Structured variation in English L-allophony

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

The alternation between light and dark L in many varieties of English is often called an example of categorical allophony, with light [l] in Onsets and dark [ɫ] in Codas (Tsukada et al. 2004, Oh & Gick 2005). Researchers typically derive the dark allophone from velarization or lenition of an underlying light coronal L in Codas (Sproat & Fujimura 1993, Browman & Goldstein 1995, Boersma & Hayes 2001, Bermúdez-Otero & Trousdale 2012). In this paper we expand on the finding by Hayes 2000, Wrench & Scobbie 2003, and Foulkes & Docherty 2007 that the relationships between light and dark L are typically more complex than this, both within and across speakers.

Hayes (2000) for example presented 10 speakers of American English with 17 words containing /l/, each read once with light L and once with dark L. Consultants rated each pronunciation on a 7-point scale. Hayes found significant variance in their responses but nonetheless concluded that "when averaged over all the consultants, the results formed a quite coherent pattern" (p. 9 in preprint version), which he asserted to be gradient in nature. Boersma & Hayes (2001) subsequently developed an implementation in Stochastic OT which was able to simulate Hayes's experimental results.

Though we believe that Hayes was correct in identifying a number of phonological and morphological factors that can play a role in the distribution of light and dark L, certain issues in his experimental design and assessment lead us to question his interpretation of the results. 1. It is not surprising for one to obtain gradient results if one averages over judgements made on a 7-point scale. Though Hayes's findings might indeed reflect the workings of a stochastic grammar, they might also be artefacts of the methods of elicitation and analysis—as in Armstrong et al.'s (1983) study of subjects' ratings of the "goodness" of even, odd, and prime numbers—and obscure an underlying system of categorical allophony. 2. Given Hayes's small sample size, the potential impact of outliers is large, and may obscure the underlying system. 3. Hayes notes that between-word-group differences are significant, but the ANOVA post-hoc LSD test he employs requires equal variance across groups, which he does not demonstrate. 4. By testing the ranking on the same data used to derive the ranking, Hayes does not demonstrate the generalizability of ranking to unseen data.

In order to investigate the subtler phonological dimensions of English L allophony identified by Hayes while avoiding the pitfalls above, we elicited judgements from 303 linguistically-trained speakers of English for their pronunciation of 20 words containing a variety of phonological environments. The list included most of Hayes's stimuli, and additional forms selected to test the possible effects of adjacent segmental features. For each word, subjects were asked to state whether they used light L, dark L, or both.

Building on Gallistel et al.'s (2004) finding that averaging over individuals produces deceptively gradient results that can obscure the categorical behavior of individuals, we analysed our results in terms of individual grammars and rule components rather than averaging over the outputs of all grammars. Our results revealed 255 distinct grammars. In order to determine whether these grammars revealed any underlying patterns, we coded each of the 20 stimuli for 24 binary prosodic and segmental parameters: "after V", "before dorsal C", "before Level 2 V-initial suffix", and so on. We then applied two types of computational analysis:

1. To investigate whether the respondents' grammars could be grouped into larger clusters of similar grammars, we employed Expectation Maximization (EM; Dempster et al. 1977), which (i) randomly assigns the test subjects to groups, (ii) determines the

likelihood of that grouping, and (iii) sets parameters for group assignment on that basis. Subjects are then reassigned to groups using those parameter settings. Steps (ii) and (iii) are then repeated until there is no further improvement. For our purposes, this method has the advantages over other clustering algorithms of not requiring the specification of a number of clusters in advance, and of being able to work with categorical/discrete data.

2. We employed feature selection with machine learning, using correlation-based feature subsets (CFS, Hall 1999) to determine (i) which of our 24 phonological parameters play important roles in L allophony, and (ii) whether the clusters of speakers identified by EM differ in the parameters on which they base their choice of L allophone. CFS finds features that individually best predict outcomes, and considers correlations between features, selecting informative features that make independent contributions to the attested patterns.

Our EM algorithm revealed eight main grammar clusters; CFS analysis revealed that the most common grammars had (i) all light L (9.6%), (ii) all dark L (6.6%), or (iii) the classic Onset/Rime distribution, modulo resyllabification of word-final L (53.8%). Less common patterns revealed participation of phonological factors including ambisyllabicity, vowel backness and height, Onset complexity and place of articulation, and morphological structure.

Our findings suggest that (i) the distribution of light and dark L shows extensive idiolectal variation, which (ii) displays categorical properties incorrectly seen as gradient by studies that average across subjects (Sproat & Fujimura 1993, Boersma & Hayes 2001, Recasens 2004). We argue that the attested variation follows from a model of phonological acquisition in which learners choose from a range of PLD-compatible hypotheses assembled from a limited number of phonological primitives (features, anchor points, quantifiers, and so on; cf. Nevins and Vaux 2003, Reiss 2003, Raimy 2007).

References

Armstrong, S., Gleitman, L., & Gleitman, H. 1983. What some concepts might not be. Cognition, 13, 163–308.

Bermúdez-Otero, Ricardo & Graeme Trousdale. 2012. Cycles and continua: on unidirectionality and gradualness in language change. In T. Nevalainen & E. Traugott (eds), The Oxford handbook of the history of English, 691-720. New York: OUP.

Bermúdez-Otero, Ricardo. 2007. Diachronic phonology. In Paul de Lacy (ed.), The Cambridge handbook of phonology, 497-517. Cambridge: CUP.

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

Browman, Catherine & Louis Goldstein. 1995. Gestural Syllable Position Effects in American English. In Producing Speech: Contemporary Issues, K. Harris, F. Bell-Berti, & L. Raphael (Eds.), 19-34. New York: AIP Press.

Dempster, A., N. Laird, & D. Rubin. 1977. Maximum Likelihood from Incomplete Data via the EM Algorithm. Journal of the Royal Statistical Society Series B 39:1-38.

Foulkes, P. & Docherty, G. 2007. Phonological and prosodic variation in the English of England. In Britain, D. (ed.) Language in the British Isles (2nd ed.). Cambridge: CUP.

Gallistel, C. R., P. Balsam, & S. Fairhurst. 2004. The learning curve: Implications of a quantitative analysis. Proceedings of the National Academy of Sciences 101:13124-13131.

Hall, Mark. 1999. Correlation-based feature selection for machine learning. Doctoral dissertation, University of Waikato.

Hayes, Bruce. 2000. Gradient Well-Formedness in Optimality Theory. In J. Dekkers et al., eds., Optimality Theory: Phonology, Syntax, and Acquisition, 88-120. Oxford: OUP. 1998

preprint version available at http://linguistics.ucla.edu/people/hayes/Gradient.pdf.
Nevins, Andrew and Bert Vaux. 2003. Metalinguistic, shmetalinguistic: The phonology of shmreduplication. Proceedings from the Annual Meeting of the Chicago Linguistic Society 39(1): 702-721.
Oh, S. and B. Gick. 2005. Articulatory and acoustic characteristics of English /l/ in children's speech production. Journal of the Acoustical Society of America 117(4), Pt.2: 2575.
Raimy, Eric. 2007. Precedence theory, root and template morphology, priming effects and the structure of the lexicon. https://goo.gl/GsRsVB
Recasens D. 2004. Darkness in [l] as a scalar phonetic property: implications for phonology and articulatory control. Clinical Linguistics & Phonetics 18:593-603.
Reiss, Charles. 2003. Quantification in structural descriptions: Attested and unattested patterns. Linguistic Review 20: 305-338.
Scobbie, James & A. Wrench. 2003. An articulatory investigation of word final /l/ and /l/-sandhi in three dialects of English. Proceedings of ICPhS 15:1871-74.
Sproat, R. & O. Fujimura. 1993. Allophonic variation in English /l/ and its implications for phonetic implementation. Journal of Phonetics 21, 291–311.
Tsukada, K., D. Birdsong, M. Mack, H. Sung, E. Bialystok, and J. Flege . 2004. Release Bursts in English Word-Final Voiceless Stops Produced by Native English and Korean Adults and Children. Phonetica 61:67-83.
Wrench, A. and J. Scobbie. 2003. Categorising vocalisation of English /l/ using EPG, EMA and ultrasound. Proc. of the 6th International Seminar on Speech Production, pp. 314-319.