

JOURNAL ARTICLE

Nasal place assimilation trades off inferrability of both target and trigger words

Rory Turnbull¹, Scott Seyfarth², Elizabeth Hume^{2,3} and T. Florian Jaeger^{4,5}

¹ Department of Linguistics, University of Hawai'i at Mānoa, US

² Department of Linguistics, Ohio State University, US

³ Department of Linguistics, University of Canterbury, NZ

⁴ Department of Brain and Cognitive Sciences, University of Rochester, US

⁵ Department of Computer Science, University of Rochester, US

Corresponding author: Rory Turnbull (rory.turnbull@hawaii.edu)

In English, nasal place assimilation occurs across word boundaries, such as *ten bucks* pronounced as *te[m] bucks*. Assimilation can be viewed as a reduction or loss of the assimilation target's place cue (/n/ in *ten*), and simultaneously as an enhancement of the assimilation trigger's place cue (/b/ in *bucks*) by spreading its place cue earlier in the signal. A message-oriented phonological approach predicts that assimilation is sensitive to the relative contextual inferrability of both the target and trigger words: More assimilation should be observed for more contextually predictable target words, while less assimilation should be observed for contextually more predictable *trigger* words. These predictions deviate from accounts that view assimilation solely as reduction. To test these predictions, sequences which license assimilation were extracted from a conversational speech corpus. Both categorical assimilation (based on close phonetic transcription) and gradient acoustic assimilation (based on F2) were analyzed. As predicted, assimilation was more likely both when a target like *ten* was high in predictability and when its trigger *bucks* was low in predictability. Assimilation thus serves as both reduction and enhancement, and can be used to manage redundancy in the speech signal. More broadly, this constitutes evidence for the influence of communicative pressures on phonology.

Keywords: assimilation; predictability; message-oriented phonology; nasals; English

1. Introduction

Assimilation, the sound pattern whereby one sound acquires a phonetic property from a neighbouring sound (e.g., [nb] ~ [mb]), is one of the most widespread phonological phenomena in the world's languages. Intuitively, assimilation can be accounted for as a straightforward ease-of-articulation effect where the production of two properties (e.g., [coronal]-[labial]) is simplified to one (e.g., [labial]), under the assumption that production of heterorganic clusters is articulatorily harder than the production of homorganic clusters (for discussion see e.g., Kohler, 1991; and Winters, 2003 regarding nasal-stop clusters). The explanatory power of a phonetically-grounded account of assimilation increases when we further take into account the observation that the perceptual cues to the consonants that undergo assimilation are generally weaker than are those of the consonants that trigger it (Kohler, 1991; Hura, Lindblom, & Diehl, 1992; Jun, 1996, 2004, 2011; Chang, Plauche, & Ohala, 2001; Steriade, 2001; Kawahara & Garvey, 2014; see also Ohala, 1990; Kawasaki-Fukumori, 1992; Rysling, 2017) and are hence more likely to undergo processes of reduction.

The observation that production and perceptual factors play a role in predicting assimilation is consistent with the view that sound patterns are shaped by communicative biases towards accurate message transmission and low resource cost (e.g., Lindblom, 1990; Hall, Hume, Jaeger, & Wedel, 2016). From this perspective, reduction of an inherently weakly perceptible segment in an articulatorily difficult context can be viewed as maintaining low resource cost. That is, there is little communicative gain in investing resources in the clear production of the segment since it contributes little to the probability that the word will be successfully recognized (henceforth, inferrability).

In principle, communicative biases could manifest in direct or indirect ways. The direct route involves strategic—although possibly implicit and automatic—behaviour as suggested just above: Speakers modulate phonetic and phonological content to adjust the amount of information contained in the speech signal to improve the likelihood of communicative success (Lindblom, 1990; Aylett & Turk, 2004; Scarborough, 2004; Jaeger, 2013; Buz, Tanenhaus, & Jaeger, 2016). The simplest prediction of this framework is an inverse relationship between phonetic content and word predictability. This relationship arises because, for highly contextually-predictable words, the predictability of the word itself licenses a degree of phonetic reduction on the part of the talker. That is, the talker is free to reduce such words with acceptable risk of communication failure. On the other hand, less predictable words have no such contextual cues to assist their recognition, and they therefore are produced with relative phonetic clarity. The indirect route involves filtering as part of the perception-production loop, as outlined by Pierrehumbert (2001, 2002). Under this view, unrecognizable words are not stored in a listener's exemplar clouds, and are thus not replicated in future communications. Crucially, the recognizability of a word is a function of both its phonetic clarity and its contextual predictability. Along these lines, contextually predictable but phonetically unclear word exemplars will be more likely to 'survive' and be used in many communications. Unpredictable and unclear exemplars will not survive, with the end result being a trade-off between phonetic clarity and contextual predictability (see also Wedel, 2006). For example, speakers are unlikely to reduce a coda /n/ in a less-predictable word because they have stored relatively few exemplars of this word without a clearly-produced /n/.

Regardless of the directness of these biases, several studies have found empirical support for the basic generalization that phonetic content and word predictability stand in an inverse relationship (e.g., for English, Jurafsky, Bell, Gregory, & Raymond, 2001; Bell, Brenier, Gregory, Girand, & Jurafsky, 2009; Aylett & Turk, 2004, 2006; for review, see Ernestus, 2014; Clopper & Turnbull, 2018; Jaeger & Buz, 2017). For example, reduction of English /t, d/ in word-medial onset position is more likely when the word that it occurs in has high predictability given the following word (Raymond, Dautricourt, & Hume, 2006). Similarly, the addition of phonetic material through variable vowel epenthesis in Dutch was found to be more likely in less predictable words (Tily & Kuperman, 2012).

Two predictions arise from a communicative framework such as message-oriented phonology (Hall et al., 2016).¹ The first prediction focuses on the relative contextual predictability of a word and the view that fewer acoustic cues are needed to identify a word with high predictability. Under the assumption that phonological patterns of assimilation have their roots in coarticulatory variability (Beddor, 2009) and that coarticulation can²

¹ Specifically, the class of communicative accounts we refer to here views phonological processing as part of the chain of probabilistic inferences made during speech production and perception. These accounts hold that the utility function that involves these inferences exerts subtle but consistent biases on the historical development of phonological representations and processes.

² We note that it is likely not the case that coarticulation *always* negatively affects perception of the target (see Beddor et al., 2013; Scarborough, 2013), either because some coarticulation might leave the relevant cues to the target unaffected or because listeners compensate for coarticulation (for references and over-

involve a reduction in phonetic cues of the canonical target form, we would expect place assimilation of a segmental target to be more likely when the word it occurs in is contextually more predictable. For example, the phrase *ten bucks* is more likely to be produced as *tem bucks* when the word *ten* is relatively predictable from context. While the reduction in cues to the phonetic form may lead to diminished inferrability of the word *ten*, the word's high contextual predictability supports effective recognition: While *ten* is a probable word to precede *bucks*, *temp* or other alternative word candidates that would have non-zero probability under the assimilated signal are not probable before *bucks* (for further discussion, see Flemming, 2010 and Hall et al., 2016; for evidence that contextual inferrability ameliorates the effects of phonetic reduction; see Ernestus, Baayen, & Schreuder, 2002).³

Crucially, however, assimilation does not only reduce redundancy in the target word; it also extends a cue from the assimilation trigger to earlier in the signal, thereby increasing redundancy in the trigger word. That is, assimilation enhances the cues to the trigger (Scarborough, 2013). Indeed, listeners seem to be able to take advantage of such coarticulatory cues. Otake, Yoneyama, Cutler, and van der Lugt (1996) investigated the role of nasal place of articulation on the processing of place in a following stop consonant by Japanese subjects. In a phoneme monitoring task, they found that Japanese listeners made use of place cues in a nasal consonant to anticipate the identity of a following homorganic stop (see also Beddor et al., 2013).

The second prediction of a communicative approach is thus that assimilation is *more* likely when the trigger word is *less* contextually predictable. This prediction, to our knowledge, has not been previously proposed in the literature. In the case of nasal place assimilation, assimilation causes the place feature of the trigger to be both on the trigger and on the target thereby increasing inferrability for the trigger word. Assimilation can therefore be viewed as *enhancing* cues to the trigger, as well as potentially *reducing* cues to the target. As an enhancement process, assimilation is therefore predicted to be more likely to occur with contextually *unpredictable* trigger words (to boost their likelihood of error-free transmission). This predicts the phrase *ten bucks* to be more likely produced as *tem bucks* when the word *bucks* is relatively unpredictable from context.

We note that our predictions deviate from the most widely referenced usage-based accounts. Many usage-based accounts see routinization and practice effects, rather than effective communication, as the driving factor behind language change (e.g., Bresnan & Spencer, 2013; Bybee, 2001, 2006; Diessel, 2007; Munson, 2001; Tomaschek, Tucker, Fasiolo, & Baayen, 2018; see also Jaeger & Hoole, 2011). These accounts predict that frequency is the main factor to influence assimilation (see also certain exemplar-based accounts, e.g., Pierrehumbert, 2001, 2002). That is, if a word is very common, it will be phonetically reduced—high frequency leads to high assimilation rates. However, the communicative approach, based on general assumptions about the role of probabilistic inference during communication (see footnote 1), predicts that the main factor influencing assimilation is the inferrability of a message—and thus its conditional probability in context. The probability of the assimilation trigger is as relevant as that of the assimilation target, and high trigger probability in fact will lead to low assimilation rates. This prediction

view, see Recasens, 2018; Viswanathan, Magnuson, & Fowler, 2010, although most of this literature does not address the absolute effectiveness of compensation). The prediction we present here holds for any type of coarticulation that *on average* reduces the inferrability of the target in its context.

³ At least two sources of uncertainty can be seen as contributing to the probability distribution over word candidates. First, we can think of the acoustic signal determining a probability distribution over phonological sequences (e.g., /dag/ vs. /kæt/ vs. ...). Second, we can think of each of these sequences as determining a probability distribution over words based on, for instance, implicit knowledge of phonological processes such as coronal stop deletion or final cluster reduction. By taking the weighted average of the latter (weighted by the former)—i.e., by marginalizing over the former—we obtain the relative probabilities of different words, given both the signal and the context.

thus stands in opposition to that of a routinization-based account, which would generally predict more assimilation for more probable words. The present study directly contrasts the predictions of these two types of accounts.

The predictions of the communicative account can be framed in terms of two distinct levels of representation: segments and words.⁴ That is, do these processes respond to uncertainty calculated at the level of individual words, or of individual segments? The enhancement of both kinds of units mitigate uncertainty about the intended meaning of an utterance, which is central to message-oriented phonology (Hall et al., 2016). Some theoretical treatments privilege the word level (e.g., Hall et al., 2016), others the segment level (e.g., Cohen Priva, 2012; Scarborough, 2013), and many others are agnostic (e.g., Aylett & Turk, 2004). While determining the level of representation at which these processes apply is an important step in this research program, the present study does not adjudicate between these possibilities. As both are consistent with our main hypothesis regarding a trade-off relationship between the target and trigger words, we present analyses in terms of both word-level predictability (i.e., conditional lexical probability in context) and segment-level predictability (i.e., conditional probability of a segment given the preceding segments).⁵

We test the predictions of both routinization and communicative approaches against a corpus of conversational speech. We first examine nasal place assimilation at the level of phonetic categories, using the close phonetic transcriptions provided with the corpus. On the understanding that English place assimilation is a phonetically gradient process (Ellis & Hardcastle, 2002), we also examine assimilation using a continuous acoustic measure based on formant transitions. Specifically, we investigate assimilation of word-final coronal nasals in the citation phonological form (henceforth /n/) which precede word onsets with labial or velar plosives /p, b, k, g/. Coronal nasals are thus the *target* of assimilation for the present study. Labial and velar plosives are the *triggers* of assimilation. Before a word-initial labial trigger, a coronal nasal /n/ may be assimilated to [m]. Before a word-initial velar trigger, a coronal nasal /n/ may be assimilated to [ŋ].

Figure 1 summarizes the predictions of place assimilation across a sequence of two words, e.g., *ten bucks*. Assimilation of a word-final coronal nasal to a following word-initial

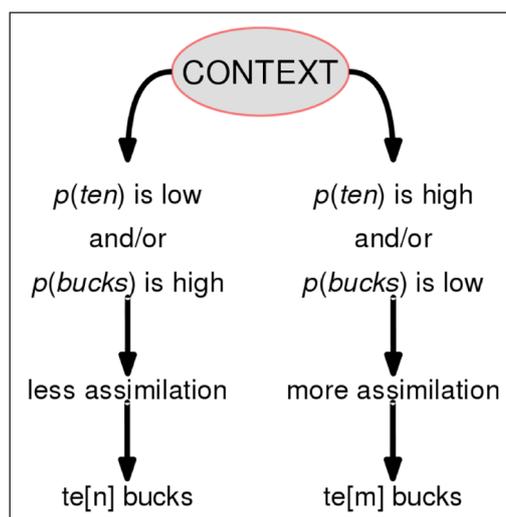


Figure 1: Predictions regarding probability of nasal place assimilation in a two-word target-trigger sequence such as *ten bucks*, formulated at the word level.

⁴ We are grateful to Uriel Cohen Priva for suggesting the analyses of within-word predictability.

⁵ Additionally, the segment-based predictability of a word's final segment—of interest in the present study—is difficult to distinguish from that word's predictability conditioned on the segments heard so far.

labial or velar stop is predicted to occur at a greater rate when the contextual predictability of the target *ten* is high. Further, assimilation is also predicted to occur at a greater rate when the predictability of the trigger *bucks* is low. Note that if place assimilation were driven by repetition or collocational strength (e.g., Bybee, 2006; Mowrey & Pagliuca 1987, 1995), greater predictability of both the target and the trigger should be associated with more assimilation. If, however, place assimilation is driven by a trade-off in the inferrability of the words in the phrase, we would expect the pattern shown in **Figure 1**. The communicative approach thus makes predictions that differ from those of other usage-based approaches to phonology.

2. Methods

2.1 Corpus data

We extracted target-trigger word sequences which license nasal place assimilation across a word boundary (i.e., bigrams such as *ten bucks*) from the Buckeye Corpus of spontaneous speech (Pitt, Dilley, Johnson, Kiesling, Raymond, Hume, & Fosler-Lussier, 2007). The Buckeye Corpus includes audio recordings of one-on-one interviews with forty native English speakers from Columbus, Ohio, conducted under the initial pretense that they were participating in a study investigating how people express their opinions about everyday topics. The interviews are approximately one hour each, and comprise roughly 300,000 words. The audio recordings are annotated with time-aligned close phonetic transcriptions that have been hand-corrected at the segment level, as well as canonical transcriptions for each word based on its dictionary transcription. These transcriptions have been used in a number of previous studies on other aspects of pronunciation variation (e.g., Raymond et al., 2006; Dilley & Pitt, 2007; Cole, Mo, & Baek, 2010; Cole, Mo, & Hasegawa-Johnson, 2010; Gahl, Yao, & Johnson, 2012; Fourtassi, Dunbar, & Dupoux, 2014; Seyfarth, 2014; Cohen Priva, 2017b; Turnbull, 2018).

For our analysis, we first identified potential target-trigger sequences meeting the following three basic criteria:

1. The first word (the *target*) had a dictionary transcription ending in a coronal nasal coda /n/ which was preceded by a vowel.
2. The second word (the *trigger*) had a dictionary transcription beginning with an oral stop with a non-coronal place of articulation (i.e., /p, b, k, g/).
3. Neither word involved a speech error or restart, or was a filled pause, and neither word had an obvious transcription error, defined as (a) a close phonetic transcription that was missing or did not match the timestamped segments or (b) a negative duration.

This yielded 1334 tokens from the entire Buckeye corpus. Next, we excluded tokens for the following reasons (in this order). **Table 1** summarizes our data and exclusions.

1. *No vowel in target's phonetic transcription*: The close phonetic transcription of the target word did not include a nucleus vowel (139 excluded tokens, 10% of the total). The majority of these exclusions (101 tokens) were tokens transcribed as having a syllabic [ŋ] nucleus in the final syllable, especially the word *can* (51 tokens) transcribed as [kŋ].
2. *No nasal in target's phonetic transcription*: The close phonetic transcription of the target word did not include a final nasal coda (18 excluded tokens, 1.3% of the total). For these purposes, a target word was included if the final nasal was transcribed as [n], [m], or [ŋ]—for example, *ten* transcribed as [tem] was allowed. A target word

Table 1: Number of tokens remaining after each exclusion criterion.

| Exclusion reason | Tokens after exclusion |
|--|------------------------|
| (All Buckeye sequences with word-final /Vn/ followed by /p, b, k, g/ onset with no speech errors, pauses, or alignment errors) | 1334 |
| No vowel in target's phonetic transcription | 1195 |
| No nasal in target's phonetic transcription | 1177 |
| No stop in trigger's phonetic transcription | 1137 |
| Short target vowel (<i>excluded for partial assimilation only</i>) | 777 |
| Diphthongal target vowel (<i>excluded for partial assimilation only</i>) | 671 |

was excluded if the coda nasal was missing or transcribed as any other segment. For example, *in* transcribed as [ɪp] (in the bigram *in particular*) was excluded. Most of these were speech errors or transcription errors.

3. *No stop in trigger's phonetic transcription:* The close phonetic transcription of the trigger word did not include an initial oral stop with the same place of articulation as the initial stop in the dictionary transcription (40 excluded tokens, 3.0% of the total). These exclusions mostly fell into two categories. First, there were 12 tokens which involved progressive nasal manner assimilation: For example, *sun bear* transcribed as [sʌn mbɛɹ]. While we would expect this kind of assimilation to also participate in the type of trade-off predicted by communicative accounts—for example, we would predict progressive manner assimilation if the predictability of the target is low and the predictability of the trigger is high—there were not enough such cases for them to be meaningfully involved in our analysis. Second, there were 26 tokens in which the trigger word was listed as *because* but the speaker produced *cause*, which was transcribed with an initial [k].

In addition, because our measure of partial assimilation is based on the target vowel's formant trajectory (cf. Dilley & Pitt, 2007), we excluded some additional tokens from the analysis of partial assimilation only:

4. *Short target vowel:* The target vowel was shorter than 50 milliseconds (360 excluded tokens, 27% of the total). Formant measurements are unreliable for short measurements, and short vowels are unlikely to have meaningful formant transitions. Fifty milliseconds is the cut-off used by the FAVE vowel measurement software (Rosenfelder, Fruehwald, Evanini, Seyfarth, Gorman, Prichard, Yuan, 2014). A large number of these excluded tokens involved the function words *can*, *in*, *even*, *one*, *when*, *been*, *on*, *then*, which are often pronounced with a very short duration and a short or unmeasurable formant trajectory.
5. *Diphthongal target vowel:* Either the dictionary transcription or the close phonetic transcription of the target vowel was listed as a diphthong (/aɪ/, /aʊ/, /ɔɪ/, /eɪ/, /oʊ/, including nasalized diphthongs) (106 excluded tokens, 7.9% of the total). Diphthongization inherently affects formant trajectories, but the realization of phonetic diphthongization is not fully predictable in American English (Fox & Jacewicz, 2009; Lehiste & Peterson, 1961).

After excluding these tokens, there were 1137 target-trigger sequences used in the analysis of categorical assimilation, and 671 target-trigger sequences used in the analysis of partial

assimilation (additional analyses confirmed that our results do not change if the analysis of categorical assimilation is also limited to the smaller data set of 671 cases).

2.2 Measuring assimilation

2.2.1 Categorical assimilation

For our analysis, we used two metrics of nasal place assimilation. The first metric was a dichotomous measure of assimilation based on the close phonetic transcriptions provided by the Buckeye Corpus annotators. If an /n/ coda was transcribed as [m] before /p, b/-initial trigger words, or as [ŋ] before /k, g/-initial trigger words, it was considered to be assimilated; otherwise it was considered to be canonical. For example, if the word *ten* in the sequence *ten bucks* was transcribed as [tem], the /n/ was considered to be assimilated. If the nasal was transcribed as [ten], it was considered to be canonical.

2.2.2 Partial assimilation

Although the segmental transcriptions in the Buckeye Corpus are hand-corrected, a wide variety of factors might influence the transcribers' subjective auditory and visual impressions of nasal assimilation, including their expectations and previous experience with the particular target-trigger sequence. Therefore, our second metric was an acoustic, continuous measure of assimilation based on the vowel formant trajectories immediately preceding the target /n/ codas. The change in the second formant frequency (F2) preceding a consonant is influenced by that consonant's place of articulation. Thus, it can be used as a gradient approximation of nasal consonant place.

We follow previous work in using F2 to identify labial or velar nasal place assimilation (Pitt & Johnson, 2003; Dilley & Pitt, 2007). If a vowel is followed by a labial consonant, F2 typically falls into the consonant onset. If a vowel is followed by a velar consonant, F2 typically rises into the consonant onset. While this pattern generally holds across vowel types, we note that (i) the magnitude of the rise or fall differs between front and back vowels; and (ii) the actual acoustics of F2 trajectories for particular vowel-stop sequences are somewhat more complicated (see e.g., Fowler, 1994), and can differ not only among vowels but also among talkers.

To manage these differences, we construct a standardized metric which corrects both for vowel context as well as talker-specific pronunciation differences. This metric has higher values when F2 at the vowel offset is unusually high or low for a particular talker and phonological context—thus suggesting place assimilation—and lower values when F2 at the vowel offset is more typical for that particular talker and phonological context.

2.2.2.1 Measuring F2

We estimated raw F2 trajectories in Hz using linear predictive coding (LPC) with the Burg algorithm, using the standard preprocessing and analysis window settings in Praat. However, it is well known that automatic formant extraction using this method can result in gross errors due to misidentification of the formant tracks (e.g., see Weenink, 2015). To avoid such errors, we used the extraction technique described in Evanini, Isard, & Liberman, 2009 (see also Evanini, 2009; Rosenfelder et al., 2014; Labov, Rosenfelder, & Fruehwald, 2013).

In this procedure, seven different sets of formant measurements are estimated for each vowel token, using seven different LPC orders (6 to 12). For each vowel token, the best LPC order is selected from the seven candidates by calculating the Mahalanobis distance between the estimated formant values at the vowel midpoint (the first two formant frequencies and bandwidths) and the average measurements for the vowel type. The average measurements

for each vowel type are taken from the Atlas of North American English (Labov, Ash, and Boberg, 2005, as provided in Rosenfelder et al., 2014). The LPC order that results in the smallest Mahalanobis distance is selected, which indicates that the formant estimates which resulted from that LPC order were most typical for the vowel type.

This method avoids gross measurement errors at the midpoint of each vowel (see discussion in Labov, Rosenfelder, & Fruehwald, 2013). However, to quantify nasal place assimilation, we require measurements at the vowel offset (i.e., the consonant onset), and gross formant tracking errors may occur also between the midpoint and vowel offset, even if the midpoint is measured correctly. For this reason, we began with the midpoint measurements, and then for each consecutive window from the midpoint to the vowel offset, we adjusted the LPC order if it lowered the Mahalanobis distance between the formant measurements in the current and following window. This ensures a relatively continuous trajectory, and helps avoid gross formant tracking errors.

2.2.2.2 Correcting offset F2 measurements for talker and context

We used the technique above to estimate F2 at the vowel offset for two datasets. The first dataset comprises the 671 target-trigger bigram tokens described above. In order to determine whether each of these measurements are typical (and thus likely precede a canonical /n/) or atypical (and thus likely precede an assimilated /n/) for each talker and vowel type, we collected a second dataset comprising /Vn/ rimes which precede vowels or /h/ in the corpus (using exclusion criteria 1–2 above). We expect that the /n/ codas in this second dataset should be realized as fully canonical, and therefore the F2 values at the vowel offsets in this dataset represent the distribution that would be *expected* for *unassimilated* /n/ given the preceding vowel and talker.

Next, we used the F2 measurements from these two datasets to calculate an F2 value for the observed target-trigger token, standardized within-talker and within-vowel.⁶ For example, in the sequence *ten bucks*, we took the observed F2 value at the vowel offset, and compared it to the mean F2 value at the vowel offset for all tokens of /ɛn/ produced by that talker, including both assimilatory (as in *ten bucks*) and non-assimilatory contexts (/Vn#V/, such as in *then again*). If the observed F2 in *ten bucks* is one standard deviation above that mean, it receives a value of 1. If the observed F2 in *ten bucks* is one-half standard deviation below that mean, it receives a value of -0.5.

This score is negative when the observed F2 at the vowel offset is below the F2 at the vowel offset that would be expected given the talker and vowel type. The score is positive when the observed F2 is above the expected F2. **Figure 2** shows the assimilation scores for the target-trigger tokens for each vowel type, divided into three categories depending on whether the corpus annotators transcribed each token as canonical [n], assimilated [m], or assimilated [ŋ]. As expected, the assimilation scores tend to be lower when the coda /n/ was transcribed as labial [m], and higher when it was pronounced as velar [ŋ].

Here we are interested in modeling the degree of partial assimilation. For velar assimilation contexts, such as *in Columbus*, these assimilation scores serve as the intended metric of partial assimilation: They are higher in categorically-assimilated productions, and closer to zero in more canonical productions. Large negative values of this metric are thus unexpected for velar assimilation contexts, which is in fact what we found (range -2.12 to 3.84, quartiles at -0.14, 0.69, 1.32 for velar assimilation contexts).

⁶ The decision to standardize F2 values within-talker and within-vowel resulted in 13 additional exclusions in the analysis of partial assimilation, for cases in which there was only a single example of a particular talker producing a particular vowel type in the relevant phonological contexts. Standardizing F2 values within-vowel only, rather than within-talker within-vowel, resulted in qualitatively the same results as those reported here.

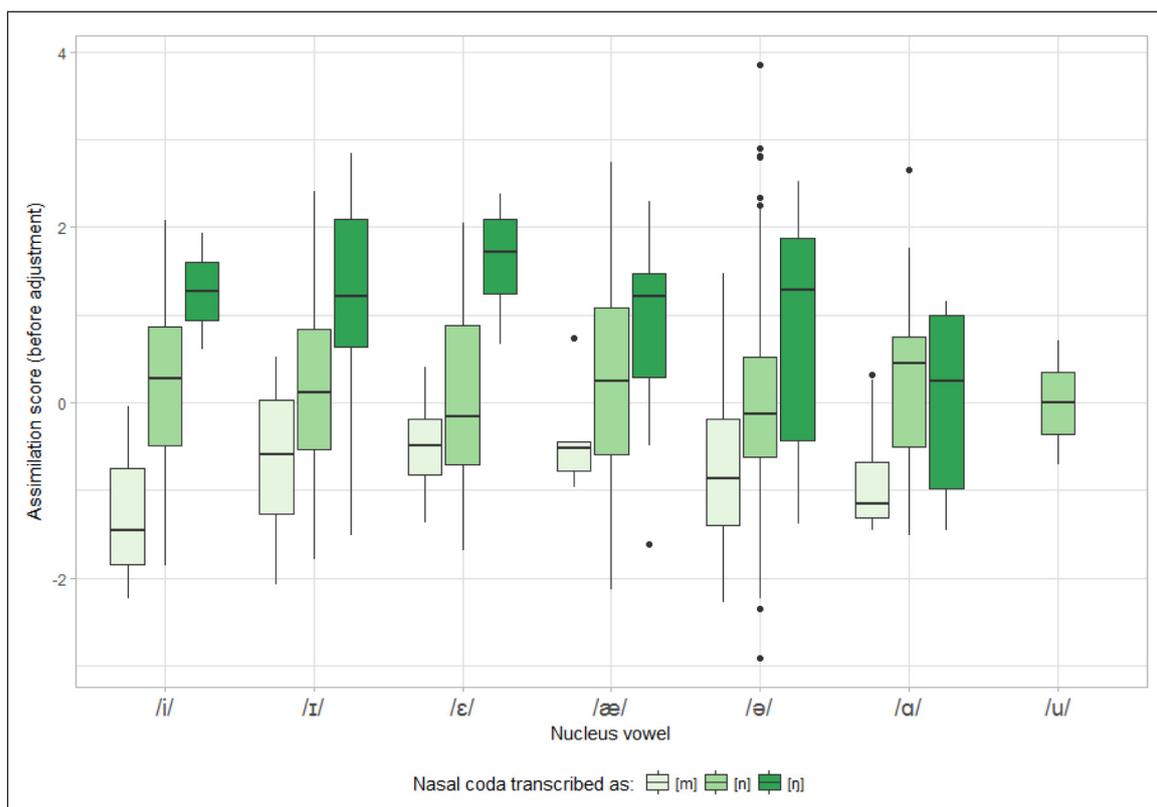


Figure 2: Assimilation scores at vowel offsets, standardized within-talker and within-vowel. The boxes show measurements before nasal /n/ codas that were transcribed as [m], [n], and [ŋ] by the corpus annotators in the left, center, and right box of each group, respectively. In each box, the center line is the median of measurements for that group; the lower and upper edges of the box are the first and third quartiles of measurements; and whiskers extend to the largest measurement that is less than 1.5 times the box height away from the edges (Wickham, 2009).

For labial assimilation contexts, such as *ten bucks*, these assimilation scores have the opposite sign: Since F2 tends to be lower preceding a labial consonant, the standardized value is lower in assimilated productions, and closer to zero in typical, canonical productions. Thus, we inverted the assimilation scores in labial contexts, so that higher values of the metric are always associated with assimilated productions, regardless of whether the trigger was labial or velar. These adjusted scores (i.e., the transformation of the standardized F2 values) served as our metric of partial assimilation.

2.2.2.3 Summary of partial assimilation metric

The assimilation scores capture assimilation well. **Figure 3** shows the assimilation scores for the target-trigger tokens for each vowel type, divided into two categories depending on whether the corpus annotators transcribed each token as being assimilated. As expected, tokens that were transcribed as phonetically assimilated have higher assimilation scores (mean = 0.81) than those transcribed as canonical [n] (mean = 0.34), and this is also true within all vowel types in the data. Validating our measure of partial assimilation, the partial assimilation scores were significantly higher for tokens transcribed in the corpus as assimilated ($t(199.1) = -4.81, p < .0001$).

At the same time, **Figure 3** also shows that there is considerable variation in the degree of F2 assimilation within tokens that received the same phonetic transcription, including considerable overlap in F2 values between phonetic categories. This confirms our motivation for the analysis of partial assimilation: Phonetic transcriptions are only part of the story. A stronger, and arguably more appropriate, test of the hypothesis that

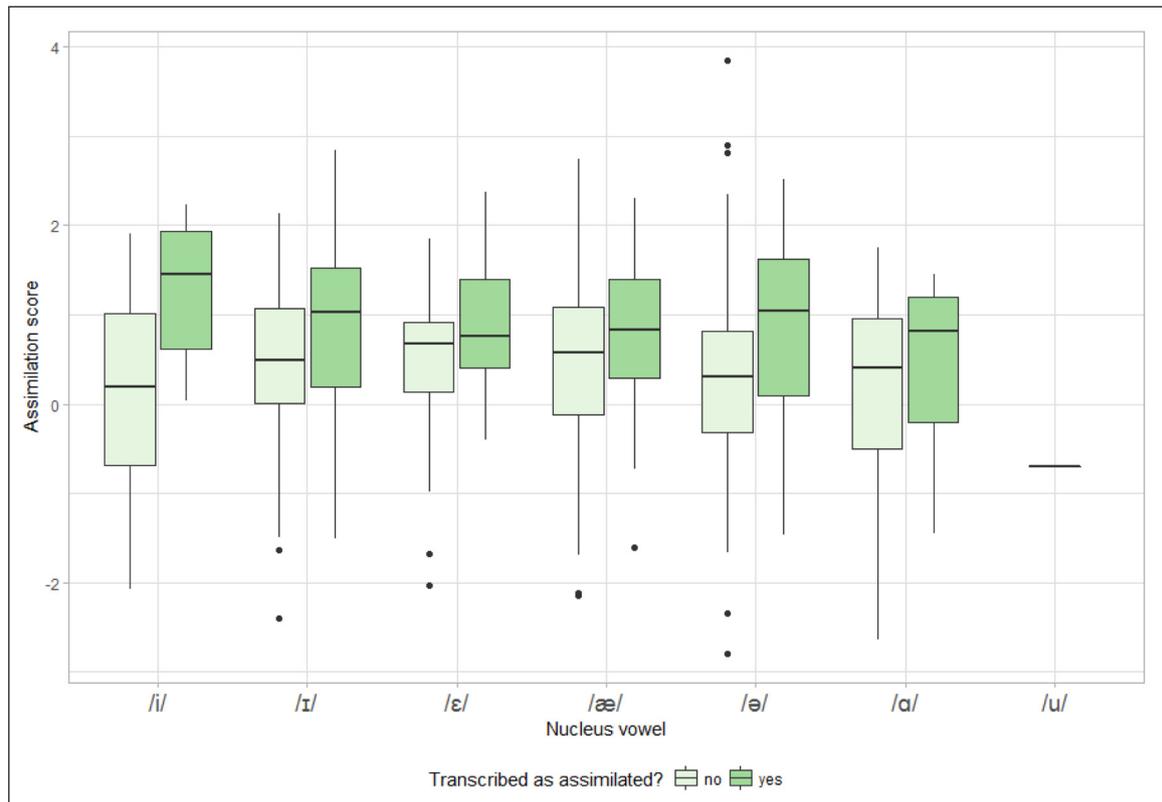


Figure 3: Assimilation scores at vowel offsets, standardized within-talker and within-vowel. Scores in labial contexts were inverted to produce the final assimilation score used in the analysis. The left box of each pair shows measurements before nasal /n/ codas that were transcribed as non-assimilated ([n]) by the corpus annotators; the right box of each pair shows measurements for those transcribed as assimilated ([m, ŋ]).

assimilation is modulated by the predictability of target and trigger words thus should analyze patterns of partial assimilation.

Next, we describe how we obtained estimates of word predictability, within-word predictability, and word frequency from spoken language databases.

2.2.3 Estimating word predictability

Message-oriented phonology predicts that there should be more nasal place assimilation when the target word is *predictable* and the trigger word is *unpredictable* (see **Figure 1**). For each target-trigger bigram token, we calculated two contextual probabilities: the bigram probability of the target word given the trigger, e.g., the backward probability $p(\text{ten}|\text{bucks})$ in *ten bucks*, and the bigram probability of the trigger word given the target, e.g., the forward probability $p(\text{bucks}|\text{ten})$ in *ten bucks*. The first measure captures how inferrable or predictable the target word is given the trigger. Our hypothesis is that if the target word is highly predictable, the target should be more likely to undergo assimilation, since there is relatively less need for the target to be produced reliably with a coronal nasal. The second measure captures how inferrable or predictable the trigger word is given the target. Our hypothesis is that if the trigger word is highly predictable, the target should be *less* likely to undergo assimilation, since there is relatively less need to enhance the perceptibility of the trigger onset's place of articulation.

Following standard practice, we can estimate the bigram predictability as experienced by a typical speaker in our data set. There are two related and competing challenges that any such estimation needs to keep in mind. The first challenge is that the distribution of words differs based on genre, register, mode, audiences, etc. Since we are analyzing the effect

of bigram predictability on assimilation in (a type of) conversational speech, it is thus preferable to obtain our bigram estimates from similar, if not the same, data. In fact, from this perspective, using only the Buckeye Corpus to estimate bigram predictability would seem best. However, this would conflict with a second challenge: The quality of estimates increases with the amount of the data that they are based on. From this perspective, it is thus desirable to use larger corpora to estimate bigram predictability.

We address these two challenges by obtaining bigram probability estimates from both the Buckeye Corpus and a much larger conversational speech corpus (Fisher Part 2, Cieri, Graff, Kimball, Miller, & Walker, 2005). The Fisher Part 2 corpus contains a similar linguistic genre as the Buckeye Corpus (informal one-on-one conversations) but contains over 30 times as many words as Buckeye (about 11,000,000 versus about 300,000 words). The 1137 target-trigger bigram tokens used for the main analyses were held out from the Buckeye Corpus, and not used to estimate bigram probabilities.⁷

The bigram probability estimates for each corpus were obtained using SRILM (Stolcke, 2002; Stolcke, Zheng, Wang, & Abrash, 2011) with the modified Kneser-Ney smoothing procedure (Chen & Goodman, 1998), which uses three discounting parameters per model. Kneser-Ney smoothing provides a reliable way to estimate probabilities for unseen words.

The two estimates for each bigram were then combined by taking a weighted average. The weight was chosen to optimize perplexity on 10% of the Buckeye Corpus, which was held out from the estimation only for the purposes of optimizing this weight parameter. This step ensures that the resulting bigrams are good estimates of words' predictability in the Buckeye Corpus—i.e., the broader linguistic context (topic distribution, register, etc.) for which we analyze assimilation.

The procedure (estimating smoothed Buckeye and Fisher language models, and combining them using an optimized weight parameter) was repeated twice: once using the two corpora as-is, to estimate forward trigger bigram probabilities, and once using the two corpora in reverse, to estimate backward target bigram probabilities. For the forward probabilities, the optimal weights were 0.514 for Buckeye data and 0.486 for Fisher data; for the backward probabilities, the optimal weights were 0.511 for Buckeye and 0.489 for Fisher. In other words, our estimates are based in roughly equal parts on the Buckeye and Fisher data (reflecting the trade-off between the two pressures described above).

We note that this estimation procedure is independent of the analyses of assimilation presented below: The smoothing and interpolation parameters were fit so as to optimize *word prediction* in held-out Buckeye data (so as to best approximate word predictability as experienced by a typical speaker in the Buckeye Corpus), not to maximize the effect of predictability on assimilation.

2.2.4 Estimating within-word predictability

Communicative accounts also predict effects of within-word segment predictability on assimilation. We thus obtained estimates of within-word predictability, defined as the probability of the target or trigger segment (i.e., word-final /n/ or word-initial /p, b, k, g/) conditioned on the preceding segments within the word (following van Son & Pols, 2003; Cohen Priva, 2008). For example, in a word like *between*, the final target /n/ is entirely predictable given the preceding string /bitwi/, because every English word that begins with the string /#bitwi/ is followed by /n/. For a word like *run*, however, the final target /n/ is much less likely, due to the existence of other words like *rough*, *rucksack*, *rudder*, and so on. These probability estimates were calculated based on the frequencies of

⁷ If the target-trigger bigram tokens are not held out from the Buckeye Corpus, the pattern of results is qualitatively similar to that presented in the present analysis.

strings of segments within the word tokens in the Buckeye Corpus mixed with the Fisher Corpus (see Section 2.2.5).

The within-word probability of word-initial trigger segments /p, b, k, g/ tends to be lower than for word-final target segments, because the trigger segment is conditioned only on a word boundary. In particular, the probability of a word-initial /p/ was calculated to be 5.2%, a word-initial /b/ was 4.7%, a word-initial /k/ was 9.6%, and a word-initial /g/ was 3.0%. The probability of word-initial /k/ is higher compared to the other word-initial segments because of the high number of relatively frequent words that begin with /k/ (*can, can't, come, etc.*), and vice-versa for /g/, with /p, b/ in between. Thus, segment-based communicative accounts predict that /k/ is less likely to trigger assimilation—since it is relatively more predictable as an onset, and thus contributes less to the identifiability of the trigger word—whereas /g/ is relatively *more* likely to do so.

2.2.5 Estimating word frequency

Estimates of unigram probability (word frequency) were obtained from a weighted mixture of the Buckeye and Fisher corpora, following the same methods for the word predictability estimates described above. Since Kneser-Ney smoothing is not applicable to unigrams, smoothed estimates for unseen words (those with zero tokens in the training data) were generated by re-dividing the total probability mass for all word types with one token in the training data equally among all types with zero or one tokens. The optimal weights were 0.967 for Buckeye and 0.033 for Fisher. Compared to the word predictability estimates obtained above, which had about equal weights for Buckeye and Fisher, our frequency estimates rely almost exclusively on the Buckeye Corpus. This is not surprising. The average number of tokens that go into a bigram predictability estimate is much smaller than the average number of tokens that go into a (unigram) frequency estimate. This means that, for word frequency estimation, the second of the two challenges described above is ameliorated. Essentially, the number of words in the Buckeye Corpus is sufficient to robustly estimate word frequency, but not quite sufficient to robustly estimate contextual predictability, such that the predictability estimation benefits from the additional data in Fisher despite the fact that Fisher has a *somewhat* different bigram distribution than Buckeye. We note that this is a *general* asymmetry between unigram frequency and bigram predictability estimates, and not specific to our study. We also note that this might create a bias *against* word predictability, in favor of word frequency, as the former estimates are based on less domain-specific (Buckeye) data (for relevant discussion, see also Cohen Priva & Jaeger, 2018).

2.2.6 Correlations among predictor variables

Following common practices, all frequency and predictability estimates were log-transformed prior to analysis. The predictor variables discussed here tend to exhibit some degree of collinearity, which can make it difficult to tease apart the individual contribution of any one variable. **Figure 4** shows correlation matrices for all six variables. The left panel depicts the data used in our analysis of categorical assimilation; the right panel depicts only the subset of data used in our analysis of partial assimilation. For both datasets, the strongest correlations were obtained between the measures of target word frequency and target word predictability and between trigger word frequency and trigger word predictability. These relationships are not surprising, given the definition of these variables. The other correlations are relatively low and follow expected patterns.

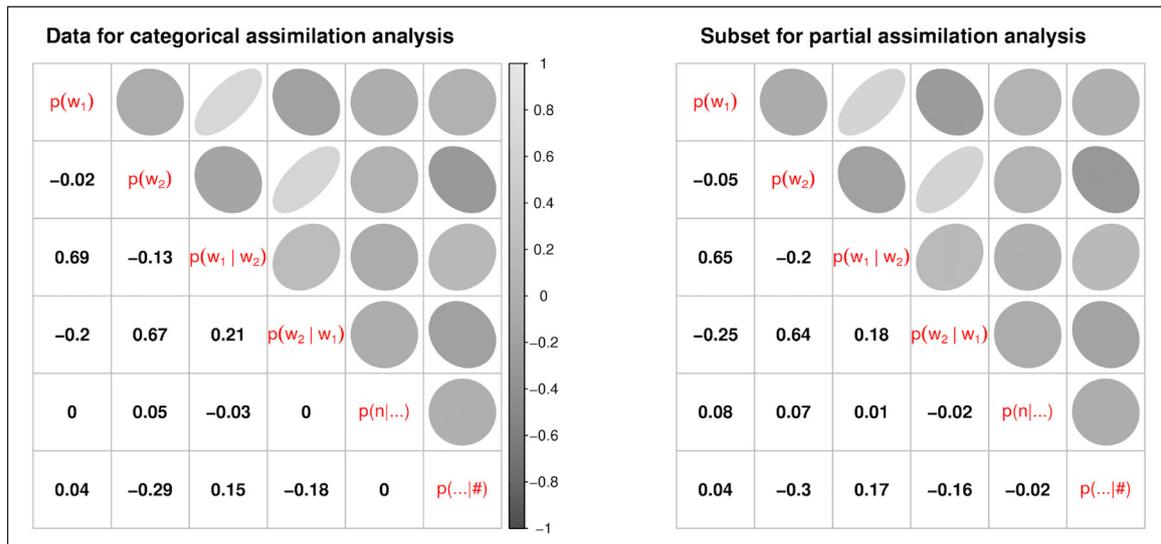


Figure 4: Correlation matrices among the six (log-transformed) predictor variables. Abbreviations: $p(w_1)$ is target word frequency; $p(w_2)$ is trigger word frequency; $p(w_1|w_2)$ is target word predictability; $p(w_2|w_1)$ is trigger word predictability; $p(n|...)$ is target within-word predictability; and $p(...|#)$ is trigger within-word predictability. Lower triangle shows correlation coefficients; upper triangle shows shaded ellipses depicting strength and direction of correlation. Left panel is the entire dataset; right panel is just the partial assimilation subset. Frequency and word predictability exhibit moderate correlations.

2.3 Analysis

We conducted two sets of analyses using multilevel (mixed-effects) regression. The first analysis set used logistic multilevel regression to predict transcribed categorical assimilation (assimilated or non-assimilated) as a dichotomous outcome for each target-trigger sequence. The second set of analyses used linear multilevel regressions to predict the acoustic measure of assimilation, as described in Section 2.2.2. Each set of regression analyses included three sets of parameters, as described in Section 2.2.3: (A) target and trigger word frequency ($p(\text{ten})$ and $p(\text{bucks})$); (B) target and trigger word predictability (conditional probabilities $p(\text{ten}|\text{bucks})$ and $p(\text{bucks}|\text{ten})$); and (C) within-word predictability (conditional segment probabilities $p(/n/|/#t\epsilon/)$ and $p(/b/|#)$).

The logistic regressions included by-speaker and by-target intercepts. The linear regressions included by-target intercepts, but not by-speaker intercepts, since the acoustic measure of assimilation was already standardized within-speaker.

Models were constructed with all seven possible combinations of the three sets of parameters: that is, one model with all three sets (frequency, lexical predictability, and within-word predictability); three models with just a single set; and three models with two sets of parameters (e.g., within-word and lexical predictability). This step allowed us to directly compare the relative explanatory power of the various predictors and to find the most parsimonious model.

We used Bayesian inference to estimate the full probability distributions for the effect sizes of these factors on nasal place assimilation. The posterior distribution of the regression parameters was estimated by Markov Chain Monte Carlo, implemented in Stan using the *rstanarm* R package for Bayesian regression modeling (R Core Team, 2016; Stan Development Team, 2016a, 2016b). We estimated the posterior density for each effect using a symmetrical, weakly-informative prior (the default in *rstanarm* version 2.17.4: a normal distribution with $\mu = 0$ and $\sigma = 2.5$, scaled by the standard deviation of each

predictor variable). By doing so, we are able to assess our relative confidence in a positive versus a negative effect for each of the crucial predictors without bias. Our predictions were that the target probability should have more density associated with a positive effect, while the trigger probability should have more density associated with a negative effect. That is, if *ten* is contextually predictable, it is more likely to be reduced via assimilation, while if *bucks* is contextually predictable, it is less likely to be enhanced via assimilation.

Model fit was evaluated by leave-one-out cross validation, which can be used to derive an expected logarithm-transformed pointwise predicted density (ELPD), an approximate measure of how accurately the model predicts the data. The ELPD is conceptually similar to the Akaike information criterion in that it can be used to compare models with different sets of predictors, and is the recommended practice for Bayesian model comparisons (Vehtari, Gelman, & Gabry, 2017). An ELPD closer to zero indicates a better model fit when considering model complexity.

3. Results

3.1 Categorical assimilation

Table 2 presents a summary of the results of the models for the analysis of transcribed nasal place assimilation. Each column shows one of the seven models calculated. Model terms are abbreviated as follows: WordFreq is target and trigger word frequency; WordPred is target and trigger word predictability; and SegPred is target and trigger within-word predictability (or ‘segmental predictability’). The model preferred by the ELPD is shown in bold, but we note that all models perform very similarly. We thus show the outcomes of all models in **Figure 5**. This is also in the spirit of Bayesian analyses in that it visualizes the uncertainty in the effects across different models.

The rows of **Table 2** depict values relating to each of the six parameters: target word frequency, trigger word frequency, target word predictability, trigger word predictability, target within-word predictability, and trigger within-word predictability. For each parameter, there are three values of interest. The first is the proportion of the posterior density in the predicted direction. Recall that our hypothesis predicts that the posterior density should be positive for the target word parameters and negative for the trigger word parameters. This proportion value therefore depicts the extent to which the results are consistent with our hypothesis: 1.0 would mean that the entire distribution is in the predicted direction, constituting the strongest possible evidence for our hypothesis; 0.5 that the distribution is exactly split around zero; and 0.0 that the entire distribution is in the opposite direction predicted, constituting the strongest possible evidence against our hypothesis.

The other two values of interest give the range of the 95% highest posterior density interval (HPDI). Broadly speaking, we can interpret this interval as having a 95% chance of including the true value of the parameter, under the assumptions made in the analysis. This interval does not have the same interpretation as a frequentist confidence interval (Morey, Hoekstra, Rouder, Lee, & Wagenmakers, 2016). Finally, the last row of **Table 2** shows the ELPD value for each of the models.

For the target, we find strong support for the predicted effects of word and within-word predictability. The majority of the posterior densities for these parameters are positive and most of the 95% HPDI values do not overlap with zero. This result is consistent with the predictions of communicative accounts: If $p(\text{ten}|\text{bucks})$ (or $p(n|\#t\epsilon)$) is high, then *ten* is contextually predictable, and thus more easily inferrable. Because of this, the phonetic cues to the coda /n/ are less important for robust communication, and nasal place assimilation (which reduces cues) is more likely.

Table 2: Results for categorical assimilation. Columns show different models. Rows show different predictors: word frequency; word predictability; and within-word predictability. For each predictor, the first row shows the posterior density compatible with the tested hypothesis (i.e., > 0 for target frequency and predictability and < 0 for trigger frequency/predictability). The best model based on Bayesian model comparison is shown in bold.

| | | WordFreq + WordPred + SegPred | WordFreq + WordPred | WordPred + SegPred | WordFreq + WordPred + SegPred | WordFreq + WordPred | WordPred + SegPred | WordFreq + WordPred + SegPred | WordFreq + WordPred | WordPred + SegPred |
|---|-----------------------|-------------------------------------|------------------------|-----------------------|-------------------------------------|------------------------|-----------------------|-------------------------------------|------------------------|-----------------------|
| Target frequency (log-transformed) | Posterior density > 0 | 0.635 | 0.760 | 0.905 | 0.854 | 0.905 | 0.854 | 0.905 | 0.854 | 0.905 |
| | (2.5% HPDI) | -0.302 | -0.223 | -0.047 | -0.077 | -0.047 | -0.077 | -0.047 | -0.077 | -0.047 |
| | (97.5% HPDI) | 0.474 | 0.558 | 0.422 | 0.374 | 0.422 | 0.374 | 0.422 | 0.374 | 0.422 |
| Trigger frequency (log-transformed) | Posterior density < 0 | 0.810 | 0.804 | 0.896 | 0.733 | 0.896 | 0.733 | 0.896 | 0.733 | 0.896 |
| | (2.5% HPDI) | -0.457 | -0.467 | -0.269 | -0.205 | -0.269 | -0.205 | -0.269 | -0.205 | -0.269 |
| | (97.5% HPDI) | 0.143 | 0.149 | 0.037 | 0.091 | 0.037 | 0.091 | 0.037 | 0.091 | 0.037 |
| Target predictability (log-transformed) | Posterior density > 0 | 0.830 | 0.576 | 0.998 | 0.966 | 0.998 | 0.966 | 0.998 | 0.966 | 0.966 |
| | (2.5% HPDI) | -0.056 | -0.118 | 0.051 | 0.008 | 0.051 | 0.008 | 0.051 | 0.008 | 0.008 |
| | (97.5% HPDI) | 0.218 | 0.149 | 0.193 | 0.147 | 0.193 | 0.147 | 0.193 | 0.147 | 0.147 |
| Trigger predictability (log-transformed) | Posterior density < 0 | 0.316 | 0.209 | 0.676 | 0.488 | 0.676 | 0.488 | 0.676 | 0.488 | 0.488 |
| | (2.5% HPDI) | -0.094 | -0.064 | -0.079 | -0.061 | -0.079 | -0.061 | -0.079 | -0.061 | -0.061 |
| | (97.5% HPDI) | 0.175 | 0.202 | 0.044 | 0.062 | 0.044 | 0.062 | 0.044 | 0.062 | 0.062 |
| Target within-word predictability (log-transformed) | Posterior density > 0 | 0.983 | 0.979 | 0.989 | 0.950 | 0.989 | 0.950 | 0.989 | 0.950 | 0.950 |
| | (2.5% HPDI) | 0.080 | 0.069 | 0.108 | 0.001 | 0.108 | 0.001 | 0.108 | 0.001 | 0.001 |
| | (97.5% HPDI) | 0.702 | 0.667 | 0.699 | 0.598 | 0.699 | 0.598 | 0.699 | 0.598 | 0.598 |
| Trigger within-word predictability (log-transformed) | Posterior density < 0 | 1.000 | 0.999 | 0.999 | 0.998 | 0.999 | 0.998 | 0.999 | 0.998 | 0.998 |
| | (2.5% HPDI) | -2.689 | -2.428 | -2.611 | -2.327 | -2.611 | -2.327 | -2.611 | -2.327 | -2.327 |
| | (97.5% HPDI) | -0.928 | -0.760 | -0.913 | -0.654 | -0.913 | -0.654 | -0.913 | -0.654 | -0.654 |
| ELPD | | -501.3 | -506.1 | -499.9 | -504.7 | -499.9 | -504.7 | -499.9 | -504.7 | -501.0 |

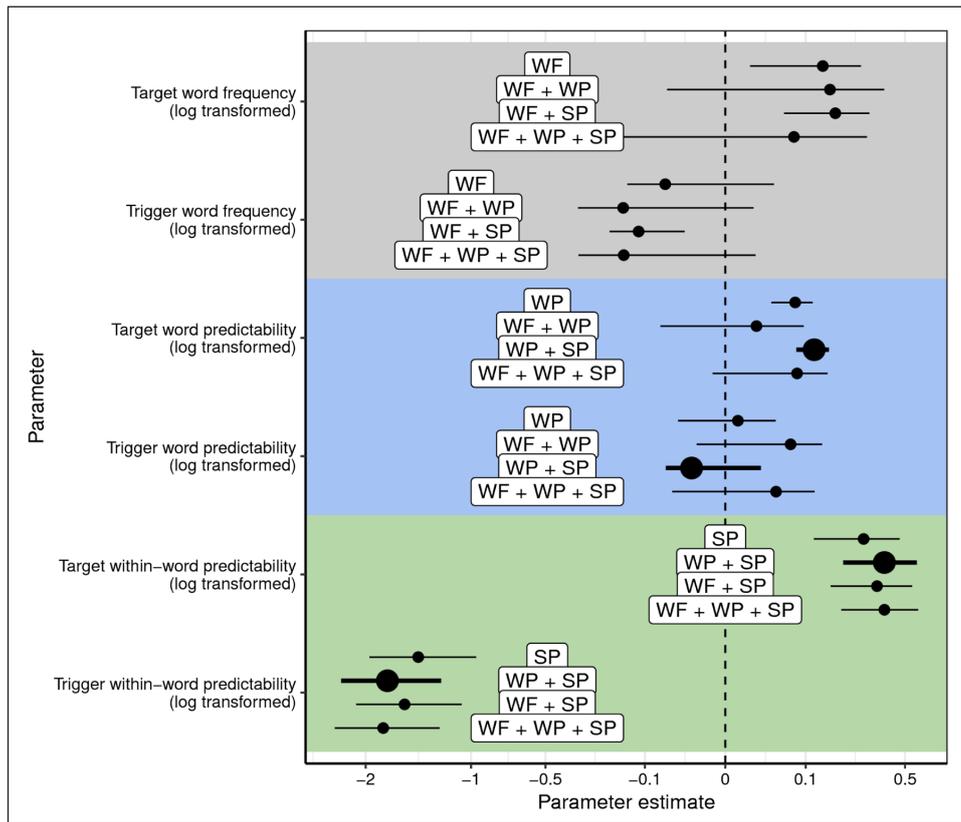


Figure 5: Parameter estimates and 95% HPDI for each parameter in each model predicting categorical assimilation. Model labelling abbreviates that of Table 2: WF: word frequency; WP: word predictability; SP: within-word predictability. The best model by Bayesian model comparison (model WP+SP) is highlighted with a larger point and thicker lines.

For the trigger, however, we only observed strong evidence in the hypothesized direction for the within-word predictability parameter. That is, if $p(b|\#)$ is high, then the /b/ of *bucks* is contextually predictable and thus there is less motivation for assimilation to occur. For the word predictability $p(bucks|ten)$, the results do not provide strong evidence in favor of either a negative or a positive effect.

We note that the evidence for effects of frequency is diminished in the models with parameters for word predictability. This is the always case for target frequency (see “WordFreq + WordPred” vs. “WordFreq” and “WordFreq + WordPred + SegPred” vs. “WordFreq + SegPred”), and sometimes for trigger frequency (“WordFreq + WordPred + SegPred” vs. “WordFreq + SegPred”). Likewise, the evidence for the effects of word predictability are diminished in all models which include parameters for frequency (models “WordFreq + WordPred” vs. “WordFreq + WordPred + SegPred” and model “WordPred” vs. “WordFreq + WordPred”). This is expected given the correlations between word frequency and predictability (see Figure 4), and highlights the need to control for both types of variables when assessing one of them (see Cohen Priva & Jaeger, 2018). Overall our analysis of categorical assimilation presents more evidence for effects of word predictability than for effects of word frequency. The evidence for effects of within-word predictability is strong in all models.

In summary, the analysis of categorical assimilation finds strong evidence for three of the four predictability effects hypothesized by communicative accounts, and no conflicting evidence. And, regardless of what combination of factors is considered, we find that the likelihood of assimilation seems to be governed not only by the inferrability of the target word, but also by that of the trigger word.

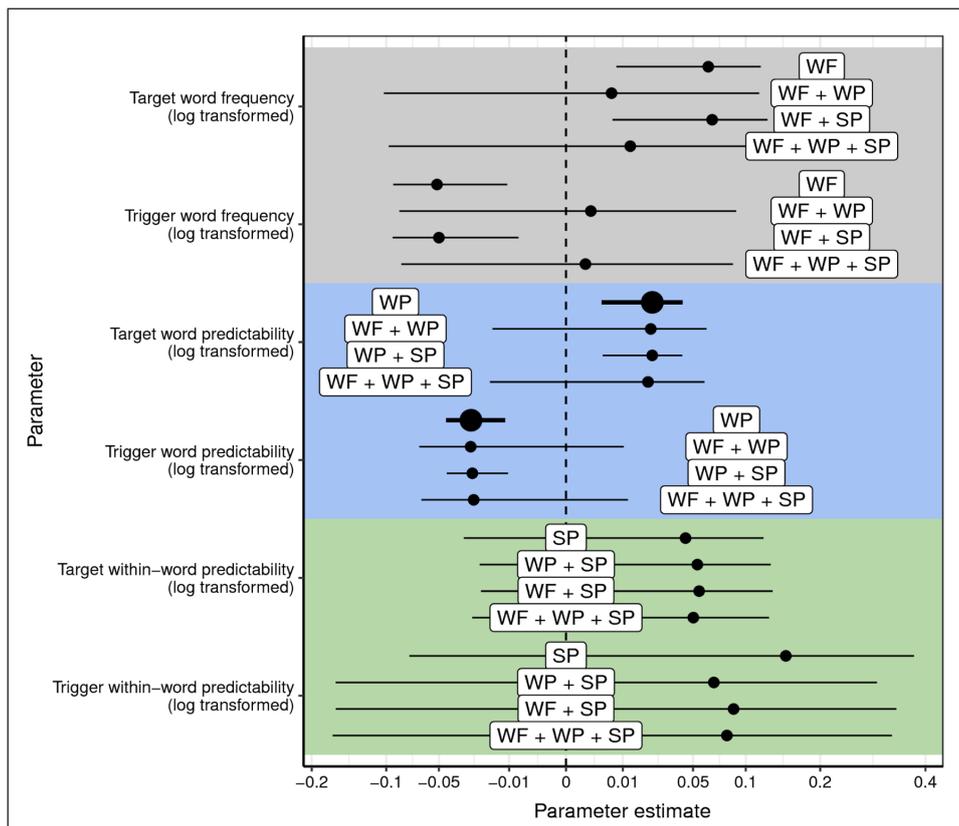


Figure 6: Parameter estimates and 95% HPDI for each parameter in each model predicting partial assimilation. Model labelling abbreviates that of Table 3: WF: word frequency; WP: word predictability; SP: within-word predictability. The best model by Bayesian model comparison (model WP) is highlighted with a larger point and thicker lines.

3.2 Partial assimilation

Table 3 and Figure 6 present a summary of the results of the models for the analysis of partial nasal assimilation. This table and figure are organized in the same way as Table 2 and Figure 5. The results partly replicate and complement the results of the analysis of categorical assimilation. We find support for effects of both frequency and predictability for both targets and triggers. Unlike for the analysis of categorical assimilation, we do not find support for within-word predictability.

As with the categorical analysis, effects of word frequency and word predictability again were diminished when included in the same model, as expected given their correlation. The ELPD suggests that there is more evidence for an effect of word predictability than frequency, although the ELPD differences among the models are modest.

The analysis of partial assimilation thus provides strong support for two of the predictability effects hypothesized by communicative accounts, and little to no conflicting evidence.

4. Discussion and conclusion

Our analyses of word-final /n/ in assimilation-licensing contexts in the Buckeye Corpus suggest a strong role of predictability in mediating nasal place assimilation. Our categorical analysis of the corpus’ phonetic transcriptions demonstrate that annotators were more likely to transcribe assimilation when the predictability of the assimilation target (e.g., *ten*) given the trigger (e.g., *bucks*) was high. Our acoustic analysis of the same tokens additionally finds that the degree of place assimilation apparent in F2 was greater when the predictability of the trigger given the target was low. Our analyses also demonstrate

Table 3: Results for partial assimilation. Columns show different models. Rows show different predictors: word frequency; word predictability; and within-word predictability. For each predictor, the first row shows the posterior density compatible with the tested hypothesis (i.e., > 0 for target frequency and predictability and < 0 for trigger frequency/predictability). The best model based on Bayesian model comparison is shown in bold.

| | | WordFreq + WordPred + SegPred | WordFreq + WordPred + SegPred | WordFreq + WordPred + SegPred | WordFreq + WordPred + SegPred | WordPred + SegPred | WordPred + SegPred | WordFreq + WordPred + SegPred | WordPred + SegPred | WordFreq + WordPred + SegPred | WordPred + SegPred |
|---|-----------------------|-------------------------------------|-------------------------------------|-------------------------------------|-------------------------------------|-----------------------|-----------------------|-------------------------------------|-----------------------|-------------------------------------|-----------------------|
| Target frequency (log-transformed) | Posterior density > 0 | 0.548 | 0.516 | 0.865 | 0.858 | | | | | | |
| | (2.5% HPDI) | -0.174 | -0.166 | -0.033 | -0.035 | | | | | | |
| | (97.5% HPDI) | 0.187 | 0.176 | 0.16 | 0.157 | | | | | | |
| Trigger frequency (log-transformed) | Posterior density < 0 | 0.496 | 0.475 | 0.883 | 0.885 | | | | | | |
| | (2.5% HPDI) | -0.142 | -0.132 | -0.12 | -0.12 | | | | | | |
| | (97.5% HPDI) | 0.143 | 0.136 | 0.02 | 0.018 | | | | | | |
| Target predictability (log-transformed) | Posterior density > 0 | 0.699 | 0.726 | 0.89 | 0.888 | | | | | | |
| | (2.5% HPDI) | -0.043 | -0.039 | -0.009 | -0.008 | | | | | | |
| | (97.5% HPDI) | 0.088 | 0.082 | 0.054 | 0.054 | | | | | | |
| Trigger predictability (log-transformed) | Posterior density < 0 | 0.755 | 0.788 | 0.948 | 0.955 | | | | | | |
| | (2.5% HPDI) | -0.09 | -0.089 | -0.055 | -0.056 | | | | | | |
| | (97.5% HPDI) | 0.036 | 0.034 | 0 | -0.001 | | | | | | |
| Target within-word predictability (log-transformed) | Posterior density > 0 | 0.743 | 0.756 | 0.747 | 0.712 | | | | | | |
| | (2.5% HPDI) | -0.074 | -0.076 | -0.08 | -0.085 | | | | | | |
| | (97.5% HPDI) | 0.18 | 0.182 | 0.18 | 0.173 | | | | | | |
| Trigger within-word predictability (log-transformed) | Posterior density < 0 | 0.374 | 0.403 | 0.354 | 0.247 | | | | | | |
| | (2.5% HPDI) | -0.313 | -0.319 | -0.3 | -0.226 | | | | | | |
| | (97.5% HPDI) | 0.47 | 0.463 | 0.482 | 0.531 | | | | | | |
| ELPD | | -908.9 | -907.3 | -907.5 | -907.2 | | | | | -905.5 | -905.2 |

that higher within-word target predictability led to more transcribed assimilation, and that higher within-word trigger predictability led to less transcribed assimilation (we return to the asymmetry in the results between categorical and partial assimilation below).

Our results also speak to the relative plausibility of different usage-based explanations of assimilation. Model comparison suggested that the models with word predictability, rather than frequency, were preferable. This result implies that the observed frequency effects are likely to be predictability effects. For reasons described in the method section, frequency estimates tend to be more robust than predictability estimates, creating a bias *against* detecting predictability effects. The fact that we nevertheless find word predictability to be a better predictor of assimilation than word frequency is thus quite informative. At first blush, this result may seem to be at odds with those of Ernestus, Lahey, Verhees, and Baayen (2006), who observed effects of frequency in Dutch voicing coarticulation. However, Ernestus and colleagues investigated effects of *collocation* frequency, which is more closely related to the type of bigram predictability measure we used here. We also note that Ernestus and colleagues—like the bulk of studies on reduction and other pronunciation variation—did not include controls for predictability in their model (for exceptions, see e.g., Ernestus, Hanique, & Verboom, 2015; Gahl & Garnsey, 2004; Gahl et al., 2012; Scarborough, 2010; Seyfarth, 2014). This omission is potentially critical, as recent computational work suggests: In the absence of predictability controls, effects of predictability are *very* likely to lead to spurious frequency effects (Cohen Priva & Jaeger, 2018). Our result thus echo Cohen Priva and Jaeger’s call for further studies that assess the effects of frequency and predictability simultaneously.

Crucially, our results are inconsistent with any account that attributes reduction or enhancement solely to routinization or repetition. Under those hypotheses, we would expect more assimilation when the target *as well as* the trigger are frequent. However, inasmuch as we found evidence for frequency effects, more frequent trigger words led to less assimilation (mirroring the results for word predictability), contrary to accounts that focus on routinization or repetition as the source of reduction.

These results are, however, consistent with a communicative approach where the phonetic form of messages (e.g., words) can be shaped by biases towards accurate message transmission and low resource cost (e.g., Buz et al., 2016; Lindblom, 1990; Hall et al., 2016). All else being equal, cues to a given word may be either reduced or enhanced depending on their potential contribution to successfully conveying the word. Along these lines, reducing the coronal place cue in words that are predictable from the following word promotes cost-effective communication; there would be little communicative gain in investing resources in the clear production of the nasal given that the word it occurs in is predictable. Conversely, increasing redundancy in the trigger word is advantageous when the word has low contextual predictability. Similar explanations can account for a large range of reduction/enhancement phenomena during language production (for a recent overview, see Jaeger & Buz, 2017). This includes implicit decisions during not only phonetic and phonological encoding, but also during morphological (e.g., *don’t* vs. *do not*, Frank & Jaeger, 2008; optional case-morphology, Kurumada & Jaeger, 2015), lexical (Mahowald, Fedorenko, Piantadosi, & Gibson, 2013) or syntactic encoding (e.g., optional argument or function word omission, Jaeger, 2010; Resnik, 1996; Wasow, Levy, Melnick, Zhu, & Juzek, 2015). Though many questions remain, a new generation of experimental and historical studies has now also begun to investigate in more depth the link between such implicit decisions during language processing and communicative effects on the phonology, lexicon, and grammar over historical time (e.g., Fedzechkina, Newport, & Jaeger, 2017; Kanwal, Smith, Culbertson, & Kirby, 2017; Sóskuthy & Hay, 2017; Wedel, Nelson, & Sharp, 2018).

One possible mechanism for the observed results relies on the relative prosodic weights of the target word and the trigger word. It is known that prosodic boundaries between adjacent words can block variable assimilation processes (Kuzla, Cho, & Ernestus, 2007); and that contextual predictability plays a role in the relative prosodic prominence of individual words, independently of information-structural considerations like focus (Brenier, Nenkova, Kothari, Whitton, Beaver, & Jurafsky, 2006; Turnbull, 2017). However, this explanation is difficult to reconcile with our specific pattern of results, where we observed high predictability targets having more assimilation but high predictability triggers having less assimilation. Any proposed prosodic mechanism would have to strengthen the prosodic boundary preceding a predictable word (to account for the trigger word pattern), and have to weaken the prosodic boundary following a predictable word (to account for the target word pattern). While this mechanism is not outside of the realm of possibility, it lacks the clear motivation and intuitive appeal of a communicative approach couched in terms of enhancing and reducing cues to word identity.

We found evidence for effects of both word and within-word predictability. The effect of the former was comparable for both categorical and partial assimilation, except that we did not find evidence that trigger word predictability affected categorical assimilation. The effect of within-word predictability differed between the two analyses to a greater extent. For categorical assimilation, within-word predictability of both the target and the trigger was strongly associated with effects in the hypothesized directions. For partial assimilation, there was no evidence for an effect of within-word predictability in any direction. These findings could indicate different mechanisms for full assimilation versus partial assimilation, although further work is needed to assess this hypothesis.

Future work is also required to tease apart possible explanations for the effects of within-word predictability. Communicative accounts that link phonological processes to a bias for robust message transmission, and thus inferrability, can be formulated with regard to different levels of representation. This includes the transmission of phonological segments (Cohen Priva, 2015) or the transmission of meaning-bearing units (e.g., words, Hall et al., 2016), or both (for discussion, see Hall et al., 2016). While the present study would seem to provide evidence for both formulations, the effects of within-word predictability can also be interpreted with reference to word (rather than segment) inferrability. In particular, the predictability of the last segment of the target word also describes the target word's predictability given the segmental context.

Another area for future work concerns the sensitivity of assimilation to its potential effects whether or not applying assimilation to the target word will make it homophonous with another word. For example, assimilation of *ten bucks* to [tɛm bʌks] is unambiguous, as /tɛm/ is not an English word; but in a sequence like *a quick run picks you up*, nasal assimilation may not necessarily enhance the likelihood of communicative success due to the presence of the word *rum* in the lexicon (Gaskell & Marslen-Wilson, 2001). Recent work suggests that minimal pairs may indeed play a role in determining articulation (Buz & Jaeger, 2016; Buz et al., 2016; Wedel et al., 2018). While the present data do not realistically allow for such an investigation (only 11% of all word sequence types have a /m/ or /ŋ/-final competitor for the target word, implying low statistical power), we believe that this prediction holds a great deal of promise.

More broadly, these results expand the set of factors shown to influence patterns of assimilation. While previous research has shown the importance of articulatory and perceptual factors, the probabilistic nature of assimilation has received comparatively little attention (but see Coetzee, 2016; Coleman, Renwick, & Temple, 2016; Dilley & Pitt, 2007; Steele, Colantoni, & Kochetov, in press). Indeed, the amount of data is rather limited, especially given that the effects we are assessing are predicted to be relatively subtle

(see Hall, Hume, Jaeger, & Wedel, 2018 among others). The finding that the contextual predictability of target and trigger words play a role in predicting place assimilation thus adds to the growing body of evidence showing that sound patterns are shaped by statistics of the lexicon (for review see Cohen Priva, 2017a, *inter alia*).

In this paper we explored the extent to which the contextual predictability of words conditions nasal place assimilation. While previous work has shown that reductive sound processes can be associated with high predictability, here we argued that assimilation—which enhances a target at the expense of a trigger—is conditioned by both high predictability of the target and low predictability of the trigger. This fits within a communicative, message-oriented approach to phonology: Assimilation shifts redundancy from the target to the trigger, and thus can be used to manage the relative inferrability of target and trigger.

Acknowledgements

We are grateful to Rebecca Morley for her assistance and guidance during earlier stages of this project, and to NZILBB members for feedback and direction. We also thank Uriel Cohen Priva and an anonymous reviewer for constructive comments on this paper. This research was partially supported by a National Science Foundation Graduate Research Fellowship to the second author (#DGE-1144086), and a National Science Foundation CAREER award (#IIS-1150028) to the fourth author. Any opinion, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

Competing Interests

The authors have no competing interests to declare.

References

- Aylett, M., & Turk, A. E. 2004. The smooth signal redundancy hypothesis: A functional explanation for relationships between redundancy, prosodic prominence, and duration in spontaneous speech. *Language and Speech*, 47(1), 31–56. DOI: <https://doi.org/10.1177/00238309040470010201>
- Aylett, M., & Turk, A. E. 2006. Language redundancy predicts syllabic duration and the spectral characteristics of vocalic syllable nuclei. *Journal of the Acoustical Society of America*, 119, 3048–3058. DOI: <https://doi.org/10.1121/1.2188331>
- Beddor, P. S. 2009. A coarticulatory path to sound change. *Language*, 85(4), 785–821. DOI: <https://doi.org/10.1353/lan.0.0165>
- Beddor, P. S., McGowan, K. B., Boland, J. E., Coetzee, A. W., & Brasher, A. 2013. The time course of perception of coarticulation. *Journal of the Acoustical Society of America*, 133, 2350–2366. DOI: <https://doi.org/10.1121/1.4794366>
- Bell, A., Brenier, J. M., Gregory, M., Girand, C., & Jurafsky, D. 2009. Predictability effects on durations of content and function words in conversational English. *Journal of Memory and Language*, 60(1), 92–111. DOI: <https://doi.org/10.1016/j.jml.2008.06.003>
- Brenier, J. M., Nenkova, A., Kothari, A., Whitton, L., Beaver, D., & Jurafsky, D. 2006. The (non)utility of linguistic features for prediction prominence in spontaneous speech. In: *Spoken Language Technology Workshop*, 54–57.
- Bresnan, J., & Spencer, J. 2013. Frequency and variation in English subject-verb contraction. Unpublished manuscript.
- Buz, E., & Jaeger, T. F. 2016. The (in)dependence of articulation and lexical planning during isolated word production. *Language, Cognition and Neuroscience*, 31(3), 404–424. DOI: <https://doi.org/10.1080/23273798.2015.1105984>

- Buz, E., Tanenhaus, M. K., & Jaeger, T. F. 2016. Dynamically adapted context-specific hyper-articulation: Feedback from interlocutors affects speakers' subsequent pronunciations. *Journal of Memory and Language*, 89, 68–86. DOI: <https://doi.org/10.1016/j.jml.2015.12.009>
- Bybee, J. 2001. *Phonology and Language Use*. Cambridge: Cambridge University Press. DOI: <https://doi.org/10.1017/CBO9780511612886>
- Bybee, J. 2006. From usage to grammar: The mind's response to repetition. *Language*, 82(4): 711–733. DOI: <https://doi.org/10.1353/lan.2006.0186>
- Chang, S., Plauche, M., & Ohala, J. 2001. Markedness and consonant confusion asymmetries. In: Hume, E., & Johnson, K. (eds.), *The role of speech perception in phonology*. New York: Academic Press.
- Chen, S., & Goodman, J. 1998. An empirical study of smoothing techniques for language modeling. *Technical Report TR-10-98*, Computer Science Group, Harvard University.
- Cieri, C., Graff, D., Kimball, O., Miller, D., & Walker, K. 2005. *Fisher English Training Part 2*. Linguistic Data Consortium, Philadelphia, PA.
- Clopper, C. G., & Turnbull, R. 2018. Exploring variation in phonetic reduction: Linguistic, social, and cognitive factors. In: Cangemi, F., Clayards, M., Niebuhr, O., Schuppler, B., & Zellers, M. (eds.), *Rethinking reduction: Interdisciplinary perspectives on conditions, mechanisms, and domains for phonetic variation*. Berlin: de Gruyter.
- Coetzee, A. 2016. A comprehensive model of phonological variation: Grammatical and non-grammatical factors in variable nasal place assimilation. *Phonology*, 33, 211–246. DOI: <https://doi.org/10.1017/S0952675716000117>
- Cohen Priva, U. 2008. Using information content to predict phone deletion. In: Abner, N., & Bishop, J. (eds.), *Proceedings of the 27th West Coast Conference on Formal Linguistics*, 90–98. Somerville, MA: Cascadilla Proceedings Project.
- Cohen Priva, U. 2012. *Sign and Signal: Deriving linguistic generalizations from information utility*. Doctoral dissertation, Stanford University.
- Cohen Priva, U. 2015. Informativity affects consonant duration and deletion rates. *Laboratory Phonology*, 60(2), 243–278. DOI: <https://doi.org/10.1515/lp-2015-0008>
- Cohen Priva, U. 2017a. Informativity and the actuation of lenition. *Language*, 93(3), 569–597. DOI: <https://doi.org/10.1353/lan.2017.0037>
- Cohen Priva, U. 2017b. Not so fast: Fast speech correlates with lower lexical and structural information. *Cognition*, 160, 27–34. DOI: <https://doi.org/10.1016/j.cognition.2016.12.002>
- Cohen Priva, U., & Jaeger, T. F. 2018. The interdependence of frequency, predictability, and informativity in the segmental domain. *Linguistics Vanguard*, 4(S2). DOI: <https://doi.org/10.1515/lingvan-2017-0028>
- Cole, J., Mo, Y., & Baek, S. 2010. The role of syntactic structure in guiding prosody perception with ordinary listeners and everyday speech. *Language and Cognitive Processes*, 25, 1141–1177. DOI: <https://doi.org/10.1080/01690960903525507>
- Cole, J., Mo, Y., & Hasegawa-Johnson, M. 2010. Signal-based and expectation-based factors in the perception of prosodic prominence. *Laboratory Phonology*, 1, 425–452. DOI: <https://doi.org/10.1515/labphon.2010.022>
- Coleman, J., Renwick, M. E., & Temple, R. A. 2016. Probabilistic underspecification in nasal place assimilation. *Phonology*, 33(3), 425–458. DOI: <https://doi.org/10.1017/S0952675716000208>
- Diessel, H. 2007. Frequency effects in language acquisition, language use, and diachronic change. *New Ideas in Psychology*, 25(2), 108–127. DOI: <https://doi.org/10.1016/j.newideapsych.2007.02.002>

- Dilley, L. C., & Pitt, M. A. 2007. A study of regressive place assimilation in spontaneous speech and its implications for spoken word recognition. *The Journal of the Acoustical Society of America*, 122(4), 2340–2353. DOI: <https://doi.org/10.1121/1.2772226>
- Ellis, L., & Hardcastle, W. J. 2002. Categorical and gradient properties of assimilation in alveolar to velar sequences: Evidence from EPG and EMA data. *Journal of Phonetics*, 30, 373–396. DOI: <https://doi.org/10.1006/jpho.2001.0162>
- Ernestus, M. 2014. Acoustic reduction and the roles of abstractions and exemplars in speech processing. *Lingua*, 142, 27–41. DOI: <https://doi.org/10.1016/j.lingua.2012.12.006>
- Ernestus, M., Baayen, R. H., & Schreuder, R. 2002. The recognition of reduced word forms. *Brain and Language*, 81, 162–173. DOI: <https://doi.org/10.1006/brln.2001.2514>
- Ernestus, M., Hanique, I., & Verboom, E. 2015. The effect of speech situation on the occurrence of reduced word pronunciation variants. *Journal of Phonetics*, 48, 60–75. DOI: <https://doi.org/10.1016/j.wocn.2014.08.001>
- Ernestus, M., Lahey, M., Verhees, F., & Baayen, R. H. 2006. Lexical frequency and voice assimilation. *The Journal of the Acoustical Society of America*, 120(2), 1040–1051. DOI: <https://doi.org/10.1121/1.2211548>
- Evanini, K. 2009. *The permeability of dialect boundaries: A case study of the region surrounding Erie, Pennsylvania*. Doctoral dissertation, University of Pennsylvania.
- Evanini, K., Isard, S., & Liberman, M. 2009. Automatic formant extraction for sociolinguistic analysis of large corpora. *Interspeech 2009*.
- Fedzechkina, M., Newport, E. L., & Jaeger, T. F. 2017. Balancing effort and information transmission during language acquisition: Evidence from word order and case marking. *Cognitive Science*, 41(2). DOI: <https://doi.org/10.1111/cogs.12346>
- Flemming, E. 2010. Modeling listeners: Comments on Pluymaekers et al. and Scarborough. In: Fougeron, C., Kühnert, B., D'Imperio, M., & Vallée, N. (eds.), *Laboratory Phonology 10*. Mouton de Gruyter, Berlin.
- Fourtassi, A., Dunbar, E., & Dupoux, E. 2014. Self-consistency as an inductive bias in early language acquisition. *Proceedings of the 36th annual meeting of the Cognitive Science Society*.
- Fowler, C. A. 1994. Invariants, specifiers, cues: An investigation of locus equations as information for place of articulation. *Perception & Psychophysics*, 55, 597–610. DOI: <https://doi.org/10.3758/BF03211675>
- Fox, R. A., & Jacewicz, E. 2009. Cross-dialectal variation in formant dynamics of American English vowels. *The Journal of the Acoustical Society of America*, 126(5), 2603–2618. DOI: <https://doi.org/10.1121/1.3212921>
- Frank, A. F., & Jaeger, T. F. 2008. Speaking rationally: Uniform information density as an optimal strategy for language production. In: Love, B. C., McRae, K., & Sloutsky, V. M. (eds.), *Proceedings of the 30th Annual Meeting of the Cognitive Science Society (CogSci08)*, 939–944. Austin, TX: Cognitive Science Society.
- Gahl, S., & Garnsey, S. M. 2004. Knowledge of grammar, knowledge of usage: Syntactic probabilities affect pronunciation variation. *Language*, 80(4), 748–775. DOI: <https://doi.org/10.1353/lan.2004.0185>
- Gahl, S., Yao, Y., & Johnson, K. 2012. Why reduce? Phonological neighborhood density and phonetic reduction in spontaneous speech. *Journal of Memory and Language*, 66(4), 789–806. DOI: <https://doi.org/10.1016/j.jml.2011.11.006>
- Gaskell, M. G., & Marslen-Wilson, W. D. 2001. Lexical ambiguity resolution and spoken word recognition: Bridging the gap. *Journal of Memory and Language*, 44(3), 325–349. DOI: <https://doi.org/10.1006/jmla.2000.2741>

- Hall, K. C., Hume, E., Jaeger, T. F., & Wedel, A. 2016. The Message Shapes Phonology. Unpublished manuscript.
- Hall, K. C., Hume, E., Jaeger, T. F., & Wedel, A. 2018. The role of predictability in shaping phonological patterns. *Linguistics Vanguard*, 4(S2).
- Hura, S. L., Lindblom, B., & Diehl, R. L. 1992. On the role of perception in shaping phonological assimilation rules. *Language and Speech*, 35, 59–72. DOI: <https://doi.org/10.1177/002383099203500206>
- Jaeger, M., & Hoole, P. 2011. Articulatory factors influencing regressive place assimilation across word boundaries in German. *Journal of Phonetics*, 39, 413–428. DOI: <https://doi.org/10.1016/j.wocn.2011.03.002>
- Jaeger, T. F. 2010. Redundancy and reduction: Speakers manage syntactic information density. *Cognitive Psychology*, 61(1), 23–62. DOI: <https://doi.org/10.1016/j.cogpsych.2010.02.002>
- Jaeger, T. F. 2013. Production preferences cannot be understood without reference to communication. *Frontiers in Psychology* 4. DOI: <https://doi.org/10.3389/fpsyg.2013.00230>
- Jaeger, T. F., & Buz, E. 2017. Signal reduction and linguistic encoding. In: Fernández, E. M., & Cairns, H. S. (eds.), *Handbook of Psycholinguistics*, 38–81. Wiley-Blackwell. DOI: <https://doi.org/10.1002/9781118829516.ch3>
- Jun, J. 1996. Place assimilation is not the result of gestural overlap: Evidence from Korean and English. *Phonology*, 13(3), 377–407. DOI: <https://doi.org/10.1017/S0952675700002682>
- Jun, J. 2004. Place assimilation. In: Hayes, B., Kirchner, R., & Steriade, D. (eds.), *Phonetically based phonology*, 58–86. Cambridge: Cambridge University Press. DOI: <https://doi.org/10.1017/CBO9780511486401.003>
- Jun, J. 2011. Positional effects in consonant clusters. In: Oostendorp, M., van Ewen, C. J., Hume, E., & Rice, K. (eds.), *The Blackwell Companion to Phonology*. Malden, MA: Wiley-Blackwell. DOI: <https://doi.org/10.1002/9781444335262.wbctp0046>
- Jurafsky, D., Bell, A., Gregory, M., & Raymond, W. 2001. Probabilistic relations between words: Evidence from reduction in lexical production. In: Bybee, J., & Hopper, P. (eds.), *Frequency and the emergence of linguistic structure*. John Benjamins, Amsterdam. DOI: <https://doi.org/10.1075/tsl.45.13jur>
- Kanwal, J., Smith, K., Culbertson, J., & Kirby, S. 2017. Zipf's Law of Abbreviation and the Principle of Least Effort: Language users optimise a miniature lexicon for efficient communication. *Cognition*, 165, 45–52. DOI: <https://doi.org/10.1016/j.cognition.2017.05.001>
- Kawahara, S., & Garvey, K. 2014. Nasal place assimilation and the perceptibility of place contrasts. *Open Linguistics*, 1(1), 17–36. DOI: <https://doi.org/10.2478/opli-2014-0002>
- Kawasaki-Fukumori, H. 1992. An acoustical basis for universal phonotactic constraints. *Language and Speech*, 35, 73–86. DOI: <https://doi.org/10.1177/002383099203500207>
- Kohler, K. J. 1991. The phonetics/phonology issue in the study of articulatory reduction. *Phonetica*, 48, 180–192. DOI: <https://doi.org/10.1159/000261883>
- Kurumada, C., & Jaeger, T. F. 2015. Communicative efficiency in language production: Optional case-marking in Japanese. *Journal of Memory and Language*, 83, 152–178. DOI: <https://doi.org/10.1016/j.jml.2015.03.003>
- Kuzla, C., Cho, T., & Ernestus, M. 2007. Prosodic strengthening of German fricatives in duration and assimilatory devoicing. *Journal of Phonetics*, 35(3), 301–320. DOI: <https://doi.org/10.1016/j.wocn.2006.11.001>

- Labov, W., Ash, S., & Boberg, C. 2005. *The atlas of North American English: Phonetics, phonology and sound change*. Walter de Gruyter. DOI: <https://doi.org/10.1515/9783110206838>
- Labov, W., Rosenfelder, I., & Fruehwald, J. 2013. One hundred years of sound change in Philadelphia: Linear incrementation, reversal, and reanalysis. *Language*, 89(1), 30–65. DOI: <https://doi.org/10.1353/lan.2013.0015>
- Lehiste, I., & Peterson, G. E. 1961. Transitions, glides, and diphthongs. *The Journal of the Acoustical Society of America*, 33(3), 268–277. DOI: <https://doi.org/10.1121/1.1908638>
- Lindblom, B. 1990. Explaining phonetic variation: A sketch of the H&H theory. In: Hardcastle, W. J., & Marchal, A. (eds.), *Speech production and speech modelling*. Kluwer, Dordrecht. DOI: https://doi.org/10.1007/978-94-009-2037-8_16
- Mahowald, K., Fedorenko, E., Piantadosi, S. T., & Gibson, E. 2013. Info/information theory: Speakers choose shorter words in predictive contexts. *Cognition*, 126(2), 313–318. DOI: <https://doi.org/10.1016/j.cognition.2012.09.010>
- Morey, R. D., Hoekstra, R., Rouder, J. N., Lee, M. D., & Wagenmakers, E. J. 2016. The fallacy of placing confidence in confidence intervals. *Psychonomic bulletin & review*, 23(1), 103–123. DOI: <https://doi.org/10.3758/s13423-015-0947-8>
- Mowrey, R., & Pagliuca, W. 1987. Articulatory evolution. In: Ramat, A. G., Carruba, O., & Bernini, G. (eds.), *Papers from the 7th International Conference on Historical Linguistics*. John Benjamins Publishing.
- Mowrey, R., & Pagliuca, W. 1995. The reductive character of articulatory evolution. *Rivista di linguistica*, 7, 37–124.
- Munson, B. 2001. Phonological pattern frequency and speech production in adults and children. *Journal of Speech, Language, and Hearing Research*, 44, 778–92. DOI: [https://doi.org/10.1044/1092-4388\(2001/061\)](https://doi.org/10.1044/1092-4388(2001/061))
- Ohala, J. J. 1990. The phonetics and phonology of aspects of assimilation. In: Kingston, J., & Beckman, M. E. (eds.), *Papers in laboratory phonology I: Between the grammar and physics of speech*, 258–275. Cambridge: CUP. DOI: <https://doi.org/10.1017/CBO9780511627736.014>
- Otake, T., Yoneyama, K., Cutler, A., & van der Lugt, A. 1996. The representation of Japanese moraic nasals. *The Journal of the Acoustical Society of America*, 100(6), 3831–3842. DOI: <https://doi.org/10.1121/1.417239>
- Pierrehumbert, J. 2001. Exemplar dynamics: Word frequency, lenition and contrast. In: Bybee, J., & Hopper, P. (eds.), *Frequency and the emergence of linguistic structure*, 137–158. John Benjamins, Amsterdam. DOI: <https://doi.org/10.1075/tsl.45.08pie>
- Pierrehumbert, J. 2002. Word-specific phonetics. In: Gussenhoven, C., & Warner, N. (eds.), *Laboratory Phonology VII*. Mouton de Gruyter, Berlin. DOI: <https://doi.org/10.1515/9783110197105.101>
- Pitt, M. A., Dilley, L., Johnson, K., Kiesling, S., Raymond, W., Hume, E., & Fosler-Lussier, E. 2007. Buckeye Corpus of Conversational Speech (2nd release) [www.buckeyecorpus.osu.edu]. Columbus, OH: Department of Psychology, Ohio State University (Distributor).
- Pitt, M. A., & Johnson, K. 2003. Using pronunciation data as a starting point in modeling word recognition. *Proceedings of the 15th International Congress of Phonetic Sciences*.
- Raymond, W. D., Dautricourt, R., & Hume, E. 2006. Word-internal /t,d/ deletion in spontaneous speech: Modeling the effects of extra-linguistic, lexical, and phonological factors. *Language Variation and Change*, 18, 55–97. DOI: <https://doi.org/10.1017/S0954394506060042>
- R Core Team. 2016. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. URL: <https://www.R-project.org/>.

- Recasens, D. 2018. Coarticulation. *The Oxford Research Encyclopedia of Linguistics*. Oxford University Press.
- Resnik, P. 1996. Selectional constraints: An information-theoretic model and its computational realization. *Cognition*, 61(1–2), 127–159. DOI: [https://doi.org/10.1016/S0010-0277\(96\)00722-6](https://doi.org/10.1016/S0010-0277(96)00722-6)
- Rosenfelder, I., Fruehwald, J., Evanini, K., Seyfarth, S., Gorman, K., Prichard, H., & Yuan, J. 2014. *FAVE (Forced Alignment and Vowel Extraction) Program Suite v1.2.2*. DOI: <https://doi.org/10.5281/zenodo.22281>.
- Rysling, A. 2017. *Preferential early attribution in segmental parsing*. Doctoral dissertation, University of Massachusetts Amherst.
- Scarborough, R. 2004. *Coarticulation and the structure of the lexicon*. Doctoral dissertation, University of California Los Angeles.
- Scarborough, R. 2010. Lexical and contextual predictability: Confluent effects on the production of vowels. In: Fougeron, C., Kühnert, B., D’Imperio, M., & Vallée, N. (eds.), *Laboratory Phonology 10*. Mouton de Gruyter, Berlin.
- Scarborough, R. 2013. Neighborhood-conditioned patterns in phonetic detail: Relating coarticulation and hyperarticulation. *Journal of Phonetics*, 41(6), 491–508. DOI: <https://doi.org/10.1016/j.wocn.2013.09.004>
- Seyfarth, S. 2014. Word informativity influences acoustic duration: Effects of contextual predictability on lexical representation. *Cognition*, 133, 140–155. DOI: <https://doi.org/10.1016/j.cognition.2014.06.013>
- Sóskuthy, M., & Hay, J. 2017. Changing word usage predicts changing word durations in New Zealand English. *Cognition*, 166, 298–313. DOI: <https://doi.org/10.1016/j.cognition.2017.05.032>
- Stan Development Team. 2016a. rstanarm: Bayesian applied regression modeling via Stan. *R package version 2.17.4*. <http://mc-stan.org/>.
- Stan Development Team. 2016b. RStan: The R interface to Stan. *R package version 2.16.2*. <http://mc-stan.org/>.
- Steele, J., Colantoni, L., & Kochetov, A. (in press). Gradient assimilation in French cross-word /n/ velar stop sequences. *Journal of the International Phonetic Association*. DOI: <https://doi.org/10.1017/S0025100317000573>
- Steriade, D. 2001. Directional asymmetries in place assimilation: A perceptual account. In: Hume, E., & Johnson, K. (eds.), *The role of speech perception in phonology*, 219–250. New York: Academic Press.
- Stolcke, A. 2002. SRILM – an extensible language modeling toolkit. *Interspeech 2002*.
- Stolcke, A., Zheng, J., Wang, W., & Abrash, V. 2011. SRILM at sixteen: Update and outlook. In: *Proceedings of IEEE Automatic Speech Recognition and Understanding Workshop 5*.
- Tily, H., & Kuperman, V. 2012. Rational phonological lengthening in spoken Dutch. *Journal of the Acoustical Society of America*, 132(6), 3935–3940. DOI: <https://doi.org/10.1121/1.4765071>
- Tomaschek, F., Tucker, B., Fasiolo, M., & Baayen, H. 2018. Practice makes perfect: The consequences of lexical proficiency for articulation. *Linguistics Vanguard*, 4(S2).
- Turnbull, R. 2017. The role of predictability in intonational variability. *Language and Speech*, 60(1), 123–153. DOI: <https://doi.org/10.1177/0023830916647079>
- Turnbull, R. 2018. Patterns of probabilistic segment deletion/reduction in English and Japanese. *Linguistics Vanguard*, 4(S2). DOI: <https://doi.org/10.1515/lingvan-2017-0033>
- van Son, R. J. J. H., & Pols, L. C. W. 2003. How efficient is speech? *Proceedings of the Institute of Phonetic Sciences*, 25, 171–184.

- Vehtari, A., Gelman, A., & Gabry, J. 2017. Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC. *Statistics and Computing*, 27(5), 1413–1432. DOI: <https://doi.org/10.1007/s11222-016-9696-4>
- Viswanathan, N., Magnuson, J. S., & Fowler, C. A. 2010. Compensation for coarticulation: Disentangling auditory and gestural theories of perception of coarticulatory effects in speech. *Journal of Experimental Psychology: Human Perception and Performance*, 36(4), 1005–1015. DOI: <https://doi.org/10.1037/a0018391>
- Wasow, T., Levy, R., Melnick, R., Zhu, H., & Juzek, T. 2015. Processing, Prosody, and Optional to. In: Frazier, L., & Gibson, E. (eds.), *Explicit and implicit prosody in sentence processing (Studies in Theoretical Psycholinguistics)*, 133–158. Cham: Springer.
- Wedel, A. 2006. Exemplar models, evolution and language change. *The Linguistic Review*, 23(3), 247–274. DOI: <https://doi.org/10.1515/TLR.2006.010>
- Wedel, A., Nelson, N., & Sharp, R. 2018. The phonetic specificity of contrastive hyperarticulation in natural speech. *Journal of Memory and Language*, 100, 61–88. DOI: <https://doi.org/10.1016/j.jml.2018.01.001>
- Weenink, D. J. 2015. Improved formant frequency measurements of short segments. *Proceedings of the 18th International Congress of Phonetic Sciences*.
- Wickham, H. 2009. *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York. DOI: <https://doi.org/10.1007/978-0-387-98141-3>
- Winters, S. 2003. *Empirical investigations into the perceptual and articulatory origins of cross-linguistic asymmetries in place assimilation*. Doctoral dissertation, Ohio State University.

How to cite this article: Turnbull, R., Seyfarth, S., Hume, E. and Jaeger, T. F. 2018 Nasal place assimilation trades off inferrability of both target and trigger words. *Laboratory Phonology: Journal of the Association for Laboratory Phonology* 9(1): 15, pp.1–27, DOI: <https://doi.org/10.5334/labphon.119>

Submitted: 01 October 2017 **Accepted:** 30 July 2018 **Published:** 20 September 2018

Copyright: © 2018 The Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CC-BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited. See <http://creativecommons.org/licenses/by/4.0/>.

]u[*Laboratory Phonology: Journal of the Association for Laboratory Phonology* is a peer-reviewed open access journal published by Ubiquity Press.

OPEN ACCESS 