**DE GRUYTER** MOUTON
DOI 10.1515/lp-2013-0015 — Laboratory Phonology 2013; 4(2): 475 – 511

Sang-Im Lee-Kim, Lisa Davidson and Sangjin Hwang

# Morphological effects on the darkness of English intervocalic /l/

**Abstract:** Articulatory and acoustic studies have provided evidence that in word-initial and word-final positions, English /l/ exhibits substantial differences in 'darkness': dark [ɫ] in word-final position is produced with a more retracted tongue dorsum and lowered tongue body than light [l] in word-initial position. The darkness of intervocalic /l/, however, is variable. While Sproat and Fujimura (1993) argue that /l/ darkness is on a continuum strongly affected by duration, Hayes (2000) maintains that the morphological status of intervocalic /l/s should affect whether they are produced as light or dark variants. In this study, ultrasound imaging is used to investigate whether the morphological affiliation of the /l/ affects the degree of tongue dorsum retraction and tongue body lowering and the acoustic characteristics of /l/ darkness. Six American English speakers produced three types of stimuli which were predicted to increase in darkness in the following order: (1) when /l/ corresponded with the onset of a suffix (e.g., flaw-less), (2) when /l/ corresponded with the final position of the stem word (e.g., tall-est), and (3) when /l/ was the final consonant of a stem word (e.g., tall). For both articulatory and acoustic measures, the predicted order was upheld. The strongest articulatory correlate of darkness was tongue body lowering, and acoustic differences were mainly manifested in F1 and normalized intensity. Phonological implications of these findings are discussed.

**Sang-Im Lee-Kim:** New York University. E-mail: sangim@nyu.edu
**Lisa Davidson:** New York University. E-mail: lisa.davidson@nyu.edu
**Sangjin Hwang:** New York University. E-mail: sangjin.hwang@gmail.com

## 1 Introduction

A wealth of phonological and phonetic research has been carried out to investigate the claim that American English /l/ has two variants that are syllabically conditioned (Jones 1956; Chomsky 1968; Ladefoged 2005; Cruttenden 2008, etc.). It has been stated in the literature that in syllable onsets, /l/ is produced as 'light' [l], whereas /l/ in syllable codas is the 'dark' [ɫ] variant. While a number of studies have already debunked the claim for a few varieties of English that light [l] and dark [ɫ] are singular entities, most researchers still agree that there is a lighter and darker end of the spectrum of /l/ production (Sproat and Fujimura 1993; Gick, Campbell, and Tamburri-Watt 2006; Scobbie and Pouplier 2010; Proctor and Walker 2012). Articulatorily, darker [ɫ] has a more retracted tongue root and lower tongue body than lighter [l]. Another articulatory difference between lighter and darker /l/s concerns the timing of the tongue tip and tongue body gestures; whereas the tongue root movement and tongue tip raising in [l] are simultaneous, the retraction of the body in darker [ɫ] precedes the raising of the tip (Giles and Moll 1975; Sproat and Fujimura 1993; Gick 2003; Scobbie and Pouplier 2010). Acoustically, lighter [l] has a higher F2 and lower F1 as compared to darker [ɫ] (Fant 1960; Lehiste 1964; Bladon and Al-Bamerni 1976; Espy-Wilson 1992; Sproat and Fujimura 1993; Recasens 2011).

While the above descriptions are most characteristic of /l/ in word-initial and word-final position, a few studies have shown that at least for some measures, the darkness of intervocalic /l/ varies depending on a number of factors (Bladon and Al-Bamerni 1976; Sproat and Fujimura 1993; Huffman 1997; Gick 2003; Gick et al. 2006). In particular, based on extensive acoustic and articulatory measurements from x-ray microbeam, Sproat and Fujimura (1993) argued that the darkness of /l/ is ultimately on a phonetic continuum. Assuming that the primary determinant of the phonetic scale of darkness is the duration of the rime containing the /l/, they examined the production of /l/ at a number of boundary environments in the context of /i-ɪ/: phrasal boundaries, VP boundaries, VP-internal breaks, compound boundaries, affix boundaries, and word-internally. Results showed that the darkness of /l/ is strongly correlated with juncture strength: the stronger the boundary, the darker /l/ is. For example, /l/ was darkest at major phrase boundaries (e.g., 'Nea/l/, equate the actors', '|') and lightest word-internally (e.g., 'Mr. Bee/l/ik wants actors', '%'), and there were intermediate degrees of darkness before VP-internal breaks (e.g., 'I gave Bee/l/ equated actors', 'P') and before affix boundaries ('The knee/l/ing men are actors', '#'), with the former being darker than the latter. As they expected, the boundary strength was reflected in the relative duration of the /Vl/ sequence; the /Vl/ at the major phrase boundary had the longest duration, whereas the /Vl/ at word-internal boundaries was produced in the shortest amount of time. The relative lightness of the word-internal /l/ is thus attributed to articulatory undershoot of the tongue root due to the limited amount of time allotted to the /l/. The scale of darkness proposed by Sproat and Fujimura (1993) is shown in (1).

(1) Phonetic continuum of dark /l/ (Sproat and Fujimura 1993). (%: word-internal intervocalic, #: productive morpheme boundary, C: compound boundary,

h: before [h], P: VP-internal break, V: VP phrase boundary, |: intonation boundary)

% < # < C < h < P < V < |

←――――――――――→

| | | |
|---|---|---|
| Boundary strength: | weaker | stronger |
| Rime duration: | shorter | longer |
| Relative darkness: | lighter | darker |

In a perceptual study following up on the idea that phonological and morphological structure may affect the darkness of /l/, Hayes (2000) hypothesized that there should be a difference in darkness between two types of intervocalic /l/s in morphologically complex words which would both be syllabified as an onset by most phonological theories: an /l/ that is the final consonant of a related stem (*feel-ing*), and an /l/ in the suffix after a morpheme boundary (*free-ly*). While Sproat and Fujimura did not systematically vary whether the /l/ belonged to the stem or suffix, Hayes (2000) made the more explicit prediction that the stem /l/ followed by a vowel-initial suffix (e.g., *feel-ing*) should have dark [ɫ] compared to forms with an /l/-initial suffix (e.g., *free-ly*). This is because the stem-final /l/ is in a paradigmatic relationship with the stem form (e.g., *feel*), which prevents the intervocalic /l/ from deviating from its stem, while such pressure is absent for the /l/-initial suffix. These conditions are a subset of the types of data that Hayes examined, which also included /l/ in a final cluster (e.g., *help* /hɛlp/), intervocalic /l/ in a monomorpheme (e.g., *mellow* /mɛloʊ/), and word-initial /l/ (e.g., *Lee* /li/).

However, Hayes departed from Sproat and Fujimura by hypothesizing that there is not a continuum of darkness in speakers' phonetic implementation, but rather categorical [l] or [ɫ] variants that vary, with different proportions, depending on morphological contexts. That is, speakers produce the dark [ɫ] variant 'more often' than the light [l] variant in the stem-final position, and therefore, listeners are more likely to accept dark variants in stem-final position than in suffix-initial position. Hayes tested this hypothesis with a perception task; he recorded himself producing various words containing /l/ with both light [l] and dark [ɫ], and asked listeners to rate both pronunciations on a scale of 1 (good) to 7 (bad). For the current study, the relevant environments in Hayes were coda /l/ in the stem form (e.g., *hole, bell*), pre-boundary /l/ (e.g., *feel-y*, *heal-ing*), and post-boundary /l/ (e.g., *free-ly*, *eye-let*). The rating scores averaged over listeners is reproduced in Table 1 below. By using the rating scale and averaging over listeners, Hayes obtained a gradient effect demonstrating that dark [ɫ] is preferred in pre-boundary position, whereas light [l] is preferred in post-boundary environments.

**Table 1:** Hayes's (2000) well-formedness judgment test. 1 represents 'good' and 7 represents 'bad'. English listeners gave intermediate rating scores for the pre-boundary form produced with either [l] or [ɫ].

| | Pronounced w/[l] | Pronounced w/[ɫ] |
|---|---|---|
| Stem form (*hole*, *bell*) | 6.60 | 1.12 |
| Pre-boundary form ((touchy) *feel-y, heal-ing*) | 3.01 | 2.01 |
| Post-boundary form (*free-ly*, *eye-let*) | 1.57 | 3.34 |

The methodology in Hayes (2000) has limitations, however. First, since Hayes recorded the stimuli himself, the variants he produced were potentially not naturally occurring. It is unknown whether they contained the articulatory and acoustic characteristics of canonical (word-final) dark and (word-initial) light /l/. Second, a perception study is not a direct investigation of the articulatory productions of English speakers. Thus, the goal of the current study is to examine whether the morphological status of intervocalic /l/ (stem-final vs. suffix-initial) affects its articulatory and acoustic realization.

In this paper, we use ultrasound imaging and acoustic measures to examine morphological effects on the production of /l/ in the stem, pre-boundary, and post-boundary environments proposed by Hayes (2000). We follow Sproat and Fujimura in expecting that speakers' articulations will indicate a phonetic continuum of darkness, and we share Hayes' hypothesis that a word's morphological status should critically bear on the articulatory implementation of /l/. Indeed, morphological boundaries have been shown to affect gestural and acoustic implementation in other languages or word types. For example, Cho (2001) showed that gestural timing between CV sequences was more stable in monomorphemic words as compared to homophonous words including a morpheme boundary (e.g., /napi/ 'butterfly' vs. /nap-i/ 'lead+Nom.'). For Scottish English, Sugahara and Turk (2009) showed that morphologically complex words were typically longer than homophonous or nearly homophonous monomorphemes (e.g., *puffing-puffin*, *laps-lapse*). Other morphological effects on acoustic implementation have been shown for Dutch (Ernestus and Baayen 2006; Pluymaekers et al. 2010).

Given these findings, we predict that the pre-boundary /l/ is intermediate in darkness, with the stem form being the darkest and the post-boundary /l/ being the lightest. In particular, the pre-boundary /l/ (e.g., *feel-ing*) is predicted to be lighter than the /l/ in the stem (e.g., *feel*) because of the phonological pressure that prefers an onset for the following vowel-initial suffix. Although pre- (e.g., *feel-ing*) and post-boundary (e.g., *free-ly*) conditions are phonologically equivalent, i.e., in prevocalic position, the former is predicted to be darker than the latter, because the pre-boundary /l/ is potentially in a paradigmatic relation with the coda /l/ in the stem form (e.g., *feel*). These complex morphological and phonological factors lead to the prediction that the relative darkness of /l/ should be reflected in both articulation and acoustics. Articulatorily, /l/ in pre-boundary position is predicted to have greater tongue dorsum retraction and/or tongue body lowering (e.g., darker in *feel-ing*) than /l/ in post-boundary position (e.g., lighter in *free-ly*). Pre-boundary /l/ (e.g., *feel-ing*) is predicted to show a lesser degree of tongue dorsum retraction or tongue body lowering than the /l/ in stem-final position (e.g., *feel*). Acoustically, F2 should be lower for the /l/ in pre-boundary position compared to /l/ in post-boundary position. F1 is expected to have the opposite directionality of F2: pre-boundary /l/ should have higher F1 than the post-boundary /l/. Again, F2 values for stem /l/ should be lowest, and, conversely, F1 values should be the highest. Following findings in Espy-Wilson (1992) and Stevens (1998), we also examine the relative intensity of the /l/. Stevens found that differences in amplitude for intervocalic /l/ as compared to vowels are much smaller when the /l/ follows a stressed syllable ('coda position') than when it precedes a stressed syllable ('onset position'). For our data, we predict that when /l/ is darker, it should have greater intensity than lighter /l/. Operationally, there should be a smaller intensity difference between a darker /l/ and the subsequent vowel than between a lighter /l/ and the following vowel.

There are also two alternative possibilities. First, one possible result is that the two intervocalic /l/s in the pre-boundary and post-boundary conditions have the same articulatory and acoustic properties. This would indicate that their phonetic implementation is not affected by their morphological status at all, but is rather due to their intervocalic position. A second possibility is that the /l/ in the pre-boundary condition is not different from the /l/ in the stem form (but both are different from the post-boundary condition). This would mean that the phonetic implementation of the pre-boundary /l/ arises solely from the morphological environment, such that speakers are retaining the articulation of the stem form (e.g., *feel*) in producing the morphologically-inflected form (e.g., *feel-ing*). A more formal interpretation of this possible finding is considered in the discussion.

While Sproat and Fujimura's (1993) study contained semantically unnatural or invented words (e.g., *the te[l]ic man, Mr. Bee[l]ik*) in order to keep the vocalic environment strictly to /i-ɪ/, we include several vowels preceding the /l/ to examine the generality of the morphological environments. Additionally, this enables us to test whether there is coarticulatory interaction between the vowel and /l/. The prediction is that phonetic differences between the stem, pre-boundary, and post-boundary environments should be less evident for preceding low back vowels as compared to high or front vowels. This is because the tongue is already low and retracted for low back vowels (cf. Recasens 2011), and it may not be possible for the tongue to further retract or lower in stem or pre-boundary position for dark [ɫ].

As discussed above, it is well-known that tongue dorsum retraction is not the only correlate of a darker /l/. The relative timing between the tongue dorsum retraction movement and the tongue tip raising gesture has also been examined, and studies have shown that these two movements for lighter /l/s are generally temporally close, whereas the tongue dorsum retraction is often earlier than tongue tip raising for darker /l/s. This is shown robustly in Sproat and Fujimura for American English using x-ray microbeam. Similar results are obtained by Scobbie and Pouplier (2010) for Scottish Standard English and Southern Standard British English speakers for the distinction between word-initial and word-final (non-ambisyllabic) /l/ using electropalatography (EPG). However, using ultrasound, results in Gick et al. (2006) did not follow the expected pattern for Western Canadian English, since there was a greater negative lag in pre-vocalic position (dorsum retraction preceding tongue tip raising) and a smaller negative lag in post-vocalic position. Gick et al. note that this finding may be due to dialectal differences between Canadian and American English, but it is also possible that the relatively slow frame rate of ultrasound is obscuring the actual timing differences between retraction and raising for these speakers.

In this study, we focus on tongue dorsum retraction and tongue body lowering as the main articulatory measures because the frame rate of our ultrasound imaging is not adequate for obtaining accurate temporal measurements of the lag between dorsum retraction and tip raising. With a frame rate of approximately 30 frames per second, we cannot obtain the fine-grained precision that is possible with either EPG or electromagnetic midsagittal articulography (EMMA) (Stone 2010), which are better suited to studies of timing relationships. Although Gick et al. (2006) do use ultrasound to examine the lag between tongue dorsum retraction and tip raising for pre-vocalic, intervocalic, and post-vocalic /l/, we feel that this imaging technique is too coarse to be used reliably to look for distinctions in an intervocalic environment differing only in morphological status. In contrast, the advantage of ultrasound is that it allows for the imaging of a greater portion of the tongue dorsum/root (as compared to both EPG and EMMA), and does not rely on the particular placement of a pellet (in the case of EMMA). Thus, while timing differences would provide a more complete articulatory picture of the interaction between morphological boundaries and articulation of /l/ (though we briefly provide some comments about tongue tip position in Section 4.1), we combine tongue dorsum retraction as our main articulatory correlate with formant and intensity measures as complementary acoustic correlates to provide converging evidence regarding the status of /l/.

# 2 Experiment

## 2.1 Participants

Participants were two female (MG, RT) and four male (AD, AN, WT, JO) New York University undergraduate students. All speakers were born and raised in New York State, and they are all monolingual native speakers of American English. Although we were not able to quantify the amount of tongue tip/blade contact, impressionistically, none of the speakers were /l/-vocalizers in the sense of a perceptibly vocalic-only production of dark [ɫ] (Hall-Lew and Fix 2012). None of the speakers reported history of speech or hearing disorders. The participants were paid for their time. This study was carried out with the approval of the Institutional Review Board of New York University.

## 2.2 Materials

The stimuli for the stem in the stem condition included five monosyllabic adjectives ending with /l/. Those words differ in the preceding vowel context: /u, o, aɪ, aʊ, ɑ[1]/. To create pre-boundary forms, the superlative suffix *-est* was attached to each stem form. For the post-boundary forms, the suffix *-less* was chosen as an /l/-initial affix that was minimally different from *-est*, especially with respect to the quality of the following vowel (i.e., [-əst] and [-ləs]).[2] Monosyllabic nouns with matched preceding vowels preceded the suffix *-less*. Table 2 contains the target words.

These vowels were chosen because they were amenable to forming phonetically matched, semantically plausible words with suffixes *-est* and *-less*. The /i/

1 Speakers varied as to whether the quality of this vowel was more [ɑ]-like or more [ɔ]-like. However, speakers were consistent within their own productions and across the three words specified for that vowel. Because the vowel was sufficiently low and back for each speaker regardless of the precise quality, and because we report individual results when appropriate, we use the label /ɑ/ for the words *tall-tallest-flawless*, as shown in Table 2.

2 Since speakers were both familiarized with the materials before the recording, and produced 10 repetitions, they consistently produced the productive suffix *-less* with a reduced vowel that was acoustically comparable to the vowel in *-est*.

**Table 2:** Target stimuli used in experiment.

| | Stem | Pre-boundary *(-est)* | Post-boundary *(-less)* |
|---|---|---|---|
| /u/ | cool | coolest | coupless |
| /o/ | droll | drollest | crowless |
| /aɪ/ | vile | vilest | eyeless |
| /aʊ/ | foul | foulest | cowless |
| /ɑ/ | tall | tallest | flawless |

vowel context, which was the only context in Sproat and Fujimura (1993), was unfortunately not included because there were no adequate words for the *-est* condition. Since the post-boundary forms were mostly formations using the productive suffix *-less*, mini-scenarios for all the target words were prepared in order to provide some context. The contexts are given in the appendix. Participants were asked to read these scenarios before the recording, and they all found the scenario natural and pronounced the productively formed post-boundary words with no disfluencies.

For the recording, semantically natural carrier phrases were created for each target word (e.g., *the coolest person*). For the stem form, the words following the /l/ always began with [h], following the stimulus design in Sproat and Fujimura (1993).[3] The vowel following the [h] was restricted to [ɛ] to match the vowel quality as closely as possible to the [ə] following pre- and post-boundary /l/ (e.g., *the cool headphones*). The whole list of phrases containing the stimuli is presented in the appendix. The list of 15 target phrases plus 2 filler phrases at the beginning and the end of the list were presented 10 times in random order.

## 2.3 Design and data collection

### 2.3.1 Ultrasound

During the recording session, participants were seated in a comfortable chair and instructed to find the most comfortable pose for themselves. Each participant's

**3** Some research indicates that a labial following consonant may be more likely to block coarticulation with the following vowel than /h/ (Scobbie and Pouplier 2010); thus, a significant difference between the stem condition and either pre- or post-boundary stimuli suggests that an effect is present and strong.

head was fit into a moldable head stabilizer (Comfort Company) that is mounted on the wall, and the head was further stabilized with a Velcro strap around the forehead (Davidson and De Decker 2005; Davidson 2006). This method, when used together with an ultrasound transducer that is fixed in place with a Bogen-Manfrotto Magic Arm, ensures that the same slice of tongue is imaged (Davidson 2012). The transducer was placed under the speaker's chin and adjusted until clear midsagittal images were obtained.

Tongue images were recorded using a Sonosite Titan portable ultrasound in the sound-attenuated booth at the Phonetics and Experimental Phonology Lab at New York University. A 5–8 MHz Sonosite C-11 transducer with a 90° field of view was set at a depth of 8.2–10 cm, depending on the speaker. The audio was collected using an Audio Technica AT-813 microphone, and was synchronized in an *AVI* movie file with the video signal from the ultrasound machine using a Canopus ADVC-1394 capture card and Adobe Premiere. The *WAV* files were then extracted from the movie file, and submitted to Praat for acoustic segmentation. This segmentation was carried out to match up the articulatory and acoustic events of interest; see below for more detail on the articulatory measure.

Since the /l/ occurs intervocalically for the pre- and post-boundary condition, the following procedure was used for acoustic segmentation. We approximated the onset of /l/ where the second formant started to decrease and the third formant started to rise. To account for the possibility that maximal tongue retraction could proceed beyond the acoustic realization of /l/, about half of the following segment, i.e., vowel [ə] for the pre- and post-boundary condition and [h] for the stem condition, was included to ensure that the frame corresponding to the maximal retraction was not missed. The tongue images captured during the acoustic realization of /l/ plus half of the following segments were extracted using Matlab. Figure 1 illustrates acoustic segmentation of /l/ and the corresponding ultrasound images.

Among the frames corresponding to the articulation of /l/, the frame capturing the most retracted tongue dorsum/upper root was selected by visual inspection. Since a small portion of the following segments [ə] (for the pre- and post-boundary condition) or [h] (for the stem condition) was included, there was usually clear forward movement of the tongue body corresponding to the vowel. One frame before this forward movement was thus chosen as the frame containing maximal retraction of /l/. Those selected frames were then submitted to EdgeTrak (Li, Kambhamettu, and Stone 2005) which automated the tracking of tongue contours by extracting x-y coordinates of the target region from the lower edge of the tongue. One hundred equidistant points were extracted for the statistical analysis in the next step. A screenshot of the tongue curve tracking in EdgeTrak is shown in Figure 2.

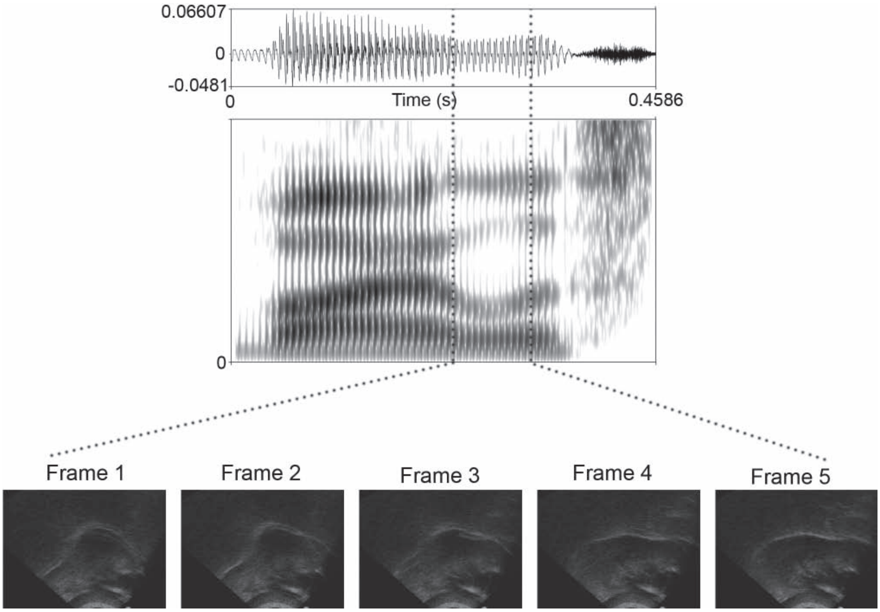


**Fig. 1:** A sample waveform and spectrogram of *vilest* produced by male speaker AN. Five ultrasound images were obtained during the segmented period. The tongue dorsum is on the left, and the tongue tip is on the right. Retraction of the tongue dorsum is evident starting at frame 3. A slight forward movement is seen between frame 4 and frame 5.

### 2.3.2 Acoustic measurement

As darkness of /l/ is present in the acoustic signal as well as in articulation, F2 and F1 of the acoustic /l/ were measured using Praat 5.1 (Boersma and Weenink 2011). Following similar measurements in previous literature, the minimum F2 occurring in the acoustic duration of /l/ was measured automatically using a Praat script. Subsequently, F1 values were measured at the same time as the F2 minimum. We call these values minimum F2 and F1 at minimum F2, respectively.

Based on the report of intensity differences among approximants in Espy-Wilson (1992), we also calculated normalized intensity, the difference in intensity between the /l/ and the following vowel /ə/ for the pre- and post-boundary conditions.[4] The minimum intensity during the acoustic realization of the /l/ was mea-

4 The stem condition was not considered for the intensity measure, since the /l/ in the stem was not followed by a vowel, and there was no intensity valley; instead, the intensity curve monotonically decreased toward the following /h/.

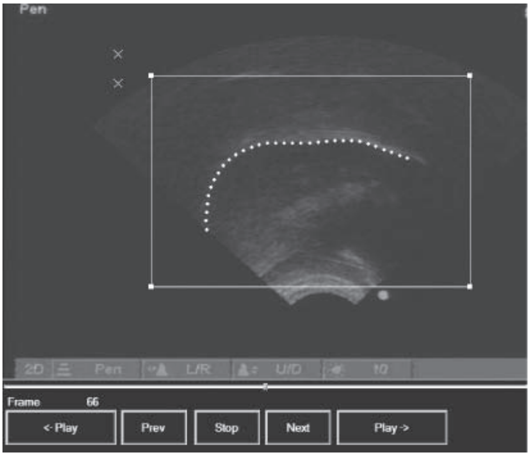


**Fig. 2:** A screenshot of EdgeTrak tracking for the /l/ at maximal retraction in *vilest* for subject MG. The white dots are the edge of the tongue created by the automated edge-tracking algorithm. The tongue tip is on the right and the tongue dorsum/root is on the left. Tongue tip is not seen in the ultrasound images due to the shadow of the jaw bone.

sured using a Praat script, which captures the intensity valley often found in the spectrogram. Because the precise boundary between the /l/ and the following vowel was sometimes difficult to determine, the intensity of the following vowel was measured 40 ms before the acoustic onset of the subsequent fricative. This measurement point is also appropriate for cases where there is not necessarily an intensity trough corresponding to the /l/, but rather relatively stable intensity from the /l/ to the following /ə/.

### 2.3.3 Statistical analyses for ultrasound

A smoothing spline ANOVA was used to compare tongue shapes. SS ANOVA is a statistical procedure for investigating similarities and differences of curve shapes (Gu 2002; Davidson 2006; Wang 2011). Using this method, smoothing splines which best fit the ten repetitions for the stimuli were obtained. Sample smoothing splines of two /l/s at maximal retraction in *vile* vs. *vilest* (a) and *coolest* vs. *coupless* (b) are presented in Figure 3 along with the raw data points of ten repetitions.

Whereas the standard ANOVA returns an *F*-value, the SS ANOVA returns parameter values for the smoothing spline that would be the best fit for all of the

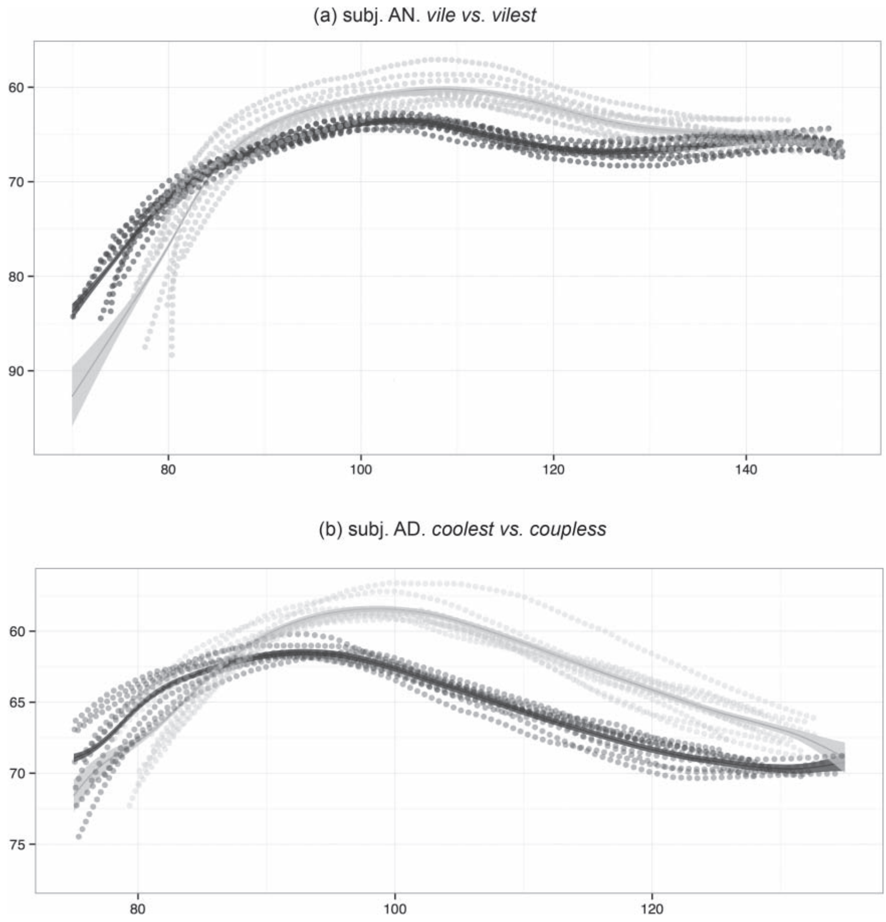


**Fig. 3:** Raw data points from ten repetitions (dotted lines) and the smoothing spline estimate (solid lines) for comparison of the tongue shape for /l/ in *vile* (in black) vs. *vilest* (in gray) for subject AN (a) and in *coolest* (in black) vs. *coupless* (in grey) for subject AD (b). Like the ultrasound images, the tongue tip is on the right and the tongue dorsum is on the left. The axes are in millimeters corresponding to the boxed-in region shown in Figure 2.

data at once and for the spline of the interaction, which represents the difference between the main effect splines and the spline that best fits all of the data. Whether the two curves are significantly different is roughly determined by comparing whether these two parameter values are within the same order of magnitude, which indicates that there is a significant difference between two shapes at least in some area. A more precise technique for determining where the differences along the tongue dorsum and tongue body lie is to construct 95% Bayesian confidence intervals around the smoothing splines. The two curves are significantly different where the confidence intervals are not overlapping. This statistical technique has been used in other studies of ultrasound imaging of the tongue (e.g., Davidson 2006; Chen and Lin 2011; Mielke et al. 2011), and we follow these papers in using the confidence intervals to determine regions of significant differences. SS ANOVA is implemented using the gss package in R (Gu 2012).

To exemplify, the smoothing splines of the *vile-vilest* and *coolest-coupless* pairs are presented in Figure 4 with 95% Bayesian confidence intervals marked by the light and dark gray shading that is superimposed on the smoothing spline. Comparing the smoothing splines and the confidence intervals around the splines in Figure 4, there is a significant difference both in the tongue dorsum and tongue body; for example, in Figure 4(b) the tongue dorsum is more retracted in *coolest* (in black) than in *coupless* (in gray), whereas the tongue body is lowered in *coolest* as compared to *coupless*. Note that because the same extent of the tongue root may not be imaged for all of the participants, we will refer to the portion of the tongue being inspected for retraction as the dorsum instead of the root.

In order to assess the backness of the tongue dorsum and lowering of the tongue body separately, we partitioned the tongue curve mathematically in thirds along the x-axis; the resulting regions roughly correspond to the tongue dorsum, tongue body, and tongue tip/blade (see Figure 5). As there is not yet an agreed upon technique for defining articulator boundaries along the tongue surface in ultrasound imaging, this is only a rough metric to be refined in future work. One concern with this method of dividing the tongue surface is that it is dependent on the field of view and angle of the transducer to some degree; with a field of view wider than 90°, it is possible that the tongue root/dorsum would be more completely imaged. It is also possible that the back third which we take to roughly correspond to the tongue dorsum/root actually contains a part of the tongue which might be more properly characterized as the tongue body. However, since the differences in the tongue dorsum/root and tongue body were generally robust, this method of defining articulator boundaries should be adequate for the question of /l/ darkness.

Since two lines sometimes crossed within one of the thirds, we implemented stricter criteria as follows. For the tongue dorsum, the difference along the y-axis between the lowest tongue point and the crossing point (represented as ‘a’ in Figure 5) was relativized with the entire range that the tongue stretches over (represented as ‘b’ in Figure 5). If two lines are not overlapping and the a/b ratio is more than two-thirds, 66.7%, we coded the case as one line being ‘more retracted’

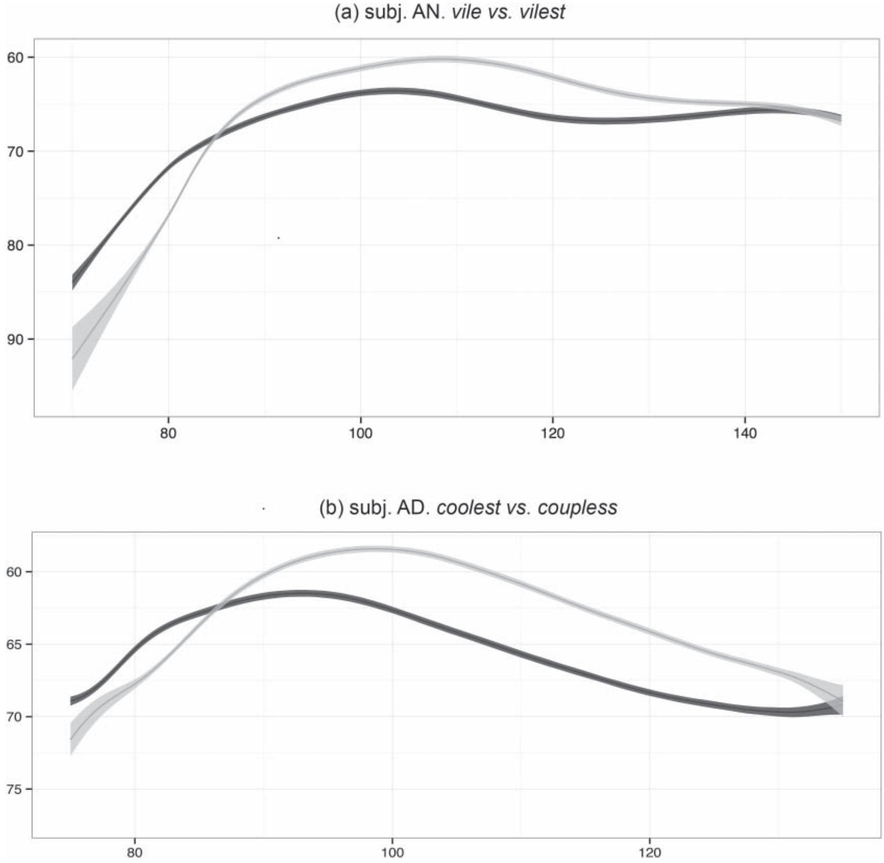


**Fig. 4:** Smoothing spline estimate and 95% Bayesian confidence interval for comparison of the mean curves for /l/ in *vile* (in black) vs. *vilest* (in gray) for subject AN (a) and in *coolest* (in black) vs. *coupless* (in gray) for subject AD (b). Like the ultrasound images, the tongue tip is on the right and the tongue dorsum is on the left.

than the other. For example, the a/b ratio in Figure 5(a) was 84%, and this case was coded as 'stem' being more retracted than 'pre-boundary'. If two lines are not overlapping and the a/b ratio is below two-thirds, the case was coded as 'inconclusive', which means neither line is conclusively more retracted than the other. The a/b ratio in Figure 5(b) was only 64%, and thus coded as inconclusive. The overlapping region was about 78% in Figure 5(c), and this case was coded as 'not significantly different'. The same criteria were used for the tongue body lowering, except that the relevant ratio was defined with reference to the x-axis, since the tongue body stretches over the x-axis.

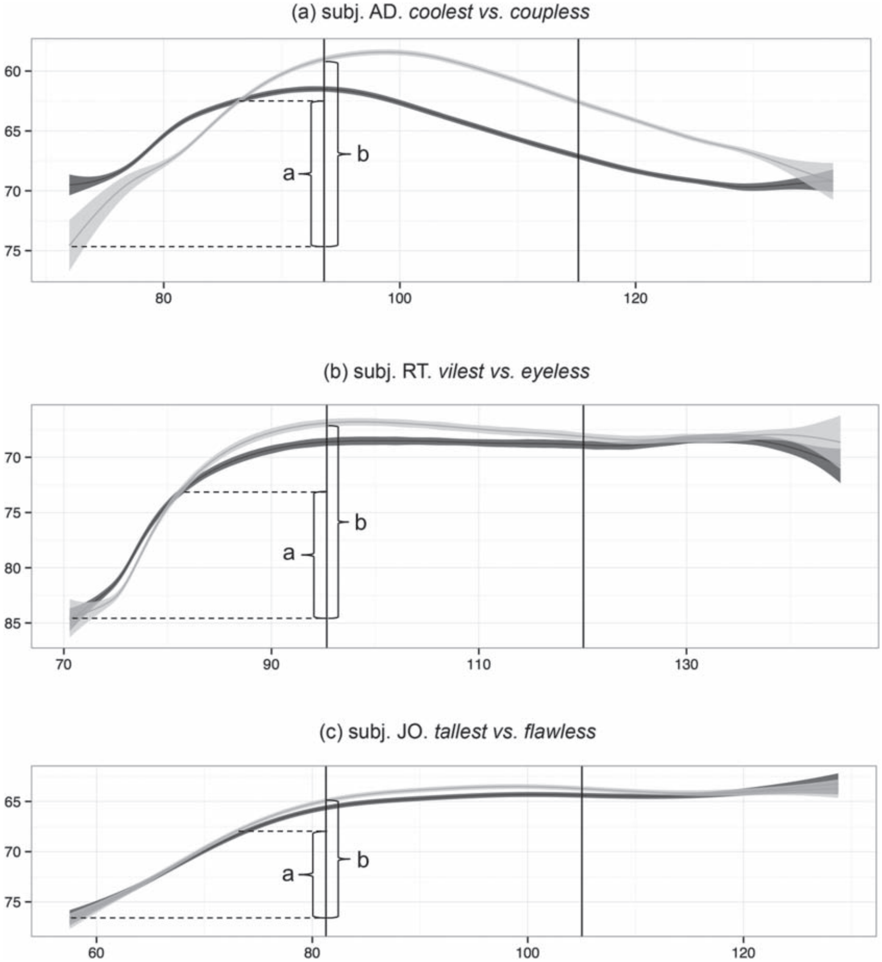


**Fig. 5:** SS ANOVA coding examples. In all examples, the dark gray lines represent the pre-boundary condition, whereas the light gray lines represent the post-boundary condition. The leftmost portion in each image is the tongue dorsum (TD), and the middle portion corresponds to the tongue body (TB). (a) Significantly different for both TD and TB, (b) TD is an 'inconclusive case' where the a/b ratio was below two-thirds and TB is significantly different, and (c) TD is not significantly different and TB is significantly different. The tongue shapes shown in this and other figures are typical representatives of the patterns found in the data.

# 3 Results

## 3.1 Articulatory results

### 3.1.1 Ultrasound results

Figure 6 exemplifies smoothing splines of the frames corresponding to the maximal retraction of /l/ of two speakers, AN and MG, in the preceding /o/ vowel context.

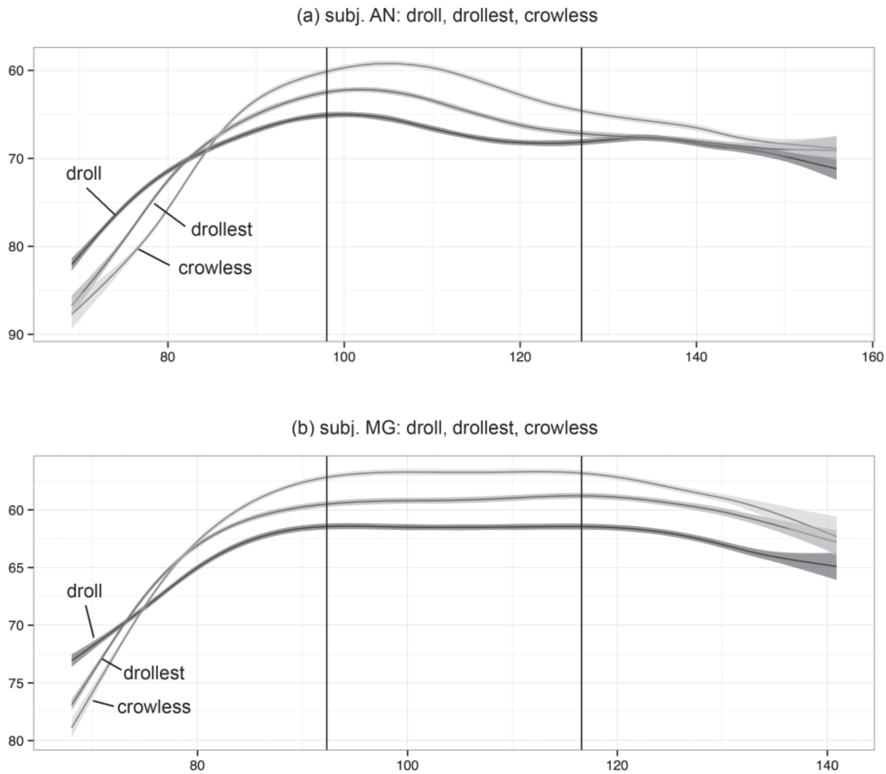


**Fig. 6:** Smoothing splines and 95% Bayesian confidence interval of /l/ of speaker AN (top) and MG (bottom) for pairs of words in the /o/ vowel context: *droll* (stem in black) vs. *drollest* (pre, in dark gray) vs. *crowless* (post, in light grey). The leftmost portion in each image is the tongue dorsum (TD), and the middle portion corresponds to the tongue body (TB). The results were coded as follows: subject AN (top): stm > pre > post (TD), stm > pre > post (TB), subject MG (bottom): pre > post, stm ∗ pre (A ∗ B = inconclusive result) (TD), stm > pre > post (TB).

**Table 3:** Tongue dorsum retraction. 'A > B' indicates that A shows greater retraction than B. The pairs in line with the prediction (the degree of retraction: stem > pre boundary > post-boundary) are highlighted in bold. A ∗ B indicates the 'inconclusive' result, where two curves cross in the middle of the relevant region.

| | AD | MG | WT | JO | AN | RT |
|---|---|---|---|---|---|---|
| /u/ | **pre > post**, stm | **stm > pre > post** | stm = **pre > post** | pre > stm, pre ∗ post | **pre > post**, stm | **pre > post**, stm |
| /o/ | post > pre > stm | **pre > post**, stm ∗ pre | **stm > pre > post** | post > pre = stm | **stm > pre > post** | post > pre > stm |
| /aɪ/ | post = pre > stm | **pre > post**, stm | **stm > pre > post** | **stm > pre > post** | **stm > pre** = post | pre ∗ post, pre > stm |
| /aʊ/ | post > pre > stm | post, **stm > pre** | post, **stm > pre** | post, **stm > pre** | stem ∗ pre, pre ∗ post | post > pre = stm |
| /ɑ/ | stm = **pre > post** | **pre > post**, stm | post, **stm > pre** | **stm > pre** = post | **stm > pre** = post | **pre > post**, stm |

For both speakers, tongue body lowering shows the same pattern: the stem condition (*droll*) is the lowest, pre-boundary intermediate (*drollest*), and post-boundary the highest (*crowless*). However, the results diverge for the tongue dorsum retraction. Speaker AN shows the predicted pattern where the stem condition (*droll*) is the most retracted, pre-boundary intermediate (*drollest*) (a/b ratio = 67%), and post-boundary the least (*crowless*) (a/b ratio = 70%), whereas speaker MG illustrates the inconclusive case where the stem condition (*droll*) crosses with the pre-boundary condition (*drollest*) (a/b ratio = 40%). For the pre-boundary /l/ (*drollest*) versus post-boundary /l/ (*crowless*) comparison, the difference is in the predicted direction: pre > post (a/b ratio = 68%). Table 3 and Table 4 summarize the results of the statistical analysis of tongue dorsum retraction and tongue body lowering, respectively, for each speaker. The notation 'A > B' indicates that A shows greater retraction or lowering than B. Those pairs that conform to the prediction are presented in bold.

Table 3 and Table 4 are further summarized in Table 5 by counting the number of pairs that are consistent with the prediction. Since the SS ANOVA compares a pair of words, i.e., stem vs. pre-boundary, and pre-boundary vs. post-boundary, the results were summarized accordingly.

**Table 4:** Tongue body lowering. 'A > B' indicates that A shows greater lowering than B. The pairs in line with the prediction (the degree of lowering: stem > pre-boundary > post-boundary) are highlighted in bold. A ∗ B indicates the 'inconclusive' result, where two curves cross in the middle of the relevant region.

| | AD | MG | WT | JO | AN | RT |
|---|---|---|---|---|---|---|
| /u/ | **stm > pre > post** | **stm > pre > post** | post > pre = stm | **stm > pre > post** | **stm > pre > post** | **stm > pre > post** |
| /o/ | **stm > pre > post** | **stm > pre > post** | **stm > pre,** pre ∗ post | **stm > pre > post** | **stm > pre > post** | **pre > post,** stm ∗ pre |
| /aɪ/ | **stm > pre > post** | **stm > pre > post** | **pre > post,** stm | **stm > pre** = post | **stm > pre > post** | stm = **pre > post** |
| /aʊ/ | **stm > pre > post** | **stm > pre > post** | **stm > pre > post** | **stm > pre > post** | **stm > pre > post** | stm = **pre > post** |
| /ɑ/ | **stm > pre,** pre ∗ post | post, **stm > pre** | stm ∗ pre, pre = post | **stm > pre > post** | **stm > pre > post** | **stm > pre** = post |

**Table 5:** Number of speakers who showed the predicted pattern for each condition.

| | Tongue dorsum retraction | | Tongue body lowering | |
|---|---|---|---|---|
| | STM > PRE | PRE > POST | STM > PRE | PRE > POST |
| /u/ | 1/6 | 5/6 | 5/6 | 5/6 |
| /o/ | 2/6 | 3/6 | 5/6 | 5/6 |
| /aɪ/ | 3/6 | 3/6 | 4/6 | 5/6 |
| /aʊ/ | 3/6 | 0/6 | 5/6 | 6/6 |
| /ɑ/ | 3/6 | 3/6 | 5/6 | 3/6 |
| | 12/30 | 14/30 | 24/30 | 24/30 |

### 3.1.2 Summary of articulatory findings

Table 5 illustrates interesting articulatory aspects of /l/ in different morphological contexts. Tongue body lowering emerges as a substantial correlate of darkness as compared to tongue dorsum retraction, which has been commonly assumed to be the most prominent feature of dark [ɫ]. The results show that tongue body lowering met the prediction most often across all speakers, whereas a relatively smaller number of such cases were attested for tongue dorsum retraction. The contrast between tongue dorsum retraction and tongue body lowering as a correlate of darkness of /l/ is less surprising when the specific vowel contexts employed in this study are taken into consideration. The vowels preceding /l/ are mostly back vowels with varying heights. It is thus likely that retraction of tongue dorsum might not be as immediately available as lowering of the tongue body, and that speakers implement the morphological differences through fine-grained manipulations of the tongue body. It is also possible that the rough division of the tongue surface is too conservative; if some portion of the tongue body is also included in our ratio measure, it could seem as if retraction is less reliable than it actually is.

Although lowering turns out to be a more reliable correlate of darkness of /l/ in the given vowel contexts, speakers sometimes seem to compensate for the lack of tongue body lowering with tongue dorsum retraction to complete the pattern; where tongue body lowering does not go along with the prediction, tongue dorsum tends to be more retracted in the predicted direction. For example, the pre-boundary vs. post-boundary comparison of speaker AD in the /ɑ/ context was inconclusive for the tongue body, since the two curves cross in the middle of the region of interest (Table 4), but the tongue dorsum instead is more retracted for the pre-boundary than the post-boundary condition (Table 3). Note that speaker AD rarely uses retraction of the tongue dorsum as a correlate of darkness, as compared to other participants.

The articulatory results suggest that the production of intervocalic /l/ is highly sensitive to the morphological boundary; /l/ before the morphological boundary (*droll-est*) is articulatorily darker than /l/ after the boundary (*crow-less*). In addition, the comparison between pre-boundary /l/ and the /l/ in the stem suggests that morphological affiliation is not the only determinant of the articulation of /l/; the stem /l/ (*droll*) is consistently darker than pre-boundary /l/ (*droll-est*). Moreover, this intermediate degree of lowering/backness of pre-boundary /l/ crucially refutes Hayes's assumption that a speaker consistently produces both dark [ɫ] and light [l], and that the gradient property of pre-boundary /l/ comes from the averaging of these categorical variants. As demonstrated in Figure 3, which is representative of all of the utterances, the individual data points show that there are not bimodal sets of /l/s for each morphological type that are fit by the spline, but rather that the spline fits a single overlapping set of utterances. Thus, the results are consistent with Sproat and Fujimura (1993) in that the phonetic implementation of darkness is a gradient continuum within subjects.

## 3.2 Acoustic results

In order to obtain complete acoustic results comparable to the articulatory results (where we examined the tongue dorsum retraction separately from tongue body

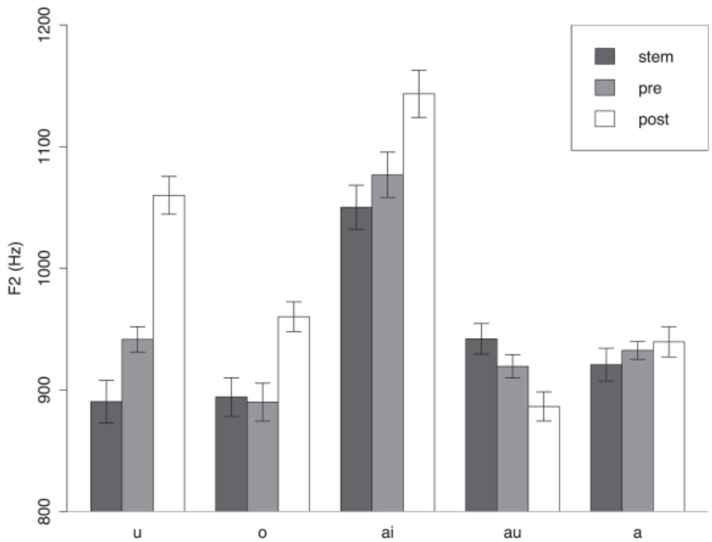


**Fig. 7:** Mean of the minimum F2 (Hz) value during the acoustic realization of /l/. Error bars indicate standard error.

lowering), we assess the acoustic properties of /l/ using F2 and F1 independently. We also examine whether changes in intensity are affected by the morphological differences.

### 3.2.1 F2 measures

Figure 7 presents the minimum F2 values measured during the acoustic realization of /l/. A two-way repeated measures ANOVA was conducted to examine the effect of the morphological contexts on the minimum F2 value. The within-subjects independent variables were morphological conditions (stem form (STEM), pre-boundary (PRE), post-boundary (POST)) and the vowel type (u, o, aɪ, aʊ, ɑ). Subjects were treated as a random factor. The dependent variable was the minimum F2 (Hz) value.

Results show that there was a significant main effect of morphological condition [$F(2, 10) = 5.733$, $p < .05$] and vowel type [$F(4, 20) = 21.757$, $p < .001$], and there was a significant interaction between them [$F(8, 40) = 7.175$, $p < .001$]. To investigate the source of interaction, separate ANOVAs were carried out for each vowel context with the morphological condition as an independent variable. For the vowel /u/, morphological condition was significant [$F(2, 10) = 14.919$, $p < .01$], and a Tukey post-hoc test revealed the following order of F2 values: POST > PRE > STEM. Morphological condition was also significant for the vowels /o/ [$F(2, 10) = 4.208$, $p < .05$] and /aɪ/ [$F(2, 10) = 5.749$, $p < .05$]. For those two vowels, however, the F2 values of stem and pre-boundary condition were not significantly different: POST > PRE = STEM. In contrast, for the vowels /aʊ/ and /ɑ/, the main effect of morphological condition did not reach statistical significance: [$F(2, 10) = 3.052$, $p$ = n.s.] and [$F(2, 10) = .229$, $p$ = n.s.], respectively.

The statistical analysis showed that the predicted pattern was most obviously attested in the non-low or front vowel (or off-glide) contexts, i.e., /u, o, aɪ/, whereas it was neutralized in the low and back vowel contexts, i.e., /a, aʊ/, suggesting a strong coarticulatory effect of the preceding vowels on the phonetic implementation of dark [ɫ].[5] Specifically, in the non-low vowel contexts, /u, o, aɪ/, the F2 value was significantly lower for the pre-boundary /l/ compared to the post-boundary /l/, which in turn is consistent with a pre-boundary /l/ that is produced with greater retraction and lowering than the post-boundary /l/. In addition, the /u/ vowel context showed that the same pattern holds between stem /l/ and pre-boundary /l/; F2 was significantly lower for the former than the latter.

A comparison between the F2 measures and tongue dorsum retraction summarized in Table 3, however, calls for an explanation of a discrepancy between articulation and acoustics. Specifically, in Table 3, the low back vowel /ɑ/ does elicit greater tongue dorsum retraction for the pre-boundary /l/ than for the post-boundary /l/, at least for three speakers, but there was no significant F2 difference between two conditions. This discrepancy is suggestive of the non-linear nature of the relation between articulation and acoustics as proposed by Stevens (1972, 1989). In his proposal for a quantal theory of speech, he argued that there are articulatorily stable regions where a certain amount of articulatory movement does not immediately result in corresponding acoustic changes. For the low back vowel /ɑ/, for example, F1 and F2 are insensitive to the change in the length of the back cavity within a certain range. The absence of the acoustic difference in /ɑ/ despite the articulatory difference may exemplify the quantal effect of this particular vowel.

---

**5** The distinction between /aɪ/ and /aʊ/ may initially be unexpected, but a closer investigation of the ultrasound images of the vowel portion revealed that a clear forward movement is found for /aɪ/ for most speakers, but not for /aʊ/. The absence of the forward movement in the production of the offglide /ʊ/ in /aʊ/ seems to make it pattern with /ɑ/. Though we did not collect lip data, we speculate that rounding of the second half of the diphthong may be enough to give rise to the percept of the offglide without substantial corresponding tongue movement.

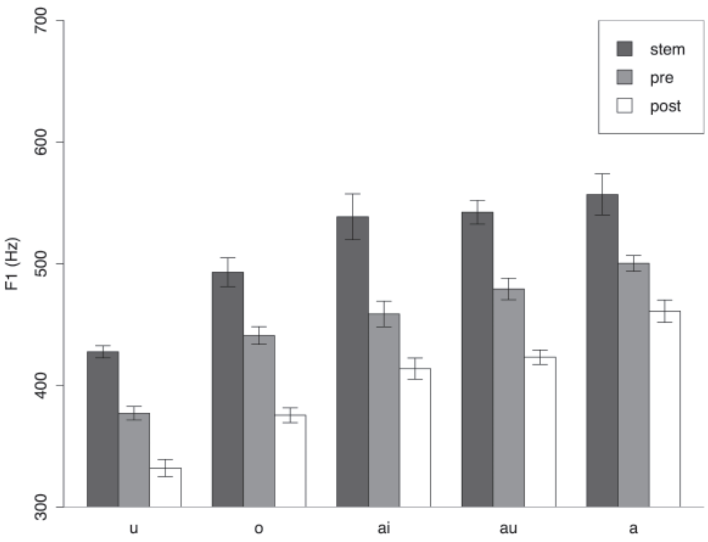


**Fig. 8:** Mean of F1 (Hz) at minimum F2. Error bars indicate standard error.

### 3.2.2 F1 measures

Subsequently, a repeated measures ANOVA with F1 as a dependent variable was carried out to test the statistical difference in F1 values. Results show that there was a significant main effect of the morphological condition [$F(2, 10) = 30.914$, $p < .001$] and the vowel type [$F(4, 20) = 15.895$, $p < .001$], but the interaction between them did not reach statistical significance [$F(8, 40) < 1$]. As predicted earlier, the stem condition showed the highest F1 value ($M = 511.6$ Hz), the pre-boundary condition intermediate ($M = 451.1$ Hz), and the post-boundary the lowest ($M = 401.4$ Hz), regardless of the vowel context. Figure 8 presents the mean F1 values measured at the time when the minimum F2 value was measured.

The results stand in contrast to the results of F2. The coarticulatory effect observed in F2 measures no longer persists in F1 measures; the acoustic implementation of relative darkness of /l/ between pre- and post-boundary /l/ arises differently not only in the non-low vowels contexts (u, o, aɪ), but also in low back vowels (aʊ, ɑ). In particular, the pre-boundary /l/ showed significantly higher F1 values than the post-boundary /l/, and F1 of the stem /l/ is highest of all, in all vowel contexts.

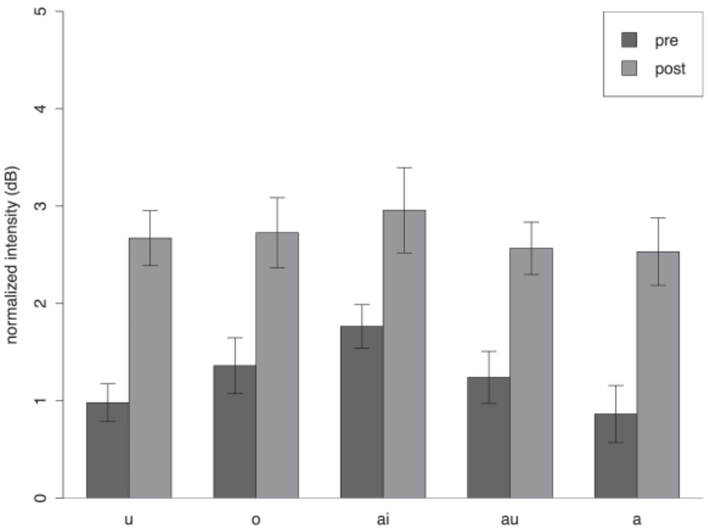


**Fig. 9:** Normalized intensity (dB). Error bars indicate standard error.

### 3.2.3 Normalized intensity

Figure 9 shows the mean of normalized intensity of pre- and post-boundary conditions in each vowel context. A paired *t*-test was conducted to examine the effect of the morpheme context on the normalized intensity, (the difference in intensity between the /l/ and the following vowel /ə/). The results show that normalized intensity of post-boundary /l/ ($M = 2.69$ dB) was significantly higher than that of the pre-boundary /l/ ($M = 1.24$ dB) [$t(299) = -8.273$, $p < .001$]. The effect was uniform for all vowel contexts.

The greater normalized intensity of the post-boundary /l/ indicates that the /l/ in this position is lighter than in pre-boundary position. In addition to the F1 measurement, the intensity results further provide evidence that the morphological affiliation of the /l/ has a significant effect on the implementation of the relative darkness of /l/ in all vowel contexts.

### 3.2.4 Summary of acoustic results

In accordance with the articulatory results, the acoustic findings suggest that the most consistent correlate of the morphological boundary differences is not necessarily lowering of F2, but rather raising of F1, which can emerge as a primary acoustic correlate of a darker [ɫ] depending on the specific vowel contexts. The current study complements the findings in Sproat and Fujimura (1993) that both tongue dorsum retraction and tongue body lowering can be active as an indicator of dark [ɫ]. Both the articulatory and the acoustic results, however, further suggest that individual vowels are instrumental in manipulating the relative contribution of tongue dorsum and tongue body. That is, speakers actively utilize tongue body lowering to attain a darker /l/ that signals the morphological boundary. This is accentuated when the vowel articulation does not allow sufficient room for tongue dorsum retraction.

Moreover, F1 differences characterize a clear three-way phonetic realization in all vowel contexts. For F1, there is a clear separation between the /l/ of the stem form and pre-boundary /l/; the F1 of stem /l/ was consistently higher than the F1 of pre-boundary /l/. Note that this was only true for the /u/ vowel context in F2 measures. This is consistent with the articulatory finding that the degree of tongue body lowering was intermediate for the pre-boundary /l/. For intensity, the difference between the intensity of the /l/ and the following vowel was significantly larger for post-boundary /l/ than for pre-boundary /l/, which is consistent with the claims that light variants of /l/ are more consonantal, whereas dark [ɫ] is more vocalic (Sproat and Fujimura 1993; Gick 1999, 2003). Again, the finding that the pre-boundary /l/ is not identical to either the stem /l/ or the post-boundary /l/, both articulatorily and acoustically, refutes Hayes's assumption that pre-boundary /l/ is implemented categorically as either dark [ɫ] or light [l].

# 4 General discussion

The present study of the production of American English /l/ in coda and intervocalic position tested the prediction that the morphological structure of a word is reflected in speakers' articulatory and acoustic implementation. Inspired primarily by both Sproat and Fujimura (1993) and Hayes (2000), we compared stem-final /l/ in pre-boundary position in a morphologically complex word (e.g., *tall-est*), /l/ in post-boundary position (e.g., *flaw-less*), and stem-final /l/ in word-final position (e.g., *tall*). It was predicted that /l/ in the pre-boundary position would be darker both articulatorily and acoustically than /l/ in post-boundary position. It was also predicted that /l/ in the word-final position in stem forms should be the darkest of all.

## 4.1 Phonetic implications

The results of this study are consistent with Sproat and Fujimura's claim that /l/ is produced on a continuum of darkness. These results also demonstrate that the morphological status of /l/ in pre- and post-boundary position is signaled by differences in phonetic implementation. For most cases, either tongue dorsum retraction or tongue body lowering indicate greater darkness for /l/ in pre-boundary condition as compared to /l/ in post-boundary condition. Likewise, F1, intensity, and (to a lesser extent) F2 are also consistent with a darker pre-boundary /l/. The stem form is the darkest with respect to the phonetic variables corresponding to darkness that were examined. Previous accounts that treat tongue body retraction as the most reliable correlate of /l/ darkness may be partially attributable to the vowels used in those studies; that is, when /l/ is in the context of high or mid front vowels, retraction will be substantial (e.g., Giles and Moll 1975; Sproat and Fujimura 1993; Scobbie and Pouplier 2010). In contrast, Gick (2003) uses a low back vowel and does not find significant differences in tongue dorsum retraction for the conditions investigated using EMMA, including /a#la/, /al#a/, and /al#h/. This finding can be contrasted with those reported by Scobbie and Pouplier (2010), who examine the same environments with EPG data, but using stimuli that contain high vowels instead of /a/. Their findings show that there were at least differences in retraction for word-initial (onset) /l/ versus word-final (ambisyllabic) /l/ and word-final /l/ before a consonant.

Further evidence for the coarticulatory effect of high front versus low back preceding vowels on the articulation of /l/ comes from Proctor and Walker (2012), who use real-time MRI data to examine the articulation of /l/ in the context of preceding and following high vowels like /i/ or low vowels like /ʌ/. The MRI data show that while the tongue tip coronal gesture stays largely the same regardless of vowel context, the tongue dorsum must move much more after /i/ than after /ʌ/. Proctor and Walker's examination of /l/ following low back vowels confirms that there is very little movement of the dorsum from the vowel articulation to the most retracted part of the /l/ articulation. Although in the present work we examine word-medial conditions, the coarticulatory effects found in our results for tongue dorsum retraction are consistent with those previous studies. In our stimuli, most of the vowels preceding /l/ are back vowels, and tongue dorsum retraction may not be readily available when the tongue is already retracted for the articulation of the vowel. Instead, the greater consistency of tongue body lowering as a correlate of /l/ darkness is evident as a result of the particular stimulus set that we tested.

One aspect of Sproat and Fujimura's (1993) findings that has not yet been explicitly discussed is their demonstration that the degree of darkness of /l/ was highly correlated with the duration of /l/; the longer the duration, the darker the

/l/ is. Since it is typically the case that major prosodic boundaries such as a VP phrase boundary elicit longer duration for articulation, /l/s in those positions are produced as relatively darker. In contrast, when /l/ is at a weak boundary, such as an affix boundary, durations are shorter, and /l/s in those positions are relatively light. This leads Sproat and Fujimura to conclude that juncture strength can be reduced to phonetic duration, and it is, therefore, the primary determinant of darkness of /l/. In our study, because of our emphasis on creating semantically plausible and morphologically matched stimuli, there were limitations on our ability to segment the /l/ from the surrounding vowels. That is, /l/ alone could not be segmented because it is intervocalic (and dark [ɫ] is even more difficult to segment than a light [l]), and for some pairs, there were even more approximants adjacent to the VlV sequence that were a further impediment to accurate segmentation (e.g., the pre-/ɑ/ /l/ in *fl̲awless*). Because Sproat and Fujimura use only an /ilɪ/ context, they were able to use rapid F2 movement to help them segment the /l/. Since we are also using back and low vowels, this rapid movement is not always present. Thus, it is difficult for us to confirm whether duration plays a role in our results.

However, a closer look at the results of Sproat and Fujimura suggests that even in their data, there is an effect of the morphological boundary that is not explicitly attributable to a difference in duration. The relevant figure in Sproat and Fujimura is reproduced below in Figure 10. In this figure, duration is marked in seconds along the x-axis, and F2-F1, a metric that encodes the relative darkness of /l/, is given along the y-axis. The lower F2-F1 is, the darker the /l/ is. On the top left corner are those word-internal /l/s that are directly relevant to our study, indicating that they are lighter than /l/s at major boundaries. The word-internal /l/s include /l/s at a productive morpheme boundary (indicated by #), such as *kneel-ing*, /l/s at a non-productive morpheme boundary (indicated by +), such as *tel-ic*, and /l/s within monomorphemic forms (indicated by %), such as Beelik. We use impressionistic groupings (shown by the ellipses in Figure 10) to indicate that the rime duration of these words can be grouped together at approximately 150–200 ms. They do not appear to be significantly different from one another, although Sproat and Fujimura do not explicitly report on this group (and we are not able to carry out statistical testing based on the figure alone). A striking difference, however, lies in their relative darkness. With a few exceptions, /l/s at a productive morpheme boundary (*kneel-ing*) (shown by the dotted ellipse in the figure) are considerably darker than /l/s in non-productive morpheme boundaries (*tel-ic*) and monomorphemic forms (*Beelik*) (shown by the solid ellipse). The coda /l/ before /h/ (indicated by ‘h’), which is equivalent to our stem condition, has substantial variability but is generally darker than the productive morpheme boundary cases (shown by the dashed ellipse).

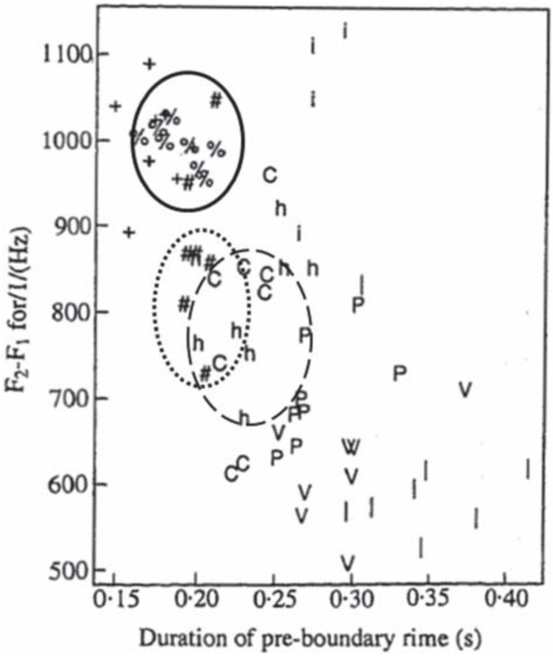


**Fig. 10:** The relation between F2-F1 (darkness parameter) and the duration of the pre-boundary rime (in seconds). This figure is reproduced from Figure 1a of Sproat and Fujimura (1993), but the ellipses have been added for illustration. See the text for description of the ellipses. #: productive morpheme boundary (*kneel-ing*), +: non-productive morpheme boundary (*tel-ic*), %: monomorpheme (*Beelik*), h: /l/ in coda before /h/ (*Mr. Neal Hikkovsky*).

The explanation for /l/ darkness solely based on duration breaks down here, since those three word-internal environments demonstrate no particular difference in duration. Yet, the /l/s in the productive-morpheme boundary (*kneel-ing*) are darker without being longer. When the effect of morphological affiliation is considered, there is a natural explanation for the pattern. The productive morpheme boundary differs from the other two environments, i.e., non-productive boundary or monomorphemic forms, in that there is a free-standing morpheme (*kneel*) which can exert morphophonological pressure such as paradigm uniformity (e.g., Benua 1997; Steriade 2000; Burzio 2005; Hall 2005). That is, the independently occurring dark /l/ in *kneel* requires the intervocalic /l/ in *kneeling* to be similar, which would elicit a darker /l/ in this environment. On the other hand, /l/s at non-productive morpheme boundaries (*tel-ic*) or in monomorphemic forms (*Beelik*) are free from such pressure, resulting in a lighter production of the /l/.

The apparent difference between morphologically complex and simple (or bound) stimuli is readily comparable to the current data. The relative darkness of the pre-boundary /l/ (*tall-est*), both articulatorily and acoustically, as compared to the post-boundary /l/ (*flaw-less*) can be attributable to their distinct morphological affiliation in relation to the free-standing stem /l/ (*tall*). As opposed to post-boundary /l/, the pre-boundary /l/ is in a paradigmatic relation with the stem /l/ (*tall*), which, as a coda, is necessarily dark. Due to this morphophonological pressure, the pre-boundary /l/ is darker than the post-boundary /l/ despite their similar phonological status as intervocalic /l/s.

Although we cannot conclusively rule out duration as a contributing factor in our results, our reconsideration of the Sproat and Fujimura data indicates that those existing data cannot be explained by duration alone. Moreover, if it turns out that there are reliable duration differences in our data, such that the pre-boundary stimuli are always longer than the post-boundary stimuli, presumably the lengthening would still be attributable to the morphological boundary. In that case, the morphological difference would be giving rise to both a duration effect and a difference in darkness, but the interpretation that morphology plays a role in distinguishing these environments would still be warranted.

To conclude this section, a brief comment about the movement of the tongue tip is warranted. While timing of tongue tip raising with respect to tongue dorsum retraction is not reliable for our data, we can report on some findings from the first third of the same plots that we use to evaluate tongue dorsum retraction and body lowering. Since we do not have palate data, we cannot know if the highest curve is actually touching the alveolar ridge. However, we can use these plots to tell whether any of the tongue tips are significantly less constricted than the others at the point of maximal retraction of the tongue dorsum (e.g., *droll* for speaker MG in Figure 6). We would predict that, at this time point, the tongue tip might be most constricted for the post-boundary stimulus (the lightest /l/, e.g., *flawless*) and least constricted for the stem (the darkest /l/, e.g., *tall*). This pattern would either indicate that tongue tip raising lags behind the dorsum retraction for darker /l/s (and so the highest position of the tongue tip would appear in a subsequent frame), or that the constriction is incomplete, both of which would be consistent with previous research. Similar to the way the data is reported in Table 5, this analysis shows that in 16/30 comparisons, the tongue tip of the stem item is significantly less constricted than the pre- and post-boundary stimuli. Two participants, MG and WT, show significantly lower tongue tip positions for the stem for all of the comparisons, and JO shows this pattern for all comparisons except /u/. For the other three participants, the majority of the comparisons show no significant differences. Only in 7/30 comparisons is the post-boundary stimulus (e.g., *flawless*) significantly more constricted than the pre-boundary stimulus (e.g., *tallest*). These cases are evenly distributed among the participants and vowels for both comparisons. These results indicate that, for the comparisons of the intervocalic stimuli, the tongue tip position is not likely to be a reliable indicator of the darkness of the /l/. For the stem word, however, which seems to contain the darkest /l/s in our study, a number of the productions conform to the expectation that tongue tip raising should either follow tongue dorsum retraction or that the tongue tip closure should be incomplete. For three of the six speakers, this pattern was nearly absolute.

## 4.2 Phonological implications

The articulatory and acoustic results of the current study indicate that pre-boundary /l/ has a phonetic implementation that is intermediate between stem and post-boundary /l/. This is in sharp contrast with Hayes's (2000) assumption that speakers produce only either light or dark allophones in both pre- and post-boundary environments, but dark [ɫ] 'more often' in pre-boundary position. According to Hayes, the percept of greater darkness of the pre-boundary /l/ thus comes from a higher proportion of the dark /l/ allophone as compared to the post-boundary /l/. This assumption was necessary for Hayes, since his analysis was specifically designed to model phenomena such as free variation. In his Strictness Band model, a precursor to Stochastic Optimality Theory (Boersma and Hayes 2001), constraints can span over a certain range on the ranking scale, allowing overlap with other constraints. Crucial constraints overlapped each other to allow free variation between two output forms. Relevant to the current study, overlapping of constraints that favor light [l] and dark [ɫ] make it possible to generate two output forms on the surface.

Hayes argues that this is a formal phonological grammar that can generate forms that are intermediate in well-formedness. Strictly speaking, this model cannot generate intermediate forms *per se*, but it ensures multiple output forms on the surface, so that the averaged value can result in intermediate values on the phonetic scale that Hayes discusses. Plots of the individual repetitions of our data, as in Figure 3, indicate that the splines used for analysis are not the average of two bimodal productions of light [l] and dark [ɫ], but rather that any articulatory differences between the stem, pre-boundary, and post-boundary utterances are systematic. The presence of intermediate forms shown in our study not only creates difficulty for Hayes's Stochastic Optimality Theory, but is generally incompatible with many formal phonological theories.

Instead, our data are compatible with the continuous range of darkness found in Sproat and Fujimura (1993), as exemplified in (1). While one interpretation of

Sproat and Fujimura's conditions might suggest that linguistic variables such as phrase boundary, word boundary, morpheme boundary, etc., condition the darkness of the /l/ that is produced, their own claim is that duration is responsible for the bulk of the findings. Yet, we believe that there is a more nuanced interaction between these linguistic variables and duration alone. Linguistic distinctions between different kinds of boundaries are often signaled by duration, so another interpretation is that duration, and hence /l/-darkness, results from linguistic factors. Still, the phonological conditioning environment and its phonetic reflex are confounded in much of their data. However, a closer look at the Sproat and Fujimura results indicates that even within their own dataset, there are darkness differences that cannot be attributable to duration alone, and that morphological boundaries may have an independent contribution. Our results are consistent with the idea that morphological boundaries can likely affect phonetic implementation aside from a pure effect of duration. They are also compatible with Cho (2001), who also shows independent morphological effects on articulatory implementation.

Since we only examine three morphological types in our data, it is tempting to say that there are just a few categorical variants of /l/, and that they are determined by the grammar. However, when situated in the larger picture, it is evident that there are many factors that give rise to gradient levels of /l/-darkness. An extensive discussion of the myriad factors affecting the implementation of /l/, and what this means for the standard view of /l/-allophones, can be found in Scobbie and Pouplier (2010). Ultimately, Scobbie and Pouplier reject both the typical notion that varieties of English contain categorical allophones and Sproat and Fujimura's argument that /l/ allophony cannot be accounted for by any type of phonological representation. Instead, they favor a gestural account for their complex set of findings from /l/ in a variety of environments, including onset position, as a geminate across a word boundary, at the end of a word followed by a vowel-initial word, and at the end of a word followed by either /b/ or /h/. Their gestural explanation relies on differences in coordination and coupling strength as a result of both different boundaries and surrounding gestures.

Following up on Scobbie and Poulier's proposal, one way in which morphology could influence gestural coordination is through the coupling relationships that are established between the gestures of the /l/ and the surrounding vowels. Coupling relationships refer to sets of gestures that are phonologically defined to be coordinated with one another, such as the tongue dorsum (TD) and tongue tip (TT) gestures for /l/ (Goldstein, Byrd, and Saltzman 2006; Saltzman et al. 2008). While we do not have adequate timing data to conclusively establish the appropriate coupling relationships for our stimuli, here we sketch out a possible organization that would lead to a different phonetic realization for our stem, pre-boundary, and post-boundary words. For stem words, like *cool*, the TT and TD gestures corresponding to the /l/ are coupled with one another, and there is no following vowel that must be coordinated with the /l/. Since previous research has shown that there is a negative lag between the TD gesture and the TT gesture, the TD gesture for /l/ is likely also coupled to the preceding vowel. Pre-boundary words, like *coolest*, have the same coupling relations as for *cool*, but because there is a following vowel, the TT gesture (only) is also coupled to that vowel. The coupling of the TD to the preceding vowel and the TT to the following vowel, in addition to the coupling of these two gestures with one another, would reflect the morphological pressure for /l/ to be simultaneously coda-like and onset-like. Finally, for post-boundary words like *coupless*, the TT and TD gestures of the /l/ remain coupled, but there is no coupling of the TD to the preceding vowel and instead, the TT is coupled with the following vocalic gesture. While the precise coupling relationships sketched here may not turn out to be the correct ones, it is nevertheless easy enough to see how differences in coupling relationships among the relevant gestures – presumably influenced by the morphological affiliation of the gestures of the /l/ – could give rise to differences in tongue dorsum and tongue body articulations, which in turn affect acoustic parameters such as formants and intensity.

### 4.2.1 Alternative phonological accounts

Though we have already argued that a categorical allophonic interpretation of our results is unlikely to bear fruit, there nevertheless have been some attempts to account for similar types of data using formal phonological frameworks that assume categorical phonemic representations. At this point, we would like to demonstrate that it is not straightforward to account for our results if they are assumed to be classically allophonic. The relevant related data are those which have been characterized as incomplete neutralization, since it can be argued that the implementation of /l/ darkness is a type of incomplete neutralization. Taking an example from the most well-studied case, final partial obstruent devoicing in Catalan, Dutch, German, or Russian (e.g., Port and Crawford 1989; Warner et al. 2004; Ernestus and Baayen 2006; Dmitrieva, Jongman, and Sereno 2010), underlyingly voiced segments retain voicing to some extent despite an ostensible devoicing rule in word-final position. Instead, devoicing is not complete due to the pressure from the underlying form, which contains a voiced obstruent. Crucially, these competing pressures ensure that neither a completely voiced nor voiceless output form is produced; instead, a phonetic trace of underlying voicing persists, resulting in incomplete neutralization. In our case, the pre-boundary /l/ is the coda of a stem form when the stem is in isolation, but most phonological theories would assume that such an intervocalic /l/ would be syllabified as an onset (e.g., onset maximization; see overview in Blevins 1995). Yet, the fact that pre-boundary /l/ is not as dark as stem /l/ indicates that its realization is affected by its dual status as both a coda and an onset; this suggests that pre-boundary /l/s are not simply syllabified as onsets.

The phenomenon of incomplete neutralization is not just limited to final obstruent devoicing. Similar phenomena have been widely observed in numerous studies, including partial voice assimilation in Russian obstruent clusters (Burton and Robblee 1997), English flapping (Herd, Jongman, and Sereno 2010; Braver 2011), and epenthetic vowel quality in Lebanese Arabic (Gouskova and Hall 2009). Our data joins the list of cases necessitating a formal account for such phenomena.

While full phonological alternative accounts of the results are beyond the scope of this paper, we can point out alternatives to Hayes's stochastic account of categorical light and dark variants that are possible phonological treatments of these results. One theoretical attempt to deal with incomplete neutralization has been proposed by Gouskova and Hall (2009) for the difference between epenthetic and lexical vowels in Lebanese Arabic. Using Optimality Theory with candidate chains (OT-CC), phonetic interpretation is assumed to be able to access the entire derivational chain, and can implement any intermediate stages. In their case, the epenthetic vowel in Lebanese Arabic is believed to be [i], but their phonetic examination revealed that the quality of the epenthetic vowel varies from reduced vowels, such as [ɨ] and [ə], to [i]. Based on this, Gouskova and Hall argue that the derivational chain must include [ɨ] and [ə], i.e., <CC, CɨC, CəC, CiC>, since these are more reduced vowels and therefore less marked epenthetic possibilities than [i]. They further argue that there is interaction between phonetic implementation and the entire derivational chain, which would result in speakers sometimes producing [ɨ], [ə], or even [i] for the epenthetic vowel. Applying this idea to the current data, the phonological derivation would involve an incrementally improving chain, starting from the most dark /l/ to the most light counterpart, e.g., *coolest*, <coo[ɫ]est, coo[ɫ~l]est, coo[l]est>. When phonetics happens to access and implement the intermediate form in the derivation, the grammar generates a phonetically intermediate outcome. Yet, the problem for this approach comes from the fact that the pre-boundary articulation did not show any evidence for the other variants that would be in the chain – dark [ɫ] or light [l] – as there should be if phonetics has access to all the intermediate stages in derivation.

Another possible analysis is to posit 'turbid' representation, as proposed by Turbidity theory (Goldrick 2000). Using Turbidity theory, van Oostendorp (2008) attempts to account for incomplete final obstruent devoicing by positing additional restrictions on GEN (the function that generates candidate output forms in Optimality Theory), such that GEN cannot alter lexical information represented as ‘projection lines’. Projection lines are defined by van Oostendorp as “an abstract, structural relationship holding between a segment and the feature” (p. 1368). For example, GEN cannot generate voiceless [t] as an output form for the underlying voiced [d]. Instead, GEN operates on ‘pronunciation lines’ so that it can generate an unfaithful output form such as devoiced [d̥]. Crucially, this output form differs from voiceless [t] in that, phonetically, it is only partially devoiced. Similarly, pre-boundary /l/ can be assumed to be ‘projected’ as a coda, maintaining the underlying morphological affiliation on the surface, but it is ‘pronounced’ as an onset by preceding a vowel by vowel-initial suffix. The /l/ ‘projected’ as a coda differs from the suffix-initial /l/ which is both ‘projected’ and ‘pronounced’ as an onset; it retains its coda property to some extent even though it is ‘pronounced’ as an onset. As a consequence, the onset /l/ retaining underlying representation of dark /l/ is realized as an intermediate dark /l/, differing from both coda /l/ and the true onset /l/.

Yet, the possible phonological analysis sketched in the Turbidity theory framework also has a drawback. While it may work for the cases examined in this study, it effectively assumes that /l/ before prosodic and morphological boundaries can be treated as categorical variants. This would not be true if there are varying levels of darkness at different boundaries (assuming that prosodic and morphological information can be separated from duration effects, at least in some cases.). A study similar to Sproat and Fujimura (1993), with a greater number of tokens, vowel environments, and morphological contexts, would be necessary to establish whether there is a phonetic continuum or a series of discrete underlying categories for /l/ production.

# 5 Conclusion

Following other instrumental studies that have examined how linguistic factors condition the darkness of /l/ in varieties of English, the results of this study suggest that morphological information must be considered among the influences on the implementation of /l/. Our findings are compatible with Sproat and Fujimura (1993), though a closer look at the duration breakdown in their paper suggests that morphological boundaries and durational influences are probably not totally confounded, as their claim would indicate. The current findings are also compatible with Hayes’s (2000) hypothesis that such paradigmatic effects should exist for /l/ darkness, but they do not confirm his claims that the overall level of /l/ darkness is due to differences in the proportion of categorical variants of light and dark /l/. These findings contribute to the growing literature both on /l/ in particular, and on the interplay between linguistic structure and how it might be signaled by gradient phonetic implementations.

**Acknowledgments:** We would like to thank the members of the Seminar in Phonetics at NYU from Fall 2009 for helping design this study and carrying out pilot recordings. We also thank the members of the NYU PEP Lab, the audience at NELS 41, and the Phlunch group at UCSC for their feedback on this research. This research was partially funded by NSF CAREER Grant BCS-0449560 to Lisa Davidson.

## References

Benua, Laura. 1997. Transderivational Identity: Phonological Relations between Words. Ph.D. dissertation, University of Massachusetts.

Bladon, R.A.W., & Ameen Al-Bamerni. 1976. Coarticulation resistance in English /l/. Journal of Phonetics 4. 137–150.

Blevins, Juliette. 1995. The syllable in phonological theory. In John Goldsmith (ed.), Handbook of Phonological Theory, 206–244. Oxford: Blackwell.

Boersma, Paul, & Bruce Hayes. 2001. Empirical tests of the Gradual Learning Algorithm. Linguistic Inquiry 32. 45–86.

Boersma, Paul, & David Weenink. 2011. Praat 5.2: Doing phonetics by computer. http://www.praat.org.

Braver, Aaron. 2011. Incomplete neutralization in American English flapping: A production study. University of Pennsylvania Working Papers in Linguistics 17(1). Article 5.

Burton, Martha, & Karen Robblee. 1997. A phonetic analysis of voicing assimilation in Russian. Journal of Phonetics 25. 97–114.

Burzio, Luigi. 2005. Sources of paradigm uniformity. In Laura Downing, T.A. Hall, & Renate Raffelsiefen (eds.), 65–106. Paradigms in Phonology Theory. Oxford: Oxford University Press.

Chen, Yu, & Hua Lin. 2011. Analysing tongue shape and movement in vowel production using SS ANOVA in ultrasound imaging. In Proceedings of the 17th International Congress of Phonetic Sciences, 124–127. Hong Kong: City University of Hong Kong.

Cho, Taehong. 2001. Effects of morpheme boundaries on intergestural timing: Evidence from Korean. Phonetica 58. 129–162.

Chomsky, Noam. 1968. The Sound Pattern of English. New York: Harper.

Cruttenden, Alan. 2008. Gimson's Pronunciation of English. London: Hodder Education.

Davidson, Lisa. 2006. Comparing tongue shapes from ultrasound imaging using smoothing spline analysis of variance. Journal of the Acoustical Society of America 120(1). 407–415.

Davidson, Lisa. 2012. Ultrasound as a tool for speech research. In Abigail Cohn, Cécile Fougeron, & Marie Huffman (eds.), The Oxford Handbook of Laboratory Phonology, 484–495. Oxford: Oxford University Press.

Davidson, Lisa, & Paul De Decker. 2005. Stabilization techniques for ultrasound imaging of speech articulations. Journal of the Acoustical Society of America 117(4), pt 2. 2544.

Dmitrieva, Olga, Allard Jongman, & Joan A. Sereno. 2010. Phonological neutralization by native and non-native speakers: the case of Russian final devoicing. Journal of Phonetics 38. 483–492.
Ernestus, Mirjam, & R. Harald Baayen. 2006. The functionality of incomplete neutralization in Dutch: The case of past-tense formation. In Louis Goldstein, Douglas Whalen, & Catherine Best (eds.), Laboratory Phonology VIII, 27–49. Berlin: Mouton de Gruyter.
Espy-Wilson, Carol. 1992. Acoustic measures for linguistic features distinguishing the semivowels /w j r l/ in American English. Journal of the Acoustical Society of America 92(2). 736–757.
Fant, Gunnar. 1960. Acoustic Theory of Speech Production. s'Gravenhage: Mouton.
Gick, Bryan. 1999. A gesture-based account of intrusive consonants in English. Phonology 16. 29–54.
Gick, Bryan. 2003. Articulatory correlates of ambisyllabicity in English glides and liquids. In John Local, Richard Ogden, & Rosalind Temple (eds.), Papers in Laboratory Phonology VI: Constraints on Phonetic Interpretation, 222–236. Cambridge: Cambridge University Press.
Gick, Bryan, Fiona Campbell, Sunyoung Oh, & Linda Tamburri-Watt. 2006. Toward universals in the gestural organization of syllables: A cross-linguistic study of liquids. Journal of Phonetics 34. 49–72.
Giles, Stephen B., & Kenneth L. Moll. 1975. Cinefuorographic study of selected allophones of English /l/. Phonetica 31. 206–227.
Goldrick, Matthew. 2000. Turbid output representations and the unity of opacity. Proceedings of North East Linguistics Society 30. 231–245.
Goldstein, Louis, Dani Byrd, & Elliot Saltzman. 2006. The role of vocal tract gestural action units in understanding the evolution of phonology. In Michael Arbib (ed.), From Action to Language: The Mirror Neuron System, 215–249. Cambridge: Cambridge University Press.
Gouskova, Maria, & Nancy Hall. 2009. Acoustics of epenthetic vowels in Lebanese Arabic. In Steve Parker (ed.), Phonological Argumentation: Essays on Evidence and Motivation, 203–225. London: Equinox Publishing.
Gu, Chong. 2002. Smoothing Spline ANOVA Models. New York: Springer.
Gu, Chong. 2012. gss: General smoothing splines. R package version 2.0-9. http://cran.r-project.org/web/packages/gss/index.html.
Hall, Tracy Alan. 2005. Paradigm uniformity effects in German phonology. Journal of Germanic Linguistics 17. 225–264.
Hall-Lew, Lauren, & Sonya Fix. 2012. Perceptual coding reliability of (L)-vocalization in casual speech data. Lingua 122(7). 794–809.
Hayes, Bruce. 2000. Gradient well-formedness in Optimality Theory. In Joost Dekkers, Frank van der Leeuw, & Jeroen van de Weijer (eds.), Optimality Theory: Phonology, Syntax, and Acquisition, 88–120. Oxford: Oxford University Press.
Herd, Wendy, Allard Jongman, & Joan A. Sereno. 2010. An acoustic and perceptual analysis of /t/ and /d/ flaps in American English. Journal of Phonetics 38. 504–516.
Huffman, Marie. 1997. Phonetic variation in intervocalic onset /l/'s in English. Journal of Phonetics 25. 115–141.
Jones, Daniel. 1956. The Pronunciation of English. Cambridge: Cambridge University Press.
Ladefoged, Peter. 2005. Vowels and Consonants: An Introduction to the Sounds of Languages. Malden, MA: Blackwell.
Lehiste, Ilse. 1964. Acoustical Characteristics of Selected English Consonants. Bloomington, IN: Indiana University.

Li, Min, Chandra Kambhamettu, & Maureen Stone. 2005. Automatic contour tracking in ultrasound images. Clinical Linguistics and Phonetics 19(6/7). 545–554.
Mielke, Jeff, Kenneth Olson, Adam Baker, & Diana Archangeli. 2011. Articulation of the Kagayanen interdental approximant: An ultrasound study. Journal of Phonetics 39(3). 403–412.
Pluymaekers, Mark, Mirjam Ernestus, R. Harald Baayen, & Geert Booij. 2010. Morphological effects on fine phonetic detail: the case of Dutch-igheid. In Cécile Fougeron, Barbara Kuhnert, Mariapaola D'Imperio, & Nathalie Vallée (eds.), Laboratory Phonology 10, 511–532. Berlin: De Gruyter.
Port, Robert, & Penny Crawford. 1989. Incomplete neutralization and pragmatics in German. Journal of Phonetics 16. 257–282.
Proctor, Michael, & Rachel Walker. 2012. Articulatory bases of sonority in English liquids. In Steve Parker (ed.), The Sonority Controversy, 285–312. Berlin: Mouton de Gruyter.
Recasens, Daniel. 2011. A cross-language acoustic study of initial and final allophones of /l/. Speech Communication 54. 368–383.
Saltzman, Elliot, Hosung Nam, Jelena Krivokapic, & Louis Goldstein. 2008. A task-dynamic toolkit for modeling the effects of prosodic structure on articulation. Proceedings of the Speech Prosody 2008 Conference, 175–184.
Scobbie, James M., & Marianne Pouplier. 2010. The role of syllable structure in external sandhi: An EPG study of vocalisation and retraction in word-final English /l/. Journal of Phonetics 38. 240–259.
Sproat, Richard, & Osamu Fujimura. 1993. Allophonic variation in English /l/ and its implications for phonetic implementation. Journal of Phonetics 21. 291–311.
Steriade, Donca. 2000. Paradigm uniformity and the phonetics-phonology boundary. In Janet Pierrehumbert and Michael Broe (eds.), Papers in Laboratory Phonology VI, 313–334. Cambridge: Cambridge University Press.
Stevens, Kenneth. 1972. The quantal nature of speech: Evidence from articulatory-acoustic data. In E. E. David & Peter Denes (eds.), Human Communication: A Unified View, 51–66. New York: McGraw-Hill.
Stevens, Kenneth. 1989. On the quantal nature of speech. Journal of Phonetics 17. 3–46.
Stevens, Kenneth. 1998. Acoustic Phonetics. Cambridge, MA: MIT Press.
Stone, Maureen. 2010. Laboratory techniques for investigating speech articulation. In William Hardcastle, John Laver, & Fiona Gibbon (eds.), The Handbook of Phonetic Sciences, 9–38. Oxford: Blackwell.
Sugahara, Mariko, & Alice Turk. 2009. Durational correlates of English sublexical constituent structure. Phonology 26. 477–524.
van Oostendorp, Marc. 2008. Incomplete devoicing in formal phonology. Lingua 118:9. 1362–1374.
Wang, Yuedong. 2011. Smoothing Splines: Methods and Applications. Boca Raton, FL: CRC Press.
Warner, Natasha, Allard Jongman, Joan A. Sereno, & Rachèl Kemps. 2004. Incomplete neutralization and other sub-phonemic durational differences in production and perception: Evidence from Dutch. Journal of Phonetics 32. 251–276.

# Appendix

The following table includes the stimuli, the phrases containing the stimuli that were read by the speakers, and the scenario sentences that were given to the speakers to contextualize the stimuli.

| Vowel | Stimuli | Phrases | Scenario sentences |
|---|---|---|---|
| /u/ | cool | the cool headphones | Alex was very happy with his new gift. The cool headphones were just what he wanted. |
| | coolest | the coolest person | John got a new skateboard on Friday. He is now the coolest person at school. |
| | coupless | a coupless revolution | Did you hear about the revolutionaries who took over the government without a coup? It was a coupless revolution. |
| /o/ | droll | the droll headmaster | The morning announcements are usually very boring. The droll headmaster sometimes makes them a little more amusing. |
| | drollest | the drollest show | The comic show was very droll. It is the drollest show I have ever seen. |
| | crowless | a crowless garden | Ever since we put the scarecrow in the garden, we haven’t seen a crow at all. We now have a crowless garden. |
| /aɪ/ | vile | a vile headline | The famous actor was angry. A vile headline in the newspaper ruined his reputation. |
| | vilest | the vilest smell | The cheese on the table smells disgusting. It has the vilest smell. |
| | eyeless | the eyeless fish | Fish living in total darkness don’t need eyes. Some of them became the eyeless fish. |
| /aʊ/ | foul | a foul headache | I haven’t been feeling very well. I’ve had a foul headache all day. |
| | foulest | the foulest mouth | That boy always curses and makes crude comments. He has the foulest mouth ever. |
| | cowless | a cowless farm | The cows wandered away from the farm. Now it’s a cowless farm. |
| /ɑ/ | tall | the tall hemlock | The trees in the forest are very old. The tall hemlock is my favorite. |
| | tallest | the tallest building | The building on the corner is taller than all of the others. It is the tallest building around. |
| | flawless | a flawless diamond | The diamond in her ring has no flaws. It’s a flawless diamond. |