Published in final edited form as:
Lang Cogn Neurosci. 2016 ; 31(5): 639–645. doi:10.1080/23273798.2015.1130234.

Dimensions of similarity in the mental lexicon

Melinda Fricke,
Psychology Department, Center for Language Science, Pennsylvania State University, 112 Moore Building, University Park, PA 16802 USA

Melissa M. Baese-Berk, and
Department of Linguistics, University of Oregon, 279 Straub Hall, 1290 University of Oregon, Eugene, OR 97403-1290 USA, mbaesebe@uoregon.edu

Matthew Goldrick
Department of Linguistics, Northwestern University, 2016 Sheridan Road, Evanston, IL 60208 USA, matt-goldrick@northwestern.edu

Abstract

During language production planning, multiple candidate representations are implicitly activated prior to articulation. Lexical representations that are phonologically related to the target (phonological neighbors) are known to influence phonetic properties of the target word. However, the question of which dimensions of phonological similarity contribute to such lexical-phonetic effects remains unanswered. In the present study, we reanalyze phonetic data from a previous study, examining the contrasting predictions of different definitions of phonological similarity. Our results suggest that similarity at the level of position-specific phonological segments best predicts the influence of neighbor activation on phonetic properties of initial consonants.

Keywords

language production; lexical-phonological neighborhood; lexical-phonetic variation; interactive models; lexical access

Introduction

Most theories of language production assume that during production planning, activation spreads to both the intended target and a set of phonologically similar non-target forms (“phonological neighbors,” e.g., *hat, sat* for target *cat;* cf. Dell, 1986 for a feedback-based account and Roelofs, 2004 for a monitoring-based account). Despite this general agreement, an important unresolved question is *which dimensions* of similarity define the set of word forms that become active: what words live in the neighborhood (Goldrick, Folk, & Rapp, 2010)? We examine this issue in a reanalysis of data from a previously published experiment investigating the effect of phonological neighbors on phonetic production (Baese-Berk & Goldrick, 2009). We compare multiple accounts that make contrasting assumptions

---

(corresponding author) mdf18@psu.edu, (814) 863 – 0126.

regarding the dimensions of similarity that define the phonological neighborhood. Our results suggest that similarity relationships among lexical representations are primarily defined over position-specific representations at the level of the phonological segment.

Background

Phonological neighbors and language production

Determining which dimensions of similarity define the lexical-phonological neighborhood has been a persistent issue for models of language production because different dimensions of structure are highly correlated. For example, phonological segments can be characterized by finer-grained phonetic features (e.g., /k/ is a voiceless velar stop), such that two words that share many segments must also share many features. Complicating the issue further, these correlated dimensions also tend to have correlated effects on processing; for example, auditory primes that share an entire segment with a target word speed word-form encoding (Schriefers, Meyer, & Levelt, 1990) as do primes that share only features (but not whole segments; Roon & Gafos, 2015).

Previous studies have demonstrated that a target word's phonological neighborhood impacts the speed with which it can be retrieved from the lexicon (Sadat, Martin, Costa, & Alario, 2014; Vitevitch, 2002), as well as its phonetic realization: lexical-phonetic effects have been documented in simple word reading (Baese-Berk & Goldrick, 2009, Experiment 1; Munson & Solomon, 2004; Wright, 2004 [cf. Gahl, 2015]), sentence reading (Scarborough, 2010), referential communication tasks (Baese-Berk & Goldrick, 2009, Experiment 2; Buz, Jaeger, & Tanenhaus, 2014) and spontaneous speech (Gahl, Yao, & Johnson, 2012). A few recent studies have explicitly examined whether certain dimensions of phonological structure have a greater influence on lexical-phonetic variation than others. Kirov and Wilson (2012) found that when two words were simultaneously presented in a referential communication task, speakers enhanced a phonetic property (voice onset time) only when their initial consonants contrasted by a single phonological feature (e.g., longer VOT for *kilt* in the context of *guilt*, but not *hilt*). Schertz (2013) found similar effects when speakers clarify productions misheard as a specific neighbor. This suggests that when attention is explicitly drawn to a particular neighbor, features—not segments—are the relevant dimension of similarity. Finally, Caselli, Caselli, and Goldberg (2015) found that in spontaneous speech, lexical-phonetic effects on whole word duration may be driven by the set of phonological neighbors differing from the target in the initial segment (e.g., *hat, sat,* for target *cat*), suggesting a critical role for the location of overlap between a target and its neighbors.

In the present paper, we revisit the results from a single-word reading task in which phonological neighbors were *not* presented to the speaker and there was no explicit communicative partner. These data allow us to examine the extent to which different dimensions of phonological structure drive implicit activation of neighbors during target processing.

Previous study: minimal pairs and voice onset time

Baese-Berk and Goldrick (2009) designed materials to examine the influence of minimal pair neighbors on word-initial consonant voice onset time (VOT, the interval between the release of a stop consonant and the onset of vocal fold vibration; Lisker & Abramson, 1964). VOT is an important cue to the contrast between voiced and voiceless stops (e.g., the distinction between neighbors like *cod* and *god*). Pairwise matching was used to control for phonetic and lexical variables that could affect the realization of word-initial VOT (e.g., word frequency, phonotactic probability) such that pairs of words differed only in their “minimal pair status”: for one set of words, changing the voicing of the initial consonant would result in a real English word (e.g., the word *cod* has a minimal pair neighbor *god*), but for the other set, no such neighbor existed (e.g., *cog* has no neighbor *gog*). In their Experiment 1, these two sets of target words were embedded in a list composed primarily of filler items, and participants were presented with one word at a time and asked to say it aloud. Importantly, the voiced counterparts to the minimal pair words were not presented (i.e., participants named *cod* but were never exposed to *god*), reducing the probability that speakers were aware of the minimal pair manipulation. Nonetheless, minimal pair words were produced with significantly longer VOT than non-minimal pair words, suggesting that the implicit activation of the minimal pair neighbor during planning affected the phonetic realization of the target words.

Current study: exploring the relevant phonological dimensions

While the two sets of words in Baese-Berk and Goldrick (2009) were selected to minimize the influence of nuisance variables, the more general phonological neighborhoods of the target words were not controlled. Consequently, the minimal pair words in the study tended to reside in structurally different neighborhoods from the words without minimal pairs. Here we investigate two phonological dimensions that were not considered in the original analysis: the total number of neighbors that differ from each target by the standard one-segment edit distance metric (Landuaer & Streeter, 1973; “total density”), and the number of neighbors that differ by a one segment substitution in each location in the word (“position-specific density”). For example, the minimal pair word *cod* has a total of 27 neighbors; eight of these neighbors are formed by a substitution of the initial consonant (“VC neighbors”, e.g., *sod*), six by a substitution of the vowel (“CC neighbors”, e.g., *kid*), and eight by a substitution of the final consonant (“CV neighbors”, e.g., *cop*)[1]. As shown in Table 1, the Baese-Berk and Goldrick stimuli differ not only in minimal pair status but also in each of these other neighborhood metrics.

In Analysis 1, we ask whether the minimal pair effect was driven by differences in overall neighborhood size. Because the minimal pair relationship under investigation in the original study concerned the voicing of the initial consonant, this analysis pits global neighborhood structure against featural contrast: is the VOT effect more closely tied to the total number of phonologically related words in the lexicon, or to the existence of a single word contrasting along a specific phonological dimension?

---

[1] The position-specific analyses here include only substitution neighbors. Analyses that included addition and deletion neighbors returned qualitatively similar results.

Analysis 2 is motivated by recent findings underscoring the importance of the location of overlap between a target and its neighbors (Caselli et al., 2015). Here, we ask whether a position-specific definition of the phonological neighborhood is a significant predictor of VOT, and if so, how the predictive power of a segment-based, position-specific metric stacks up against both minimal pair status and total density[2].

The outcome of these two analyses is directly related to the weighting of similarity dimensions in the mental lexicon. If minimal pair status is the best predictor of VOT, then the featural level of representation must have a privileged role during speech planning and articulation, even in an "offline" context. In contrast, if total density—based on a one-*segment* edit distance—is the best predictor, this would suggest that all segmental relationships to the target word are weighted approximately equally. Finally, if one of the position-specific density metrics best predicts VOT, this would suggest that the set of words that are most relevant at the moment of articulation is defined by phonemic overlap with the target in a specific location in the word.

Analysis 1: the one-segment edit distance vs. the phonological feature

The data set analyzed here is the same one analyzed in Baese-Berk and Goldrick (2009). The reader is referred to that paper for further methodological details. To compare the predictive power of minimal pair status versus total neighborhood density, we used linear mixed-effects regression (lme4 package version 1.1-7; Bates, Maechler, Bolker, & Walker, 2014) in R (version 3.1.3; R Core Team, 2015) to model log-transformed VOT while controlling for other variables known to affect VOT. All continuous predictors were centered and scaled, and log-transformed where appropriate. We began with a model that included only random intercepts by participant ($n = 22$) and by word ($n = 94$). We then added predictors in a stepwise fashion, retaining only those that significantly improved model log likelihood: initial consonant identity (/p/ vs. /t/ vs. /k/; cf. Lisker & Abramson, 1964), then repetition number (first vs. second vs. third[3]; cf. Fowler & Housum, 1987), and finally vowel height (high vs. low; cf. Klatt, 1975), all dummy coded. Likelihood ratio comparisons (Barr, Levy, Scheepers, & Tily, 2013) confirmed that each of these predictors was significant with all other predictors in the model. Adding mean segmental probability, mean biphone probability, initial biphone probability, or word frequency did not significantly improve the model[4].

Next, we added minimal pair status (contrast coded) to the control model, verifying that it was a significant predictor when all other variables were taken into account, and when random by-participant slopes for the effect of minimal pair status were also included (Barr et al., 2013)[5]. As expected, minimal pair words had significantly longer VOT than non-

---

[2] Contrary to Caselli et al. (2015), the present analyses do not distinguish between monomorphemic and morphologically complex neighbors. They also do not consider the role of neighbor orthographic similarity, which, as a reviewer points out, is an additional factor that could be at play.

[3] A reviewer asks whether the any of the filler words were neighbors of targets (Jacobs, Yiu, Watson, & Dell, 2015). Additional analyses showed that very few neighbors appeared as fillers, and this factor was not a significant predictor of VOT.

[4] A reviewer asks whether the VOT effect could be attributed to changes in overall speech rate. An earlier set of analyses found that the VOT effect remains significant when the rime duration of each word is included as a covariate (Fricke, 2013).

[5] All results reported here included by-participant slopes for the critical predictors, but we note that the same results obtained whether these were included or not.

minimal pair words ($\beta = 0.049$, $t = 2.85$, $\chi^2(1) = 8.06$, $p < .01$). Similarly, total neighborhood density was a significant predictor of VOT when added to the control model ($\beta = 0.003$, $t = 2.90$, $\chi^2(1) = 8.27$, $p < .01$).

Next, we used pairwise model comparisons to determine whether either predictor could account for any variance that the other could not. We added *both* minimal pair status and total density to the control model, including random by-participant slopes for both predictors. We then compared this model to one that included only total density, and to one that included only minimal pair status. Both of these comparisons were marginally significant; when added on top of minimal pair status, total neighborhood density marginally improves model log likelihood ($\beta = 0.002$, $t = 1.89$, $\chi^2(1) = 3.66$, $p = .06$), and the reverse is also true ($\beta = 0.034$, $t = 1.80$, $\chi^2(1) = 3.33$, $p = .07$).

This result suggests two possible interpretations. First, given the marginal improvement contributed by total density over minimal pair status and *vice versa*, it is possible that both global neighborhood structure and featural contrast contribute to the VOT effect observed in Baese-Berk and Goldrick (2009). However, a second possibility is that total density and minimal pair status each capture some aspect of a third, correlated variable; in Analysis 2 we ask how these predictors fare when compared with a density metric that takes into account the position of contrast.

Analysis 2: position-specific segment substitution

Analysis 2 followed the same logic, and included the same control variables, as Analysis 1. We first tested the significance of three position-specific density metrics: the number of neighbors created by substituting the initial, medial, vs. final segments of the target word[6]. Only the number of neighbors differing in the initial segment (“VC neighbors”) significantly improved model log likelihood; words with many vs. few such neighbors were produced with significantly longer VOTs (Figure 1; $\beta = 0.008$, $t = 3.76$, $\chi^2(1) = 13.4$, $p < .001$). The other two metrics were not significant ($\chi s^2(1) < 1.3$, $ps > .25$)[7].

We then used pairwise model comparisons to determine whether the number of VC neighbors contributed additional predictive power over either of the two significant neighborhood metrics from Analysis 1. When the VC neighbors metric is added to a model that includes minimal pair status, model log likelihood increases significantly ($\beta = 0.007$, $t = 2.79$, $\chi^2(1) = 7.80$, $p < .01$), but the reverse is not true ($\beta = 0.023$, $t = 1.21$, $\chi^2(1) = 1.50$, $p = .22$). Likewise, the number of VC neighbors contributes significant predictive power over the total number of neighbors ($\beta = 0.010$, $t = 2.24$, $\chi^2(1) = 5.03$, $p = .02$), but the reverse is not true ($\beta = -0.001$, $t = -0.54$, $\chi^2(1) = 0.30$, $p = .58$). This finding indicates that minimal pair status and total density are each significant predictors on their own because each captures a separate dimension of the VC neighbors metric: both the location of contrast and the number

---

[6] Most words (66 of 94) had a single consonant in final position. Qualitatively similar results were observed when restricting analysis to this subset.

[7] As a reviewer points out, the greater range of VC neighbors in this data set may have played a role here. Future research should explicitly investigate this question.

of similar neighbors are contributing factors to the VOT lengthening observed in Baese-Berk and Goldrick (2009). Table 2 provides the final fitted model.

Discussion

Which dimensions of phonological similarity define the set of neighbors that contribute to lexical-phonetic variation? The answer from the present data is that a position-specific, segment-based metric provides the best definition. The next sections briefly consider the ways in which this result relates to previous findings in phonological processing, and the questions it raises for models of language production.

The special status of the phonological segment

By comparing minimal pair words with non-minimal pair words, the design of the study in Baese-Berk and Goldrick (2009) implicitly assumed that the phonological feature constituted a privileged determinant of the phonological neighborhood. The present reanalysis demonstrates, however, that a segmentally based definition was better able to account for the observed effects. This is consistent with previous work suggesting a role for abstract, segmental representations in speech production (Goldrick & Rapp, 2007; Oppenheim & Dell, 2008; Roelofs, 1999; Reilly and Blumstein, 2014). However, such results do not discount the importance of featural representations during the production planning process more generally; rather, they suggest that such representations do not strongly constrain activation during retrieval of phonological structure from the lexicon. Furthermore, as discussed in the introduction, it is likely that featural relationships between words do play a role in listener-oriented modifications to speech (Kirov and Wilson, 2012; Schertz, 2013; see also Buz et al., 2014).

Position-specific overlap

Because word-initial position and the location of the dependent variable were confounded in this dataset, the present results cannot speak to the relative importance of these factors. As noted above, however, Caselli et al. (2015) found that lexical-phonetic effects on whole-word duration were also influenced by the number of neighbors differing in the initial consonant: Caselli et al. found shortening of overall word duration (as did Gahl et al., 2012, in the same corpus), whereas here, we find lengthening of initial consonant duration. An important question for future research concerns the direction of these durational effects. One possible source of divergence is the word position affected: some positions (e.g., onset) may show lengthening, while others may show shortening (Goldrick, Vaughn, & Murphy, 2013). Future studies should decouple the issue of position-specific overlap from the dependent variable under investigation.

Studies in other empirical domains have also suggested position of overlap is a critical determinant of neighborhood structure (but see Reilly & Blumstein, 2014). Goldrick et al. (2010) found that whole-word substitution errors arising within lexical access were sensitive to position-specific overlap, particularly in the initial position. Zamuner (2009) found a disproportionate number of words sharing vowel and coda in children's vocabularies (relative to adult vocabularies), suggesting that overlap in these positions facilitates word learning.

Mechanisms underlying lexical effects on phonetic production

A longstanding debate in language production research concerns the mechanisms that relate lexical variables, such as word frequency and neighborhood density, to a word's articulatory realization. Our results can be accommodated by a number of theoretical frameworks that reference properties of both language production and perception (Aylett & Turk, 2004, 2006; Baese-Berk & Goldrick, 2009; Lindblom, 1990; Pierrehumbert, 2002). The present findings place critical constraints on these theories. During single-word production, without an interlocutor present, featural contrast between minimal pair words in the lexicon does not give rise to phonetic enhancement along the corresponding dimension of contrast. Rather, it is relationships at the segmental level that give rise to lexically driven phonetic effects in this context. Whether such effects should be attributed to production-internal or perception-oriented mechanisms is a question for future research.

Conclusion

Whereas previous studies of lexical-phonetic effects in language production have typically assumed equal influence from all words differing from the target according to a one-segment edit distance, the present analyses demonstrate that some aspects of phonological structure are weighted more heavily than others during language production. The fact that not all neighbors are created equal necessitates a reconsideration of the ways in which non-target lexical items affect the production process. Importantly, this opens the door to new research questions within *all* models of language production. While the mechanisms that relate lexical activation to phonetic variation remain to be explored, our findings serve to constrain possible directions this exploration may take.

Acknowledgments

A preliminary version of these analyses appeared in M. Fricke's Ph.D. dissertation at the University of California, Berkeley. She gratefully acknowledges Susanne Gahl, Keith Johnson, Judy Kroll, Giuli Dussias, and Matt Carlson for comments on earlier versions of this work. The authors also thank Ariel Goldberg and an anonymous reviewer for comments and suggestions that have significantly improved the paper.

This writing of this paper was supported by NSF Grant SMA-1409636 to M. Fricke, J. F. Kroll and P. E. Dussias, by NIH Grant HD053146 to J. F. Kroll, and by NIH Grant HD077140 to M. Goldrick.

References

Aylett M, Turk A. The Smooth Signal Redundancy Hypothesis: A functional explanation for relationships between redundancy, prosodic prominence, and duration in spontaneous speech. Language and Speech. 2004; 47(1):31–56. doi:10.1177/00238309040470010201. [PubMed: 15298329]

Aylett M, Turk A. Language redundancy predicts syllabic duration and the spectral characteristics of vocalic syllable nuclei. Journal of the Acoustical Society of America. 2006; 119(5):3048–3058. doi: 10.1121/1.2188331. [PubMed: 16708960]

Baese-Berk M, Goldrick M. Mechanisms of interaction in speech production. Language and Cognitive Processes. 2009; 24:527–554. doi:10.1080/01690960802299378. [PubMed: 19946622]

Barr DJ, Levy R, Scheepers C, Tily HJ. Random effects structure for confirmatory hypothesis testing: Keep it maximal. Journal of Memory and Language. 2013; 68:255–278. doi:10.1016/j.jml.2012.11.001.

Bates, D.; Maechler, M.; Bolker, B.; Walker, S. lme4: Linear mixed-effects models using Eigen and S4.. R package version 1.1-7. 2014. http://CRAN.R-project.org/package=lme4

Buz, E.; Jaeger, TF.; Tanenhaus, MK. Contextual confusability leads to targeted hyperarticulation.. In: Bello, P.; Guarini, M.; McShane, M.; Scassellati, B., editors. Proceedings of the 36th Annual Meeting of the Cognitive Science Society. Cognitive Science Society; Austin, TX: 2014. p. 1970-1975.

Caselli NK, Caselli MK, Cohen-Goldberg AM. Inflected words in production: Evidence for a morphologically rich lexicon. The Quarterly Journal of Experimental Psychology. 2015:1–53. doi: 10.1080/17470218.2015.1054847.

Dell GS. A spreading-activation theory of retrieval in sentence production. Psychological Review. 1986; 93(3):283–321. doi:10.1037/0033-295X.93.3.283. [PubMed: 3749399]

Fowler CA, Howsum J. Talkers’ signaling of “new” and “old” words in speech and listeners’ perception and use of the distinction. Journal of Memory and Language. 1987; 26:489–504. doi: 10.1016/0749-596X(87)90136-7.

Fricke, M. Phonological encoding and phonetic duration (Unpublished doctoral dissertation). University of California; Berkeley: 2013.

Gahl S. Lexical competition in vowel articulation revisited: Vowel dispersion in the Easy/Hard database. Journal of Phonetics. 2015; 49:96–116. doi:10.1016/j.wocn.2014.12.002.

Gahl S, Yao Y, Johnson K. Why reduce? Phonological neighborhood density and phonetic reduction in spontaneous speech. Journal of Memory and Language. 2012; 66:789–806. doi:10.1016/j.jml.2011.11.006.

Goldrick M, Rapp B. Lexical and post-lexical phonological representations in spoken production. Cognition. 2007; 102:219–260. doi:10.1016/j.cognition.2005.12.010. [PubMed: 16483561]

Goldrick M, Folk JR, Rapp B. Mrs. Malaprop's neighborhood: Using word errors to reveal neighborhood structure. Journal of Memory and Language. 2010; 62:113–134. doi:10.1016/j.jml.2009.11.008. [PubMed: 20161591]

Goldrick M, Vaughn C, Murphy A. The effects of lexical neighbors on stop consonant articulation. Journal of the Acoustical Society of America. 2013; 134:EL172–EL177. doi:10.1121/1.4812821. [PubMed: 23927221]

Jacobs CL, Yiu LK, Watson DG, Dell GS. Why are repeated words produced with reduced durations? Evidence from inner speech and homophone production. Journal of Memory and Language. 2015; 84:37–48. doi:10.1016/j.jml.2015.05.004. [PubMed: 26089592]

Kirov, C.; Wilson, C. The specificity of online variation in speech production.. In: Miyake, N.; Peebles, D.; Cooper, RP., editors. Proceedings of the 24th Annual Meeting of the Cognitive Science Society. Cognitive Science Society; Austin, TX: 2012. p. 587-592.

Klatt DH. Voice onset time, frication, and aspiration in word-initial consonant clusters. Journal of Speech, Language, and Hearing Research. 1975; 18:686–706. doi:10.1044/jshr.1804.686.

Landauer TK, Streeter LA. Structural differences between common and rare words: Failure of equivalence assumptions for theories of word recognition. Journal of Verbal Learning and Verbal Behavior. 1973; 12(2):119–131. doi:10.1016/S0022-5371(73)80001-5.

Lindblom, B. Explaining phonetic variation: A sketch of the H&H theory.. In: Hardcastle, WJ.; Marchal, A., editors. Speech production and speech modeling. Kluwer; Dordrecht: 1990. p. 403-439.doi:10.1007/978-94-009-2037-8_16

Lisker L, Abramson AS. A cross-language study of voicing in initial stops: Acoustical measurements. Word. 1964; 20(3):384–422. doi:10.1080/00437956.1964.11659830.

Munson B, Solomon NP. The effect of phonological neighborhood density on vowel articulation. Journal of Speech, Language, and Hearing Research. 2004; 47:1048–1058. doi: 10.1044/1092-4388(2004/078).

Oppenheim GM, Dell GS. Inner speech slips exhibit lexical bias, but not the phonemic similarity effect. Cognition. 2008; 106:528–537. doi:10.1016/j.cognition.2007.02.006. [PubMed: 17407776]

Pierrehumbert, JB. Word-specific phonetics.. In: Gussenhoven, C.; Warner, N., editors. Laboratory Phonology VII. Mouton de Gruyter; Berlin: 2002. p. 101-140.doi:10.1515/9783110197105.101

R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing. Vienna, Austria: 2015. http://www.R-project.org/

Reilly M, Blumstein SE. Effect of sound similarity and word position on lexical selection. Language, Cognition and Neuroscience. 2014; 29(10):1325–1341. doi:10.1080/23273798.2014.917193.

Roelofs A. Phonological segments and features as planning units in speech production. Language and Cognitive Processes. 1999; 14:173–200. doi:10.1080/016909699386338.

Roelofs A. Error biases in spoken word planning and monitoring by aphasic and nonaphasic speakers: Comment on Rapp and Goldrick (2000). Psychological Review. 2004; 111:561–572. doi:10.1037/0033-295X.111.2.561. [PubMed: 15065924]

Sadat J, Martin CD, Costa A, Alario FX. Reconciling phonological neighborhood effects in speech production through single trial analysis. Cognitive Psychology. 2014; 68:33–58. doi:10.1016/j.cogpsych.2013.10.001. [PubMed: 24291531]

Scarborough, RA. Lexical and contextual predictability: Confluent effects on the production of vowels.. In: Goldstein, L.; Whalen, DH.; Best, CT., editors. Laboratory phonology 10. De Gruyter Mouton; Berlin, New York: 2010. p. 557-586.

Schertz J. Exaggeration of featural contrasts in clarifications of misheard speech in English. Journal of Phonetics. 2013; 41:249–263. doi:10.1016/j.wocn.2013.03.007.

Schriefers H, Meyer AS, Levelt WJM. Exploring the time course of lexical access in language production: Picture-word interference studies. Journal of Memory and Language. 1990; 29(1):86–102. doi:10.1016/0749-596X(90)90011-N.

Roon KD, Gafos AI. Perceptuo-motor effects of response-distractor compatibility in speech: Beyond phonemic identity. Psychonomic Bulletin & Review. 2015; 22(1):242–250. doi:10.3758/s13423-014-0666-6. [PubMed: 24865282]

Vitevitch MS. The influence of phonological similarity neighborhoods on speech production. Journal of Experimental Psychology: Learning, Memory, and Cognition. 2002; 28(4):735–747. doi:10.1037/0278-7393.28.4.735.

Wright, R. Factors of lexical competition in vowel articulation.. In: Local, J.; Ogden, R.; Temple, R., editors. Papers in laboratory phonology VI. Cambridge University Press; Cambridge: 2004. p. 26-50.

Zamuner TS. The structure and nature of phonological neighbourhoods in children's early lexicons. Journal of Child Language. 2009; 36(01):3–21. doi:10.1017/S0305000908008829. [PubMed: 18808731]

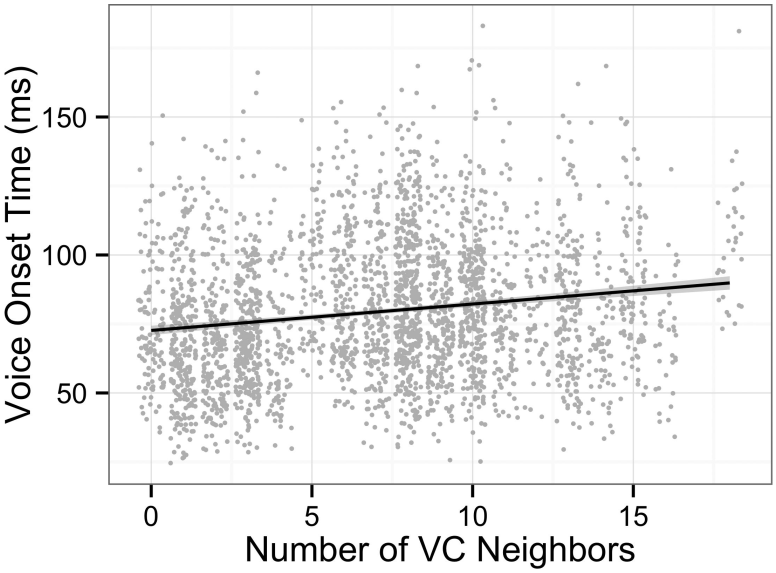

Figure 1.

The effect of position-specific neighborhood density on VOT production in the Baese-Berk and Goldrick (2009) data set, with a best-fit regression line. The number of VC neighbors has been jittered for plotting purposes.

Table 1

Differences in neighborhood structure for minimal pair words versus non-minimal pair words in Baese-Berk and Goldrick (2009). Each cell gives the mean, standard deviation, and range for the number of neighbors as defined by each metric (see text for details). The bottom row gives the point-biserial correlation coefficient for minimal pair versus non-minimal pair words, for each metric.

Word Type	**Total Density**	**VC Neighbors**	**CV Neighbors**	**CC Neighbors**
Minimal Pair	23.7 (8.0)	9.0 (3.6)	6.1 (2.3)	5.1 (3.2)
(*n*=47 words)	7 – 36	3 – 18	1 – 10	0 – 10
No Minimal Pair	16.2 (9.1)	5.1 (4.1)	5.4 (2.6)	3.5 (3.2)
(*n*=47 words)	2 – 33	0 – 13	0 – 10	0 – 9
point-biserial correlation	0.40	0.45	0.14	0.24

Table 2

The final fitted model reanalyzing data from Baese-Berk and Goldrick (2009). Removal of outliers with standardized residuals greater than +/− 2.5 standard deviations resulted in the removal of 46 data points, leaving a total of 2,917 data points.

Linear mixed model fit by REML [‘lmerMod’]				
Formula: log(VOT) ~ (VC.c | Subject) + (1 | Word) + Consonant + VHeight + Repetition + VC.c				
Data: refit				
REML criterion at convergence: −311.6				
Scaled residuals:				
Min	1Q	Median	3Q	Max
−2.75221	−0.64441	0.04967	0.67506	2.62264
Random effects:				
Groups	Name	Variance	Std.Dev.	Corr
Word	(Intercept)	4.464e-03	0.066812	
Subject	(Intercept)	3.899e-02	0.197458	
	VC.c	9.724e-06	0.003118	−0.42
	Residual	4.793e-02	0.218919	
Number of obs: 2917, groups: Word, 94; Subject, 22				
Fixed effects:				
	Estimate	Std. Error	t value	
(Intercept)	−2.504168	0.064306	−38.94	
Consonantp	−0.184735	0.085891	−2.15	
Consonantk	0.108975	0.020487	5.32	
VHeightlow	−0.034567	0.016736	−2.07	
Repetition2	−0.023891	0.009929	−2.41	
Repetition3	−0.039662	0.009936	−3.99	
VC.c	0.008464	0.002035	4.16	