle and the prosodic hierarchy is not explored further in the present paper, but for a full discussion see Bermúdez-Otero (2012) and Turton (2014a, p. 63–66).

One of the emergent benefits of the life cycle is that it can account for, or even predict, the presence of rule scattering, the situation whereby two diachronically related rules coexist in the synchronic grammar (Bermúdez-Otero, 2015). The opposite assumption to rule scattering would be that, when a gradient phonetic rule stabilizes into a categorical phonological one, the phonological rule fully replaces all trace of the phonetic process. An example will help to explain this idea in more detail. Zsiga's (1995) study of palatalization in English using electropalatography (EPG) demonstrates the process of rule scattering nicely. Zsiga considered the process whereby /s/ sounds more like [ʃ] in /s + j/ clusters in English, i.e., it becomes palatalized. She measured the articulation of /s + j/ clusters both word-internally, e.g., *pressure* and across word-bounies, e.g., *press you*. In both cases, there was audible palatalization present, but the articulations looked very different in the EPG palate traces. In the word-internal *pressure*-type tokens, the palate trace was identical to the [ʃ] in *shoe*. However, for the token which was palatalized across word boundaries in the *press you* condition, the trace was different, showing an articulation similar to an [s] overlaid with a [j]. That is, in the word internal condition, the articulators were moving

to a regular [ʃ], but across word-boundaries the articulators were still heading towards the separate [s] and [j] targets. This is an example of rule scattering: A word-internal rule which categorically palatalizes /s/ to [ʃ] resulting in featural change coexists in the same grammar as a process which creates gradient palatalization by gestural overlap. Again, the idea behind rule scattering is that when a rule enters a higher component of the grammar, it does not completely stop its application at the lower level. Innovative phonological rules do not replace the phonetic rules from which they emerge, but typically coexist with them. Although it is fair to say that speakers of present-day English simply have an underlying [ʃ] in *pressure*, this emerged diachronically from gradient palatalization which is still evident in the word final *press you* contexts.

We can make the same claim for /l/-darkening, as Sproat and Fujimura's data also provide evidence for rule scattering. The data in **Figure 1** are consistent with the idea of darkening originating as a duration-sensitive gradient phonetic process due to extra dorsal lag word-finally, the very phonological environment which is susceptible to pre-boundary lengthening. Over time, this becomes reanalyzed by learners as a categorical process, which is no longer sensitive to duration. Crucially, as explained above with palatalization, the original duration-sensitive gradient process of phonetic implementation coexists in the grammar as well as the newer categorical process.

Entangled with the ideas of the life cycle is evidence of lenition trajectories in different varieties of a language. A lenition trajectory refers to the output typology of a sound change which results in several allophonic variants of a phoneme at varying stages of consonantality. In English, dialects show evidence of lenition trajectories in all kinds of consonantal phonological processes. The trajectory of /r/ in present day Standard British English serves as an example of this kind of distribution. We find full consonantal /r/ word initially (e.g., *red*), fully vocalized /r/ phrase-finally (e.g., *far* [fɑ:], but word-final prevocalically we have an in-between variant, which is lenited but still retained (e.g., *far away* [fɑ:ɻweɪ]; see Bermúdez-Otero, 2011: 18 for a thorough overview of the articulatory and acoustic correlates of word-final prevocalic /r/ lenition). /r/-vocalization represents a more advanced stage of the trajectory than /r/-lenition. Diachronically, successive steps in a lenition trajectory give rise to a series of separate phonological rules entering the grammar one after the other. Synchronically, older rules affect milder types of lenition and have narrower cyclic domains, that is, they apply at higher levels, e.g., /r/ lenition applying at the word level. Conversely, the younger rules affect more drastic types of lenition, building further on the previous forms, e.g., /r/-vocalization applying at the phrase level.

There is plenty of evidence for the same kind of distribution for /l/, with dark variants representing the older, milder lenition in word-final prevocalic position and vocalized variants, representing the younger, harsher lenition, occurring phrase-finally. However, this is not as well documented, perhaps because vocalization is not as widespread, or it is not a feature of RP. In fact, many dialects of English are reported as having full-vocalization of /l/, usually post-vocalically, both in the UK (Johnson & Britain, 2007; Scobbie & Pouplier, 2010; Scobbie & Wrench, 2003; Tollfree, 1999) and the US (Ash, 1982; Hall-Lew & Fix, 2012). In England, /l/-vocalization is associated with working-class Cockney English (Wells, 1982: 313–315), or London English in general, although this realization seems to be spreading (Britain, 2009; Wright, 1989). For these varieties, /l/-vocalization is often described as being accompanied by strong labialization, so that /l/ sounds more like [ʊ] or [w]. Its presence in Scottish varieties is also well-documented (Scobbie & Wrench, 2003; Wrench & Scobbie, 2003).

The difference between a dark and vocalized /l/ can not always be sharply defined, with previous findings indicating that what is referred to as dark [ɫ] in American English

can often be vocalized, by articulatory standards, (i.e., there is no tongue-tip contact) especially in faster speaking rates, and when /l/ is followed a low vowel (Giles & Moll, 1975). This touches on the debates of categoricity vs. gradience: /l/-vocalization is treated as a categorical phenomenon by most sociolinguistic and phonological analyses, and as either categorical or gradient by phonetic ones. It is clear that the idea of having gradient and categorical factors of the same phonological process is not new, and for some phonological processes these combinatory factors are discussed in a non-controversial manner. However, it seems that for /l/-darkening, the presence of both categorical and gradient effects has not yet been considered or accepted. In contrast, such an approach to the distributional facts seems to be well documented in discussions of /l/-vocalization (Hardcastle & Barry, 1989; Kerswill & Wright, 1990; Scobbie & Pouplier, 2010; Scobbie & Wrench, 2003; Wrench & Scobbie, 2003; Wright, 1989). Perhaps this is because the articulatory difference between a light and vocalized /l/ is more distinct, or perhaps it is because vocalization has a clear definition that we can categorize as vocalized or not vocalized based on tongue-tip contact, whereas darkness is often judged in comparison with the typical light variant.

Thus far, it has been argued that previous studies have reduced the variation in /l/-darkening to a false dichotomy between purely categorical and purely gradient effects, when the evidence points to both operating within the grammar. This raises interesting questions for phonetics-phonology interactions. We are comfortable in assigning separate categories if the phonetics are suitably distinct. In the next section, diagnostics for how we might decide whether two sets of phonetic realizations are truly distinct are outlined.

3 Methodology

Speakers of RP, London, Manchester, Newcastle and Belfast English were recorded producing /l/ in the phonological contexts shown in **Table 4** (henceforth referred to by example token). The contexts were chosen to elicit a wide range of canonical onset and coda /l/s, in order to address the issue of categoricity and gradience, which cannot be reliably investigated with just two or three environments.

Subjects were recorded reading lists of words and phrases containing the target stimuli on a Mindray DP2200 ultrasound machine (frame rate 60 fps deinterlaced), with acoustics recorded through a Audio-Technica ATR-3350 microphone. A probe stabilization headset was used, as designed by Articulate Instruments, and tested at Queen Margaret University for reliability (Scobbie et al., 2008). /l/s were flanked by high front vowels in

Table 4: The ten phonological environments studied with example tokens.

Context	Example
initial	***l**eap*
foot-initial	*be**l**ieve*
suffix	*free-**l**y*
intervocalic	*he**l**ix*
morpheme boundary	*pee**l**-ing*
internal boundary	*pee**l**-instrument*
word-final internal boundary	*hea**l**#* V
word-final phrase boundary	*hea**l**#*, V
final	*pee**l***
word-final pre-consonantal	*pee**l** bananas*

all contexts, and for each context there were two sentences and five repetitions, resulting in a target of 100 /l/s per speaker.[3] The data were collected using Articulate Assistant Advanced (AAA; Articulate Instruments Ltd 2012). Sound files were exported out for acoustic segmentation in Praat, before being imported back to AAA where splines were hand-drawn for all frames within /l/ boundaries and flanking vowels. Spline coordinates (over 42 points) were extracted for contextual comparison and mean midpoint values plotted in R's ggplot2 package (Wickham, 2009). Midpoints only are analyzed in this paper, which are defined acoustically as the midpoint of the steady state of the /l/, as shown in **Figure 3**. Although not all tokens were as clear as **Figure 3** in exhibiting an obvious measurement point of preference, identifying relatively stable formants within the region of the liquid was always possible, as found by previous acoustic studies of laterals (Carter & Local, 2007; Kirkham, 2015; Nance, 2014).

Due to the frame rate of the ultrasound machine used for the present study, a temporal investigation proves problematic. Although between three and six frames are usually available for each token, imaging is often poor in the peripheral frames in the majority of speakers used in this dataset. It is rare that all of these can be reliably interpreted and rapid gestural movements may be missed by the coarse monitoring, resulting in a limited analysis. This is not to say that basing a temporal analysis on 3–6 frames would be problematic for all studies, and would certainly be possible given consistently clear imaging. Although a temporal analysis would be useful, as shown by studies such as Sproat and Fujimura (1993), the dorsal displacement alone suffices for many of the UK speakers in this experiment. This argument is revisited in Section 4.2, with support from temporal data for the RP speaker.

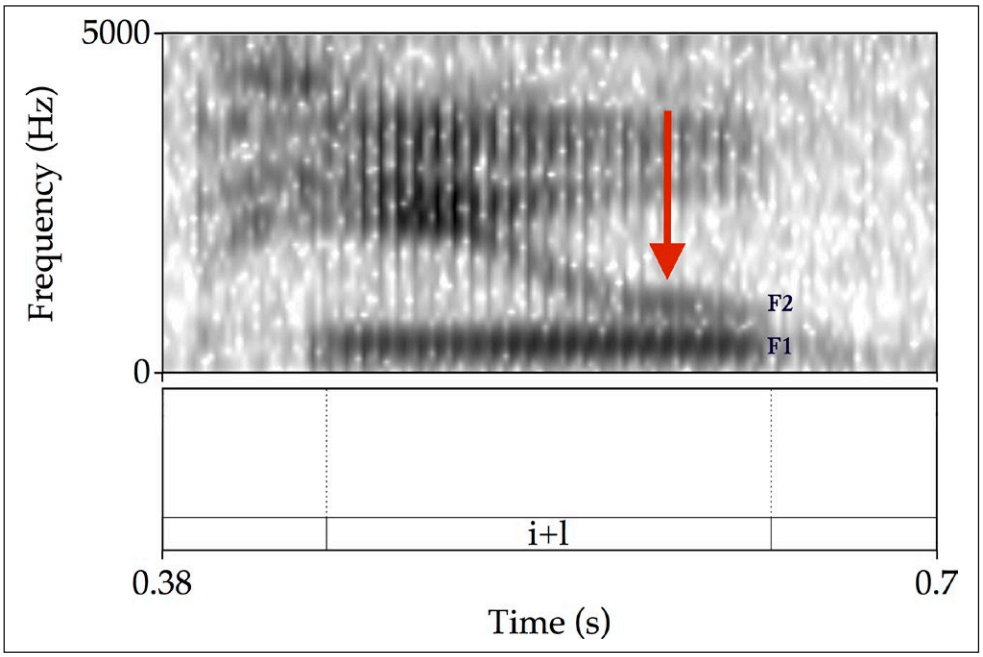

Figure 3: The acoustic midpoint of word-final /l/ for the RP speaker, defined as the midpoint of the steady state. This RP speaker shows a more abrupt transition between the [i] and [ɬ] than for most speakers, but it is still not clear where the boundary should be placed. Nevertheless, the midpoint of the steady state guarantees a definite measure of the lateral.

[3] The details of the sentences and repetitions can be found in Appendix A. Note that not all speakers reach the 100 token target due to poor quality imaging in a minority of elicitations. In some circumstances, particular sentences were dropped altogether for a speaker, as unexpected pronunciations resulted in the high front vowel flanking being breached.

Significant differences are shown through non-overlapping LOESS confidence intervals (similar to the SS ANOVA method used by Davidson 2006). Correlations with duration are performed using the rime duration. This is due to the difficulty of choosing a reliable segmentation point between /l/ and the preceding vowel, and to stay consistent with the Sproat and Fujimura study. Unlike Sproat and Fujimura (1993), the duration is log transformed here, to avoid skew in the scale.

The patterns in the data are supported by a Principal Components Analysis (henceforth PCA) of the spline coordinate data. PCA is a way of boiling down large amounts of data in order to extract the main correlates of variance, and has been used previously with articulatory analyses (Johnson, 2011; Stone, 2005). As AAA calculates spline coordinates across 42 fan points, the output gives up to 84 data points from the x and y coordinates. Conducting a PCA allows us to account for the main area (s) of variation, rather than attempting to deal with 84 separate values.

Figure 4 shows the loadings of the PCA computed on /l/ midpoints for the RP speaker. The black central line shows the average tongue contour for this speaker (tongue-tip on the right, tongue-root on the left), and the solid red lines show the maximum and minimum tokens of the first principal component (PC1). All splines fall within this range. Therefore, a high PC1 represents a light /l/ (advanced tongue body, higher tongue tip) and a low PC1 value represents a darker variant (retracted tongue root and lower tongue tip). The second principal component (PC2) fails to account for any meaningful variation in this dataset, as can be seen from the dashed blue lines in **Figure 4**. Whereas PC1 accounts for 89% of the variation here, PC2 covers just 5%. Therefore, PC2 is omitted from the analysis following the recommendation of Baayen (2008, p. 130), who advises that anything at 5% or below is not significant, as well as to discount any instances where there is a clear discontinuity in the variation accounted for. Although a separate PCA is run for each individual speaker, all speakers show this same pattern: High PC1s reflect the speaker's lightest /l/s, and PC2 is always non-significant.

The PCA used in the present study does not result in separate scales for height or backness (i.e., two PCs). See Turton (2014a: 107, 2015) for a full discussion of such issues, where it is argued that one PC can be preferable in terms of simplicity, and that assigning one quantitative value to a tongue spline that can be clearly interpreted is more valuable than, e.g., 9 PCs of ultrasound pixels where there is no concrete mapping from tongue to PC.

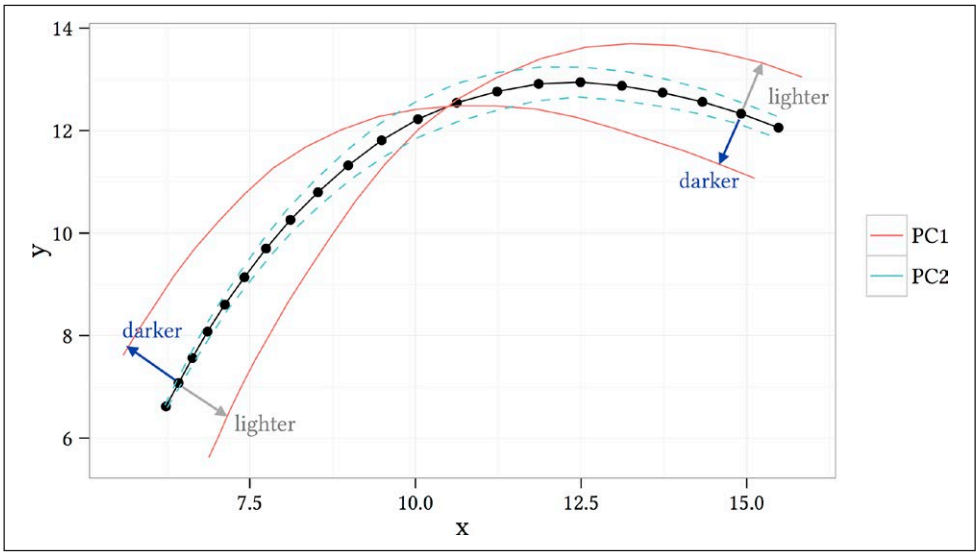

Figure 4: PCA loadings on midpoints /l/s for RP speaker.

The PCA results are also used to monitor the relationship between darkening and duration. Simple linear models are used to test this relationship, with regression lines plotted to demonstrate a significant regression equation for a particular speaker. Pearson's r is used to measure the strength of correlation between duration and darkness, and r^2 for goodness of fit (the closer to 1, the stronger the correlation or the better the fit).

As is typical in the presentation of ultrasound tongue imaging data visualization, the figures in Section 4 show the tongue-tip on the right side of the image, and the tongue-root on the left side. The images are not necessarily to scale. For visualization purposes, /l/s in word-final positions (i.e., contexts 6–10) are plotted with dashed lines to distinguish them from non-word-final /l/s, which are plotted with solid lines. This division is also found in the scatterplots, with non-word-final /l/ represented by solid circles, and word-final /l/ with hollow triangles. This divisional choice is purely for visualization and, although it often reflects an articulatory difference, it is not intended to reflect pre-assumed phonological categories.

3.1 Empirical diagnostics for categoricity

We have seen that existing studies which find evidence for gradience can sometimes overlook the possibility of categoricity. Although in some cases we may not be able to tell for absolute certain whether a phonological process is categorically conditioned in the speaker grammar, we can outline some diagnostics which we would hope to find in the empirical data from which categoricity can be inferred.

Previous work has touched on the debate of distinguishing between processes under phonetic or phonological control (e.g., Bermúdez-Otero, 2015; Cohn, 1993; Ellis & Hardcastle, 2002; Holst & Nolan, 1995; Kerswill, 1985; Myers, 2000; Nolan et al., 1996; Strycharczuk, 2012; Strycharczuk & Simon, 2013; Wright & Kerswill, 1989) which I will attempt to build on here. Following Turton (2015), I use the following diagnostics for categoricity:

1. Articulatory discontinuity between two (or more) sets of splines.
2. Articulatory consistency within these sets.
3. Statistical bimodality in the quantitative data.

Points 1 and 2 above are fairly straightforward, and could be based purely on qualitative observation of the spline data in many cases. However, by attaching a PC to each spline, the pattern can be observed on a quantitative scale, which makes for easier visualization. Point 3 is addressed by using Hartigan's dip test statistic (Hartigan & Hartigan, 1985), which is a measure of unimodality and bimodality in a given dataset. Using R's diptest package (Maechler, 2013), this test is used to investigate bimodality in the PC1 results. Hartigan's dip test works by assessing potential 'dips' in the distribution and outputs a p-value reflecting whether or not the dip is due to chance (cut-off for significance is $p < 0.05$). The dip test is conducted on the overall PC1 values.

4 Results

4.1 RP

The results from the RP speaker in **Figure 5** corroborate the descriptions in the existing literature for the variety (e.g., Cruttenden, 2008) in the distribution between phrase-level onset and coda /l/s. This speaker's onset /l/s display a typically light realization, with an advanced tongue tip and fronted tongue root. In contrast, the coda tokens show a retracted tongue root, lowered tongue body and reduced tongue tip gesture. Again, the right side of the spline plots show the front of the tongue and the left side the root of the tongue. Visually, the pattern supports categoricity by the first two diagnostics, and this pattern is accentuated by the boxplots in **Figure 6**. A significant Hartigan's diptest

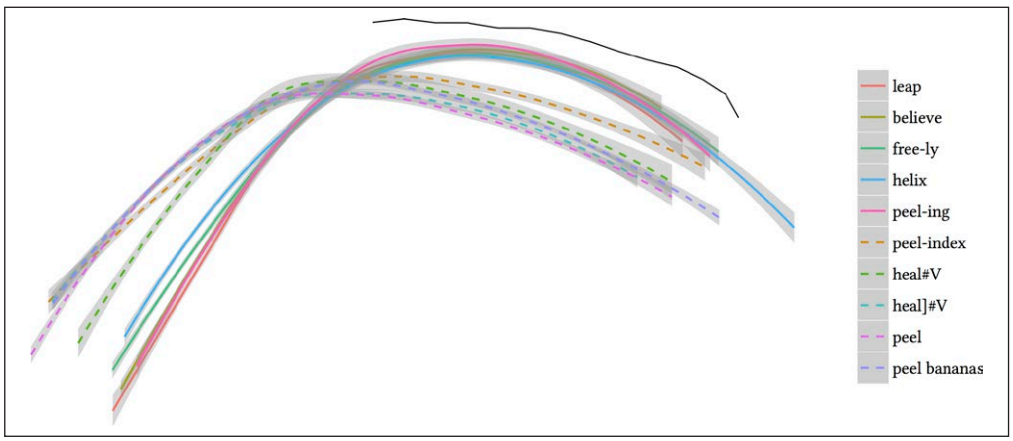

Figure 5: RP splines at /l/ midpoint across phonological context (see text for full description of spline patterns).

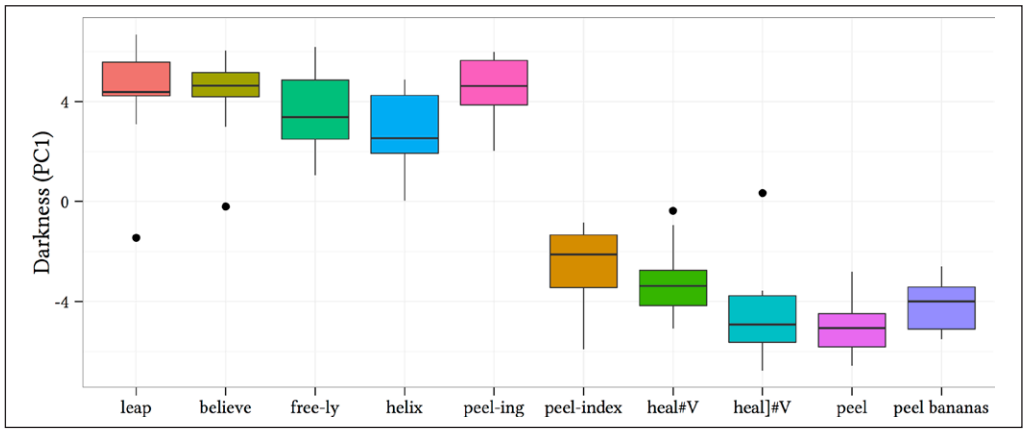

Figure 6: RP speaker's PC1 values across phonological context.

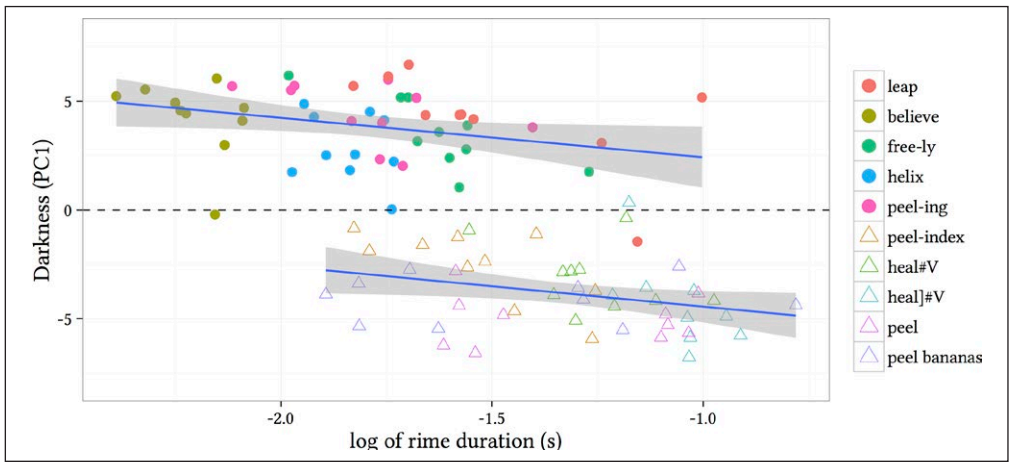

Figure 7: RP's speaker's PC1 values by log of rime duration.

confirms this observation statistically ($D = 0.08$; $p < 0.0001$), suggesting that the RP speaker has a categorical allophonic distinction between light and dark /l/.

Recall that some have claimed that duration is solely responsible for darkness, with longer /l/s being darker. **Figure 7** attempts to observe this relationship between darkness and duration in the RP data. Also recall that Yuan and Liberman (2009, 2011)

found a correlation for the dark variants only. For the RP speaker, we can observe a weak correlation across both light and dark categories ($r = 0.3$). Running a simple linear regression on the observed categories separately, to see if duration can be used to predict darkness (PC1), gives us significant regression equations but with a very low adjusted r^2 value, indicating that duration does not convincingly account for levels of darkness for this speaker (light: $F(1, 48) = 4.8$, $r^2 = 0.07$, $p = 0.03$; dark: $F(1, 48) = 4.96$, $r^2 = 0.08$, $p = 0.03$), although there seem to be some low-level effects.

4.2 A sidenote on temporal analyses

As discussed previously, a temporal investigation proves problematic for the dataset at hand, for most speaker data. However, data from the RP speaker suggest that a temporal analysis does not always represent the primary articulatory correlates for all varieties of English. As **Figure 8** of the RP speaker shows, regardless of gestural ordering, the splines during the course of the /l/ tend not to move much. The main facts reflect that the initial and final /l/s are an entirely different shape throughout the course of the articulation, and the tongue dorsum retraction remains stable. It is not the case that initial and final tokens consist of the the phasing of gestures temporally over the course of the /l/, at least not for this speaker. Of course, this is not to say that temporal information is not preferable. As Lin (2011) demonstrates, the phasing of the tongue tip and tongue dorsum gestures in American English is not only informative in production of /l/s, but also in perception, showing that synchrony between the two gestures actually aids listeners in their identification of onset /l/. This, alongside work such as Browman and Goldstein (1995) and Gick (2003), demonstrates that, for many varieties it may be extremely useful to have this kind of data. However, **Figure 8** suggests that the midpoint is perfectly suitable to indicate the articulation of the /l/ in RP. As far as we can see, there are no covert gestures that the midpoint measurement is missing. Be that as it may, future research into gestural phasing in RP and other varieties spoken in England is needed to reliably compare the temporal findings with American English.

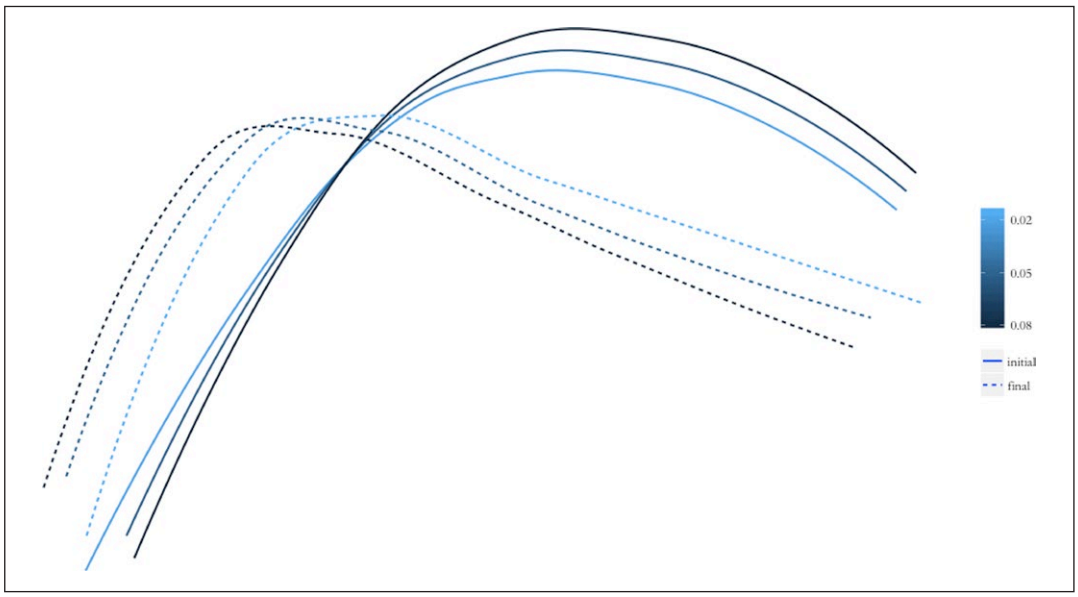

Figure 8: Tokens of initial and final RP splines over the course of the /l/. The lighter the spline, the earlier it was articulated during the course of the /l/.

4.3 London

Figure 9 shows the splines from the London informant. As with many varieties spoken in the South-East of England, London is well-known for having vocalized /l/s in word-final or phrase-final position. Ultrasound is not ideal for dealing with the articulatory differences between dark and vocalized /l/s, as it can not reliably monitor tongue-tip contact. However, this speaker's final /l/s are auditorily obviously vocalized in word-final position (i.e., one can unquestionably hear it in the sound files), making it easy to distinguish between light and lenited variants. As **Figure 9** shows, this speaker has the same distribution as RP, providing evidence for Diagnostics 1 and 2, albeit with a smaller magnitudinal difference between initial and final positions. One thing which may seem surprising from **Figure 9** is the small difference between the two allophonic distributions. There is some backing/raising of the tongue dorsum in final position, but only a small amount. The distinction between the light and dark /l/s in the RP speaker in **Figure 5** is greater than the distinction between light and vocalized for the London speaker. One might expect a vocalized /l/ to be more articulatorily backed than a regular dark realization, resulting in a bigger difference between initial and final /l/ for a vocalizer than a mere darkener. This has been found in other studies, such as Wrench and Scobbie (2003), who found that their 'impressionistically obvious' vocalizer (also from South-East England) had the most extreme magnitude difference than the other speakers. However, they were considering tongue tip height only in this EMA study, the part of the tongue which gives the worst and most unreliable image in ultrasound data. It poses a question for future work on /l/-vocalization, then, to discover whether the backing of the tongue is an articulatory correlate of vocalization at all, or whether it is merely small movements of the tongue tip which creates the difference between that and a regular dark [ɬ]. The difficulty in distinguishing between dark and vocalized /l/ on the spectrogram (Hall-Lew & Fix, 2012) might suggest the latter. However, it is not just the tongue tip realization which defines the difference between a dark and vocalized /l/. Firstly, London vocalized /l/ does not just lose its alveolar contact, but is accompanied by lip rounding (Wells, 1982: 95). In addition, vocalized variants may vary from dark ones in terms of lateral bracing of the sides of the tongue.[4] Although British vocalized variants have been described as retaining some form of narrowing of the tongue (Ladefoged & Maddieson, 1996, p. 193) or a 'saddle' type shape (Wrench & Scobbie, 2003), it is not clear to what extent this bracing compares to an onset /l/, and whether or not the sides of the tongue would be lower in such cases.

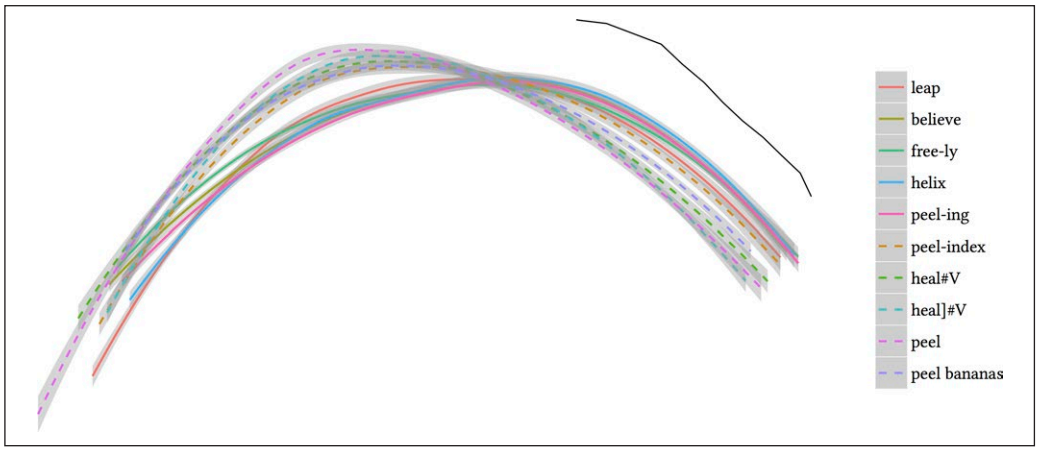

Figure 9: London speaker's splines at /l/ midpoint across phonological context.

[4] With thanks to an anonymous reviewer for pointing this out.

Although beyond the scope of this study, these additional factors demonstrate the need for an all-encompassing approach analysis of /l/, accounting for both mid-sagittal and coronal planes, as well as lip-rounding.

Nevertheless, it is difficult to compare magnitude in this way between speakers with ultrasound tongue imaging. It raises the issue of inter-speaker comparison and normalization in ultrasound studies and other articulatory work. The PCA works well for comparing intra-speaker phonological contexts, but for comparing the extent of extreme magnitude between, for example, the London speaker with the RP speaker it is not the best method. Differences in vocal tract size, as well as position of the probe varying from subject to subject mean that the two raw datasets are not comparable. Despite the perceived smaller difference for the London speaker, the boxplots in **Figure 10** show that the two distributions are not overlapping. The PCA in **Figure 10** provides added confidence to Diagnostics 1 and 2, as does auditory judgement. A significant Hartigan's dip statistic ($D = 0.06$; $p = 0.02$) confirms the observed pattern quantitatively: There seems to be a bimodal distribution in the London speaker's /l/s.

Finally, considering durational effects on London English /l/, **Figure 11** shows a similar effect to the Sproat and Fujimura data revisualized in **Figure 1** (note that high tip delay means a darker variant in **Figure 1**, whereas a high PC1 is a lighter variant, so the y axis here is reversed in comparison to **Figure 1**). Light /l/s are somewhat clustered together at the top ($r = -0.28$), whereas darker variants on the bottom panel show a strong correlation

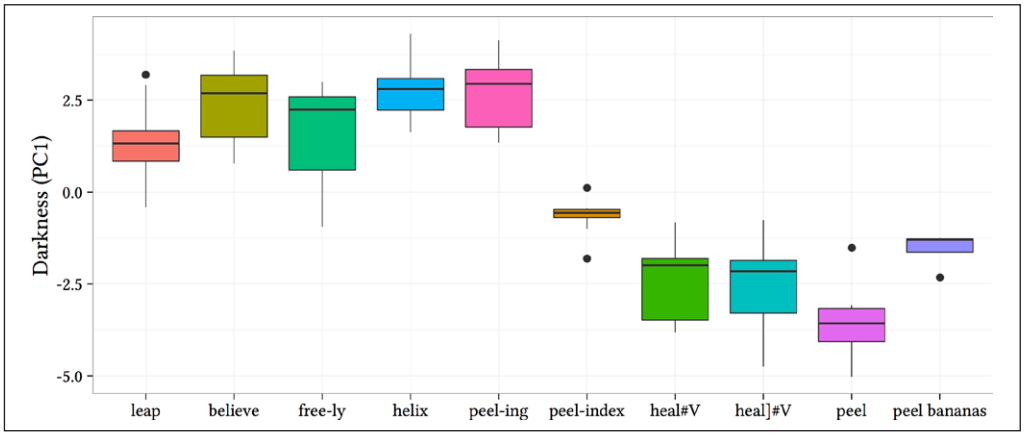

Figure 10: London speaker's PC1 values across phonological context.

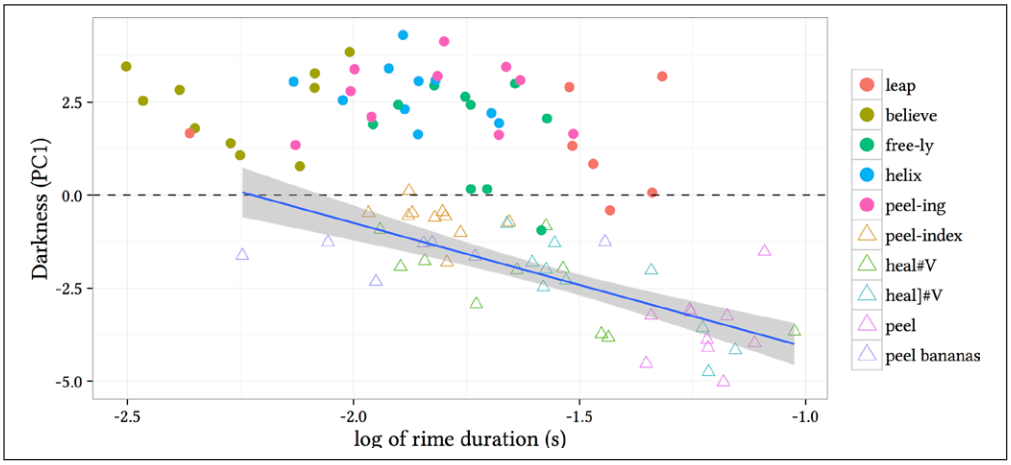

Figure 11: London speaker's PC1 values by log of rime duration.

with duration (r = –0.75). Simple linear regressions of the light and dark categories separately confirm this further, with only the dark model producing a significant regression equation and a good r^2 fit (light: $F(1, 45) = 2.87$, $r^2 = 0.04$, $p = 0.09$; dark: $F(1, 45) = 55.2$, $r^2 = 0.54$, $p < 0.0001$). Duration is a significant predictor for the dark model only.

This speaker provides strong evidence that categorical allophonic distributions can be overlaid by gradient phonetic effects in the same grammar. The London speaker has a categorical distinction between word level onset and coda /l/s, with an extra gradient durational effect overlaid for the coda /l/s. It seems that the London speaker shows a much stronger influence of phonetic duration than found for RP, which is unsurprising given RP's conservative nature. It is clear that London is further advanced on the /l/ lenition trajectory (i.e., it has vocalized /l/s), and so one might expect a more conspicuous pattern of phonologization of such gradient phonetic effects in this dialect.

4.4 *Manchester*

Manchester English was originally selected as a variety of interest due to its descriptions in the existing literature, with the general claim being that /l/s are dark in all positions (Cruttenden, 2008; Kelly & Local, 1986). There is acoustic evidence that initial and final /l/s have a small but significant difference in quality, but generally it seems that even word initial /l/s are acoustically dark (Carter, 2002; Kelly & Local, 1986). We could deduce from these descriptions that Mancunians have just one category of /l/, and the splines in **Figure 12** would seem to corroborate this.

The PCA in **Figure 13** allows us to observe this distribution from a quantitative perspective. One immediate observation is that this speaker does not show the linearly trend like RP and London. The *freely*-type tokens are a little out of line here, which is very likely due to this speaker's very lax pronunciation of the *happ*Y vowel. Studies of Manchester English have shown that speakers exhibit an opposite pattern to the widely reported phenomenon of *happ*Y-tensing found in the South of England, whereby the *happ*Y vowel lowers and backs in phrase-final position to become more like [ɛ̈] (Baranowski & Turton, 2015, 2016; Turton & Ramsammy, 2012). That means, for this speaker, /l/ is not flanked by two high front vowels in this position and the word was a poor choice of example. Kirkham and Wormald (2015) also found this in their study of /l/ in nearby Sheffield. Overall, visual inspection of **Figure 13** is probably sufficient to deduce that this distribution is not bimodal, but nevertheless this is confirmed by the dip test (D = 0.02; $p = 0.97$).

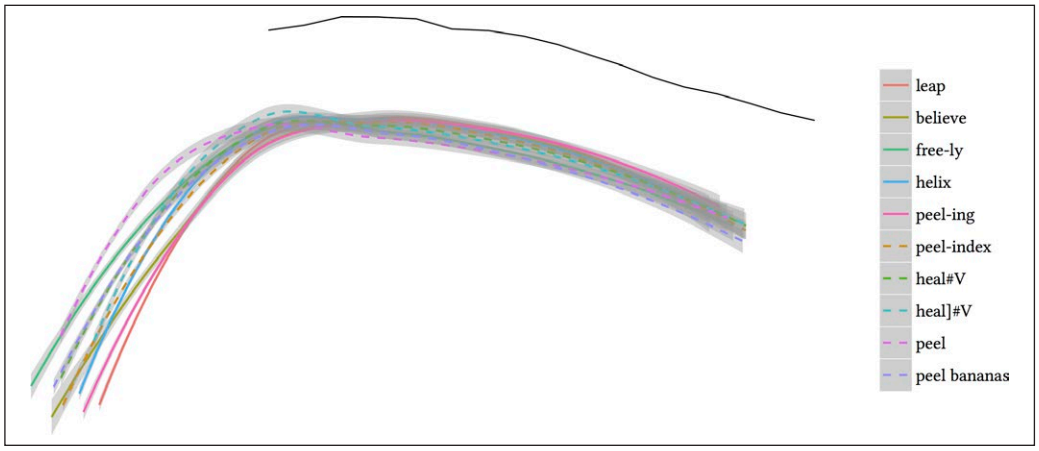

Figure 12: Manchester splines at /l/ midpoint across phonological context.

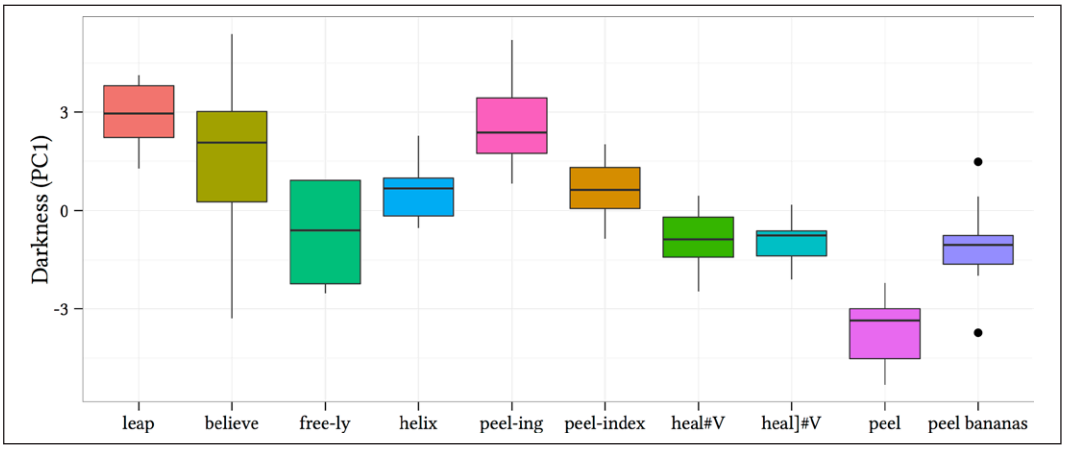

Figure 13: Manchester speaker's PC1 values across phonological context.

Despite the lack of an obvious initial vs. final distinction in this variety, there is one phonological context which seems to show consistent backing/darkening compared to all others. Note the small difference and the lack of confidence interval overlap in tongue-root backing between phrase-final *peel*-type tokens and the other contexts in **Figures 12** and **13**. **Figure 14** shows the individual splines for each of the 10 *leap*-type and 10 *peel*-type tokens in more detail. The extra tongue root retraction in the *peel*-type tokens is convincing and consistent, and shows no overlapping confidence intervals with the initial tokens. However, this small but significant difference cannot be seen to the same extent for the *peel bananas*-type tokens, demonstrating this is not just a phrase-level coda darkening process. Although the London speaker showed a somewhat small (but significant) difference between initial and final /l/s, the other phonological environments patterned with the expected extreme depending on whether the /l/ was in a typical onset or coda position. With the Manchester speaker, the distribution does not match this: Word-initial and word-final /l/s are significantly different, but the other phonological environments pattern somewhere in between these two extremes.

The data for London and to a lesser extent RP, also show that phrase final *peel*-type tokens have a lower PC1 than the *peel bananas*-type tokens, i.e., they are darker. This makes sense from a durational point of view, given pre-boundary lengthening, the long-observed phenomenon that strong prosodic boundaries result in lengthened realizations (Lehiste, 1980, p. 7). In contrast, the /l/ in *peel bananas* needs to move on to the next segment, and will not have as much articulation time for the cumbersome dorsal gesture to reach its full potential. If a categorical allophonic distinction for the Manchester speaker is unconvincing, given all of the other coda /l/ environments do not pattern with phrase final *peel*-type tokens, then perhaps duration could be causing the extra tongue root retraction. **Figure 15** shows darkness against the duration of the rime, with smoothers added to the word-final contexts where we would expect darkening. The only context which shows a strong correlation with duration is for the *peel*-type tokens ($r = 0.9$), as seen by the slope and the narrow confidence intervals. Running a simple linear regression across the whole sample (as it would not make sense to impose two categories for this speaker) gives us a significant regression equation of rime duration on PC1 ($F(1, 96) = 6.39$, $r^2 = 0.05$, $p = 0.01$), but this significance drops out as soon as the *peel*-type tokens are removed from the dataset, suggested that they alone are carrying this pattern. For this speaker, it would seem that there is only one category for /l/, and phrase final tokens show the added gradient phonetic effect of duration. Future investigation to determine this for certain includes temporal analysis and also the interaction of /l/-darkening with other

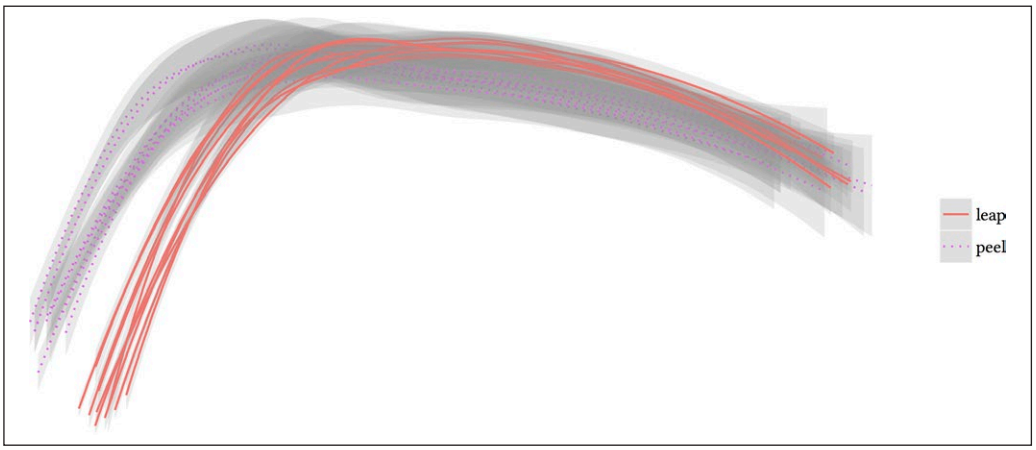

Figure 14: Manchester speaker's individual splines at /l/ midpoint in initial and final positions.

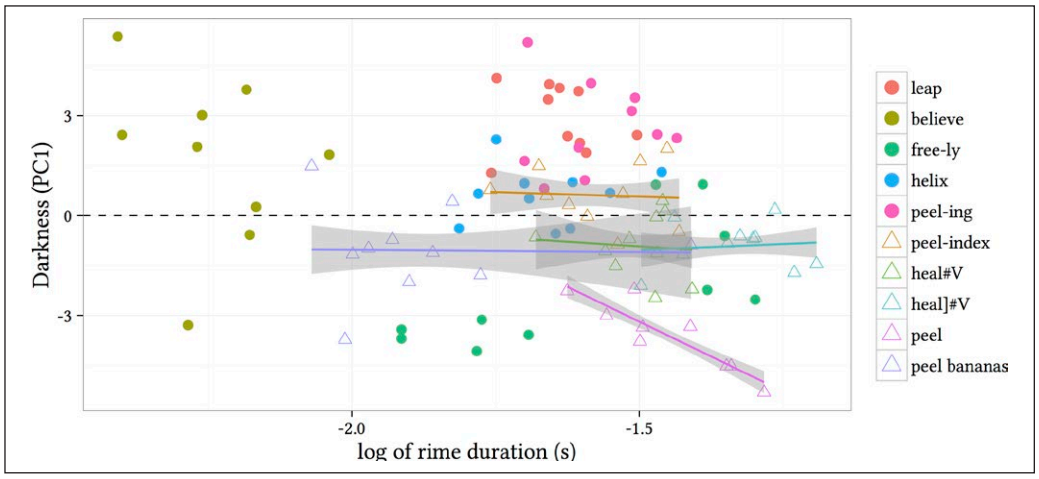

Figure 15: Manchester speaker's PC1 values by log of rime duration.

processes found in Manchester, such as /u/-fronting. Turton and Baranowski (2015) show how /u/ fronts before [ɬ] in Manchester, a context in which fronting is prohibited in most varieties of English, arguing that this is potentially evidence for only one category of /l/. Perhaps more in-depth temporal analysis would reveal categorical patterns of gestural phasing, as found for some American English varieties which also have acoustically dark [ɬ]s everywhere. However, for now, we cannot claim a categorical allophonic pattern for this speaker.

4.5 Newcastle

In opposition to descriptions of North-West varieties of English such as Manchester, the North-East of England (where Newcastle is located) is said to have light /l/s in all phonological contexts. **Figure 16** shows the midpoint splines for the Newcastle speaker. Although the splines are very tightly clustered, it is possible to observe a small but distinct split between certain environments. Close inspection of the splines shows that the first five phonological environments are not significantly different from one another in terms of confidence interval overlap, and they pattern together as the lightest variants. The final three phonological environments are also not significantly different from one another and pattern together as the darkest variants. This may be difficult to observe in **Figure 16**, but the combined box and jitter plot in **Figure 17** accentuates this pattern.

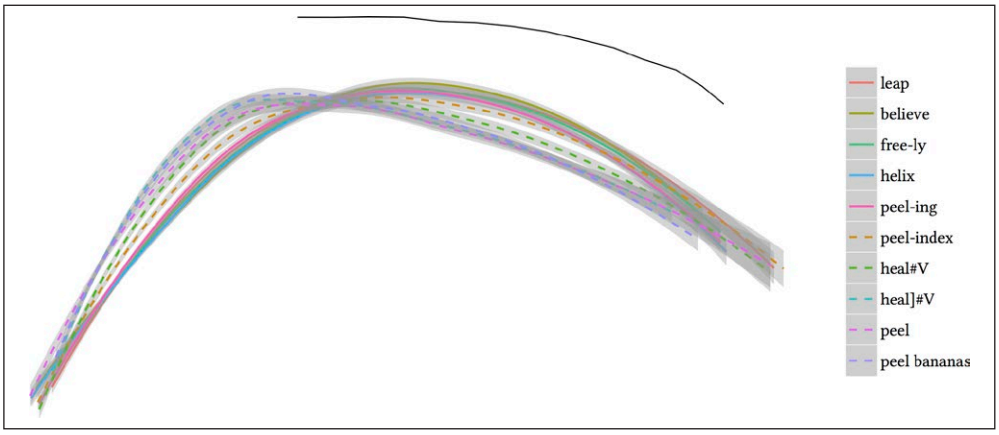

Figure 16: Newcastle splines at /l/ midpoint across phonological context.

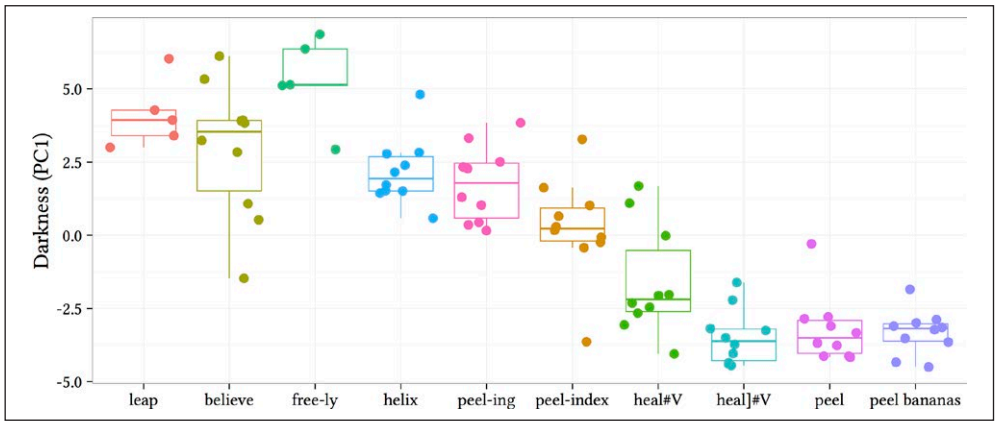

Figure 17: Jitter plot of Newcastle PC1 values across phonological context.

The visualization of the data in **Figure 17** shows the intermediate nature of the middle phonological environments for this speaker, where /l/ is word-final and prevocalic. Initially, this seems as though we cannot claim categoricity based on Diagnostic 1, as there is not an articulatory disconnect, which may indicate that Newcastle /l/s are indeed all similar. However, the question of variance vs. gradience arises when we observe the pattern more closely through the placement of the individual points. Rather than these two phonological contexts showing an intermediate realization, the in-between nature is more down to variation, that is, the speaker uses light variants some of the time and darker variants at other times and the mean value makes it look as though it is intermediate. Thus far, speakers have been consistent within a phonological environment, but the Newcastle speaker is more variable. For example, in *healVP*-type sentences, such as *I sent Neil interesting emails*, the Newcastle speaker would sometimes produce a short prosodic break before the noun phrase, and other times would not. As shown by Cho et al. (2014), a prosodic boundary prevents the resyllabification of a final consonant across word boundaries, and Lin's (2011) ultrasound investigation also demonstrates that prosodic conditioning has consequences for /l/ realization. It was not clear in advance how speakers would manipulate prosodic boundaries in these sentences, meaning that the potential resyllabification of an /l/ into the onset of the following word was unpredictable. Either way, the pattern is difficult to diagnose. The evidence for categoricity is weak, but it does not look like a complete lack of categories either. Hartigan's dip statistic just misses the cut-off for significance ($D = 0.05$; $p = 0.07$), confirming that this case is very much borderline, but also that we cannot make clear claims of categoricity.

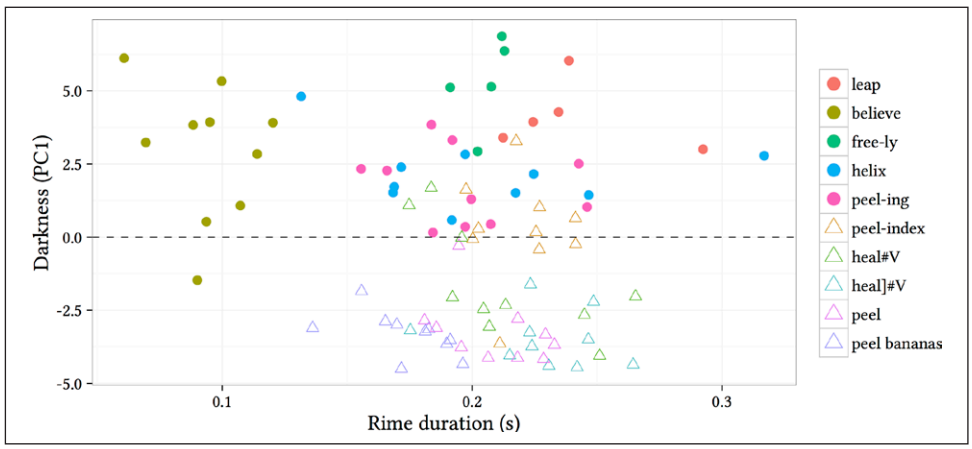

Figure 18: Newcastle PC1 values against rime duration (note that the peel-index type tokens pattern more with the circles for this speaker).

The relationship between duration and darkness for this speaker is visualized in **Figure 18**. Correlations are close to zero for both assumed categories[5] (light $r = 0$, dark $r = -0.05$). The linear models are non-significant, and show a very poor fit, with r^2 values approaching zero (light: $F\,(1, 38) = 0.043$, $r^2 = -0.03$, $p = 0.84$; dark: $F\,(1, 48) = 0.05$, $r^2 = -0.02$, $p = 0.81$). Again, a lack of clear categories in this variety is accompanied by a lack of a durational effect. We have seen that Newcastle /l/s are all very similar, but not completely the same. Following Scobbie's Caveat (after Bermúdez-Otero, 2010; Scobbie, 1995), if phonological processes need not be neutralizing, then evidence of gradience is not evidence of absence of a categorical effect. This speaker represents a case whereby a temporal analysis would be useful in order to make any further claims.

4.6 Belfast

Moving to another reportedly 'all light' dialect, we turn to an Irish English variety, that of Belfast English. Irish English /l/ has been described as "strikingly light in all positions" (Wells, 1982, p. 431). This description is common in the existing literature (Hickey, 2005: 272; Hughes, Trudgill, and Watt, 2012), although some sociolinguistic work has suggested this might be changing for speakers in Northern Ireland (McCafferty, 1999) where Belfast is located (although McCafferty's reports were for Derry). However, the evidence here suggests that Belfast English does seem to show next to no variation in /l/ depending on phonological environment, as **Figure 19** shows (indeed, **Figure 19** makes the Newcastle splines look allophonic in contrast).

The token *believe*, where the /l/ occurs in foot-initial position, has a slightly lowered tongue body and is slightly backer than the others. This can be seen more clearly in **Figure 20**. For all of the other speakers in the experiment, this token stands as one of the lightest of all contexts, so it is unusual that the speaker with the consistent light [l] pattern would have a darker realization in this prosodically strong position. However, on listening to the tokens and comparing the acoustics, it is clear that this token is subject to pre-stress contraction (Zwicky, 1972: 283). Often justified by fast speech rates, this process gives [bl̥i:v] for *believe*, with no audible vowel and a darker realization of the /l/ (Huffman, 1997: 118). These /l/s are also much shorter in duration (see **Figure 21** below), as expected from cluster consonants (Lehiste 1980: 18).

[5] Note that the cut-off between light and dark has been placed after the first six environments for this speaker. It is acknowledged that the treatment of morphological and prosodic boundaries may vary from dialect to dialect (Bermúdez-Otero, 2011; Turton, 2014a). Rather than impose 'categories' of light and dark, this paper seeks to derive possible splits in category realization in a data-informed way. If this speaker has any categories at all, the line *would* be drawn after the first six contexts.

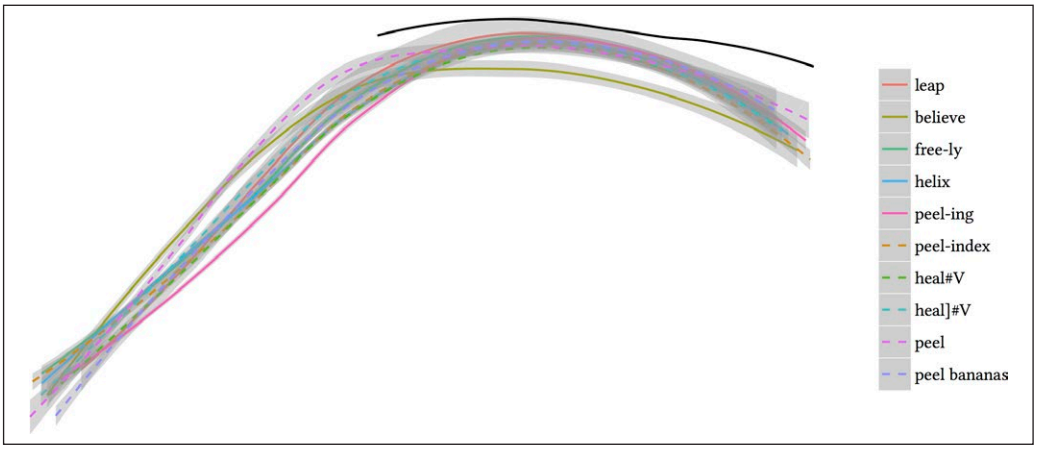

Figure 19: Belfast splines at /l/ midpoint across phonological context.

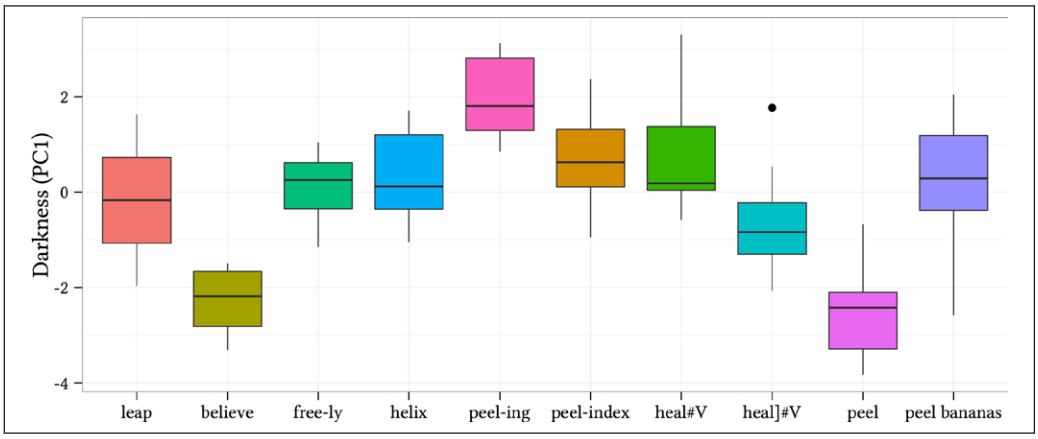

Figure 20: Belfast speaker's PC1 values across phonological context.

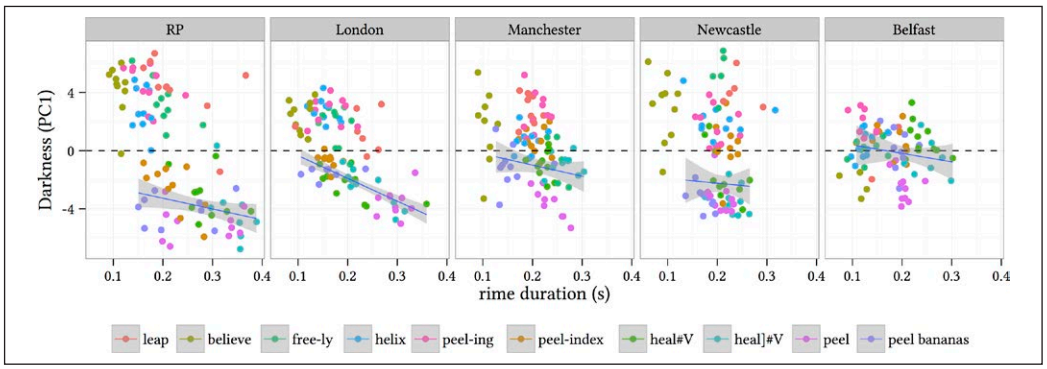

Figure 21: PC1 against rime duration for all speakers.

There is no evidence of a more distinct light/dark pattern, although phrase-final *peel*-type tokens do show significantly more retraction. This can be best observed in the **Figure 20** PCA plot, as the number of contexts in the spline plot makes it difficult to spot. This is unsurprising, considering we have seen this extra retraction in phrase-final tokens for almost all speakers, but in particular it is the same distribution found for the 'all dark' Manchester speaker. Hartigan's dip statistic is non-significant ($D = 0.04$; $p = 0.66$) as expected, confirming the lack of bimodality. The Belfast pattern once again provides support for the idea that /l/-darkening originates as a side-effect of duration which may exists in all dialects, whether they have a categorical distinction emerging from this or not.

Due to an overall absence of variability in the Belfast data, the PCA here is arguably not as meaningful when conducting an intra-speaker analysis. Duration shows no significant effect on darkness across the board ($F(1, 81) = 2.28$, $r^2 = 0.02$, $p = 0.14$), but it is possible we are missing something because of the reduced power in the PCA on midpoint splines. This highlights the issue of running the PCA on tongue contour data, and for future research more work is needed to take account of the entire ultrasound image through pixel PCA or velocity measurement, as utilized in more recent articulatory studies (Carignan et al., 2016; Moisik, 2013; Pouplier & Hoole, 2013; Strycharczuk & Scobbie, 2015). Here we are just making intra-/l/ comparisons, and if there is only a small amount of variation, the PCs of the tongue midpoints provide little information. Thus, rather than interpreting the Belfast rime duration data in isolation, it seems it will be more useful to view it in comparison with all speakers, which is the subject of the next section.

4.7 Comparing speakers

Figure 21 shows PC1 values plotted against rime duration across all five speakers in this experiment, in order to exemplify cross-dialectal differences. The final five environments (or final four for Newcastle) have been fit with a simple linear regression line to show the relationship with duration (the first five environments, or 'light' environments were not included as there was no significant effect of duration for these tokens for any speaker). **Figure 21** shows that tokens from Belfast, Manchester and Newcastle have a smaller durational range than in RP and London and thus the fitted lines do not look as informative. The pattern here is consistent with what we would expect if duration effects are best observed in varieties which have an obvious allophonic distinction between two categories. This provides further argumentation against duration being intrinsically linked to phonetic darkness, and instead provides support for a situation where extra durational effects occur on top of existing phonological patterns.

Table 5 summarizes the evidence in the data for a categorical split between light and dark allophones by considering the three categoricity diagnostics outlined in Section 3.1: Articulatory disconnect between two sets of splines (or PC values), continuity within those sets, and a significant bimodal statistic for the dip test. Evidence of gradience in the data is assessed through whether or not duration is a significant factor in the linear models. Overall, **Table 5** shows that the presence of two allophones of /l/ in English is not uniform across varieties. As has been reported for some studies of American English, we have varieties with no evidence of a categorical distinction for these midpoint splines. However, we have some varieties with clear patterns of categoricity. What is interesting here is that, in

Table 5: Indicators of categoricity and gradience in the data.

	RP	London	Manchester	Newcastle	Belfast
indicators of categoricity					
1. articulatory disconnect	X	X	×*	?	×
2. continuity within sets	X	X	×*	?	×
3. bimodal dip test	X	X	×	×	×
indicators of gradience					
4. Significant durational effect?	X (weak)	for [ɬ]	×*	×	×

* Categorical effects hold between ***peel***-type tokens and everything else, but not the traditional light/dark environments.

such varieties, when duration plays a strong role in predicting darkness, this is only for the allophonically dark (or vocalized) tokens. Arguably, this can even be seen in Manchester, where our speaker has a distinction for phrase-final *peel*-type tokens only.

These patterns are consistent with a situation where /l/-darkening begins as a gradient side-effect of duration occurring only phrase-finally. Articulatorily, this makes sense, as the dorsal gesture has enough time to fulfil its maximum. It could be that the Manchester speaker sits at this level. This may be phonologized (perhaps by the next generation of speakers) at the phrase level over time, showing a relationship with duration: Longer variants are darker. Over time, this darker realization may be stabilized, resulting in a reanalysis of the phonetically dark articulation as a separate allophone. This may be applied to all coda realizations of /l/, and is the situation we observe for the London speaker. In this particular case, the coda /l/ becomes lenited even further, losing its tongue tip gesture and moving further along the lenition trajectory. Although not discussed at length in this paper, we know from various studies of darkening that the process can become word level, meaning that /l/ darkens in the coda at the word level, even if a vowel follows, giving a dark [ɫ] in *heal it* (see Turton, 2014b, 2016 for a full discussion). Reports from other varieties (e.g., Boersma and Hayes, 2001; Hayes, 2000; Olive, et al. 1993; Turton, 2014a) show evidence that the process can become conditioned morphologically, resulting in dark tokens in words like *healing* where /l/ precedes a stem-suffix boundary, but not *helix* where the /l/ is monomorphemic. Such varieties demonstrate that darkening can move up from being a word-level process to a stem-level process, in line with the predictions of the life cycle of phonological processes.

5 Conclusion

/l/-darkening in English, and its subsequently related processes of lenition, shows effects of categorical allophony, phonetic gradience, dialectal variation, and morphosyntactic sensitivity. The current investigation has shown evidence of some of these effects across speakers from different dialects of English. Accordingly, the analyses conducted using ultrasound tongue imaging contribute to several ongoing debates in phonology, phonetics, and language variation and change. It has been shown that previous approaches to /l/ allophony have underestimated the diversity of the phenomenon and that these patterns show a great deal of orderliness if considered from the viewpoint of the life cycle of phonological processes. We need a theory that can account for the evidence that categorical darkening domains may differ in size between dialects. The effects of /l/-vocalization, in addition, may coexist as a lenition process with darkening or replace it altogether. Modular theories such as the life cycle of phonological processes can make sense of such facts, with rule scattering and domain narrowing accounting for the coexistence of categorical and gradient effects, as well as lenition trajectories accounting for the presence of /l/-vocalization alongside darkening. A modular approach such as the life cycle predicts the existence of synchronically overlaid gradient and categorical effects, where the latter arise diachronically from the former by stabilization. The data here provide evidence for this, in line with the findings of Yuan and Liberman (2009, 2011) and against the arguments made by Sproat and Fujimura (1993). The life cycle can also account for existence of dialectal microtypologies within a language providing and additional explanatory layer to a basic modular approach, as we would expect different varieties to be at different stages in the series of micro-level sound changes (Ramsammy, 2015).

The results also challenge arguments about morphology-phonetics interactions based on the assumption that /l/-darkening is purely gradient. In addition, varieties which do not provide compelling evidence for categorical allophony also fail to show a clear effect of duration (as found for some studies of American English, such as Lee-Kim,

et al. 2013) showing that duration cannot account for darkness in all varieties, even in the potential absence of categories. Further work in variationist linguistics is required to test to what extent such speaker patterns are representative of a variety as a whole, or whether inter-speaker variation in terms of both categorical and fine-grained variation can be more idiolectal (we know, for example, that speakers of the same variety show different strategies of /r/ articulation; Delattre & Freeman, 1968; Lawson et al., 2008; Mielke et al., 2010). In addition, temporal dynamics of /l/ patterns may give us further insight into intra and inter-dialectal variation. Nevertheless, the patterns found for RP, Manchester, London and Belfast reflect the descriptive sociolinguistic literature very accurately, suggesting that the general picture presented by the present paper is a good representation of these varieties.

Overall, the phonetics-phonology interface effects in /l/-darkening patterns have thus far been reduced to a false dichotomy between either a purely categorical, or purely gradient approach, when in fact, both can exist within the same grammar. As a result, the debates surrounding whether darkening is categorical or gradient are not always fully informed. Moreover, phonetic studies dismissing the presence of categoricity may miss the opportunity to observe such patterns if a wide range of phonological contexts are not taken into consideration. The wide range of dialectal diversity, for which this paper provides only a small subset, shows a great deal of orderliness if considered from the viewpoint of the life cycle.

Acknowledgements

This research was funded by the UK Arts and Humanities Research Council grant number AH/I015108/1 and would not have been possible without the speakers who volunteered to participate. With thanks to Ricardo Bermúdez-Otero, Silke Hamann, and an anonymous reviewer for their extensive comments and insightful suggestions, as well as Laurel MacKenzie and Jim Scobbie for lengthy discussion and debate. I am grateful to Maciej Baranowski, Josef Fruehwald, Michaela Hejna, Yuni Kim, Jeff Mielke, Michael Ramsammy, and Patrycja Strycharczuk for discussion and guidance. I would also like to thank audiences at the 21st Manchester Phonology Meeting, BAAP, 2014 and NWAV 20 for their excellent questions and feedback.

Competing Interests

The author has no competing interests to declare.

References

Articulate Instruments Ltd. 2012. *Articulate Assistant Advanced Ultrasound Module User Guide, Version 2.14*. Articulate Instruments Ltd.

Ash, S. 1982. *The vocalization of /l/ in Philadelphia*. Doctoral dissertation, University of Pennsylvania.

Baayen, H. 2008. *Analyzing Linguistic Data: A Practical Introduction to Statistics using R*. Cambridge: CUP. DOI: https://doi.org/10.1017/CBO9780511801686

Baranowski, M. and Turton, D. 2015. Manchester English. In: Hickey, R. (Ed.) *Researching Northern Englishes*. Amsterdam and Philadelphia: John Benjamins. DOI: https://doi.org/10.1075/veaw.g55.13bar

Baranowski, M. and Turton, D. 2016. *Unstressed vowel backing in Manchester English*. Paper given at the 7th Northern English Workshop, Edinburgh, 14th April 2016.

Bermúdez-Otero, R. 1999. *Constraint interaction in language change: quantity in English and Germanic [Opacity and globality in phonological change]*. Doctoral dissertation, University of Manchester and Universidad de Santiago de Compostela.

Bermúdez-Otero, R. 2007. Diachronic phonology. In: de Lacy, P. (Ed.), *The Cambridge handbook of phonology* (p. 497–517). Cambridge: Cambridge University Press, Diachronic phonology. DOI: https://doi.org/10.1017/CBO9780511486371.022

Bermúdez-Otero, R. 2010, 2 December 2010. *Morphologically conditioned phonetics? Not proven.* Paper presented at On Linguistic Interfaces II, Belfast.

Bermúdez-Otero, R. 2011. Cyclicity. In: Oostendorp, M. v, Ewen, C. J, Hume, E. and Rice, K. (Eds.), The Blackwell companion to phonology (Vol. 4, p. 2019–2048). Malden, MA: Wiley-Blackwell.

Bermúdez-Otero, R. 2012. *Traces of change in synchronic phonology: English syllabification and the life cycle of lenition.* Workshop at IV Seminário Internacional de Fonologia, Porto Alegre, 26 April 2012.

Bermúdez-Otero, R. 2015. Amphichronic explanation and the life cycle of phonological processes. In: Honeybone, P. and Salmons, J. C. (Eds.), *The Oxford handbook of historical phonology* (p. 374–399). Oxford: Oxford University Press.

Bermúdez-Otero, R. and Trousdale, G. 2012. Cycles and continua: on unidirectionality and gradualness in language change. In: Nevalainen, T. and Traugott, E. C. (Eds.), *The Oxford handbook of the history of English* (p. 691–720). Oxford: Oxford University Press. DOI: https://doi.org/10.1093/oxfordhb/9780199922765.013.0059

Boersma, P. 2011. A programme for bidirectional phonology and phonetics and their acquisition and evolution. In: Benz, A. and Mattausch, J. (Eds.), *Bidirectional optimality theory* (p. 33–72). Amsterdam: John Benjamins. DOI: https://doi.org/10.1075/la.180.02boe

Boersma, P. and Hayes, B. 2001. Empirical tests of the gradual learning algorithm. *Linguistic Inquiry, 32*(1), 45–86. DOI: https://doi.org/10.1162/002438901554586

Britain, D. 2009. One foot in the grave. Dialect death, dialect contact, and dialect birth in England. *International Journal of the Sociology of Language 2009*, 121–155. DOI: https://doi.org/10.1515/IJSL.2009.019

Browman, C. P. and Goldstein, L. 1995. Gestural syllable position effects in American English. In: Bell-Berti, F. and Raphael, L. J. (Eds.), *Producing speech: contemporary issues. for Katherine Safford Harris* (p. 19–33). New York: American Institute of Physics.

Carignan, C., Mielke, J. and Dodsworth, R. 2016. Tongue trajectories in North American English /æ/-tensing. *The future of dialects: Selected papers from Methods in Dialectology XV, 1*, 313.

Carter, P. 2002. *Structured variation in British English liquids.* Unpublished doctoral dissertation, University of York.

Carter, P. 2003. Extrinsic phonetic interpretation: spectral variation in English liquids. In: John Local, R. O. and Temple, R. (Eds.), *Phonetic interpretation: Papers in laboratory phonology vi* (p. 237–252). Cambridge: Cambridge University Press.

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. DOI: https://doi.org/10.1017/S0025100307002939

Cho, T., Yoon, Y. and Kim, S. 2014. Effects of prosodic boundary and syllable structure on the temporal realization of CV gestures. *Journal of Phonetics*, *44*, 96–109. DOI: https://doi.org/10.1016/j.wocn.2014.02.007

Chomsky, N. and Halle, M. 1968. *The sound pattern of English*. New York: Harper & Row.

Cohn, A. 1993. Nasalisation in English: phonology or phonetics. *Phonology*, *10*, 43–81. DOI: https://doi.org/10.1017/S0952675700001731

Cruttenden, A. (Ed.). 2008. *Gimson's pronunciation of English* (7th ed.). London: Hodder.

Davidson, L. 2006. Comparing tongue shapes from ultrasound imaging using smoothing spline analysis of variance. *The Journal of the Acoustical Society of America*, *120*, 407. DOI: https://doi.org/10.1121/1.2205133

Delattre, P. and Freeman, D. C. 1968. A dialect study of American r's by x-ray motion picture. *Linguistics, 44*, 29–68. DOI: https://doi.org/10.1515/ling.1968.6.44.29

Ellis, L. and Hardcastle, W. J. 2002. Categorical and gradient properties of assimilation in alveolar to velar sequences: evidence from EPG and EMA data. *Journal of Phonetics*, *30*(3), 373–396. DOI: https://doi.org/10.1006/jpho.2001.0162

Gick, B. 2003. Articulatory correlates of ambisyllabicity in English glides and liquids. In: Local, J., Ogden, R. and Temple, R. (Eds.), *Phonetic interpretation: papers in laboratory phonology vi.* Cambridge: Cambridge University Press.

Giles, S. B. and Moll, K. L. 1975. Cinefluorographic study of selected allophones of English /l/. *Phonetica*, *31*, 206–227. DOI: https://doi.org/10.1159/000259670

Hale, M. and Reiss, C. 2000. "Substance abuse" and "dysfunctionalism": current trends in phonology. *Linguistic Inquiry*, *31*, 157–169. DOI: https://doi.org/10.1162/002438900554334

Halle, M. and Mohanan, K. P. 1985. Segmental phonology of Modern English. *Linguistic Inquiry, 16*, 57–116.

Hall-Lew, L. and Fix, S. 2012. Perceptual coding reliability of (l)-vocalisation in casual speech 33 data. *Lingua, 122*(7), 794–809. DOI: https://doi.org/10.1016/j.lingua.2011.12.005

Hardcastle, W. and Barry, W. 1989. Articulatory and perceptual factors in /l/ vocalisations in English. *Journal of the International Phonetic Association, 15*(02), 3–17. DOI: https://doi.org/10.1017/S0025100300002930

Hartigan, J. A. and Hartigan, P. 1985. The dip test of unimodality. *The Annals of Statistics*, 70–84. DOI: https://doi.org/10.1214/aos/1176346577

Hayes, B. 2000. Gradient well-formedness in Optimality Theory. In: Dekkers, J., Leeuw, F. v. d. and Weijer, J. v. d. (Eds.), *Optimality Theory: phonology, syntax, and acquisition* (p. 88–120). Oxford: Oxford University Press.

Hickey, R. 2005. Dublin English: current changes and their motivation. In: Foulkes, P. and Docherty, G. (Eds.), *Urban Voices: Accent Studies in the British Isles*. London: Arnold. DOI: https://doi.org/10.1075/veaw.g35

Holst, T. and Nolan, F. 1995. The influence of syntactic structure on [s] to [ʃ] assimilation. *Papers in Laboratory Phonology, IV*, 315–3333. DOI: https://doi.org/10.1017/CBO9780511554315.022

Huffman, M. K. 1997. Phonetic variation in intervocalic onset /l/s in English. *Journal of Phonetics, 25*(2), 115–141. DOI: https://doi.org/10.1006/jpho.1996.0038

Hughes, A., Trudgill, P. and Watt, D. 2012. *English accents and dialects: an introduction to social and regional varieties of English in the British Isles.* Routledge.

Huwaldt, J. A. 2010. *Plot Digitizer.* Available online at: plotdigitizer.sourceforge.net.

Johnson, K. 2011. *Quantitative methods in linguistics.* John Wiley & Sons.

Johnson, W. and Britain, D. 2007. L-vocalisation as a natural phenomenon: Explorations in sociophonology. *Language Sciences, 29*(2), 294–315. DOI: https://doi.org/10.1016/j.langsci.2006.12.022

Kelly, J. and Local, J. 1986. Long-domain resonance patterns in English. In *Institute of electronic engineers: Proceedings of the international conference on speech input/output* (p. 304–308).

Kerswill, P. 1985. A sociophonetic study of connected speech processes in Cambridge English: an outline and some results. *Cambridge papers in phonetics and experimental linguistics, 4*(1987), 25–49.

Kerswill, P. and Wright, S. 1990. The validity of phonetic transcription: Limitations of a sociolinguistic research tool. *Language Variation and Change, 2*(03), 255–275. DOI: https://doi.org/10.1017/S0954394500000363

Kiparsky, P. 1982. *Explanation in phonology*. Dordrecht: Foris. DOI: https://doi.org/10.1515/9783111666242

Kiparsky, P. 1988. Phonological change. In: Newmeyer, F. J. (Ed.), *Linguistics: the Cambridge survey* (Vols. 1, Linguistic theory: foundations, p. 363–415). Cambridge: Cambridge University Press.

Kiparsky, P. 2006. The amphichronic program vs. evolutionary phonology. *Theoretical Linguistics*, *32*(2), 217–236. DOI: https://doi.org/10.1515/TL.2006.015

Kirkham, S. 2015. Ethnicity and phonetic variation in Sheffield English liquids. *Journal of the International Phonetic Association*.

Kirkham, S. and Wormald, J. 2015. *Acoustic and articulatory variation in British Asian English liquids*. Proceedings of the XVII International Congress of Phonetic Sciences, Glasgow.

Ladefoged, P. and Maddieson, I. 1996. *The sounds of the world's languages*. Oxford: Blackwell.

Lawson, E., Stuart-Smith, J. and Scobbie, J. M. 2008. Articulatory insights into language variation and change: Preliminary findings from an ultrasound study of derhoticization in Scottish English. *University of Pennsylvania Working Papers in Linguistics, 14*(2), 1524–1549.

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. DOI: https://doi.org/10.1515/lp-2013-0015

Lehiste, I. 1964. Acoustical characteristics of selected English consonants. *Indiana research centre in anthropology, folklore and linguistics, 34*, 10–50.

Lehiste, I. 1980. Phonetic Manifestation Of Syntactic Structure In English. *Annual Bulletin of the Research Institue of Logopedics and Phoniatrics*, *14*, 1–27.

Lin, S. 2011. *Production and perception of prosodically varying inter-gestural timing in American English laterals.*Unpublished doctoral dissertation, University of Michigan.

Maechler, M. 2013. diptest: Hartigan's dip test statistic for unimodality – corrected code [Computer software manual]. Retrieved from http://CRAN.R-project.org/package = diptest (R package version 0.75–5).

McCafferty, K. 1999. Derry: between Ulster and local speech–class, ethnicity and language change. In: Foulkes, P. and Docherty, G. (Eds.), *Urban voices: accent studies in the British Isles* (pp. 246–264). London: Arnold.

Mielke, J., Baker, A. and Archangeli, D. 2010. Variability and homogeneity in American English /r/ allophony and /s/ retraction. *Laboratory phonology, 10*, 699–730.

Moisik, S. R. 2013. *The epilarynx in speech.* Unpublished doctoral dissertation, University of Victoria.

Myers, S. 2000. Boundary disputes: the distinction between phonetic and phonological sound patterns. In: Burton-Roberts, N. Carr, P. and Docherty, G. (Eds.), *Phonological knowledge: conceptual and empirical issues* (p. 245–272). Oxford: Oxford University Press.

Nance, C. 2014. Phonetic variation in Scottish Gaelic laterals. *Journal of Phonetics, 47*, 1–17. DOI: https://doi.org/10.1016/j.wocn.2014.07.005

Nolan, F., Holst, T. and Kühnert, B. 1996. Modelling [s] to [ʃ] accommodation in English. *Journal of Phonetics, 24,* 113–137. DOI: https://doi.org/10.1006/jpho.1996.0008

Olive, J. P., Greenwood, A. and Coleman, J. 1993. *Acoustics of American English speech: a dynamic approach.* New York: Springer Verlag.

Pierrehumbert, J. 2002. Word-specific phonetics. In: Gussenhoven, C. and Warner, N. (Eds.), *Laboratory phonology 7,* (p. 101–139). Berlin: Mouton de Gruyter. DOI: https://doi.org/10.1515/9783110197105.101

Pouplier, M. and Hoole, P. 2013. *Comparing principal component analysis of ultrasound images with contour analyses in a study of tongue body control during German coronals.* Paper presented at Ultrafest 6, Queen Margaret University, Edinburgh.

Ramsammy, M. 2015. The life cycle of phonological processes: accounting for dialectal microtypologies. *Language and Linguistics Compass, 9*(1), 33–54. DOI: https://doi.org/10.1111/lnc3.12102

Scobbie, J. M. 1995. What do we do when phonology is powerful enough to imitate phonetics? comments on Zsiga. *Phonology and phonetic evidence, papers in Laboratory Phonology I, 303*. DOI: https://doi.org/10.1017/CBO9780511554315.021

Scobbie, J. M. and Pouplier, M. 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), 240–259. DOI: https://doi.org/10.1016/j.wocn.2009.10.005

Scobbie, J. M., Stuart-Smith, J. and Lawson, E. 2008. *Looking variation and change in the mouth: developing the sociolinguistic potential of ultrasound tongue imaging.* Final Research Report for ESRC Project RES-000-22-2032. Queen Margaret University.

Scobbie, J. M. and Wrench, A. A. 2003. An articulatory investigation of word final /l/ and /l/-sandhi in three dialects of English. *Proceedings of the International Congress of Phonetic Sciences, 15*, 1871–1874.

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

Stone, M. 2005. A guide to analysing tongue motion from ultrasound images. *Clinical linguistics & phonetics, 19*(6–7), 455–501. DOI: https://doi.org/10.1080/02699200500113558

Strycharczuk, P. 2012. Sonorant transparency and the complexity of voicing in Polish. *Journal of Phonetics, 40*(5), 655–671. DOI: https://doi.org/10.1016/j.wocn.2012.05.006

Strycharczuk, P. and Scobbie, J. M. 2015. *Velocity measures in ultrasound data. Gestural timing of post-vocalic /l/ in English.* Proceedings of the XVII International Congress of Phonetic Sciences International Conference of Phonetic Science, Glasgow.

Strycharczuk, P. and Simon, E. 2013. Obstruent voicing before sonorants. the case of West-Flemish. *Natural Language & Linguistic Theory, 31*(2), 563–588. DOI: https://doi.org/10.1007/s11049-013-9189-5

Tollfree, L. 1999. South-east London English: discrete versus continuous modelling of consonantal reduction. In: Foulkes, P. and Docherty, G. (Eds.), *Urban voices: Accent studies in the British Isles* (p. 163–184). London: Arnold.

Turton, D. 2014a. *Variation in English /l/: Synchronic reflections of the life cycle of phonological processes.* Unpublished doctoral dissertation, University of Manchester.

Turton, D. 2014b. Some /l/s are darker than others: accounting for variation in English /l/ with ultrasound tongue imaging. *Penn Working Papers in Linguistics, 20*(2).

Turton, D. 2015. *Determining categoricity in English /l/-darkening: A principal component analysis of ultrasound spline data.* Proceedings of the XVII International Congress of Phonetic Sciences International Conference of Phonetic Science 2015, Glasgow.

Turton, D. 2016. Synchronic stratum-specific rates of application reflect diachronic change: morphosyntactic conditioning of variation in English /l/-darkening. *Papers in Historical Phonology, 1*. DOI: https://doi.org/10.2218/pihph.1.2016.1697

Turton, D. and Baranowski, M. 2015. *Absence of a blocking r[ʏɬ]?: the presence of /u/-fronting before /l/ in Manchester.* Paper given at UKLVC, University of York, 1st–3rd September 2015.

Turton, D. and Ramsammy, M. 2012. */ɪ, ə/-lowering in Manchest[ʌ]: Contextual patterns of gradient variability.* Paper given at the 20th Manchester Phonology Meeting. Manchester, 24th–26th May 2012.

Wells, J. C. 1982. *Accents of English*. Cambridge: Cambridge University Press. DOI: https://doi.org/10.1017/CBO9780511611759

Wickham, H. 2009. *ggplot2: elegant graphics for data analysis*. Springer Publishing Company, Incorporated. DOI: https://doi.org/10.1007/978-0-387-98141-3

Wrench, A. A. and Scobbie, J. M. 2003. Categorising vocalisation of English /l/ using EPG, EMA and ultrasound. In: *Proceedings of the 6th International Seminar on Speech Production* (pp. 314–319).

Wright, S. 1989. The effects of style and speaking rate on /l/-vocalisation in Local Cambridge English. *York Papers in Linguistics, 13*, 355–365.

Wright, S. and Kerswill, P. 1989. Electropalatography in the analysis of connected speech processes. *Clinical linguistics & phonetics, 3*(1), 49–57. DOI: https://doi.org/10.3109/02699208908985270

Yuan, J. and Liberman, M. 2009. Investigating /l/ variation in English through forced alignment. In: *INTERSPEECH 2009: 10th Annual Conference of the International Speech Communication Association*, Brighton, UK, 6–10 September 2009 (p. 2215–2218).

Yuan, J. and Liberman, M. 2011. /l/ variation in American English: A corpus approach. *Journal of Speech Sciences, 1*(2), 35–46.

Zsiga, E. C. 1995. An acoustic and electropalatographic study of lexical and postlexical palatalization in American English. In: Connell, B. and Arvaniti, A. (Eds.), *Phonology and phonetic evidence: papers in laboratory phonology IV* (p. 282–302). Cambridge: Cambridge University Press. DOI: https://doi.org/10.1017/CBO9780511554315.020

Zwicky, A. 1972. Note on a phonological hierarchy in English. *Linguistic change and generative theory, 275*, 301.

How to cite this article: Turton, D 2017 Categorical or gradient? An ultrasound investigation of /l/-darkening and vocalization in varieties of English. *Laboratory Phonology: Journal of the Association for Laboratory Phonology* 8(1): 13, pp. 1–31, DOI: https://doi.org/10.5334/labphon.35

Published: 25 May 2017

]u[*Laboratory Phonology: Journal of the Association for Laboratory Phonology* is a peer-reviewed open access journal published by Ubiquity Press.

OPEN ACCESS