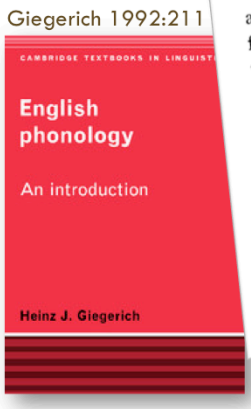
STRUCTURED VARIATION IN ENGLISH L-ALLOPHONY

OCP 12 Jan 2018

Bert Vaux, Cambridge University
Emily Prud'hommeaux, Boston College

Introduction: a textbook example...

Giegerich 1992:211



8.2.1 ***Complementary distribution of allophones: clear and dark /l/***

In many accents of English, for example RP, the lateral phoneme /l/ has two major realisational variants, or allophones: a 'clear' one, transcribed simply as [l], and a 'dark' one – [ɫ]. Both are laterals with alveolar contact; the articulatory difference between the two is that in the former the back of the tongue is lowered while in the latter it is raised towards the velum or retracted towards the uvula (without making contact in either case). Here are some examples for RP:

(1) a. 'clear' [l]: *l*ull, *l*ip, *l*ow, b*l*ind, sp*l*ice, ye*ll*ow, foo*l*ish
b. 'dark' [ɫ]: lu*ll*, hi*ll*, poo*l*, he*l*p, so*l*ve, e*l*bow, litt*l*e

[...]

Recalling the principles of syllabification that were discussed in chapter 6, we note that all the cases of clear [l] in (1a) are sited in syllable onsets while the instances of dark [ɫ] occur in syllable rhymes.

Also Steriade 1988, Rubach 1996, Huffmann 1997, Donegan 2002, Carter and Local 2007 (for some varieties), Johnson and Britain 2007...

English /L/ is traditionally considered to display a literally textbook case of allophony, with a single underlying lateral approximant surfacing in two categorically different ways for latero-typical speakers. One allophone is referred to as "light" or "clear" L, and the other as "dark" or "hollow" L. We can see here [POINT TO SLIDE] that Giegerich believes, like many others, that light L occurs in Onsets and dark L in Rimes. But this is not the only popular picture of L allophony: [NEXT]

Other popular generalizations

Onset vs Coda	Bladon and Al Bamerni 1976, Bladon and Nolan 1977, Wrench and Scobbie 2003, Tsukada et al. 2004, Oxley et al. 2006, Scobbie and Pouplier 2010, Yavaş 2011, Yuan and Liberman 2011
σ-initial vs σ-final	Browman and Goldstein 1995, Pegg and Werker 1997, Krakow 1999, Carter 2003, Nakamura 2009, Zhou 2009
light prevocalically	Trager and Smith 1951, Jones 1960, Giles and Moll 1975, Roach 1991, Cruttenden 1994, Gesuato 1996, Narayanan et al. 1997, Lunn et al. 1998, Wells 1998
dark postvocalically	Matthews 1938, Campbell and Gick 2003
light foot-initially	Jensen 1993, 2000; Olive et al. 1993, Hayes 2000 (some speakers)
word-initial vs word-final	Lee-Kim et al. 2013
light word-initially	Delattre 1971
dark word-finally	Stoddart et al. 1999, Plag et al. 2015

Many see the light and dark L divide in terms of Onset vs Coda, syllable-initial vs. syllable-final, or foot-initial vs. elsewhere, while others believe the variation is conditioned by segmental factors: prevocalic, preconsonantal, word-initial, or word-final. [NEXT]

A selection of non-starters

- "/l/ is realised as [l̥] after word-initial voiceless consonants, as [ɬ] in word-final position, and as [l] elsewhere" (Plag et al. 2015:46)
  - he[ɬ]p, e[ɬ]bow, app[l̥]y, etc.
- RP has "clear /l/ before vowels and dark /ɫ/ elsewhere" (Cruttenden 1994:84)
  - mi[l]ion, twink[ɬ]ing (the verb, for some)
- "the lateral /l/ is velarised when after a vowel or before a consonant at the end of a word" (Ladefoged 2006:76)
  - a[l]ive, twink[ɬ]e…

Many of these analyses can be quickly ruled out by relatively uncontroversial data, such as Plag et al.'s generalization that /l/ is realised as voiceless after word-initial voiceless consonants, dark in word-final position, and light elsewhere. This fails to account for the appearance of dark L in coda clusters, devoiced light L in non-initial clusters, and so on. Once we factor out analyses of this type… [NEXT]

Coda vs Rime

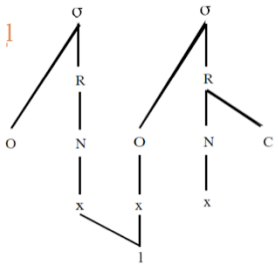
- /l/ → [ɬ] in Coda?
  - problem: syllabic [ɬ̩] (Steriade 1988)
  - "SSBE syllabic /l/ is always dark" (Gimson 1989, Roach 1991, Spencer 1995)

We are left with two viable competitors for the phonological component of L variation: darkening in Coda vs. darkening in Rime.
Steriade 1988 argues for the Rime analysis on the basis of the fact that syllabic Ls, which by hypothesis are in the Nucleus rather than the Coda as shown here for "apple", are dark. If darkening applied only in Codas, we would incorrectly predict the syllabic L in "apple" to be light.
However...[NEXT]

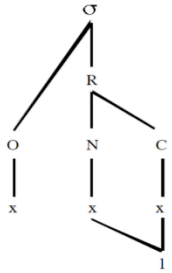
Coda vs Rime

- /l/ → [ɫ] in Rime?
  - problem: syllabic [l] in SSBE (Barry 2000)
  - "When an affix beginning with a vowel is added or the next word begins with a vowel… the lateral may become clear, either syllabic or non-syllabic." (Cruttenden 1994)

(a) light l̩



(b) dark ɫ̩



Barry 2000 points out that not all syllabic Ls are dark for all speakers. In an electropalatographic study he finds empirical support for Cruttenden's 1994 statement that syllabic Ls may be light when immediately followed by a vowel-initial affix or word, as in "tackling" or "middle of". Barry provides various arguments for these Ls still being syllabic, which rules out the possibility of attributing the lightness to these Ls having simply resyllabified as Onsets. As shown in (a), Barry proposes that they do in fact attach to the following Onset, which is responsible for their lightness. However, they remain attached to the Nucleus as well, accounting for their syllabicity. Syllabic Ls that remain dark have the representation in (b) for Barry. Armed with the representations in (a) and (b), Barry proposes that one can account for the range of L distributions in terms of Coda Darkening, but not Rime darkening. [NEXT]

canonical articulations of /l/

articulation	light [l]	dark [ɫ]	referemces
tongue tip	midsagittal apical or laminal gesture towards an anterior target, usually the alveolar ridge (or upper teeth)		Wrench and Scobbie 2003
		reduced apical gesture	Strycharczuk and Scobbie 2015
	alveolar and lateral linguo-palatal contact	primarily alveolar contact	Lehman and Swartz 2000
	tighter constriction between tongue and palate		Giles and Moll 1975
tongue body	fronted and raised	predorsum lowered	Recasens 2011, Lee-Kim et al. 2013
		postdorsum retracted	Giles and Moll 1975, Recasens 2011, Kirkham 2017
		velar or pharyngeal constriction	Delattre 1971, S&F 1993, Lin and Demuth 2015
		increased gesture magnitude	Strycharczuk and Scobbie 2015
timing	TT&TB synchronous	TB before TT	S&F 1993, Browman and Goldstein 1995

The last piece we require before formulating our monolithic rule of L allophony is a more precise notion of what the allophones actually are. Though many articulatory components of light and dark L have been identified in the literature, as summarized here [POINT TO SLIDE], since the influential X-ray microbeam studies by Sproat and Fujimura's in 1993 and Browman and Goldstein in 1995, scholars have tended to view English laterals as a combination of two lingual gestures, a forward movement of the tongue tip and a lowering and retraction of the tongue body, with the two being synchronous in light L, and the dorsal gesture preceding that of the tongue tip in dark L. The phasing relations can be seen in the following figure from Browman and Goldstein: [NEXT]

Browman and Goldstein 1995

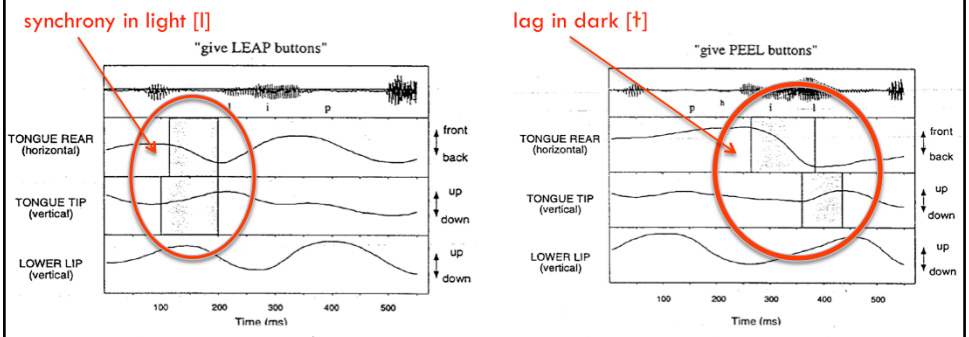
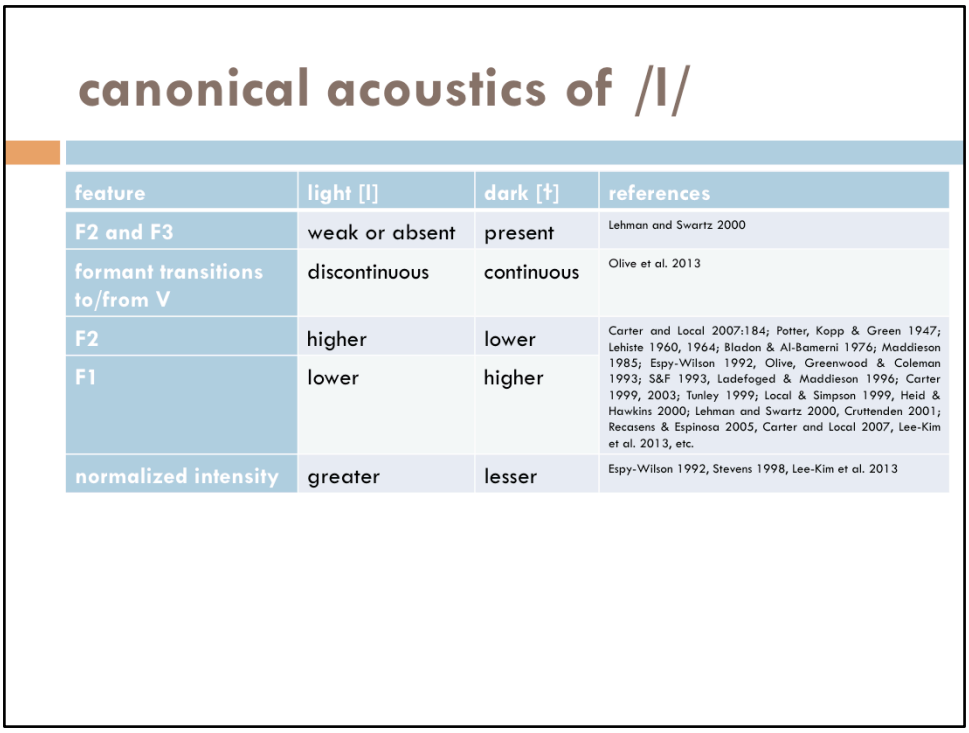
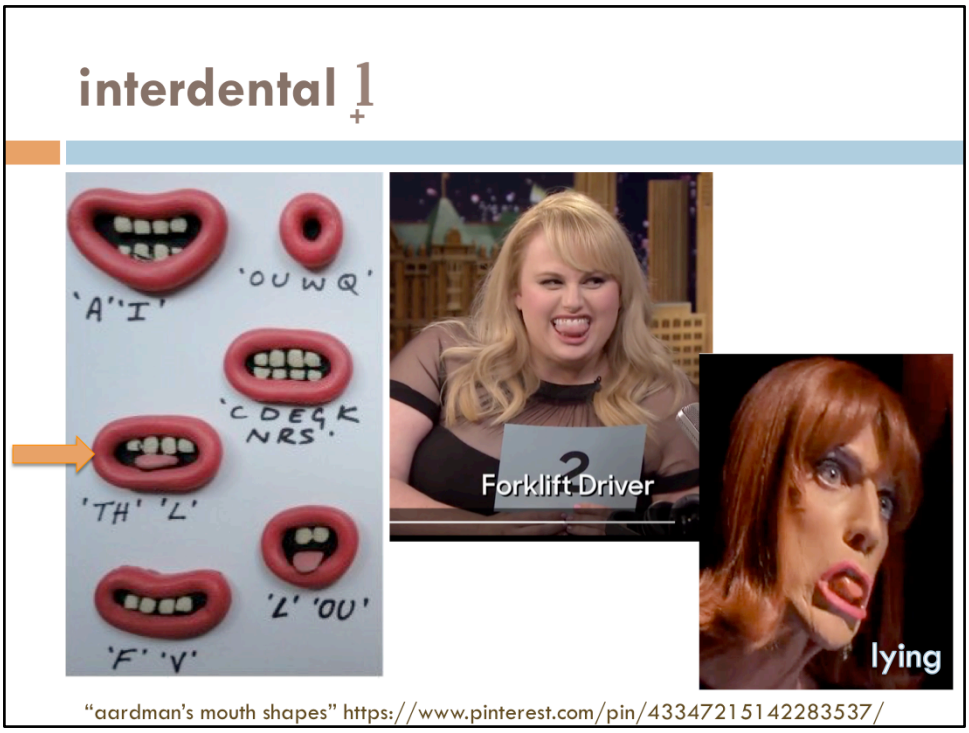


FIGURE 3. X-ray microbeam pellet data showing difference in relative timing of tongue tip raising and tongue dorsum (tongue rear) retraction for initial and final [l]. Tongue tip pellet is the frontmost pellet in the outlined image in Figure 1; tongue rear pellet is in the position indicated by the arrow in that image. (a) "leap" (b) "peel."

Here we can see that in the single speaker investigated, a midwestern American female, the apical and dorsal gestures roughly coincide in the light L of LEAP, whereas the apical lags significantly behind the dorsal in the dark L of PEEL. [NEXT]



Several acoustic consequences of these articulations have been identified in the literature, as shown here [POINT TO SLIDE], but in practice researchers tend to focus their attentions on the second formant.
[NEXT]



Despite the numerous phonetic differences between light and dark L reviewed here, since the time of Sproat and Fujimura 1993 many investigators have run with their idea that there is only one L in English, not two, and this single L category varies according to a range of morphosyntactic, prosodic, and other factors, with the variation taking the form of gradient changes in the phasing and magnitude of the apical and dorsal gestures that characterize L.
In this context it may be relevant that many humans actually have a different system of L opposition that may not submit as easily to a gradient phasing analysis. The most common, in my experience, involves a light L which is interdental rather than alveolar or dental, as we can see here with Wallace and Gromit, Rebel Wilson, and Miss Coco Peru in the film "Trick". [NEXT]

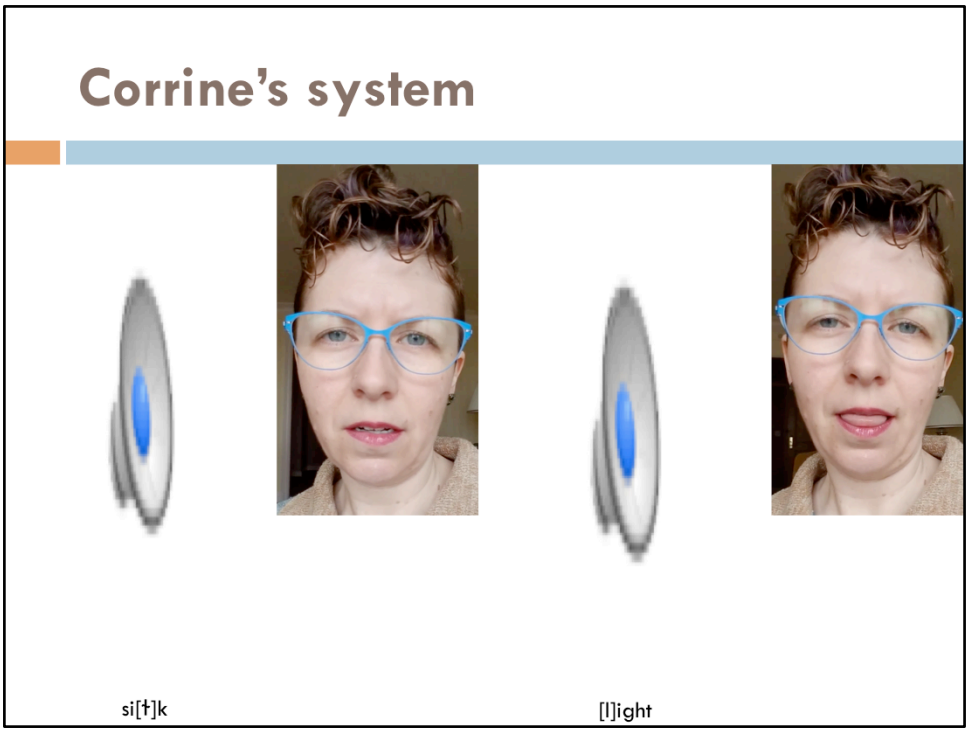
[pictures: Rebel Wilson on Tonight Show with Jimmy Fallon; from Coco: "So... There I was... lying in the middle of my bed with an eye full of cum" (https://www.youtube.com/watch?v=kJeAwTN2XEg)]

apicolabial L

Another surprisingly common implementation of light L is apicolabial, as made famous by Britney Spears. Here she can be seen producing the L of “all I need” in the video for her song “Sometimes”—note that she appears to cliticize the pronoun “I” here and syllabify the L as its Onset.

For interdental and apicolabial L speakers, there appears to be a relatively clear-cut distinction between two types of L: a light allophone with obvious tongue protrusion beyond the teeth, and a more conventional dark allophone. The distinction can bee seen clearly with my old student Corrine Occhino: [NEXT]

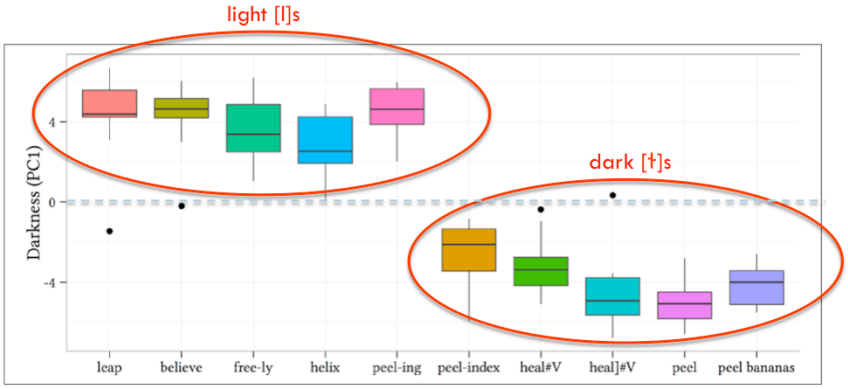
[other image is from A Place in the Sun Home or Away, Dorset vs Croatia]



Here we have images of Corrine producing "silk" with a fairly conventional dark L, and "light" with a visibly different interdental light L. We will return to Corrine's L system later, but for now the key point is that we appear to be dealing in such cases with relatively clear-cut categorical allophony, rather than subtly gradient phonetic variation. [NEXT]

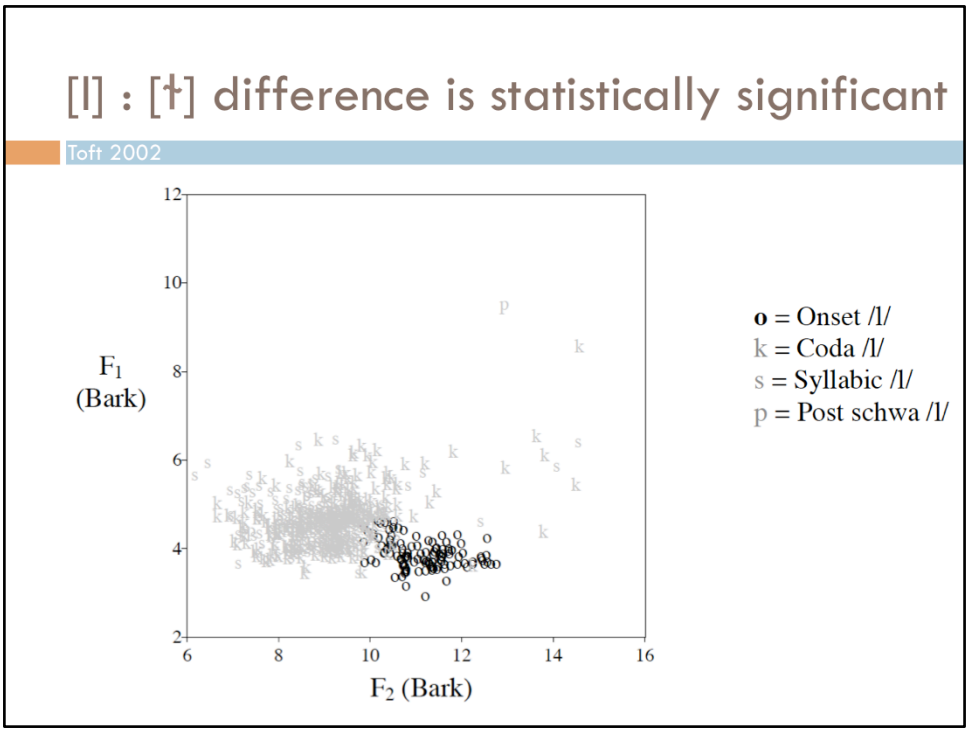
L allophony is categorical in RP

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



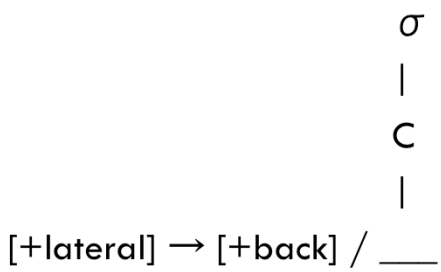
Turton, Danielle. 2017. Categorical or gradient? An ultrasound investigation of /l/-darkening and vocalization in varieties of English. Laboratory Phonology 8.1:1-31.

This is not to say that more conventional L systems cannot also be categorical. Turton has found using ultrasound, for example, that her RP speaker shows a clear-cut categorical opposition in the articulation of light Ls and dark Ls, as you can see here [POINT TO SLIDE]. [NEXT]



Toft 2002 measured the mean formant values produced by six female speakers of British English for four different /l/ allophones: word-initial onset, word-final coda, syllabic L, and post-schwa-L. She found a statistically significant difference between the onset /l/ allophone and the other three variants, with the Onset L having a significantly lower F1 and higher F2. [NEXT]

a traditional RBP-LP analysis



(post-lexical)

On the basis of what we have seen so far, we can postulate a working analysis of what I call the Monolithic L system, which as we shall see later appears to exist in just over half of our sample of 303 speakers. In this system, there is a single underlying lateral phoneme /l/, which surfaces as such when attached to a syllable Onset. When syllabified as a Coda, on the other hand, it undergoes a Darkening process. In terms of traditional rule-based Lexical Phonology this process is post-lexical and therefore can be fed by earlier processes of cliticization and resyllabification which can have the effect of attaching to Onsets some Ls that were previously Codas or Nuclei. [NEXT]

Halle and Mohanan 1985

configuration	light L	dark L
morpheme-internal	Mr Beelik	
level 1 boundary	the beel-ic men	
level 2 boundary	the beel-ing men	
compound, level 3 boundary	the beel equator's amazing	
word boundary		Beel equates the actors

(examples taken from S&F 1993)

In the mid-80s, when Halle and Mohanan begin to formalize the workings of L Darkening in extensive and explicit detail within a Lexical Phonology framework, problems with the monolithic analysis began to appear. Likely because neither Halle nor Mohanan was a native speaker of an L-Darkening variety of English, they based their analysis on the information in Kenyon and Knott's 1944 pronouncing dictionary of American English, and to a lesser extent Jones's Outline of English Phonetics, originally published in 1918. On the basis of these sources, they identified L Darkening as a post-lexical rule that could be bled by a rule of L Resyllabification which they assigned to lexical Level 4 because it followed Level 1 and Level 2 affixation as well as compounding, but preceded the phrasal phonology.
After this point some linguists began to realize that not all speakers of English had the same distribution of light and dark L in these contexts. For example, Nolan et al. 1991 found that 67% (4/6?) of young Cambridge speakers and 90% of old Cambridge speakers produced light L in "call Andy", contrary to Halle and Mohanan's system. Turton 2014 found the same light L treatment, at least with vowel-initial clitics, with her RP speaker. [NEXT]

Problem: variation

- "variation in English /l/ realization has been underestimated in the existing literature, and [...] we can observe phonetic, phonological, and morphosyntactic conditioning when accounting for a representative range of phonological environments across varieties." (Turton 2017)

Sproat and Fujimura 1993 explored the production of L by four Americans and one americanized Brit, namely Sproat, in these and additional morphosyntactic contexts, and found that the picture had changed and was also more complex than what Halle and Mohanan worked with. More recently our understanding of the dimensions along which L can vary phonologically has been enhanced significantly by Hayes in 2000 and by Bermúdez-Otero and his associates in a series of studies beginning in 2007. On the basis of these Turton 2017 concludes that "variation in English /l/ realization has been underestimated in the existing literature, and we can observe phonetic, phonological, and morphosyntactic conditioning when accounting for a representative range of phonological environments across varieties." Where do we go from here? [NEXT]

Systematizing variation

- Hayes 2000: average over ratings by 10 listeners, create gradient metagrammar
- Bermúdez-Otero, Trousdale, Turton, Yuan, and Liberman: factor out phonetic variation, identify structured phonological systems in what remains
  - B-O, Trousdale, and Turton: 4 systems, connected to frequency values generated by Hayes
  - Yuan and Liberman 2011: one system extracted from automated acoustic analysis of 50 years' worth of Supreme Court recordings

Hayes 2000 expanded Sproat and Fujimura's rich set of conditioning environments into a total of 17 that he tested on a larger set of 10 individuals, over which he averaged to create a single gradient-yet-phonological metagrammar. Bermúdez-Otero and his associates took Sproat and Fujimura's data in a different direction, separating out phonetic from phonological effects and then identifying four distinct phonological systems which they interpret as four stages in Bermudez-Otero's Life Cycle of Phonological Processes. Yuan and Liberman started with a different database, generated from automated acoustic analysis of 50 years' worth of audio recordings of Supreme Court transactions, but converged on a rule scattering analysis similar to Bermudez-Otero's, in which a single phonological system of L Darkening in Rimes co-exists with various phonetic processes.
Though each of these studies represents a significant advance in our understanding of L allophony, we feel that each can be improved on in various ways.
Hayes, for example, arguably creates a non-existent straw grammar by averaging over the judgements of 10 individuals rather than identifying the internal workings of each individual system. The same can be said for Yuan and Liberman, who average over 50 years' worth of unidentified outputs of an unidentified set of speakers covering two or three generations. Bermudez-Otero and his associates are in our opinion on the right track in identifying the structure of individual grammars with great precision, but their classification of the range of L systems into four types misses out on a large number of attested and attestable systems, such as the one possessed by Corrine:...
[NEXT]

Corrine's system

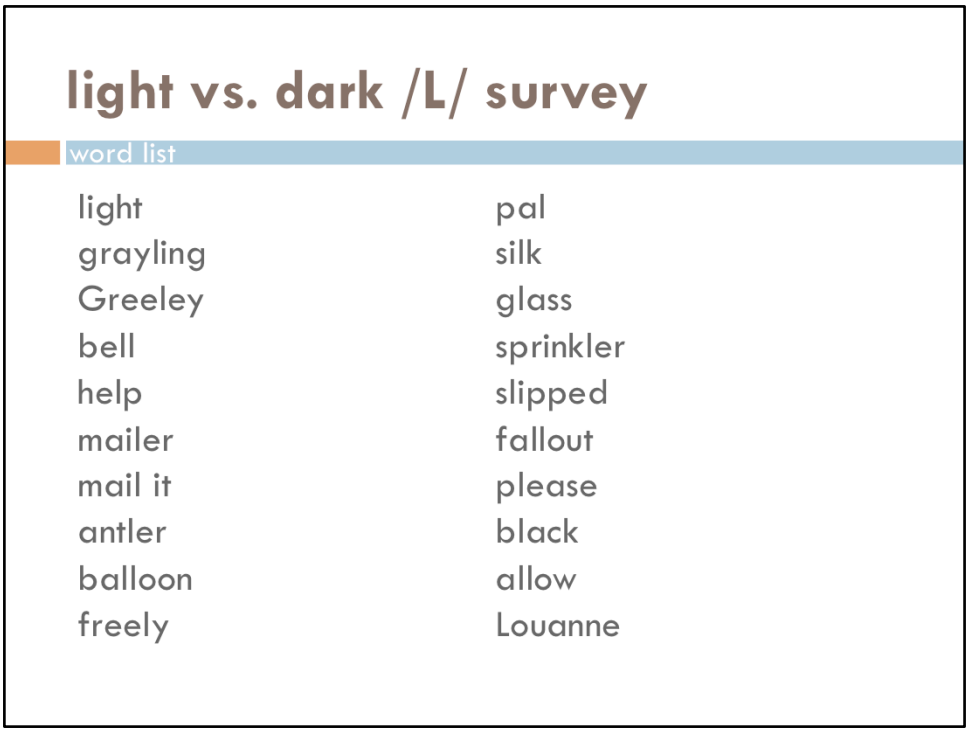
a. forms consistent with classic Onset/Coda distinction:
   - [l]: light, Louanne, please, antler, balloon, Greeley, freely
   - [ɫ]: twinkle, silk
b. no compound- or phrase-level resyllabification except with clitics
   - [ɫ]: fallout, whale office, the whale offered
   - [l]: mail it
     - cf. B-O 2007:10, Selkirk 1996 on cliticization of weak object pronouns to verbal PrWds
c. coda clusters
   - light before labials, dark before dorsals, determined by backness of preceding V before coronals

As I mentioned earlier, Corrine's system is relatively easy to investigate as its two allophones are quite distinct from one another articulatorily and visually. At first blush her system seems to be a straightforward one with light L in Onsets and dark L in Codas, as in (a) and (b). Closer inspection turns up complications that do not sit well with Sproat and Fujimura's or Bermudez-Otero's taxonomies, though, such as the behavior of Coda clusters containing L, as in (c). We hypothesized that careful inspection of a large number of individual L systems should reveal a wide range of subtle and unexpected nanovariations of this sort, resulting from essentially free combination of the phonological variables that have been proposed in the literature to date: morphosyntactic and prosodic differences, vowel backness, and so on.
In 2004 we implemented an online survey designed to address these concerns.
[NEXT]

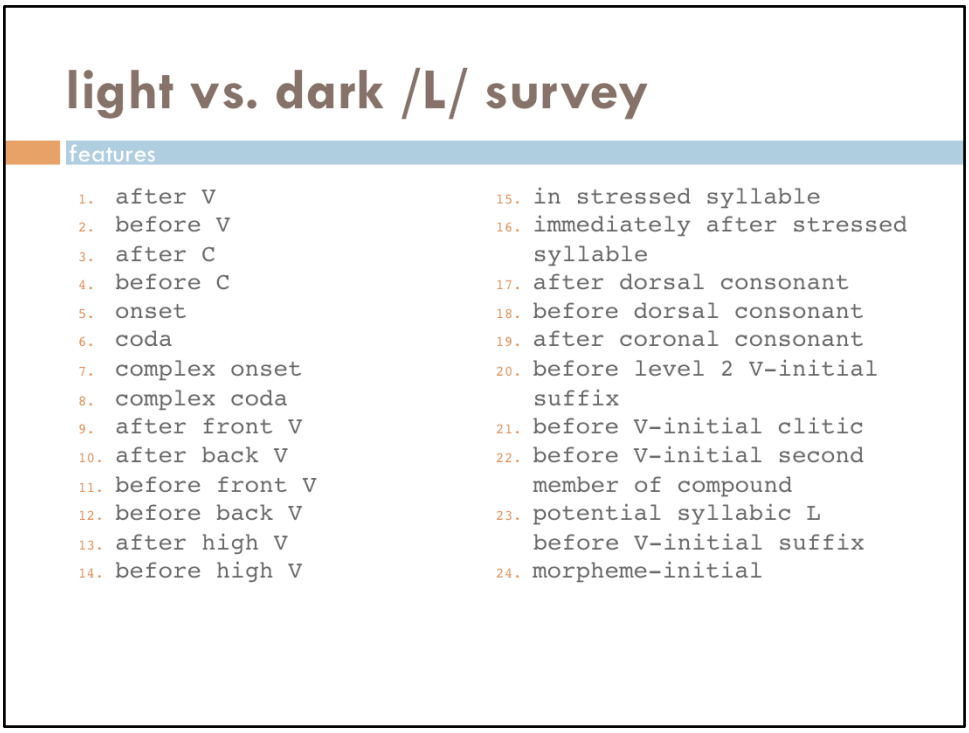
light vs. dark /L/ survey

overview

- 20 words displaying a wide variety of phonological environments: partial Hayes word list with additional forms to test effects of adjacent segmental features.
- 303 linguistically sophisticated speakers of English answered all of the questions.
- Subjects asked to state whether they use light, dark, or both in each word.
- 24 prosodic and segmental parameters investigated for each word.
- Survey ran from 28 April 2004 until 6 March 2008, when the host site became defunct.



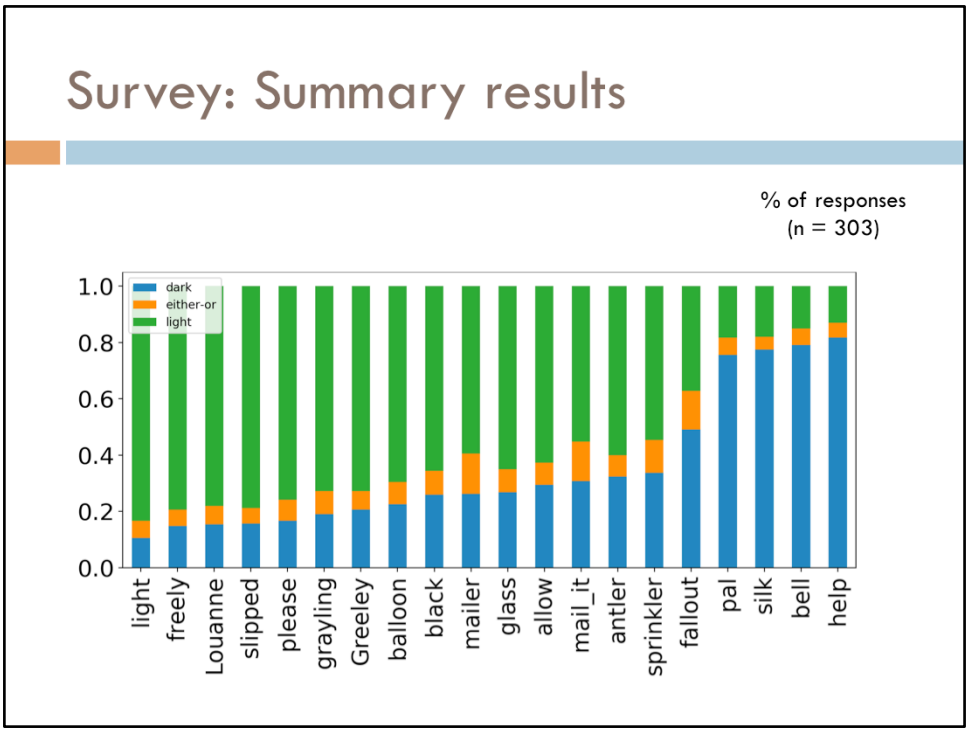
Here are the 20 words we used...[NEXT]



and here are the 24 phonological environments that we examined via our set of 20 words. [NEXT]

Part 2

survey results and analysis



Although coda L tends to be dark and onset L tends to be light, but there are speakers who do not follow this pattern, including roughly 20% of speakers who use or can use light L in word final codas. Also seems to be evidence of features at play beyond just where the L is located in a word or syllable (e.g., antler, allow, fallout).

Numerous possible grammars

- 303 speakers: 255 distinct grammars.
- Significant inter-speaker variation for most environments tested, even "canonical" onset and coda.
- Can these many individual grammars be grouped into larger clusters of similar grammars?

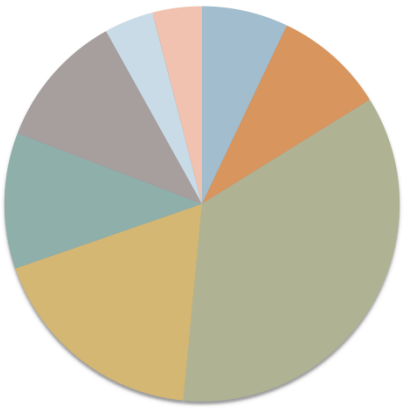
Clustering grammars via EM

expectation maximisation (EM)

- EM process (Dempster et al. 1977):
  1. Initialize: randomly assign subjects to groups.
  2. Determine likelihood of that clustering given responses, and set parameters for group assignment based on that.
  3. Reassign subjects to groups using those parameters.
  4. Repeat steps 2 and 3 until no further improvement.
- Advantages:
  - Doesn't need number of clusters specified in advance.
  - Works with categorical/discrete data.
  - Relatively easy to understand.
- EM groups the speakers subjects into 8 clusters.

FYI or if someone asks, EM does not know about the phonological features. It only knows what each speaker chose for each word.
I used the free software called Weka to do all the clustering and feature selection work.

Eight grammar clusters



1 Generally dark L.

2 Generally light L.

3 Onset-coda rule, with willingness to resyllabify across word boundaries.

4 Onset-coda rule, without willingness to resyllabify across boundaries.

5 L quality affected by features of adjacent segments.

6

7 Willing to accept either L in many different environments.

8

Features associated with clusters

correlation-based feature subset (CFS)

- Do different clusters of speakers use different features for deciding between light and dark L?
- Feature selection with machine learning: **CFS** (Halle, 1998).
  - Finds features that individually best predict outcomes.
  - Considers correlations between features.
  - High correlation = redundancy, so it picks informative features that make independent contributions.
  - Important for our features (e.g., `before high vowel` will always co-occur with `before vowel`).

This is how I worked the data. I looked at one cluster at a time. I took each word/speaker pair in that cluster and its 24 binary feature values and what kind of L that speaker picked. This set of data vectors (i.e., 20 per speaker) is what gets fed to CFS, and then CFS finds the features that best predict which type of L should be used for speakers in that cluster. CFS doesn't bother with redundant features, which is great in our case because many of our features are highly correlated, e.g., before V and before high V.

Clusters 1&2: All or nothing

- **Cluster 1**
  - Observations: almost **always dark L**. Occasional instances of light L in *grayling, Greeley,* and *freely*.
  - Key Features: `after front vowel, after high vowel.`
- **Cluster 2**
  - Observations: **light L almost exclusively**. Dark L appears only in *antler* and *sprinkler*, and then only rarely.
  - Feature: `speaker ID`. No apparent phonological basis for always choosing light L, just idiosyncratic.

There are people who always choose one or the other.

Clusters 3&4: Traditional

- **Cluster 3**
  - Observations: canonical onset-coda rule **with resyllabification** of word-final /l/ before V-initial words, though free variation in *mail it, fallout*.
  - Feature: `before vowel`
- **Cluster 4**
  - Canonical onset-coda rule **without resyllabification** of word-final /l/ before V-initial words.
  - Features: `after vowel, before vowel, before consonant, onset, coda, morpheme-initial`.

What I mean is that all /L/ are light except for those in word-final coda position.

Clusters 5&6: Adjacent segments

- **Cluster 5**
  - Observations: dark L in coda and **when adjacent to back vowel** (*mailer, antler, balloon, sprinkler, fallout, allow, Louanne*).
  - Features: `before vowel, onset, complex onset, after back vowel, before front vowel, after high vowel, before high vowel.`
- **Cluster 6**
  - Dark L in coda, in **clusters**, and in the words *fallout* and *allow*.
  - Features: `before vowel, before high vowel, in stressed syllable, morpheme-initial.` (`Morpheme initial` and `before vowel` correlate highly with complex onset.)

For features for cluster 6: Recall that CFS looks for non-correlated sets of features. A negative value of morpheme-initial will be highly correlated with a positive value for complex onset; complex onset is also highly correlated with before vowel, since an /l/ will always be the last member of a complex onset. Hence, the notion of complex onset is likely being captured by these other features.

Clusters 7&8: Not too picky

- **Cluster 7**
  - Observations: light L before high vowels; dark L in coda position; both elsewhere.
  - Features: `after high vowel, before high vowel, morpheme-initial.`
- **Cluster 8**
  - Observations: light L pretonically; dark L in codas; both elsewhere.
  - Features: `before vowel, onset.`

Classification of L quality

- Can we determine a particular speaker's L preference for a particular word given the described set of features for that word?
- Baseline evaluation: always guess light L because it is more common in our data.
- Compare this to a classifier built using:
  - only the phonological features
  - above plus the speaker's ID number and the cluster number that speaker was assigned to

Random forest classifier accuracy

- Baseline F1: 0.48
- Phonological features alone: 0.59
- With speaker ID and cluster info: 0.71

- Cluster membership and speaker ID combine with phonological features to produce a reasonably accurate prediction of an individual's L preferences.

- Most helpful features: `coda, onset, before_V, clusterID, before_C, complex_coda, before_high_V, morpheme_initial`

Nice to note that the most helpful features include (1) syllable/word position, (2) cluster membership, and (3) segment features like high V

Feature selection

correlation-based feature subset (CFS)

- Do different clusters of speakers use different features for deciding between light and dark L?
- If so, which of our 24 features are important? Two ways to figure this out
  1. Human observation of patterns.
  2. Feature selection with machine learning: **CFS** (Halle, 1998).
     - Finds features that individually best predict outcomes.
     - Considers correlations between features.
     - High correlation = redundancy, so it picks informative features that make independent contributions.
     - Important for our features (e.g., `before high vowel` will always co-occur with `before vowel`).

This is how I worked the data. I looked at one cluster at a time. I took each word/speaker pair in that cluster and its 24 binary feature values and what kind of L that speaker picked. This set of data vectors (i.e., 20 per speaker) is what gets fed to CFS, and then CFS finds the features that best predict which type of L should be used for speakers in that cluster. CFS doesn't bother with redundant features, which is great in our case because many of our features are highly correlated, e.g., before V and before high V.

Comparison with previous studies

"Free variation" in Hayes (2000)

- Hayes: ambisyllabic environment of /l/ in *Greeley* "evokes free variation between light and dark"; supported by the close listener scores (1.8 for light /l/, 2.7 for dark /l/).
- If it's free variation, we expect "both light and dark" choice to be quite frequent in our survey.
- In fact, frequency of "both" for *Greeley* (7.3%) is essentially average (7.6%). Very few acknowledge using both forms.

  ⟹ **Perception ratings ≠ choices in production.**

B-O's 4 stages vs our results

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

		leap	*helix*	*peel-ing*	*heal it*	*peel*	
24%	RP	[l]	[l]	[l]	[l]	[ɫ]	Cruttenden (2008)
9%	Am. Eng. 1	[l]	[l]	[l]	[ɫ]	[ɫ]	Sproat and Fujimura (1993)
4%	Am. Eng. 2	[l]	[l]	[ɫ]	[ɫ]	[ɫ]	Olive et al. (1993)
2%	Am. Eng. 3	[l]	[ɫ]	[ɫ]	[ɫ]	[ɫ]	Hayes (2000)

Emily	[l]	[l]	[l]	either	[ɫ]
Corrine	[l]	[l]	[ɫ]	[l]	[ɫ]

light, Greeley, mailer, mail_it, bell
everyone else 187
Other popular choices:
22243 15 (mail_it is either or)
22222 15 (they are all light L)
22323 13 (mailer=dark, mail_it=light, bell=dark)
23323 9 (light and mail_it=light)
22443 7 (mailer, mail_it=either-or
23233 7 (light, mailer light, others dark)
33333 6 (all dark L)
23333 6 (all dark L except light)

B-O's 4 stages vs our results

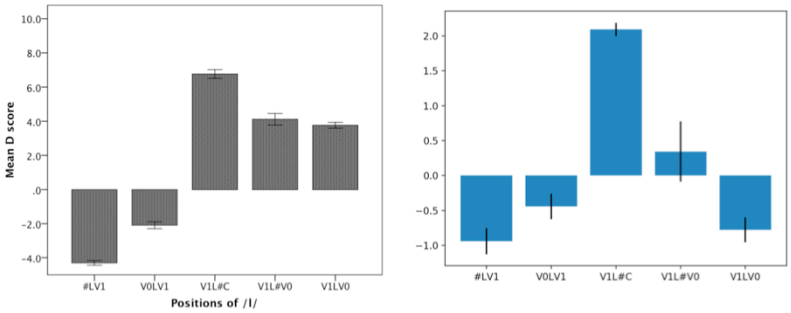
Table 3: /l/-darkening in different morphosyntactic environments. Adapted from Bermúdez-Otero (2007).

	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)

- These systems below were more common in our data than Am. Eng. 2 or Am. Eng. 3

[l]	[l]	[l]	either	[ɫ]
[l]	[l]	[l]	[l]	[l]
[l]	[l]	[ɫ]	[l]	[ɫ]

comparison with Yuan&Liberman 2011



#LV1	light, Louanne, glass, slipped, please, black
V0LV1	balloon, allow
V1L#C	pal, silk, bell, help
V1L#V0	mail_it, fallout
V1LV0	mailer, Greeley, grayling, freely

D score = log(prob of light L) − log (prob of dark L). To make something equivalent to what Yuan and Liberman had, I took log (percent light) − log (percent dark). The big difference between our results and theirs is that the V1LV0 examples, where you have an intervocalic L that can be syllabified in different ways, tend to be LIGHT in our survey, but DARK in the Y&L acoustic analysis. I suspect that at one reason is that in our data, three of the four examples of this cross morpheme boundaries. Y&L don't mention how often this happens in their data, but I suspect it's not as often as 75% of the time (e.g., more things like yellow, willing, etc.) They argue, I think, that morphosyntactic stuff doesn't matter − it's just syllables. (I think.) But I don't see anything in the paper where they actually look at the morphology stuff.

Conclusions

- first large-scale survey and review of L nanosystems
- refined picture of phonological variables that can factor into L systems
- speakers assemble more complex and varied L systems in a more nuanced yet systematic way than previously thought
- corrective to hypo- and hyper-variationist perspectives on L allophony

To conclude, we have suggested that it is important to study the phonological systems of individuals both in detail, as is done for example by Oxley et al. 2006 and Turton 2017), and on a large scale, as is done by Yuan and Liberman 2011. Consequently we conducted the first large-scale investigation of L nanosystems, which we believe leads to a more refined picture of the range of phonological variables that can factor into L systems, and the ways in which they can combine. Our findings suggest that English L allophony is not monolithic, but rather shows extensive phonetic and phonological nano- and micro-variation. In our online survey of 303 speakers, for example, we found both order and diversity, with 8 main clusters of L system but hundreds of other variants.

Our findings also suggest that the nature of L allophony in English is not so simple as a single phonological process, a combination of a phonological process and a phonetic one (Yuan and Liberman 2011), or a single prosodically-organized hierarchy of four processes (B-O). Instead, while there does appear to be a rich array of phonological variables available, as Sproat and Fujimura 1993 and Hayes 2000, etc. suggest, but speakers assemble systems from these variables in a more nuanced yet systematic way than previously thought, as we saw for example with Corrine's system and its unexpected subdivision of Coda clusters.

We hope that large-scale investigation of nanovariation can serve as a corrective to hyper and hypovariationist perspectives on L allophony.

The term **Hypovariationist** refers to the fact that most studies of L allophones focus on extremely small numbers of speakers, e.g. 1 for Browman and Goldstein 1995 and Gick 2003, 2 for Loevenbruck et al. 1999 and Nakamura 2009, 3 for Giles and Moll 1975 and Barry 2000, and 5 in Sproat and Fujimura 1993. The small numbers are understandable given the imaging techniques employed in these studies. HOWEVER, given the abundant evidence we have seen for extensive micro- and nano-variation in the realization of L, we should be careful in assigning too much weight to publications that seek to generalize to English from studies of