Open Library of Humanities

Purse, Ruaridh & Fruehwald, Josef & Tamminga, Meredith. 2022. Frequency and morphological complexity in variation. *Glossa: a journal of general linguistics* 7(1). pp. 1–34. DOI: https://doi.org/10.16995/glossa.5839

Frequency and morphological complexity in variation

Ruaridh Purse, University of Michigan, US, rupurse@umich.edu

Josef Fruehwald, University of Kentucky, US, josef.fruehwald@uky.edu

Meredith Tamminga, University of Pennsylvania, US, tamminga@ling.upenn.edu

Broad interest in probabilistic aspects of language has reignited debates about a potential delineation between the shape of an abstract grammar and patterns of language in use. A central topic in this debate is the relationship between measures capturing aspects of language use, such as word frequency, and patterns of variation. While it has become common practice to attend to frequency measures in studies of linguistic variation, fundamental questions about exactly what linguistic unit's frequency it is appropriate to measure in each case, and what this implies about the representations or processing mechanisms at play, remain underexplored. In the present study, we compare how three frequency measures account for variance in Coronal Stop Deletion (CSD) based on large-scale corpus data from Philadelphia English: whole-word frequency, stem frequency, and conditional (whole-word/stem) frequency. While there is an effect of all three measures on CSD outcomes in monomorphemes, the effect of conditional frequency is by far the most robust. Furthermore, only conditional frequency has an effect on CSD rates in *-ed* suffixed words. Thus, we suggest that frequency effects in CSD are best interpreted in terms of stem-conditional predictability of a suffix or word-edge. These results lend support to the importance of asking these fundamental questions about usage measures, and suggest that contemporary approaches to frequency should take morphological complexity into account.

1 Introduction

Systematic, intraspeaker linguistic variation has long played a central role in understanding the architecture of the grammar and its relationship to everyday language use. Early variationist sociolinguistic work posited that generative phonological rules (as in, e.g., Chomsky & Halle 1968), could be assigned probabilities to capture the patterns of variation observed in conversational speech (see, e.g., Labov 1969). In the decades since, modeling variation has become an increasingly important goal of linguistic theory across phonology, morphology, and syntax. In phonology, for example, a number of frameworks have been developed to account for probabilistic variation as a core feature of their formalism (e.g., Stochastic OT, Boersma & Hayes 2001; Noisy Harmonic Grammars, Coetzee & Pater 2011; Maximum Entropy Grammars, Hayes & Wilson 2008). Similar developments in morphology and syntax (e.g. Competing Grammars, Kroch 1994) speak to the relevance of variability at all levels of the grammar.

At the same time, this increasing interest in modeling probabilistic aspects of language has renewed longstanding debates about the dividing line, if there is one, between grammar and language use. One of the core empirical phenomena that has fueled these debates is the sensitivity of variation to the frequency with which linguistic elements are used. In everyday experience, language users encounter different sounds, affixes, words, and syntactic structures at different rates. Highly frequent words, for example, may be used and heard many times every day, while infrequent words occur far more rarely. Unsurprisingly, it seems that these dramatic differences in usage frequency may influence many aspects of how people perceive and produce language. For example, generally speaking, the more frequent a word is, the more quickly it will be recognized (Brysbaert et al. 2017), the shorter its duration will be (Aylett & Turk 2004; Gahl & Garnsey 2004), and the more likely it will be to exhibit various types of phonetic reduction (Hooper 1976; Brown 2009; Diaz-Campós & Gradoville 2011). To capture some of these facts, a number of models have proposed encoding lexical frequency information directly or indirectly into phonological representations and processes; to give two quite different examples, Coetzee & Kawahara (2013) scale the faithfulness constraint weights in a Noisy Harmonic Grammar by lexical frequency, while in Exemplar Theoretic models of variation, frequency effects emerge via the storage of episodic traces in lexical representations (Pierrehumbert 2002, *inter alia*). But another possibility is that the influence of frequency on variation reflects not grammatical or representational differences according to frequency, but rather the operation of online mechanisms of processing or articulation. This is the perspective of models of cascading activation (Goldrick 2011), in which processing at the stage of lexical access has downstream effects on articulation. Frequency effects, then, are a centrally-important testing ground for understanding the relationship between grammatical systems and language use.

In light of these theoretical consequences of frequency-type effects, it becomes important to address fundamental questions about what frequency *is* and how it is measured.[1] In this paper, we consider the shape these questions may take in the context of morphologically complex words. Of course, many different linguistic units could be considered to occur at different frequencies in many different kinds of contexts, but here we focus on the complexity that can be found even at the level of what we might call lexical frequency. Even when limiting our consideration to frequency measures for a relatively basic linguistic unit like the word, we still encounter the question: frequency of what, exactly?

In many psycholinguistic models, for example, the unit of lexical representation is the lemma, which contains the syntactic properties of a word and is shared by morphological relatives with the same root (Roelofs 1992). In such models, *jump* and *jumped* and *jumping* would all have the same lemma, with inflection added outside the lexicon. If frequency information is stored lexically and the lemma is the unit of lexical representation, then lemma frequency (i.e. the summed frequency of all words containing the *jump* lemma, or STEM FREQUENCY, as we call it here, might be more relevant to linguistic behavior than surface whole-word frequency (where *jump*, *jumped*, and *jumping* would each have distinct frequency values).

In sociolinguistic research, it is more common to see the effect of WHOLE-WORD FREQUENCY on variable performance investigated. The use of whole-word frequency has theoretical underpinnings in more austere forms of Exemplar Theory which propose that morphological abstraction is not a stored component of speakers' knowledge, but rather online analogization of word-forms in an associative network (Bybee 2002). As such, the whole-word form is the most relevant linguistic unit on which to hang frequency estimates in these models. There is also a methodological convenience to whole-word frequency: it is easily estimated from corpus data without the need for lemmatization.

Another possibility is that frequency effects on linguistic variation are actually driven by the predictability of words. Higher frequency words are more predictable, and therefore may be subject to greater compression and reduction (Lindblom 1963; Aylett & Turk 2004; Turnbull 2015) (see §2.2). While both lemma and whole-word frequency may contribute to the predictability of a word, so too may the *relative* frequency of a word form within its inflectional and derivational paradigm, which we are calling CONDITIONAL FREQUENCY. There are, of course, many other contextual factors over which predictability could be computed (both by language users and by researchers). Here we focus on the predictability of the suffix (or lack thereof) given the stem

[1] In asking how frequency is measured, we are not concerned with comparing different frequency estimates for a single linguistic unit, but rather *which* of many co-present linguistic units' frequency is relevant for a particular phenomenon. For debates on different ways to estimate frequency norms for the same linguistic unit, see Brysbaert & New (2009) and Hay et al. (2015).

because this is an area of active research in psycholinguistics whose connection to the literature on sociolinguistic variation is relatively underexplored (Kuperman et al. 2007; Cohen 2015; Tomaschek et al. 2019). Another reason to focus on conditional frequency here rather than some other, simpler contextual predictability measures (such as probability given the previous or subsequent word) is that such measures have been investigated previously and not found to strongly predict outcomes in the variable we focus on in this study (Jurafsky et al. 1998; 2001).

This paper is an investigation into how these three frequency measures (stem frequency, whole-word frequency, and conditional frequency) relate to a well-studied linguistic variable in English, Coronal Stop Deletion (CSD). CSD is a variable process of consonant cluster reduction, probabilistically conditioned by factors such as rate of speech, phonological context, and morphological structure, that has also been shown to be associated with word frequency (typically whole-word frequency has been used). CSD provides a particularly fertile territory in which to explore these frequency measures (and the mechanisms of perception/production they're related to) because *-ed* suffixation in words like *missed* or *jumped* allows the separability of stem, whole-word, and conditional frequency effects. These frequency measures can then also be examined in morphologically simplex words allowing CSD, such as *mist*, which we will refer to as "monomorphemes." CSD is also a useful testing ground for our frequency questions because we already have a good understanding of a range of other factors that are associated with deletion.

While weighing our three frequency measures' relative ability to capture variance in the CSD data will allow us to offer some practical methodological suggestions, the more important point we will make is that the different frequency estimates are *not* simply interchangeable proxies for the same basic phenomenon. This becomes apparent when we consider how different frequency measures interact with other factors conditioning the variation. While the different measures necessarily correlate with one another in some contexts (see §4, **Figure 1**), our CSD case study shows that the measures interact quite differently with words' morphological structure. When we look at whole-word frequency, higher frequency monomorphemes are more likely to undergo CSD, but there is no evidence for such a relationship with *-ed* suffixed words (Guy 2019). A parallel qualitative result is found for stem frequency as well. In contrast, conditional frequency appears to have a strong and similar association with *both* monomorphemes and *-ed* suffixed words. Looking at both a single model for all morphological classes as well as separate models for each, we find that conditional frequency greatly improves the model fit, even when the best model fit also includes one or more of the other frequency measures.

We argue that these results should lead us to understand the different frequency measures as different in kind, capturing different mechanisms that may affect linguistic variation. More specifically, if we adopt the theoretical interpretations that we have already briefly suggested, in which stem frequency approximates variable ease of lexical access, whole word frequency captures the long-term accumulation of reduction in the word form, and conditional frequency is a proxy for the variable predictability of word forms, our finding of a strong and consistent influence of conditional frequency points to an important role for predictability in CSD. However, this interpretation also suggests that there is no one simple effect of “word frequency” that can be expected to have a uniform influence on different phenomena; in other words, our results should not be interpreted as showing that conditional frequency is the “correct” frequency measure to use in the study of variation across the board. Rather, we conclude that the question of how different frequency measures relate to any given phenomenon is an empirical one: different variable phenomena may turn out to be more or less sensitive to the different mechanisms that these measures tap into. As a methodological issue, then, the selection of a frequency measure to use in quantitative analysis ought to be a considered one. More interestingly, though, empirical evidence about what kind of frequency is most closely associated with the use of linguistic variable phenomena of various types can be brought to bear on the question of what factors most heavily influence the production of those variables. This can, in turn, enrich our understanding of the place of variation in grammatical systems and the interaction between grammar and use.

2 Background

2.1 Variable CSD

CSD is, in simplest terms, the variable deletion of word-final coronal stops following consonants in English (e.g., *old* versus *ol’*). This variable has been of enduring interest to sociolinguists, as evidenced by the size of the literature that describes it. One reason for this sustained interest is the grammatical conditioning that the variable exhibits, such that stops in words of different morphological classes are deleted at different rates. The basic pattern is that coronal stops are more commonly deleted at the end of monomorphemes (e.g. *mist*) than when they comprise an *-ed* past tense suffix (e.g. *missed*). This pattern is almost universally replicated across studies of Englishes around the world (Fasold 1972; Labov 1972; Wolfram & Christian 1976; Guy 1980; Patrick 1991; Santa Ana 1991; Lim & Guy 2005; Gut 2007; Hazen 2011; Walker 2012; Hansen Edwards 2016; Baranowski & Turton 2020; cf. Tagliamonte & Temple 2005).

Researchers have posited a range of competing explanations, both formal and functional, for the morphological conditioning of CSD (see, for example, Guy 1996 for discussion). Of particular note for our interests in this paper are approaches that appeal to the different functions of the stops in these word types. An *-ed* suffix marks past tense information and its absence can give rise to homophony with the present form of the same verb. But stops at the end of monomorphemes contain no particular grammatical information and may even be predictable from the preceding string. Therefore, lower rates of CSD in *-ed* suffixed forms may have something to do with a functional pressure to preserve informative structure (following e.g. Kiparsky 1972). The functional argument is thematically similar to the argument that reduction phenomena occur when speakers exploit predictable or redundant elements of form. In both, speakers preserve unpredictable and therefore informative structure, with perhaps the difference that the functional CSD story frames the morphological information encoded in the *-ed* suffix as especially important, rather than simply important because it is an unpredictable ending to a stem. However, despite the attractiveness of a functional story, and even though multiple contextual measures of word and biphone probability are good predictors of reduction in some cases (Jurafsky et al. 1998), CSD has only been found to be sensitive to decontextualized lexical frequency (Jurafsky et al. 2001). Of course, frequency itself is not straightforwardly captured with a single measure, and a central premise of the current study is that there is more than one plausible unit whose frequency matters.

All accounts of the grammatical conditioning of CSD are complicated by other morphological categories whose rates of CSD are intermediate between monomorphemes and regular past forms. These include semiweak verbs where past tense is marked both with a vowel change and a coronal stop (e.g. *kept*), negative contractions (e.g. *didn't*), past forms where the coronal stop already present in the stem is devoiced (e.g. *sent*), and the irregular past form *went*. As a further complication, the 'monomorphemic' category has traditionally included a number of words that are actually morphologically complex, but where the word-final coronal stop is tautomorphemic with the preceding segment. For example, a superlative form like *biggest* was labelled as a 'monomorpheme', despite the fact that *-est* is a suffix, because no morphological boundary breaks up the word-final /st/ cluster. In the present study, we follow MacKenzie & Tamminga (2021) in refining this class to include only true monomorphemes (as set out in that paper), and focus on the well-established difference between this class and regular past forms. In this way, we compare forms that are unambiguously morphologically simplex and complex.

While CSD is typically described in terms of a discrete distinction between 'present' and 'absent' realizations, this perspective has been called into question. Temple (2009; 2014) remarks that the phonetic conditioning and acoustic implementation of CSD resemble that of any other connected speech process. Moreover, articulatory data has revealed that covert tongue tip raising is the norm for inaudible coronal stops (Browman & Goldstein 1990; Purse 2019), in some cases observing CSD-like morphological conditioning on the magnitude of tongue tip raising (Purse & Turk 2016; Purse 2019). These results, of course, raise questions about the proper representation of CSD, but these are beyond the scope of the current paper. Insofar as CSD is the outcome of a gradient phonetic process of lingual reduction, we might expect it to behave like other classical lenition phenomena in being sensitive to lexical frequency, as we discuss in the following subsection. Indeed, this perspective is consistent with a handful of studies finding that, for CSD overall, frequent words have higher rates of deletion than infrequent words (Bybee 2002; Jurafsky et al. 2001; Tamminga 2016, cf. Walker 2012). Other studies, however, report that lexical frequency conditions CSD differently according to morphological class (Myers & Guy 1997; Guy 2019).

This possible interaction between frequency and morphological structure, and its implications for theories of frequency and variation in form, is a central topic that we probe in this paper. In the following two subsections, we review first the literature on how frequency influences the phonetics and phonology of word forms, and second the literature on how frequency relates to morphological representation and processing.

2.2 Frequency and variation in form

Frequency, specifically whole-word frequency, is associated with variation in phonetic and phonological form in many cases. In general, frequent whole-words tend to be pronounced faster, and in more lenited or reduced forms, than infrequent whole-words. This is relevant insofar as we conceive of CSD as an example of lenition, and we generally expect phonetic reduction and lenition to be intimately related to duration (Lindblom 1963). However, in laboratory studies, evidence for the precise details of the relationship between duration and frequency is somewhat mixed. Wright (1979) claims that rare words are spoken as much as 24% slower than common words, but some subsequent studies have failed to replicate these effects between matching segments (Damian 2003; Mousikou et al. 2015; cf. Kawamoto 1999). Laboratory studies are also not entirely aligned in terms of how phonetic reduction and lenition are sensitive to frequency. In one articulatory study, Lin et al. (2014) find that tongue tip activity is generally reduced in highly frequent words, but Tomaschek et al. (2013; 2014) find that the magnitude of vowel gestures is highly sensitive to segmental context and may only be compressed for frequent words with phonologically short vowels. In contrast with these laboratory studies, corpus studies on more spontaneous speech reliably find that frequent whole-words are produced with shorter durations (Aylett & Turk 2004; Gahl 2008), and with more centralised vowels (Munson & Solomon 2004) than infrequent whole-words.

Beyond gradient phonetic properties like duration, there exist a number of variables where the apparent rate of discrete variants[2] is correlated with lexical frequency. This is particularly well exemplified by work on varieties of Spanish. Highly frequent Spanish whole-words are more likely to exhibit intervocalic /d/ deletion (Bybee 2002; Diaz-Campós & Gradoville 2011), /r/ deletion (Diaz-Campós & Carmen 2008), vowel coalescence (Alba 2006), /s/ lenition and deletion (Brown & Cacoullos 2003; Brown 2009; File-Muriel 2009), and less likely to feature /ʒ/ devoicing than infrequent whole-words. In English, too, schwa deletion (Hooper 1976), yod retention (Phillips 1981; 1984), and alveolar word-final *-in'* for the ING variable (Tamminga 2016; Forrest 2017), have all been found to be more common in frequent whole-words than infrequent whole-words. In a more general approach that is not limited to specific sociolinguistic variables,

---

[2] While they are categorized in discrete terms, for many of these variables the question of whether they arise in the phonetics or phonology is not settled.

Turnbull (2018) compares phonological and phonetic transcriptions in corpora of English and Japanese and computes the segment deletions necessary between the underlying and surface forms. He finds that whole-word frequency (among other predictability measures) conditions the rate of segment deletion, and concludes that these patterns mirror those of phonetic reduction.

Outlying results notwithstanding, it seems generally true that frequent words are more susceptible to compression and 'weakening' of their pronunciations. Explanations for this kind of reduction phenomenon fall into three main theoretical camps (Clopper & Turnbull 2018), two of which link production effects to robust results that frequent words are recognised more quickly and accurately in perception experiments (e.g. Howes 1957; Savin 1963; Taft & Hambly 1986; Connine 1990; Dupoux & Mehler 1990). (1) 'Listener-oriented' accounts (e.g Lindblom 1990; Aylett & Turk 2004) explain production effects in terms of word predictability, and the optimization of the speech signal in order to maximize communicative efficacy while minimizing effort. In other words, speakers use tacit knowledge that frequent words are easier to perceive and attenuate the articulatory effort spent on them. (2) For 'talker-oriented' accounts (e.g. Baese-Berk & Goldrick 2009), frequency effects arise as part of the cognitive mechanisms of speech production. Just as in perception, infrequent word forms have a higher threshold for activation during production, and properties of timing and magnitude of activation during retrieval are passed on to properties of timing and articulation in the phonetic implementation. (3) Finally, there are 'passive' perspectives, in which word frequency directly shapes the mental representation of words, rather than creating on-line production pressures. A notable example of this kind of perspective is Exemplar Theory (Pierrehumbert 2002; Bybee 2002), in which a persistent leniting bias affects all words, but high frequency words—which are encountered most often—most quickly accumulate exemplars with compressed and 'weakened' pronunciations. While the frequency measures we discuss in more detail in §3.2 are correlated with each other (e.g. a word with a high stem frequency is likely to also have a high whole-word frequency), each one is likely more indicative of one of these theoretical mechanisms being at play than the others. For example, an effect of stem frequency is more likely to be indicative of a talker-oriented account than a listener-oriented or a passive account. This is discussed in more detail below.

2.3 Morphology and frequency

We now turn to a brief examination of the relationship between frequency and morphological structure, with reference to both sociolinguistic and psycholinguistic results that highlight possible frequency–morphology interactions. As we have mentioned, there is already some reason to believe that frequency and morphological structure interact in how they condition CSD itself. Myers & Guy (1997) report, based on data from two Philadelphian speakers, that there is a robust effect of whole-word frequency among monomorphemes, but no such effect among

-ed suffixed words. Similarly, Bayley (2014) finds a small effect of whole-word frequency that is limited to monomorphemes in San Antonio Chicano English CSD. Interactions between lexical frequency and grammatically-defined conditioning contexts in sociolinguistics have also been reported for morphosyntactic variables. Erker & Guy (2012) find that lexical frequency has an 'amplification' effect on the grammatical conditions influencing null subjects in Spanish: effects of verb regularity, verb semantics, subject person/number, and utterance tense/mood/aspect are small or nonexistent among low frequency verbs, but very significant among high frequency verbs. An interesting question we return to in our discussion in §5.1 is whether reported frequency/grammatical context interactions are the same kind of effect for CSD and null subjects.

The relevance of morphological structure for word processing has led to more widely recognized interactions in this domain. There is some evidence that morphologically complex words are generally recognised faster than monomorphemic words of equal length and frequency (Fiorentino & Poeppel 2007), but highly frequent complex words are disadvantaged if the suffix is also highly frequent (Balling & Baayen 2008). Moreover, it has been suggested that the frequency-bearing unit most appropriate to capture variance in word recognition latencies depends on the morphological complexity of the word (Vannest et al. 2011). Morphologically complex words are recognised at speeds that vary according to lemma—or "base"—frequency, while monomorphemic words' recognition speeds are best accounted for with whole-word—or "surface"—frequency.

In addition to basic frequency/morphology interactions in behavioral reaction times, there is also a growing body of work making inferences about what level of representation is active at a given point in the timecourse of spoken word recognition based on what kind of frequency measure correlates best with neural activity during processing. Specifically, a number of MEG studies find neurological activity to be most strongly correlated with measures of morphological structure, including lemma frequency and the transition probability between stem and suffix, at around 170ms (Solomyak & Marantz 2009; 2010; Zweig & Pylkkänen 2009; Lewis et al. 2011; Fruchter et al. 2013) and again at around 350ms (Solomyak & Marantz 2009) following exposure to visual word stimuli. These results are taken as evidence for word recognition making reference to smaller morphological units, since these frequency measures associated with activation levels reflect the frequency of those sub-word units. While this literature has typically discussed these sub-lexical units in terms of decomposition (see also Embick et al. forthcoming on the nature of decomposition), we do not believe it is necessary to endorse a particular view on whether morphologically complex words are decomposed *per se* in order to draw similar inferences about the relevance of sub-lexical structure for variation in morphologically complex words. Even models that posit whole-word episodic storage in the lexicon allow for morphological relationships to emerge from patterns of phonetic and semantic overlap (Bybee 2002); these relationships may in principle influence variable outcomes.

Among the studies that do apply this strategy of comparing frequency measures to explore the role of morphological structure in production, one interesting result that has emerged is evidence of ‘paradigmatic enhancement’ effects. As well as the basic effect whereby frequent items are realized (and recognized) faster as a result of their predictability or ease of retrieval, some words with a high frequency compared to morphologically related words within the same paradigm are reinforced and pronounced with *less* phonetic reduction. An intuitive way to conceptualize this idea is in terms of speaker confidence, such that speakers are reassured that they are ‘correct’ in selecting the most relatively frequent form and do not hold back in production (Kuperman et al. 2007). Originally, paradigmatic enhancement was proposed to explain effects in ‘pockets of uncertainty’ between functionally equivalent forms that directly compete for use in the same position, like Dutch compound linking morphemes (Kuperman et al. 2007) and variable Russian agreement suffixes (Cohen 2015). This ‘pocket of uncertainty’ aligns fairly closely with mainstream variationist conceptions of the linguistic variable, and indeed we see parallel results in the effect of variant frequency on variant duration in French variable schwa (Bürki et al. 2011). More recently, however, research on paradigmatic probability has been extended to explain variation in pronunciation across paradigmatically related words that are not in direct competition, with evidence for both the enhancement (Schuppler et al. 2012; Tucker et al. 2019; Tomaschek et al. 2019; 2021; Bell et al. 2021) and reduction (Hanique & Ernestus 2011; Smith et al. 2012; Ben Hedia & Plag 2017; Plag & Ben Hedia 2017) of more relatively frequent forms. The present study represents, among other things, a contribution to this literature that may help reconcile these seemingly contradictory results.

3 Data and methods

3.1 Corpus and coding

For this paper, data are taken from the Philadelphia Neighborhood Corpus of LING560 Studies (PNC) (Labov & Rosenfelder 2011). This corpus is comprised of sociolinguistic interviews conducted by students in a graduate-level sociolinguistics course at the University of Pennsylvania. Recordings were made between 1973 and 2012, and generally last about an hour. This study uses a sample of interviews from 118 white speakers found in working-class Irish-American and Italian-American neighborhoods. Speaker birth years span from 1888 to 1991, and the speakers are roughly balanced in terms of binary gender (66 women, 52 men). All interviews have been transcribed and had this transcription forced-aligned with the corresponding audio file using the FAVE suite (Rosenfelder et al. 2011).

CSD outcomes were hand-coded according to auditory and spectrographic cues. A Praat script[3] was used to search for tokens and play a short corresponding excerpt for researcher evaluation. A

[3] Code available at https://github.com/JoFrhwld/FAAV/blob/master/praat/handCoder.praat.

number of decisions were made in order to restrict the dataset to straightforward cases that are consistently found to be eligible for CSD across its vast literature. Only words whose underlying forms end in coronal stops that are immediately preceded by consonants were considered. Instances of glottalization and palatalization were counted as /t,d/ retention.[4] Tokens preceding a stop, non-sibilant fricative, or affricate with a coronal place of articulation (i.e. /t,d,θ,ð,t͡ʃ,d͡ʒ/) were excluded, as well as tokens with both a preceding /n/ and following /s/. These contexts are particularly susceptible to processes that would neutralize the distinction between deleted and undeleted word-final coronal stops.[5] Words in which a final coronal stop was preceded by /r/ (e.g. *part*, *card*) were excluded as it has been suggested that these stops are ineligible for deletion, at least in Philadelphia English (Cofer 1972). The word *and* was excluded entirely, since it has been analyzed as an exceptional case with multiple underlying representations (Neu 1980; Guy 2007). Irregular past forms (e.g. *kept*) and negative contraction forms (e.g. *wasn't*) were also excluded to focus on a more straightforward comparison between the most common morphological categories. In addition, we follow MacKenzie & Tamminga (2021) in further restricting the 'monomorphemic' category to include only true monomorphemes, excluding superlative forms (e.g. *biggest*), agentive forms (e.g. *specialist*), and deverbal nominalized forms (e.g. *management*), among others.[6] This method yielded 8,912 word-final /t,d/ tokens, coded as belonging to monomorphemic (e.g. *act*) or regular past[7] (e.g. *jumped*) word forms.

3.2 Frequency measures

In concrete terms, the goal of this study is to evaluate how different frequency-related measures may be associated with variable CSD. In particular, we are interested in whether it is the frequency of the whole-word, the frequency of some smaller constituent, or indeed the frequency relationship between the whole-word and its component parts, that best predicts CSD outcomes. To that end, we compare how well three different measures, calculated from values in the $\text{SUBTLEX}_{\text{US}}$ Corpus, account for variance in the CSD variable. These three measures, which we introduced briefly in §1, do not exhaust all possible relationships between the frequency of different strings or units and CSD, but they do capture several distinct perspectives on how frequency measures might be relevant to the variable at hand.

[4] This is the usual decision for CSD studies on American English. It has recently been suggested that British English glottal replacement of /t/ blocks CSD (Baranowski & Turton 2020), but the exclusion of glottalized cases should only enhance the basic morphological effect since the contexts most favouring glottalization (/nt#/, /lt#/) do not occur in *-ed* suffixed forms.

[5] For example, quasi-gemination across word boundaries makes it very difficult to distinguish between *last time* and *las' time*.

[6] The true monomorphemes do not noticeably differ from a more traditional 'monomorphemic' category in terms of their sensitivity to different frequency measures.

[7] The 'regular past' category includes all preterite, perfect, and passive forms featuring an *-ed* suffix.

Our first such measure, whole-word frequency, is extracted from the FREQlow values in SUBTLEX$_{US}$: the raw number of times that a word appeared in the corpus in lower case. This measure, or a similar one, is the most widely used in linguistics, but it has some quirks. For example, in SUBTLEX$_{US}$, as in other corpora, frequency norms are calculated according to orthographic strings. This means that homographs have the same FREQlow value whether or not they are phonologically or morphologically related. However, whole-word frequency basically approximates the frequency of a surface phonological form. This measure was natural log-transformed and centred with the mean at zero.

We call our second measure stem frequency.[8] For this measure, we manually extracted and calculated the sum of all the whole-word frequencies for words that share the same stem as words in the data. We were careful to only add the frequency of the relevant parts of speech. For example, the calculation of the stem frequency for monomorphemic directional *left* does not include occurrences of verbal *left* or its morphological relatives such as *leftovers*. The stem frequency for monomorphemic *left* was calculated from its own, part-of-speech-corrected whole-word frequency, plus the whole-word frequencies for *lefty*, *lefties*, *lefts*, *leftist*, *leftists*, and *lefter*. This measure was also log-transformed and centered.

The third measure is conditional frequency. Conditional frequency is computed from the other two measures; the whole-word frequency is divided by the stem frequency. Quantitatively speaking, conditional frequency is a proportion, bounded by 0 and 1. In other words, conditional frequency approximates the frequency of a particular word among its morphological relatives.

3.3 Statistical modeling

The primary methodology used in this paper is comparison of mixed effects logistic regression models using the `lme4` package (Bates et al. 2015) in R (R-Core-Team 2015). We set up a baseline model, which included fixed effects for following segmental context (pause; vowel; consonant eligible for stop resyllabification, e.g. /t#r/; or consonant ineligible for stop resyllabification, e.g. /t#l/, sum coded), grammatical class (monomorphemic versus regular past, sum coded) and speech rate (vowels-per-second in a 7 word window, by-speaker z-score normalized), and a random intercept for speaker. We retained all these predictors in all subsequent models. Fixed effects for preceding phonological context could not be included without inducing a convergence error. From the baseline, we constructed models with all possible combinations of the three lexical frequency measures as fixed effects, including a model that included all three measures. We then performed paired likelihood ratio tests on nested models, and compard the AIC and BIC of each model. We rely on these global goodness-of-fit criteria as they are more robust to the multicollinearity between the frequency measures than coefficient estimates are.

[8] Similar measures to our stem frequency measure have been called lemma frequency in previous literature. However, lemma frequency typically only includes inflectionally related words that share a stem. Since we count both inflectionally and derivationally related words that share a stem, we opted for a different name.

4 Results

A central goal of this article is to compare multiple measures which are not only arithmetically related, but also attempt to capture similar (if not identical) aspects of how words are represented and processed. Therefore, before assessing the relative contributions of each of these frequency measures on CSD outcomes, we must explore the relationship between them. **Figure 1** shows scatterplots indicating how words in each morphological class are distributed across the frequency measures, taken pairwise. Each plot, and the Pearson's correlation test results with which it is labelled, are generated from a 'dictionary' version of our data, with one entry for each unique word along with its values for each frequency measure according to SUBTLEX$_{US}$.

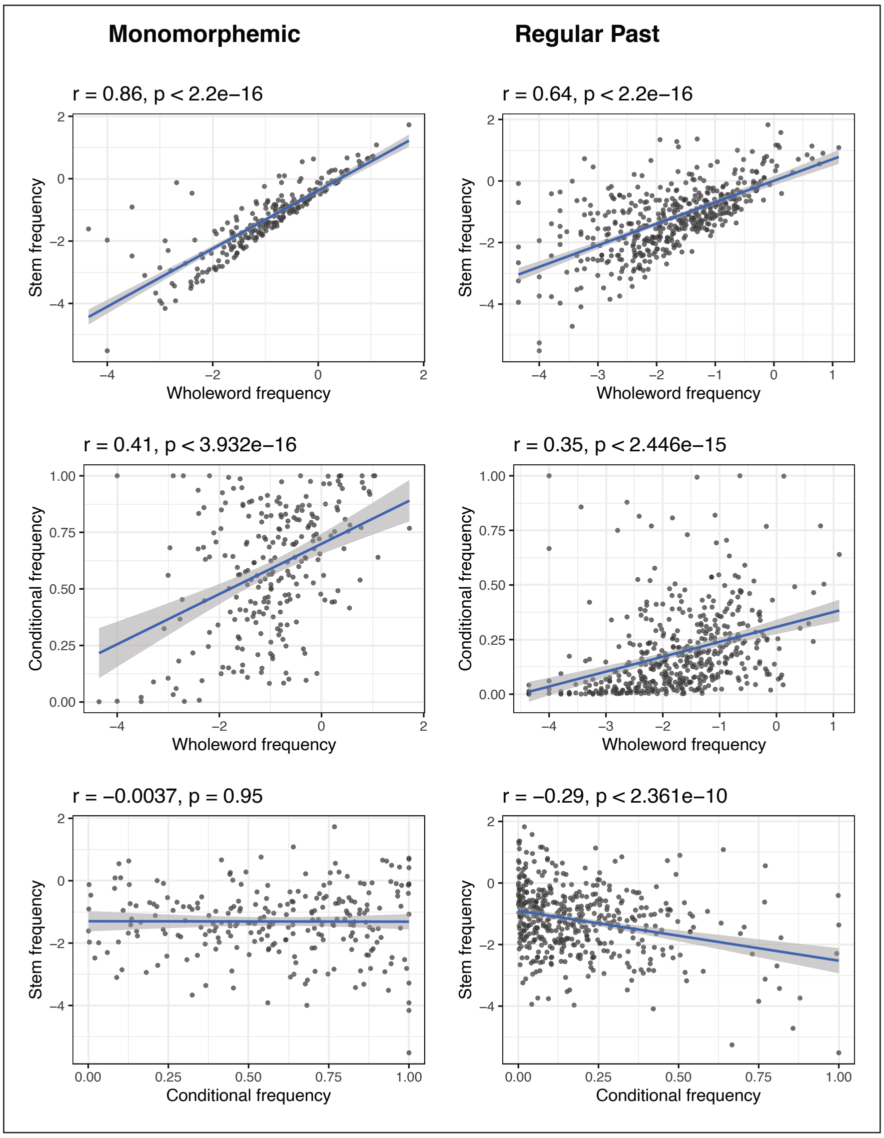

Figure 1: Relationship between frequency measures for monomorphemic (left) and *-ed* suffixed (right) words.

As is evident from **Figure 1**, monomorphemic words have quite different frequency properties to regular past *-ed* suffixed words. In both word types, there is a positive correlation between whole-word and stem frequency, and a hard border where a word's stem frequency must, by definition, be greater than or equal to its whole-word frequency. Each word's whole-word frequency value itself contributes to the stem frequency value, along with the whole-word frequencies of morphologically related forms. This means that the stem frequency cannot be lower than the whole-word frequency. The contribution of whole-word frequency values to stem frequency values is also linked to the positive correlation between whole-word and stem frequency, which is especially strong in monomorphemes. As whole-word frequency increases, the corresponding component part of stem frequency also increases. On one hand, this correlation means that it will be difficult to compare how well whole-word and stem frequency predict CSD outcomes, especially for monomorphemes. On the other hand, from a practical methodological perspective, it is useful to know that whole-word and stem frequency can, at least for monomorphemes, be used more or less interchangeably.

Monomorphemes and regular past forms differ in particular in their conditional frequency distributions. While monomorphemes are distributed fairly evenly, the majority of regular past forms have a very low conditional frequency. Reflecting on the properties of these word types, this might not be entirely unexpected. By definition, regular past forms are verbal, and implicate a whole paradigm of differently-inflected verb forms whose whole-word frequencies contribute to the stem frequency value. As a result, the regular past form often makes up only a small part of the stem frequency. On the other hand, the monomorphemic class includes words from a number of parts of speech that differ in the types of morphological relatives that occur.[9] Since whole-word and stem frequency form the numerator and denominator in the calculation of conditional frequency, respectively, we can expect a positive relationship between conditional and whole-word frequency and a negative relationship between conditional and stem frequency. Sure enough, the directions of these relationships is borne out, but the correlations between conditional and stem frequency are far shallower than the other cases. In fact, a Pearson's correlation test finds no relationship between conditional and stem frequency for monomorphemes; the line is practically flat. These results parallel the non-correlation between the closely related measures of lemma frequency and "paradigmatic probability" found by Tomaschek et al. (2021).

The investigation of correlations between the different frequency measures gives us confidence that it is reasonable to include both conditional frequency and stem frequency as predictors in a single model. Conversely, we should be wary of multicollinearity effects in models with other pairs of frequency measures. For the sake of completeness, we include all possible combinations of frequency measures in our model comparison analysis, but note that some improvements to model fit are likely to be artifacts of the relationship between measures.

---

[9] The existing literature on CSD gives us no reason to expect that part of speech is a important dimension for the variable.

4.1 Modeling both monomorphemes and past tense

In order to probe which frequency measure best captures variance in CSD, we compared a series of logistic regression models predicting CSD outcomes. The baseline model does not contain any frequency measures but does include the fixed effects for speech rate, grammatical class, and following segmental context, plus a random intercept for speaker. The subsequent models add all possible combinations of the three frequency measures to this baseline model. We use likelihood ratio tests to assess whether each additional level of complexity (i.e. each additional frequency measure) was warranted as a significant improvement over the nested smaller models.[10]

In addition to likelihood ratio tests, each model's fixed AIC (Akaike Information Criterion), BIC (Bayesian Information Criterion) and log-likelihood statistics were recorded. While the log-likelihood is inevitably improved by adding additional complexity to a model, the AIC and BIC penalize model complexity at the same time as evaluating a model's ability to account for variance. This is especially true of the BIC, whose penalty for additional complexity is proportional to the number of observations, and frequently disagrees with the AIC in favour of a simpler model. Together, these information criteria provide the clearest evaluation of these models, indicating in particular where multiple frequency measures do not account for a sufficient amount of variance to justify their inclusion. **Figure 2** shows the degree to which models with various combinations of frequency measures reduce the AIC and BIC, compared to a baseline model with no frequency measures.

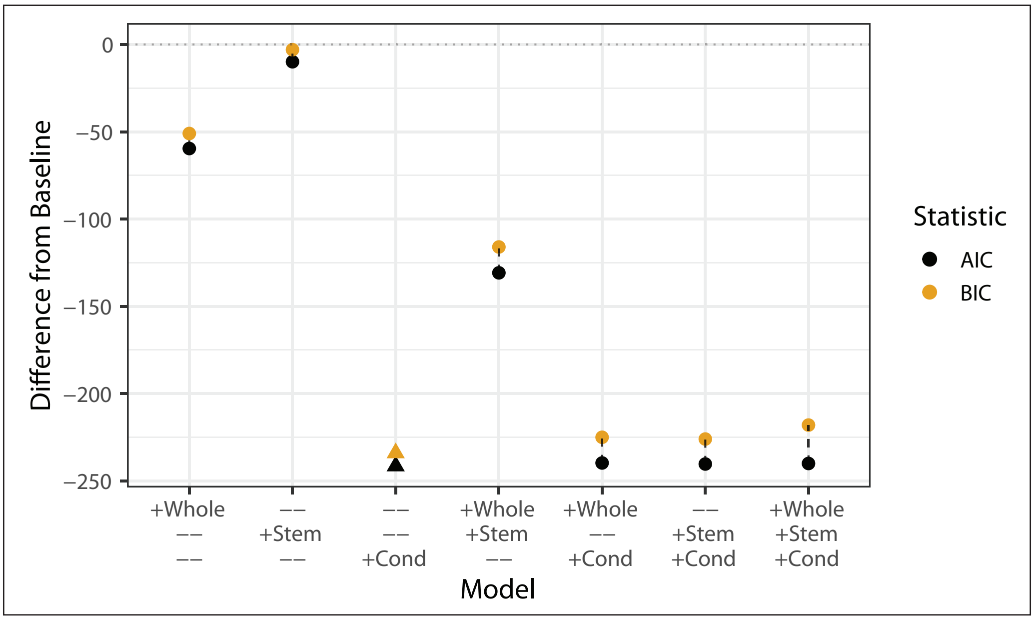

Figure 2: Information criteria reduction from baseline comparing models of full dataset (triangles = most reduced).

[10] The full results of model comparison can be found in Appendix A.

Including each of the three frequency measures, individually, yields information criteria statistics that are somewhat reduced compared to the baseline model. This result is reinforced by significant likelihood ratio tests ($p < .001$) in each case. However, the reduction in both AIC and BIC that is attained from the addition of conditional frequency far outstrips that of the other measures. In fact, the addition of conditional frequency provides a large reductions in both AIC and BIC regardless of any other frequency measures already included in a model. The model comparison also suggests that the combination of stem frequency and whole-word frequency in a single model is a significant improvement over just one of these measures. However, we cannot rule out the possibility that this is an artifact of the strong correlation between these measures causing inflation of their estimated effects. In addition, neither stem nor whole-word frequency significantly improves any model that already includes an effect of conditional frequency. This is demonstrated by likelihood ratio tests ($p > .05$), and the fact that these measures do not account for enough additional variance to counteract the penalty for model complexity that occurs in either the AIC or the BIC.

The initial model comparison results point to a need to reconsider how frequency is accounted for in linguistic variation. In particular, the success of conditional frequency over other measures in terms of accounting for variance suggests that the interplay between word frequency and morphological structure within the lexicon is important and underexplored. Morphological structure is particularly relevant for a variable like coronal stop deletion, since it has repeatedly been reported that coronal stops at the end of monomorphemes are more likely to be deleted than coronal stops that constitute *-ed* suffixes (Guy 1980; 1991). This basic difference is controlled for with the main effect of grammatical category in each of the models in **Figure 2**. However, the effect of morphology may be more complicated still, as **Figure 3** shows. In the top-left panel of the figure, we replicate previous findings that only monomorphemes are sensitive to whole-word frequency, and not regular past forms (Guy 2019). This result strengthens our confidence in the interaction between frequency and morphological category for CSD, because compared to previous reports it is based on a significantly larger dataset with more narrowly defined morphological categories. Unsurprisingly, this interaction also holds for the closely-related stem frequency in the top-right panel. In addition to replicating previous reports for CSD, we note that these results resemble the "amplification" effect described by Erker & Guy (2012), in which the effect of grammatical class is stronger at high frequencies, and may not exist at all between low frequency words. However, in all of the CSD studies, including ours, the slope of the line for regular past forms does not significantly deviate from 0, suggesting that any apparent amplification of the morphological effect does not affect the morphological categories evenly, but rather is driven by differences between high- and low-frequency monomorphemes.[11]

[11] Interactions of morphological class with whole-word frequency and stem frequency are fairly significant when they are added to models, but they are always heavily penalized in model comparison.

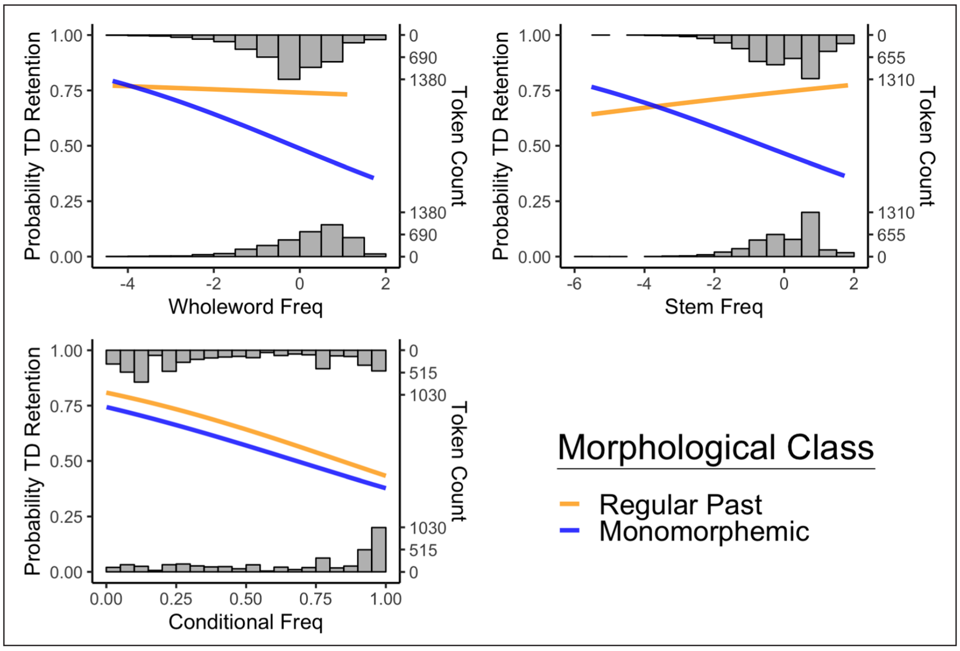

Figure 3: Observed CSD outcomes according to each frequency measure and morphological class.

Compared to the results for whole-word and stem frequency, the results for conditional frequency are striking. Here, not only is there an effect for both monomorphemes and regular past forms, but the lines are almost parallel. This helps to explain why conditional frequency was so highly favored by the model comparison for the combined data in **Figure 2** with sum-coded morphological categories. Furthermore, recall from **Figure 1** that the relationship between conditional frequency and the other measures was weaker than the relationship between whole-word and stem frequency; the robust conditional frequency effect observed here accounts for a portion of the variance in CSD outcomes that is virtually untapped by controlling for just whole-word or stem frequency.

The differences in the effects of the frequency measures between morphological categories is not captured by the regression models we have been discussing, because they do not include any interaction terms targeting the non-independence of frequency and grammatical category. As a result, the best models we've presented so far, which combine data from regular past and monomorphemic words (sum-coded), will compromise between the two. In other words, a frequency measure that might be best for one group of words will be penalized if it is inappropriate for another. This raises questions about the performance of frequency measures within morphological categories, which are not addressed by the models we have presented so far. Therefore, in the following subsections we divide the data by morphological class and test the different frequency predictors within each word type.

4.2 Modeling only monomorphemes

We begin by adopting the same method of model comparison as described for the full dataset, implemented over a subset of the data containing only monomorphemes. Once again, all models include fixed effects for speech rate and following segmental context, and a random intercept for speaker, but since all the words are monomorphemic no morphological category predictor is included.

In **Figure 4**, we can see that the picture for monomorphemes alone is very similar. As in the models for the full dataset, all the frequency measures significantly improve model fit over the baseline when they are added individually, but conditional frequency outperforms the other measures and improves every model to which it is added. These results are reinforced by likelihood ratio tests ($p < .001$). In this case, the addition of stem frequency provides a slightly more obvious reduction in AIC and BIC than was observed for the full dataset, but it is still the smallest in magnitude out of the three frequency measures. Once again, we see that the combination of stem and whole-word frequency outperforms either measure on its own, but this is very likely an artifact of the especially strong multicollinearity between these measures for monomorphemes.

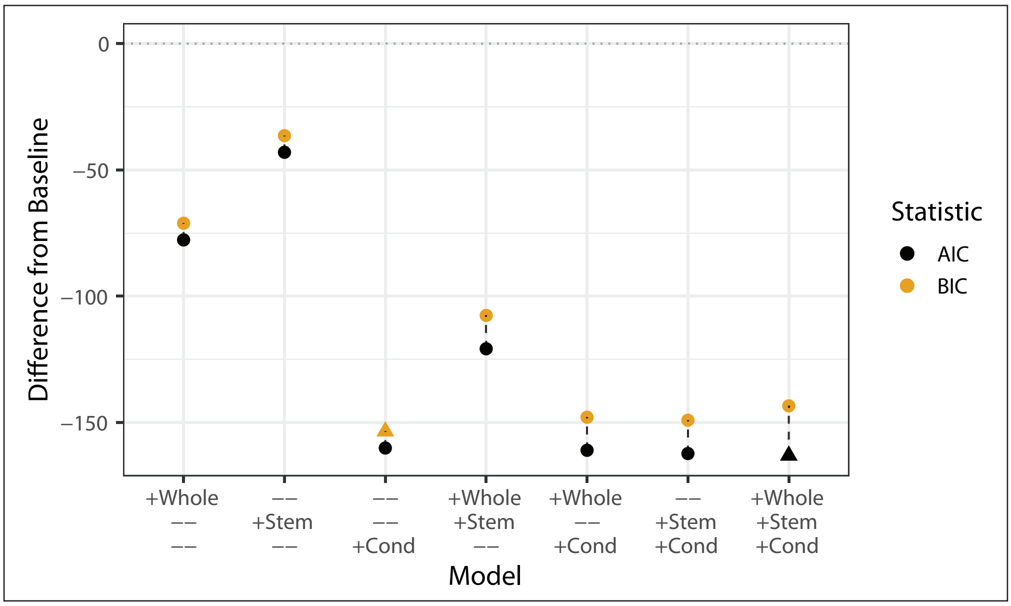

Figure 4: Information criteria reduction from baseline comparing models of monomorpheme subset (triangles = most reduced).

In terms of the models that best reduce the information criteria, the results for the monomorpheme models are slightly less straightforward than for the full dataset in that the AIC and BIC disagree. Once again, the BIC is lowest for the model with just conditional frequency in addition to the baseline effects. However, the AIC is lower in the models containing at least one other frequency measure in addition to conditional frequency, and lowest in the model with all three measures. This suggests the other measures do capture enough variance in monomorphemes to outperform the relatively small penalty for additional model complexity that is applied in the computation of AIC. This seems especially true for stem frequency, which significantly improves the fit of every model it is added to according to likelihood ratio tests ($p < .05$). This includes all models with conditional frequency and/or whole-word frequency already present. In contrast, likelihood ratio tests do not show whole-word frequency to significantly improve models with conditional frequency already present. This is likely due, in large part, to the complete absence of a correlation between conditional and stem frequency for monomorphemes, such that they do not compete to account for the same variance.

4.3 Modeling only past tense

Just like for monomorphemes, we conducted the same method of model comparison for the regular past forms alone. Again, all models include a fixed effect of speech rate and following segmental context, a random intercepts for speaker. According to **Figure 3**, only conditional frequency appears to have the expected frequency effect for this group, but model comparison allows us to observe the interplay between the different measures when they are included in different combinations. The AIC and BIC values for each of the regular past models are plotted in **Figure 5**.

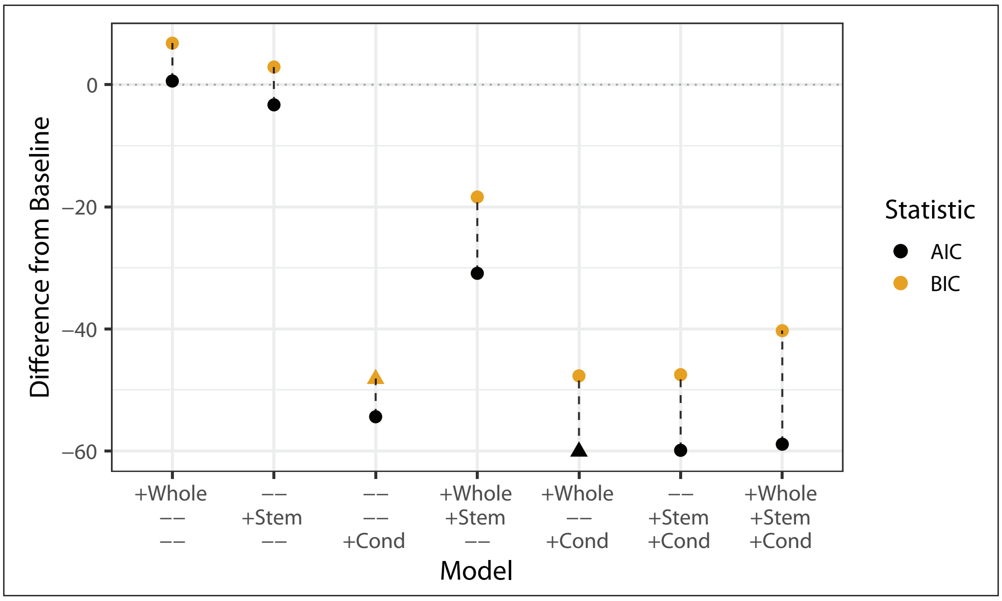

Figure 5: Information criteria reduction from baseline comparing models of complex form subset (triangles = most reduced).

Unsurprisingly, conditional frequency once again introduces a large reduction in both the AIC and BIC of every model it is added to, as well as a significant improvement in terms of likelihood ratio tests ($p < .001$). Unlike for the full and monomorpheme datasets, not all of the frequency measures improve the baseline model when they are added individually. The addition of whole-word frequency does not account for enough variance to overcome the penalty for model complexity in either the AIC or BIC, and does not significantly improve model fit according to a likelihood ratio test ($p > .1$). Stem frequency, on the other hand, does marginally reduce the AIC and significantly improve model fit according to a likelihood ratio test ($p < .05$), but the magnitude of its improvement is still less than the penalty applied by the BIC for introducing additional complexity to the baseline model. Once again, the combination of both whole-word and stem frequency apparently reduces both the AIC and BIC by a fair amount compared to the baseline model. Even though the correlation between whole-word and stem frequency is weaker for regular past forms than for monomorphemes, it is still strong enough that this effect is likely to be an artifact of multicollinearity, especially given how poorly both whole-word and stem frequency perform individually.

Like for the monomorpheme models, the AIC and BIC disagree as to the optimal model for regular past forms. For the third time, the model with conditional frequency alone is favored by the BIC, and additional frequency measures are penalized for unnecessary complexity. However, this time, the AIC is minimized in the model with both conditional and whole-word frequency. This is despite the fact that whole-word frequency performed poorest when it was added to the baseline model individually, *and* the fact that it is more strongly correlated with conditional frequency than stem frequency is, for the regular past forms.

5 Discussion

As we have discussed at some length in Section 2, the frequency of a word—or other linguistic unit—is associated with differences in the way it is perceived, produced, or even represented in the mind of the perceiver or producer. As such, even when it is not a study's primary concern, contemporary studies in various subfields of linguistics take steps to control for some relevant measure of frequency. However, several possible such measures are available, and their different properties are relatively underexplored. Moreover, the complex interplay between the frequency of different sub-lexical units and the morphological structure of words is rarely considered, especially within the quantitative analysis of linguistic variables like CSD.

With respect to these questions, there are two clear results to take away from §4, which this section will discuss in turn. First, whole-word frequency (and to a lesser extent, stem frequency) is a significant predictor of CSD in monomorphemes but not in regular past tense forms. The direction of the effect within monomorphemes is as expected for reduction phenomena in general, with more CSD in higher-frequency whole-words. Second, both monomorphemes and past tense forms are highly sensitive to conditional frequency, again in the direction of more CSD with higher conditional frequency. Conditional frequency, therefore, has both a stronger and more pervasive across-the-board effect on CSD than the more familiar whole-word frequency measure. In the following subsections, we discuss these two results in light of their theoretical implications.

5.1 Interaction between whole-word/stem frequency and morphology

Whole-word frequency and stem (or 'base' or 'lemma') frequency are the measures of frequency most commonly incorporated into studies in contemporary sociolinguistics and psycholinguistics. For our data, it turns out that these measures are very highly correlated, and correspondingly predict extremely similar patterns of CSD across different subsets of the data. On the assumption that these frequency measures would also correlate this strongly throughout the lexicon (not just for our subset of CSD words), we offer the methodological recommendation that whole-word frequency, which is considerably more straightforward to implement than stem frequency, will be at least as effective as stem frequency for capturing frequency-related variance in other linguistic variables. In other words, for researchers who simply want to incorporate a reasonable frequency control into studies that are primarily aimed at investigating other phenomena, it will not be worth the effort to operationalize a stem frequency measure. With regard to the specific pattern found using these measures, we observe a main effect of whole-word and stem frequency on CSD outcomes for the monomorphemes—coronal stops are more likely to be deleted at the end of frequent monomorphemes than infrequent monomorphemes—but not for regular past forms. An equivalent interaction between morphological category and whole-word frequency has also been reported for other CSD datasets (Myers & Guy 1997; Bayley 2014), but never at such a large scale, or corroborated with the same finding for stem frequency. In our data, as in these previous studies, the effect of frequency within monomorphemes is similar to that which has been observed for a number variable lenition and reduction phenomena, specifically that highly frequent and therefore highly predictable and/or highly practised words are pronounced with more reduced and lenited forms. However, additional explanation is required for why the same effect is not straightforwardly found for whole-word or stem frequency among regular past forms.

A potential avenue for explanation comes from Erker & Guy (2012), who report a similar interaction between whole-word frequency and grammatical category in the rate of subject personal pronoun omission in Spanish. In their data, the effects of verb regularity, verb semantics, subject person/number, and utterance tense/mood/aspect are small or nonexistent among low (whole-word) frequency verbs, but large among high frequency verbs. Thus, whole-word frequency is taken to 'amplify' the effect of these grammatical categories. The proposed reason for this is that speakers need a certain amount of experience with a word in order for the effects of its grammatical category to be learned and reproduced, either as emergent from the particular contexts in which words of that category appear or as a more abstract property that entails a particular rate of some variant. This aligns with a 'passive' perspective on frequency effects in that it is the mental representation of words that is implicated, rather than any on-line mechanism. Our results are also consistent with this idea of amplification: among high frequency words, the rate of CSD is far higher in monomorphemes than in regular past forms, but there is very little difference between low frequency monomorphemes and regular past forms in terms of rates of CSD.

On the other hand, a deficiency of the amplification story is that, at least for CSD, grammatical categories are treated more or less like arbitrary labels for words. In reality, monomorphemes and regular past forms differ in terms of morphological complexity, which may explain what we observe in terms of sensitivity (or lack thereof) to measures of frequency. Morphological complexity has two relevant properties as pertains to frequency. The first is that of informativity: while coronal stops at the end of monomorphemes are often highly predictable and contain no additional disambiguating information about the word, coronal stops at the end of regular past forms constitute a suffix that marks past tense. Moreover, when this suffix is deleted, regular past forms are always homophonous with a present or infinitival form of the verb. These are some of the primary concerns of linguists who ascribe a 'functional' motivation to grammatical patterns of CSD, arguing that deletion is avoided in cases where it would eliminate important past tense information (e.g. Kiparsky 1972).

The second relevant property of morphological complexity is that it entails pieces (whether independently-represented or emergent from shared phonology and semantics) being shared across words. That is, not only does CSD target an informative suffix when it applies in regular past forms, it targets the *same* suffix identity for every lexical item in the grammatical category. Given that we are asking about the frequency of different linguistic units, we are forced to consider whether the relevant frequency measure for this kind of word might not be of the whole-word or the stem, but the *-ed* suffix itself. Of course, a raw measure of this kind would amount to a single (high) frequency value, and would not be particularly useful for explaining the basic effect of grammatical category, never mind differences between words within a single category. Therefore, instead of considering the frequency of a suffix overall, it may be more fruitful to consider the frequency of a suffix (or the absence of a suffix) under certain conditions. This is what is achieved by our conditional frequency measure.

5.2 Main effect of conditional frequency

What we have called 'conditional frequency' is the proportion of instances of a stem that are realized as a certain whole-word. Unlike for whole-word and stem frequency, we find strong effects of conditional frequency on predicting CSD outcomes in all of our regression models.

For regular past forms, conditional frequency corresponds to the decontextualized probability of the *-ed* suffix as an ending for a given stem. *-ed* suffixed forms that are common, relative to other reflexes of the same stem, are less likely to retain the coronal stop that marks this suffix. We can consider this result in light of the functionalist framing that is sometimes used to describe the main effect of grammatical category, which states that coronal stops are less likely to be deleted when they encode important grammatical information, i.e. an *-ed* suffix. As previously mentioned, the functional analysis has a great deal in common with listener-oriented perspectives of reduction/lenition phenomena (Lindblom 1990; Aylett & Turk 2004; Jurafsky et al. 2001), which also describe the preservation of unpredictable and therefore informative structure, but the functional account makes specific reference to the grammatical information encoded by an *-ed* suffix rather than more general properties of word or segment probability. Our findings show that even a functionally important coronal stop, representing the past tense suffix, is frequently deleted when that suffix is a highly paradigmatically frequent—and therefore highly predictable—ending to the stem.

While *-ed* suffixed words lend themselves to an intuitive interpretation of the conditional frequency measure, it may be more difficult to conceptualize a similar effect in monomorphemes. If the effect of conditional frequency is to be explained in terms of how predictable a suffix is given a stem, why would we see the same effect for no suffix at all? Indeed, as well as the frequency of an *-ed* suffix given a stem, conditional frequency in regular past forms is also equivalent to the frequency of an underlying coronal stop in this context. The conditional frequency of *kicked* is both the rate at which $\sqrt{\text{KICK}}$ is used in the past or passive and the rate at which /kɪk/ is followed by an underlying word-final /t/ with no intervening word boundary. In other words, both morphological and phonological levels of representation are captured with the same measure. Conversely, for the monomorphemes in this study, all or most of the words that are morphologically related to them also have underlying representations that contain the relevant coronal stop. The conditional frequency of *act* does not capture the rate at which a /t/ appears with the stem $\sqrt{\text{ACT}}$, because that /t/ is part of the reflex of the stem itself and therefore also occurs in *acts*, *acting* and *actor*. Instead, conditional frequency in monomorphemes only corresponds to the rate at which the stem occurs with a coronal stop in word-final position, as opposed to being followed by additional phonological material within the word. If we are to reconcile this with the predictability view that works well for *-ed* suffixed forms, we can think of the conditional frequency effect in monomorphemes in terms of edge marking. Stems that do not commonly appear with a word-final coronal stop, relative to other possible forms where the stem is combined with various suffixes, are more likely to retain this coronal stop due to hyperarticulation at the word edge. In other words, stem-conditionally predictable word endings promote deletion, just as stem-conditionally predictable suffixes favor deletion.

Our results, that high conditional frequency corresponds to a high rate of coronal stop deletion, conflict with some recent findings of 'paradigmatic enhancement' effects. This is the class of results where the most common reflexes of a particular word or morpheme are found to be phonetically reinforced rather than reduced. These effects are framed from both speaker-oriented and passive perspectives. They are commonly interpreted in terms of speakers articulating common reflexes of a morpheme with increased confidence, suggesting an on-line pressure to reduce in cases where the speaker is unconfident. At the same time, speaker confidence itself has been explained as the result of extensive motor practice, allowing these words to be executed with enhanced kinematic skill (Tomaschek et al., 2018), suggesting an evolution of the specific representation or implementation associated with a word that is not generated on-line. However, comparison between paradigmatic enhancement findings and our own results is not straightforward. As we have already discussed in this section, the conditional frequency measure captures different facts about the coronal stops in monomorphemes versus *-ed* suffixed forms. Our results for monomorphemes lend themselves to a comparison with findings regarding the pronunciation of stem vowels with various suffixes (Tucker et al. 2019; Tomaschek et al. 2021), and perhaps even more pertinently with those concerning the pronunciation of consonants in the component pieces of compound nouns (Bell et al. 2021). All three of these studies find evidence of reinforcement when a stem or word is followed by a common ending; the present study finds the opposite. Instead, we show that monomorphemes whose stems typically occur in that form, with no additional suffix, are more likely to exhibit CSD than monomorphemes whose stems are more commonly suffixed.

In the case of *-ed* suffixed forms, our results can be more straightforwardly compared to research on the pronunciation of affixes themselves in terms of their relationship to a given stem (Kuperman et al. 2007; Hanique & Ernestus 2011; Smith et al. 2012; Schuppler et al. 2012; Cohen 2015; Ben Hedia & Plag 2017; Plag & Ben Hedia 2017; Tomaschek et al. 2019). In these studies and in terms of the regular past forms in ours, the frequency of the affix itself, as attached to a given stem, is what is compared to the frequency of the same word/stem with other affixes or with no affix at all. While some studies of this type look at functionally equivalent affixes in direct competition (Kuperman et al. 2007; Cohen 2015), our past tense form study aligns with the many others comparing the frequency of one affix as an ending to a stem to the frequency of the whole paradigm (Hanique & Ernestus 2011; Cohen 2015; Tomaschek et al. 2019). Like these studies, using a different suffix in place of *-ed* will no longer denote the past tense. This means the conditional frequency of an *-ed* suffixed form does not capture the same 'pocket of uncertainty', where a language user could use more than one form to convey the same thing, that was considered to be so important in many early paradigmatic enhancement results. Correspondingly, our results indicate more reduction when *-ed* is a frequent ending to a stem, aligning in particular with Hanique & Ernestus's (2011) result of greater reduction and deletion of word-final /t/ in Dutch irregular past verb forms when it is frequent within the paradigm, as opposed to Schuppler et al.'s (2012) result of greater word-final /t/ *retention* in Dutch 3SG present verb forms when this form is more frequently used than the 1SG form of the same verb. However, this pocket of uncertainty is surely to be found at every site of a sociolinguistic variable. And certainly, Bürki et al. (2011) find a comparable enhancement effect such that French variable schwa (e.g. *fenêtre* [f(ə)nɛtʀ] 'window') is longer in words that appear relatively more frequently with schwa compared to without it. In other words, even though we find no evidence of paradigmatic enhancement effects in this study, we might predict that future studies would find such effects corresponding to *variant frequency*, e.g. enhancement that is negatively correlated with the rate of CSD for a given word, such that more commonly retained stops have a reinforced pronunciation when they are retained. What our present results for conditional frequency do point towards is an effect of suffix predictability, and a corresponding effect of word edge predictability, that are correlated with CSD rates. These effects are ultimately different reflexes of the same, listener-oriented, goal to signal that the listener should not expect another suffix.

6 Conclusions

Whether researchers are directly exploring frequency effects or are trying to control for frequency in their statistical model, we must consider which frequency measure to use, how that measure relates to the purported mechanism of its influence, and how it may interact with morphological or other linguistic structure. While whole-word and stem frequency are the most commonly used frequency measures, our model comparisons showed that conditional frequency is a strong predictor of CSD. This result suggests that greater consideration should be given to conditional frequency as a predictor in the study of phonological variation. From a purely methodological perspective, it far outperforms whole-word and stem frequency in terms of accounting for variance in the dataset as a whole as well as in subsets restricted to words of just one grammatical category. Moreover, we see that conditional frequency is, at most, relatively weakly correlated with the other measures in this study (**Figure 1**), allowing it to be used alongside other frequency measures without a high risk of multicollinearity issues. In fact, in this study the addition of whole-word and stem frequency do little to improve the accuracy of any model that already contains conditional frequency.

But the value of giving greater attention to conditional frequency is not purely methodological. We have interpreted conditional probability in terms of the predictability of either an *-ed* suffix (for morphologically complex CSD words) or a word boundary (for monomorphemes), given the stem. Under that interpretation, coronal stops are more likely to be retained when they are associated with word endings that have low stem-conditional predictability. The relatively high importance of conditional probability therefore suggests an important role for listener-oriented considerations in the explanation of the frequency/lenition relationship. However, our results go beyond basic functional accounts that involve avoiding the omission of grammatical information by showing that even key grammatical information can be elided when it is highly predictable. At the same time, the robust interaction we found between whole-word and stem frequency measures and morphological category indicates that basic word predictability measures may be insufficient for cases where phonetic or phonological variation extends across morphological boundaries. It appears that, at least for a phenomenon like CSD, the predictability measures that matter most are ones that are relative to the internal structure of words and their morphological relatives. Exactly how speakers and listeners make predictions across word forms, and how far explanations appealing to the consequences of this kind of predictive behavior can take us in understanding pronunciation variation, remains to be seen. We certainly invite replications of our methods on other CSD datasets, as well as other variables, to help shed light on this question. Regardless, we hope these findings encourage greater nuance in the use of and interpretation of frequency and predictability effects in the study of language variation and change.

Additional file

The additional file for this article can be found as follows:

- **Appendix.** Appendix containing (A) full results from nested logistic regression model comparisons, and (B) individual logistic regression model summaries. DOI: https://doi.org/10.16995/glossa.5839.s1

Competing interests

The authors have no competing interests to declare.

References

Alba, Matthew C. 2006. Accounting for variability in the production of Spanish vowel sequences. In Sagarra, Nuria & Toribio, Almeida Jacqueline (eds.), *Selected proceedings of the 9th hispanic linguistics symposium,* 273–285.

Aylett, Matthew & Turk, Alice. 2004. The smooth signal redundancy hypothesis: A functional explanation for relationships between redundancy, prosodic prominence and duration in spontaneous speech. *Language and Speech* 47. 31–56. DOI: https://doi.org/10.1177/00238309040470010201

Baese-Berk, Melissa M. & Goldrick, Matthew. 2009. Mechanisms of interaction in speech production. *Language and Cognitive Processes* 24(4). 527–554. DOI: https://doi.org/10.1080/01690960802299378

Balling, Laura Winther & Baayen, Harald. 2008. Morphological effects in auditory word recognition: Evidence from Danish. *Language and Cognitive Processes* 23(7–8). 1159–1190. DOI: https://doi.org/10.1080/01690960802201010

Baranowski, Maciej & Turton, Danielle. 2020. TD-deletion in British English: New evidence for the long-lost morphological effect. *Language Variation and Change* 32. 1–23. DOI: https://doi.org/10.1017/S0954394520000034

Bates, Douglas & Mächler, Martin & Bolker, Ben & Walker, Steve. 2015. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software* 67(1). 1–48. DOI: https://doi.org/10.18637/jss.v067.i01

Bayley, Robert. 2014. The role of frequency in phonological and morphological variation. In *New ways of analyzing variation 43*.

Bell, Melanie J. & Ben Hedia, Sonia & Plag, Ingo. 2021. How morphological structure affects phonetic realisation in English compound nouns. *Morphology* 31(2). 87–120. DOI: https://doi.org/10.1007/s11525-020-09346-6

Ben Hedia, Sonia & Plag, Ingo. 2017. Gemination and degemination in English prefixation: Phonetic evidence for morphological organization. *Journal of Phonetics* 62. 34–49. DOI: https://doi.org/10.1016/j.wocn.2017.02.002

Boersma, Paul & Hayes, Bruce. 2001. Empirical tests of the gradual learning algorithm. *Linguistic Inquiry* 32. 45–86. DOI: https://doi.org/10.1162/002438901554586

Browman, Catherine & Goldstein, Louis. 1990. Tiers in articulatory phonology, with some implications for casual speech. In Beckman, Mary E & Kingston, John (eds.), *Papers in Laboratory Phonology I: Between the grammar and the physics of speech,* 341–376. DOI: https://doi.org/10.1017/CBO9780511627736.019

Brown, Earl K. 2009. The relative importance of lexical frequency in syllable- and word-final /s/ reduction in Cali, Colombia. In Collentine, Joseph, García, Maryellen, Lafford, Barbara & Marcos Marín, Francisco (eds.), *Selected proceedings of the 11th Hispanic Linguistics symposium,* 165–178.

Brown, Esther L. & Cacoullos, Rena Torres. 2003. Spanish /s/. A different story from beginning (initial) to end (final). In Núñez-Cedeño, Rafael, López, Luis & Cameron, Richard (eds.), *A Romance perspective in language knowledge and use,* 22–38. Amsterdam: John Benjamins. DOI: https://doi.org/10.1075/cilt.238.05bro

Brysbaert, Marc & Mandera, Paweł & Keuleers, Emmanuel. 2017. The word frequency effect in word processing: An updated review. *Current Directions in Psychological Science* 27(1). 45–50. DOI: https://doi.org/10.1177/0963721417727521

Brysbaert, Marc & New, Boris. 2009. Moving beyond Kučera and Francis: A critical evaluation of current word frequency norms and the introduction of a new and improved word frequency measure for American English. *Behavioral Research Methods* 41(4). 977–90. DOI: https://doi.org/10.3758/BRM.41.4.977

Bürki, Audrey & Ernestus, Mirjam & Gendrot, Cédric & Fougeron, Cécile & Frauenfelder, Ulrich Hans. 2011. What affects the presence versus absence of schwa and its duration: A corpus analysis of French connected speech. *Journal of the Acoustical Society of America* 130(6). 3980–3991. DOI: https://doi.org/10.1121/1.3658386

Bybee, Joan. 2002. Word frequency and context of use in the lexical diffusion of phonetically conditioned sound change. *Language Variation and Change* 14. 261–290. DOI: https://doi.org/10.1017/S0954394502143018

Chomsky, Noam & Halle, Morris. 1968. *The sound pattern of English*. New York: Harper & Row.

Clopper, Cynthia G. & Turnbull, Rory. 2018. Exploring variation in phonetic reduction: Linguistic, social, and cognitive factors. In Cangemi, Francesco, Clayards, Meghan, Niebuhr, Oliver, Schuppler, Barbara & Zellers, Margaret (eds.), *Rethinking reduction: Interdisciplinary perspectives on conditions, mechanisms, and domains for phonetic variation*, 25–72. De Gruyter Mouton. DOI: https://doi.org/10.1515/9783110524178-002

Coetzee, Andries & Kawahara, Shigeto. 2013. Frequency and other biases in phonological variation. *Natural Language and Linguistic Theory* 31. 47–89. DOI: https://doi.org/10.1007/s11049-012-9179-z

Coetzee, Andries W. & Pater, Joe. 2011. The place of variation in phonological theory. In Goldsmith, John, Riggle, Jason & Yu, Alan (eds.), *The Handbook of Phonological Theory. 2nd Edition,* 401–434: Blackwell. DOI: https://doi.org/10.1002/9781444343069.ch13

Cofer, Thomas M. 1972. *Linguistic variability in a Philadelphia speech community:* University of Pennsylvania dissertation.

Cohen, Clara. 2015. Context and paradigms: two patterns of probabilistic pronunciation variation in Russian agreement suffixes. *Mental Lexicon* 10(3). 313–338. Publisher: John Benjamins Publishing Company. DOI: https://doi.org/10.1075/ml.10.3.01coh

Connine, Cynthia M. 1990. Word familiarity and frequency in visual and auditory word recognition. *Journal of Experimental Psychology: Learning, Memory and Cognition* 16. 1084–1096. DOI: https://doi.org/10.1037/0278-7393.16.6.1084

Damian, Markus F. 2003. Articulatory duration in single-word speech production. *Journal of Experimental Psychology: Learning, Memory and Cognition* 29(3). 416–431. DOI: https://doi.org/10.1037/0278-7393.29.3.416

Diaz-Campós, Manuel & Carmen, Ruiz-Sánchez. 2008. The value of frequency as a linguistic factor: The case of two dialectal regions in the Spanish speaking world. In *Selected proceedings of the 4th workshop on spanish sociolinguistics.*

Diaz-Campós, Manuel & Gradoville, Michael. 2011. An analysis of frequency as a factor contributing to the diffusion of variable phenomena: Evidence from Spanish data. In Ortiz-López, Luis A. (ed.), *Selected proceedings of the 13th Hispanic Linguistics symposium,* 224–238. Cascadilla Proceedings Project.

Dupoux, Emmanuel & Mehler, Jacques. 1990. Monitoring the lexicon with normal and compressed speech: Frequency effects and the prelexical code. *Journal of Memory and Language* 29. 316–335. DOI: https://doi.org/10.1016/0749-596X(90)90003-I

Erker, Daniel & Guy, Gregory R. 2012. The role of lexical frequency in syntactic variability: variable subject personal pronoun expression in Spanish. *Language* 88(3). 526–557. DOI: https://doi.org/10.1353/lan.2012.0050

Fasold, Ralph. 1972. *Tense marking in Black English. a linguistic and social analysis.* (Urban Language Series 8). Center for Applied Linguistics.

File-Muriel, Richard. 2009. The role of lexical frequency in the weakening of syllable-final lexical /s/ in the Spanish of Barranquilla. *Hispania* 92. 348–360.

Fiorentino, Robert & Poeppel, David. 2007. Compound words and structure in the lexicon. *Language and cognitive processes* 22(7). 953–1000. DOI: https://doi.org/10.1080/01690960701190215

Forrest, Jon. 2017. The dynamic interaction between lexical and contextual frequency: A case study of (ing). *Language Variation and Change* 29. 129–156. DOI: https://doi.org/10.1017/S0954394517000072

Fruchter, Joseph & Stockall, Linnaea & Marantz, Alec. 2013. MEG masked priming evidence for form-based decomposition of irregular verbs. *Frontiers in Human Neuroscience* 7. 798. DOI: https://doi.org/10.3389/fnhum.2013.00798

Gahl, Susanne. 2008. *Time and thyme* are not homophones: the effect of lemma frequency on word durations in spontaneous speech. *Language* 84(3). 474–496. DOI: https://doi.org/10.1353/lan.0.0035

Gahl, Susanne & Garnsey, Susan. 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

Goldrick, Matthew. 2011. Using psychological realism to advance phonological theory. In Goldsmith, John, Riggle, Jason & Yu, Alan C.L. (eds.), *The Handbook of Phonological Theory, Second Edition,* chap. 19, 631–660. Blackwell. DOI: https://doi.org/10.1002/9781444343069.ch19

Gut, Ulrike. 2007. First language influence and final consonant clusters in the new Englishes of Singapore and Nigeria. *World Englishes* 26(3). DOI: https://doi.org/10.1111/j.1467-971X.2007.00513.x

Guy, Gregory R. 1980. Variation in the group and the individual: The case of final stop deletion. In Labov, William (ed.), *Locating language in time and space,* chap. 1, 1–36. New York: Academic Press.

Guy, Gregory R. 1991. Explanation in variable phonology: An exponential model of morphological constraints. *Language Variation and Change* 3. 1–22. DOI: https://doi.org/10.1017/S0954394500000429

Guy, Gregory R. 1996. Form and function in linguistic variation. In Guy, Gregory R., Feagin, Crawford, Schiffrin, Deborah & Baugh, John (eds.), *Towards a social science of language: Papers in honor of william labov,* vol. 1, 221–. John Benjamins Publishing Co. DOI: https://doi.org/10.1075/cilt.127

Guy, Gregory R. 2007. Lexical exceptions in variable phonology. *University of Pennsylvania Working Papers in Linguistics* 13(2). 109–119.

Guy, Gregory R. 2019. Variation and mental representation. In Lightfoot, David W. & Havenhill, Jonathan (eds.), *Variable properties in language: their nature and acquisition,* chap. 11, 129–140. Georgetown University Press. DOI: https://doi.org/10.2307/j.ctvfxv99p.15

Hanique, Iris & Ernestus, Mirjam. 2011. Final /t/ reduction in Dutch past-participles: The role of word predictability and morphological decomposability. In *Interspeech 2011: 12th Annual Conference of the International Speech Communication Association*, 2849–2852. DOI: https://doi.org/10.21437/Interspeech.2011-713

Hansen Edwards, Jette G. 2016. Sociolinguistic variation in asian englishes. *English World-Wide. A Journal of Varieties of English* 37(2). 138–167. DOI: https://doi.org/10.1075/eww.37.2.02han

Hay, Jennifer B. & Pierrehumbert, Janet B. & Walker, Abby & LaShell, Patrick. 2015. Tracking word frequency effects through 130 years of sound change. *Cognition* 139. 83–91. DOI: https://doi.org/10.1016/j.cognition.2015.02.012

Hayes, Bruce & Wilson, Colin. 2008. A maximum entropy model of phonotactics and phonotactic learning. *Linguistic Inquiry* 39(3). 379–440. DOI: https://doi.org/10.1162/ling.2008.39.3.379

Hazen, Kirk. 2011. Flying high above the social radar: Coronal stop deletion in modern Appalachia. *Language Variation and Change* 23. 105–137. DOI: https://doi.org/10.1017/S0954394510000220

Hooper, Joan. 1976. Word frequency in lexical diffusion and the source of morphophonological change. In Christie, William M., Jr. (ed.), *Current progress in historical linguistics,* 95–105. Amsterdam: North-Holland.

Howes, Davis H. 1957. On the relation between the probability of a word as an association and in general linguistic usage. *The Journal of Abnormal and Social Psychology* 54(1). 75–85. DOI: https://doi.org/10.1037/h0043830

Jurafsky, Daniel & Bell, Alan & Fosler-Lussier, Eric & Girand, Cynthia & Raymond, William. 1998. Reduction of English function words in Switchboard. *Proceedings of ICSLP* 98(7). 3111–3114. DOI: https://doi.org/10.21437/ICSLP.1998-801

Jurafsky, Daniel & Bell, Alan & Gregory, Michelle & Raymond, William D. 2001. Probabilistic relations between words: Evidence from reduction in lexical production. In Bybee, Joan & Hopper, Paul (eds.), *Frequency and the emergence of linguistic structure,* 229–254. Amsterdam: John Benjamins. DOI: https://doi.org/10.1075/tsl.45.13jur

Kawamoto, Alan H. 1999. Incremental encoding and incremental articulation in speech production: Evidence based on response latency and initial segment duration. *Behavioral and Brain Sciences* 22. 48–49. DOI: https://doi.org/10.1017/S0140525X99331770

Kiparsky, Paul. 1972. Explanation in phonology. In Peters, Stanley (ed.), *Goals of linguistic theory,* 189–227. Cinnaminson, NJ: Prentice-Hall.

Kroch, Anthony. 1994. Morphosyntactic variation. In Beals, K. (ed.), *Papers from the 30th regional meeting of the chicago linguistic society*, 190–201. Chicago Linguistic Society.

Kuperman, Victor & Pluymaekers, M. & Ernestus, Mirjam & Baayen, R. Harald. 2007. Morphological predictability and acoustic duration of interfixes in Dutch compounds. *The Journal of the Acoustical Society of America* 121(4). 2261–2271. DOI: https://doi.org/10.1121/1.2537393

Labov, William. 1969. Contraction, deletion, and inherent variability of the English copula. *Language* 45(4). 715–762. DOI: https://doi.org/10.2307/412333

Labov, William. 1972. Some principles of linguistic methodology. *Language in Society* 1(1). 97–120. DOI: https://doi.org/10.1017/S0047404500006576

Labov, William & Rosenfelder, Ingrid. 2011. The Philadelphia Neighborhood Corpus of LING 560 studies, 1972-2010. With support of NSF contract 921643.

Lewis, Gwyneth & Solomyak, Olla & Marantz, Alec. 2011. The neural basis of obligatory decomposition of suffixed words. *Brain and Language* 118–127. DOI: https://doi.org/10.1016/j.bandl.2011.04.004

Lim, Laureen T. & Guy, Gregory R. 2005. The limits of linguistic community: Speech styles and variable constraint effects. *University of Pennsylvania Working Papers in Linguistics* 10(2). 157–170.

Lin, Susan & Beddor, Patrice Speeter & Coetzee, Andries W. 2014. Gestural reduction, lexical frequency, and sound change: A study of post vocalic /l/. *Laboratory Phonology* 5(1). 9–36. DOI: https://doi.org/10.1515/lp-2014-0002

Lindblom, Björn. 1963. Spectrographic study of vowel reduction. *Journal of the Acoustical Society of America* 35. 1773–1781. DOI: https://doi.org/10.1121/1.1918816

Lindblom, Björn. 1990. Explaining phonetic variation: A sketch of the H&H theory. In Hardcastle, William .J. & Marchal, Alain (eds.), *Speech production and speech modelling,* 403–439. Kluwer Academic Publishers. DOI: https://doi.org/10.1007/978-94-009-2037-8_16

MacKenzie, Laurel & Tamminga, Meredith. 2021. New and old puzzles in the morphological conditioning of coronal stop deletion. *Language Variation and Change* 33. 217–244. DOI: https://doi.org/10.1017/S0954394521000119

Mousikou, Petroula & Strycharczuk, Patrycja & Turk, Alice & Scobbie, James. 2015. Morphological effects on pronunciation. In *Proceedings of the 18th international congress of the phonetic sciences.*

Munson, Benjamin & Solomon, Nancy Pearl. 2004. The effect of phonological neighborhood density on vowel articulation. *Journal of speech, language, and hearing research* 47(5). 1048–1058. DOI: https://doi.org/10.1044/1092-4388(2004/078)

Myers, James & Guy, Gregory R. 1997. Frequency effects in variable lexical phonology. *University of Pennsylvania Working Papers in Linguistics* 4(1). 215–228.

Neu, Helen. 1980. Ranking of constraints on /t,d/ deletion in American English: A statistical analysis. In Labov, William (ed.), *Locating language in time and space,* chap. 2, 37–54. New York: Academic Press.

Patrick, Peter L. 1991. Creoles at the intersection of variable processes: (td)-deletion and past-marking in the Jamaican mesolect. *Language Variation and Change* 3(2). 171–189. DOI: https://doi.org/10.1017/S095439450000051X

Phillips, Betty S. 1981. Lexical diffusion and Southern *tune, duke, news*. *American Speech* 56. 72–78. DOI: https://doi.org/10.2307/454480

Phillips, Betty S. 1984. Word frequency and the actuation of sound change. *Language* 60(2). 320–342. DOI: https://doi.org/10.2307/413643

Pierrehumbert, Janet B. 2002. Word-specific phonetics. In *Laboratory Phonology VII*, 101–139. Berlin: Mouton de Gruyter. DOI: https://doi.org/10.1515/9783110197105.1.101

Plag, Ingo & Ben Hedia, Sonia. 2017. The phonetics of newly derived words: Testing the effect of morphological segmentability on affix duration. In Arndt-Lappe, Sabine, Braun, Angelika, Moulin, Claudine & Winter-Froemel, Esme (eds.), *Expanding the Lexicon: Linguistic Innovation, Morphological Productivity, and the Role of Discourserelated Factors,* Berlin, New York: de Gruyter Mouton. DOI: https://doi.org/10.1515/9783110501933-095

Purse, Ruaridh. 2019. The articulatory reality of coronal stop ‘deletion’. In *Proceedings of the 19th International Congress of Phonetic Sciences.*

Purse, Ruaridh & Turk, Alice. 2016. ‘/t,d/ deletion’: Articulatory gradience in variable phonology. Paper presented at Laboratory Phonology 15, Cornell University, Ithaca, NY.

R Core Team. 2015. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing Vienna, Austria.

Roelofs, Ardi. 1992. A spreading-activation theory of lemma retrieval in speaking. *Cognition* 42. 107–142. DOI: https://doi.org/10.1016/0010-0277(92)90041-F

Rosenfelder, Ingrid & Fruehwald, Josef & Evanini, Keelan & Yuan, Jiahong. 2011. *FAVE program suite [forced alignment and vowel extraction]*. University of Pennsylvania. fave.ling.upenn.edu.

Santa Ana, Otto. 1991. *Phonetic simplification processes in the English of the barrio: a cross-generational sociolinguistic study of the Chicanos of Los Angeles:* University of Pennsylvania dissertation.

Savin, Harris B. 1963. Word frequency effects and errors in the perception of speech. *Journal of Verbal Learning and Verbal Behavior* 9. 292–302.

Schuppler, Barbara & Dommelen, Wim A. van & Koreman, Jacques & Ernestus, Mirjam. 2012. How linguistic and probabilistic properties of a word affect the realization of its final /t/: Studies at the phonemic and sub-phonemic level. *Journal of Phonetics* 40(4). 595–607. DOI: https://doi.org/10.1016/j.wocn.2012.05.004

Smith, Rachel & Baker, Rachel & Hawkins, Sarah. 2012. Phonetic detail that distinguishes prefixed from pseudo-prefixed words. *Journal of Phonetics* 40(5). 689–705. DOI: https://doi.org/10.1016/j.wocn.2012.04.002

Solomyak, Olla & Marantz, Alec. 2009. Lexical access in early stages of visual word processing: A single-trial correlational MEG study of heteronym recognition. *Brain and Language* 108. 191–196. DOI: https://doi.org/10.1016/j.bandl.2008.09.004

Solomyak, Olla & Marantz, Alec. 2010. MEG evidence for early morphological decomposition in visual word recognition: A single-trial correlational meg study. *Journal of Cognitive Neuroscience* 22. 2042–2057. DOI: https://doi.org/10.1162/jocn.2009.21296

Taft, Marcus & Hambly, Gail. 1986. Exploring the cohort model of spoken word recognition. *Cognition* 22. 259–282. DOI: https://doi.org/10.1016/0010-0277(86)90017-X

Tagliamonte, Sali & Temple, Rosalind. 2005. New perspectives on an ol' variable: (t,d) in British English. *Language Variation and Change* 17. 281–302. DOI: https://doi.org/10.1017/S0954394505050118

Tamminga, Meredith. 2016. Persistence in phonological and morphological variation. *Language Variation and Change* 28(03). 335–356. DOI: https://doi.org/10.1017/S0954394516000119

Temple, Rosalind. 2009. (t,d): the variable status of a variable rule. *Oxford University Working Papers in Linguistics, Philology and Phonetics* 12. 145–170.

Temple, Rosalind. 2014. Where and what is (t,d)? A case study in taking a step back in order to advance sociophonetics. In Celata, Chiara & Calamai, Silvia (eds.), *Advances in sociophonetics,* 97–136. Amsterdam: John Benjamins. DOI: https://doi.org/10.1075/silv.15.04tem

Tomaschek, Fabian & Plag, Ingo & Ernestus, Mirjam & Baayen, R. Harald. 2019. Phonetic effects of morphology and context: Modeling the duration of word-final S in English with naïve discriminative learning. *Journal of Linguistics* 57(1). 123–161. Publisher: Cambridge University Press. DOI: https://doi.org/10.1017/S0022226719000203

Tomaschek, Fabian & Tucker, Benjamin V. & Fasiolo, Matteo & Baayen, R. Harald. 2018. Practice makes perfect: The consequences of lexical proficiency for articulation. *Linguistics Vanguard* 4(S2). DOI: https://doi.org/10.1515/lingvan-2017-0018

Tomaschek, Fabian & Tucker, Benjamin V. & Ramscar, Michael & Baayen, R. Harald. 2021. Paradigmatic enhancement of stem vowels in regular English inflected verb forms. *Morphology* 31(2). 171–199. DOI: https://doi.org/10.1007/s11525-021-09374-w

Tomaschek, Fabian & Tucker, Benjamin V. & Wieling, Martijn & Baayen, R. Harald. 2014. Vowel articulation affected by word frequency. In *Proceedings of the 10th ISSP, Cologne.* 425–428.

Tomaschek, Fabian & Wieling, Martijn & Arnold, Denis & Baayen, R. Harald. 2013. Word frequency, vowel length and vowel quality in speech production: An ema study of the importance of experience. In *Proceedings of the Annual Conference of the International Speech Communication Association, INTERSPEECH*. DOI: https://doi.org/10.21437/Interspeech.2013-347

Tucker, Benjamin V. & Sims, Michelle & Baayen, R. Harald. 2019. Opposing forces on acoustic duration. Tech. rep. psyarxiv.com/jc97w. Publisher: PsyArXiv. DOI: https://doi.org/10.31234/osf.io/jc97w

Turnbull, Rory. 2015. Patterns of individual differences in reduction: Implications for listener-oriented theories. In *Proceedings of the 18th international congress of the phonetic sciences,* International Phonetics Association.

Turnbull, Rory. 2018. Effects of lexical predictability of patterns of phoneme deletion/reduction in conversational speech in English and Japanese. *Linguistics Vanguard* 4(S2). DOI: https://doi.org/10.1515/lingvan-2017-0033

Vannest, Jennifer & Newport, Elissa L. & Newman, Aaron J. & Bavelier, Daphne. 2011. Interplay between morphology and frequency in lexical access: The case of the base frequency effecy. *Brain Research* 1373. 144–159. DOI: https://doi.org/10.1016/j.brainres.2010.12.022

Walker, James A. 2012. Form, function, and frequency in phonological variation. *Language Variation and Change* 24. 397–415. DOI: https://doi.org/10.1017/S0954394512000142

Wolfram, Walt & Christian, Donna. 1976. *Appalachian speech*. Arlington, VA: Center for Applied Linguistics.

Wright, Charles E. 1979. Duration differences between rare and common words and their implications for the interpretation of word frequency effects. *Memory and Cognition* 7(6). 411–419. DOI: https://doi.org/10.3758/BF03198257

Zweig, Eytan & Pylkkänen, Liina. 2009. A visual M170 effect of morphological complexity. *Language and Cognitive Processes*, 412–439. DOI: https://doi.org/10.1080/01690960802180420