Multiple Measures of L-Vocalization

Lauren Hall-Lew
University of Edinburgh

Sonya Fix
New York University
Overview

• What is L-vocalization?
• How can we measure it?
• How consistent is impressionistic coding of L-vocalization across multiple raters?
  
  What factors impact this reliability?
• How might perceptual studies help inform the continuing process of developing an acoustic measure for L-vocalization?
L-vocalization as a sociolinguistic variable

• *In U.S. English*:
  

• *In the UK, Australia, & New Zealand*

Articulatory Variants

• Vocalized L perceptually, acoustically, and articulatorily resembles a back vowel, semi-vowel, or voiced glide (Hardcastle & Barry 1989)

• L-vocalization of coda /l/: “the lack of articulatory closure?”
  – L = a complex segment, lots of available variation
  – EPG evidence for maintained lateral articulation (Keating, p.c.)
Articulatory Variants

Vocalization may best be seen as just one point on a continuous/gradient range of L articulation

more consonantal $\rightarrow$ more vocalic

light L $\rightarrow$ dark L $\rightarrow$ vocalized L $\rightarrow$ deleted L

“The decision as to whether a given alveolar pattern is to be regarded as an /l/ articulation or a full /l/ realization is to a certain degree subject to the ultimately arbitrary criteria governing the classification of borderline cases.” (Hardcastle & Barry 1985)
Articulatory Methods

*Electropalatography (EPG)* (Hardcastle & Barry 1989; Wright 1989; Wrench & Scobbie 2003; Scobbie & Pouplier 2010)

*Electromagnetic articulography (EMA)* (Wrench & Scobbie 2003)

*Magnetic resonance imaging (MRI)* (Gick, Kang & Whalen 2002)

*Microbeam* (Sproat & Fujimora 1993)

*Ultrasound* (Gick 2002; Wrench & Scobbie 2003; Scobbie, Stuart-Smith & Lawson 2008)
Acoustic Measures

• Vocalization is notoriously difficult to measure acoustically, since velar /l/ shares a vocalic articulatory gesture with /ɔ/ (Sproat & Fujimura 1993)

• Acoustic methods also are notoriously elusive for capturing velarization vs. vocalization, particularly across all vocalic environments (see, e.g., Timmins and Stuart-Smith 2004).
'male' (_side) L-full

'scale' (_#) Vocalized
Acoustic Measures

- Dodsworth, Plichta & Durian 2006
  - $\Delta$ formant amplitude (preceding vowel vs. L)
  - Fix (2008) & Hall-Lew (2010) were not able to replicate the auditory/acoustic correlation to the same degree of accuracy
  - Not a method that is easily adaptable for syllabic L
  - Not well-controlled for various context effects
Impressionistic Coding of L

• The majority of studies operationalize vocalization via perceptual coding
  – Typically a 3- or 4-level Likert scale
  – Range from L with alveolar contact/closure to fully-vocalized or deleted L

(Ash 1982; McElhinny 1999; Horvath & Horvath 2002; Timmins et al., 2004; Fix 2004; Dodsworth 2006; Stuart-Smith et al. 2006; Durian 2008)
Motivations for cross-coder perception study

- Coders report a high degree of difficulty and doubt when coding L
- Impressionistic coding forces a perceptually gradient variable to fit discrete categories
- Subjectivity in coding based on coders’ use of and exposure to L vocalization in their own native dialects
- Subjectively in coding from speaker to speaker based on a given speaker’s range of L articulation.
- Vocalized variant can take several forms: back rounded vowel, a voiced glide, a schwa, or nothing.
Motivations for cross-coder perception study

• Study design (Yaeger-Dror, et al., 2009)

• Goal to mimic the perceptual coding process sociolinguists undergo when collecting vocalization data from interview recordings

• Recordings from two speech communities; respondents from various backgrounds

• Questions: How consistent are coders?
Pilot Studies

Pilot tasks ‘A’ & ‘B’

• 62 and 57 tokens respectively

• Pilot coders:
  • 1 new coder, both A & B
  • 1 new coder only A
  • 1 new coder only B

• All 3 pilot coders were asked to provide only one code for each token after listening to it in both syllable and sentence environments

• Stimuli were grouped by speaker
Token Selection

Tokens gathered from sociolinguistic interviews from 2 regionally and ethnically distinct communities in the U.S. that variably vocalize L:

- **Lauren’s data**: Chinese American and Japanese American speakers in San Francisco, California
- **Sonya’s data**: African Americans and white South Midland vernacular speakers with strong contact with African Americans in Columbus, Ohio
Token Selection

30 tokens selected from each regional data set across equal distributions of linguistic environments and across a range of 10 female speakers (5 speakers per regional data set, 60 tokens total)

Tokens selected based on their occurrence in the speech stream (first ones selected first); background noises, etc., minimized when possible

Speakers selected based roughly on researcher impression about their extent of use of vocalized L
Token Selection

Preceding environments:

- L / lax front vowels__ (BIT, BET; N = 14)
- L / tense front vowels__ (BEAT, BAIT; N = 6)
- L / high back rounded vowels__ (BOOT, PUT, N = 10)
- L / mid back rounded vowels__ (BOAT; N = 10)
- L / low vowels__ (BOT, BOUGHT; N = 10)
- L / syllabic (consonant__) (BOTTLE; N = 10)
Excluded Tokens

- durations of V+/l/ below 60ms
- ambisyllabic /l/
- known variable tokens like *palm, folk*, etc.
- obvious instances of /l/-insertion

*i.e.*, those tokens which would also be discarded from a typical quantitative analysis of vocalization
Token Presentation

Internet-based survey

www.surveygizmo.com
Token Presentation

Two contexts:

• ‘sentence’
  ▪ the syllable containing the /l/ is presented with several syllables of speech preceding and following it, roughly approximating the intonational phrase in which it’s embedded

• ‘syllable’
  ▪ the syllable containing the /l/ is presented with only one syllable of speech preceding and one syllable of speech following it
Study Design

• Developed based upon the a 3-stage pilot study with 5 (total) sociolinguists

• 60 tokens
  • Region: 30 San Fran, 30 Columbus, OH
  • 10 speakers total (5 from each region)
  • 6 tokens for each speaker, 1 token from each preceding vowel environment

– Tokens semi-randomized
Perceptual Coding Task

We coded tokens individually for both syllable and sentence environments (which sometimes differed, for us) according to this 4-point scale:

1. Definitely /l/-ful
2. Some /l/ vocalization
   (but more /l/ realization than vocalization)
3. Stronger vocalization of /l/
   (more vocalization than realization)
4. Definitely vocalized
Aim of Research

This is a study of the perception of coda-/l/ vocalization in varieties of US English. We hope to find out more about what factors go into the way phoneticians code /l/ vocalization auditorily, comparing these results to acoustic measures at a later stage. Your taking this survey will be an important contribution to the study of the social and phonetic variation of English liquids.

Participants' tasks

You will be asked to complete the following tasks:
1) basic demographic questionnaire (about 3-5 minutes)
2) training exercise (about 5 minutes)
2) listening and coding survey (about 20 minutes)
First, we will train you.

Listen to the following stimuli. Note that for each word, there are two sound files — the first is of the /l/ token with the words or syllables immediately preceding and following the token. The second is a longer sound clip of the token in the context of its intonational phrase.

The sound clips on this page will all play in succession. You can pause the player at any time while a sound clip is playing (but not when it’s transitioning between clips). To listen to a particular token multiple times, interrupt the parade of sounds by clicking again on that token’s link. You can play the sounds as many times as you like.

The following stimuli are examples of fully-realized (or clear) /l/. We would then give a coding rating of “1” (fully-realized /l/) or perhaps “2” (some /l/ vocalization but /l/ more realized than vocalized). Please listen to these examples carefully, as they come from dialect areas that you may not be accustomed to.

For each stimuli, consider the following scale that you will use to code in the next portion of this survey:

1: Fully-realized /l/
2: Some vocalization of /l/ (but more /l/ realization than vocalization)
3: Stronger vocalization of /l/ (but some /l/ realization)
4: Fully-vocalized /l/

ALL

- because ALL our
- never noticed growing up because ALL our friends were Asian so

Comments (optional):

HELPS

- which HELPS out
- which HELPS out

Comments (optional):
**Coder Training**

**Perception of L-Vocalization Training**

Now that you've had chance to listen to some of the clear /l/ from our samples, it's time to listen to some /l/ vocalization.

(Again, the sound clips on this page will all play in succession. You can pause the player at any time while a sound clip is playing (but not when it's transitioning between clips). To listen to a particular token multiple times, interrupt the parade of sounds by clicking again on that token's link. You can play the sounds as many times as you like.)

The following stimuli are examples of strong /l/ vocalization. We would them give a coding rating of "4" (fully-vocalized /l/). Please listen to these examples carefully, as they come from different dialect areas, and it may be that none of them sound like the kind of vocalization you are used to hearing in your area.

For each stimuli, consider the following scale that you will use to code in the next portion of this survey:

1: Fully-realized /l/
2: Some vocalization of /l/ (but more /l/ realization than vocalization)
3: Stronger vocalization of /l/ (but some /l/ realization)
4: Fully-vocalized /l/

---

For each stimuli, consider the following scale that you will use to code in the next portion of this survey:

1: Fully-realized /l/
2: Some vocalization of /l/ (but more /l/ realization than vocalization)
3: Stronger vocalization of /l/ (but some /l/ realization)
4: Fully-vocalized /l/

**FULL**

- a FULL time
- I can get a FULL time job

Comments (optional):

---

**WHOLE**

- the WHOLE city
- seemed to think the WHOLE city was burned

Comments (optional):
Sample token from the online task

Perception of L-Vocalization
Coding Task: 1 of 60

1. BOTTLED

- the BOTTLED water
- water, the BOTTLED water that I had

   1 - Fully realized /l/
   2 - Some /l/ vocalization (but /l/ more realized than vocalized)
   3 - Stronger /l/ vocalization (but some /l/ realization)
   4 - Vocalized /l/

Comments:
Participant Recruitment

- 30 linguists (phoneticians, sociophoneticans, and sociolinguists) were invited by email to participate in the coding task.

- 23 respondents completed the task.
  - All but 2 respondents are L1 English speakers.
  - Most respondents are from the US and UK.
Results

• Our survey yielded perceptual codes for 60 tokens for 23 participants = 1380 tokens
Frequency of Code by Preceding Vowel

- 1 (L-ful)
- 2 (Somewhat L-ful)
- 3 (Somewhat Vocal)
- 4 (Vocalized L)

No. of responses div. by No. of tokens

<table>
<thead>
<tr>
<th>Code</th>
<th>BOT, BOUGHT</th>
<th>BEET, BAIT</th>
<th>BIT, BET</th>
<th>BOAT</th>
<th>BOOT, PUT</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 (L-ful)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2 (Somewhat L-ful)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3 (Somewhat Vocal)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4 (Vocalized L)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
San Franciscan speakers coded as less vocalized than Columbus speakers

- 4 (Vocalized L)
- 3 (Somewhat Vocal)
- 2 (Somewhat L-ful)
- 1 (L-ful)

No. of x Code given to Token (1,380 total)
Each of the 60 tokens according to how vocalized it was perceived to be, averaging across respondents.
Respondent Differences

• Differences between respondents
  – Wilcoxon signed-rank test for each respondent’s codes vs. the average codes
  – Any respondent with $p < 0.05 = \text{significantly different coding pattern than the average}$
  – Result: 1/5 of the differed from the mean

• In what ways do different respondents differ from the mean?
Average Score, Rounded
Low correlation with response average (lower vocalization ratings)

Type 1: Underestimating vocalization
Low correlation with response average (more vocalization ratings)

- Type 2: Overestimating vocalization
High correlation with rounded response average

Overall, complete agreement at the extreme ends of vocalization ratings
Average Vocalization Rating vs. Interrater Agreement

Token Standard Deviation vs. Average Vocalization Rating
Average Standard Deviation by Preceding Vowel

- Lower standard deviations = more interrater agreement
- Tokens containing rounded back vowels [u U o] were rated more reliably than tokens in other vocalic environment
Individual Comments

[Not every respondent commented at all, some commented on individual tokens, most gave feedback at the end of the survey]

• General feedback:
  – difficulty of the task (e.g., “ack, so hard!”)
  – validity of a 4-level discretization
  – ability to actually match our tokens to an L-ful ~ vocal continuum
  – effects of other Ls in the recording
  – effects of other dialect variables in the recording
Individual Comments

• Perceptual coding comments:
  – “I think that my coding is definitely affected by the coloring of surrounding segments”
  – “I think the vowel quality affected my judgment on some of them”
  – “I feel like longer durations led me to say less /l/ vocalization”

• Study design comments:
  – some wished for even longer segments; others wanted to zoom into even shorter segments
  – people appreciated being able to comment on every token
Individual Comments

• Sample of token comments:
  - “I'm not sure I can tell the difference between /aw/ and /awl/”
  - “the /l/ is actually strong, but there's rounding”
  - “I'm being thrown off by the lack of a diphthong here”
  - “I think I may be using syllable length in distinguish between the presence and absence of /l/”
  - “maybe mediated by the fact that the preceding vowel is laxed?”

• Interesting ethnicity comment for one particular token:
  (1) “sounds like she is rapping”
  (2) “Is this person Asian?”
Individual Comments

• Comments on the more difficult tokens:
  - “too short to hear”
  - “there's quite a bit of noise on this recording”
  - “the next /l/ is way too close to evaluate this”
  - “it's hard for me to tell at this rate of speech”
  - "so creaky. hard to tell."

• If this were a traditional perception survey, the exclusion of hard-to-tell stimuli would be part of the experimental design
  - But as a study mimicking sociolinguistic coding, we tried to get at the methods that have been used to obtain vocalization data in previous & on-going variationist studies
  - We can't exclude all of the difficult tokens we have, because that would severely truncate any dataset!
Summary of Perception Task

• Linguists are reliable and in agreement with one another with respect to those coda-Ls that are the most L-full & the most vocalized
  – Suggesting that perceptual coding is a valid measure for L-vocalization studies

• However, 1/5 of the coders diverged from the mean
  – Those that diverged not only did so in disparate directions,
  – but there was no discernable social reasons for why those particular coders were the ones who diverged
Summary of Perception Task

• *Good news:* linguists from both vocalizing and non-vocalizing regions, as well as non-native English speakers, seem to be able to reliably code for L-vocalization

• *Bad news:* those tokens that are *the most sociolinguistically interesting* to us are the ones that people agreed on *the least*
  
  – Not surprising, but a formidable challenge to L-vocalization research
Implications for Acoustic Measurements

Point 1: Vocalization must be measured with respect to the identity of the preceding vowel

- Non-vocalized L has a strong co-articulatory effect on the preceding vowel (West 1999); vocalization should equal a lack of co-articulatory effect
- Velar L has the formant structure of /ɔ/, at least in some dialects (Gick, Kang, & Whalen 2002), so
  - formant position alone cannot distinguish vocalized from non-vocalized, especially for back, rounded vowels, and
  - co-articulatory effects will have different acoustic properties according to preceding environment
Implications for Acoustic Measurements

Point 2: The method will depend on dialect differences in the realization of vocalized /l/

- San Francisco speakers don’t vocalize to schwa; Columbus speakers do
- San Francisco speakers don’t vocalize after front vowels; Columbus speakers do

• Regional differences also complicated L’s conditioning of vowel production
  - Those are going to differ from region to region as well
Considerations for an Acoustic Measurement

Ideally, an acoustic measurement of L vocalization will be adaptable per *internal linguistic environment* as well as per *community* with respect to linguistic and social factors by taking into account:

1. phonological environment
2. L’s various co-articulatory effects on preceding segments in the community in question (*e.g.*, conditioning mergers)
3. vocalized realization (*e.g.*, /uʷ/, /w/, schwa)
4. Range of L production in L in the community investigated
Final Considerations for an Acoustic Measurement

• An acoustic method can only be as good at differentiating ‘vocal’ from ‘non-vocal’ as the human ear is good at distinguishing the two

• We want to measure differences in L realization that index social differences

• Measures may need to vary from study to study
  – How fine-grained can differences in L realization be before they stop being able to carry social meaning?
Conclusions

• Increased interest in L-vocalization as a sociolinguistic variable encourages us to pursue more reliable methods of measurement and coding

• Auditory, acoustic, and articulatory methods are variably appropriate for different data collection methods

• An ideal way forward in measuring L variability may involve a careful triangulation of these methods
Thanks!

Comments and Questions?
lauren.hall-lew@ed.ac.uk
sonya.fix@nyu.edu
References


Fix, Sonya. 2004. /l/ vocalization and racial integration of social networks: Sociolinguistic variation among whites in a Columbus Ohio community. Poster presented at NWAV 33, Ann Arbor, MI.


References


Articulatory Variants

*Light L*

Occurs in syllable-onset position in most varieties of American English

*Dark L*

Occurs most often in syllable coda position in most varieties of American English
Following Consonant

• Many studies have found the following consonant to be an important factor (Hardcastle & Barry 1985; Recasens 1996)

• In our data we have more coda tokens than coda-cluster tokens, and among the coda-cluster ones, almost all of them are in the environment preceding a coronal (with the exception of two tokens before /f/)