



Phonetic correlates of prefixedness in four varieties of English

Sarah Hawkins*, Centre for Music and Science, Faculty of Music, University of Cambridge, Cambridge, UK, sh110@cam.ac.uk

Jen Hay, Department of Linguistics, University of Canterbury, Christchurch, New Zealand; New Zealand Institute of Language, Brain and Behaviour, University of Canterbury, Christchurch, New Zealand

Rachel Smith, Department of English Language and Linguistics, University of Glasgow, Glasgow, UK

Jennifer Stuart-Smith, Department of English Language and Linguistics, University of Glasgow, Glasgow, UK

Robert Fromont, New Zealand Institute of Language, Brain and Behaviour, University of Canterbury, Christchurch, New Zealand

*Corresponding author.

Studies show that phonetic detail signals morphological structure. Previous studies of English prefixes examined one dialect and speech style, and/or measured only duration. This study examines phonetic detail of prefix and nonprefix *mis-* and *dis-* (e.g., *discolour*, *discover*) in four dialects of natural conversational English: New Zealand, USA (Ohio), UK (Glasgow and Liverpool). Principal component (PC) analyses reduced 14 acoustic variables to five dimensions of variation: [s] duration, [s] spectral shape, vowel duration, vowel quality and [s] ‘low-frequency cutoff’ frequency. We assessed whether these dimensions are affected by prefixedness, in a binary or gradient fashion, and by control predictors of word frequency, early/late lexical stress, the fourth phoneme, and *mis-dis-*. The dialects used different combinations of PCs to signal prefix status, with each PC used by at least one dialect. Vowel duration is the most statistically robust and generalizable effect. Whether or not the fourth phoneme is a stop is also influential. All PCs showed binary effects of prefix status. Within prefixes, vowel quality and [s] low-frequency cutoff showed gradient effects, covarying with judged degree of prefixedness. The observed patterns reflect how each dialect adjusts these syllables’ prosodic weight to conform with its dialect-specific speech rhythm and patterns of syllable reduction.



1. Introduction and background

This paper examines natural conversational speech in a range of varieties of English to ask whether the *mis* in a prefixed word like *mistimes* exhibits subtle pronunciation differences from the *mis* in an unprefixed word like *mistakes*. Linguists, lexicographers and phoneticians have long recognized that when a word-initial phoneme or phoneme sequence functions as a prefix in English, it is typically pronounced differently from that same phoneme or phoneme sequence when it begins a word but is not a prefix (Raffelsiefen, 1999). For example, the *mis* and *dis* in prefixed words like *mistimes* and *discolour* are pronounced differently from the *mis* and *dis* in unprefixed words like *mistakes* and *discover*: both are weak syllables, but other things equal, the morphological distinction accords a heavier prosodic beat to the prefix. But therein lies a problem: the relative prominence of this first syllable is affected by many factors in addition to prefix status, and prefix status itself, in any given word, is subject to flux over time. Consequently, while the distinction is widely accepted, there has been much debate both about criteria for defining prefix status, and about phonetic realisation of the prefix/nonprefix distinction.

Raffelsiefen (1999) offers a comprehensive and authoritative discussion. In line with other prosodic phonologists (Aronoff & Sridhar, 1983; Cohn & McCarthy, 1998; McCarthy & Prince, 1993; Selkirk, 1982; Sugahara & Turk, 2009; Szpyra, 1989), she concludes that the presence of a prosodic boundary between the putative prefix and the first syllable of the rest of the word confirms its status as a ‘p-word’ and hence an independent prefix. The boundary’s presence or absence produces contrasting syllabification between prefixes and nonprefixes. The properties criterial to achieving p-word status are structures below the word: foot, syllable, mora. These correlate with one another. P-words align with stems that comprise independent words, but not with Selkirk’s notion of root (Selkirk, 1984, cited in Raffelsiefen, 1999: p.199). Other theoretical approaches achieve the same type of distinction. For example, Lexical Phonology assigns prefixes to different levels of derivation (Siegel, 1974, Bauer, 1983). Updated versions of traditional Firthian Prosodic Phonology (e.g., Local, 1995; Ogden et al., 2000) combine strong/weak metrical stress, phonological weight (determined by syllable structure) and ambisyllabicity to describe phonetic detail of each segment in the syllable. Smith, Baker and Hawkins (2012, Appendix A) describe their application to words beginning with *mis* and *dis*.

Whilst there are differences between these theoretical descriptions, what is relevant for this study is that all agree that there is some distinction. Despite some exceptions (Hanique & Ernestus, 2012; Strycharczuk, 2019; Strycharczuk & Scobbie, 2016), empirical studies of a wide range of affixes confirm that morphological status can affect phonetic realization (e.g., Engemann & Plag, 2021; Hay, 2003; Plag & Ben Hedia, 2018; Plag, Homann & Kunter, 2017; Rose, 2017; Seyfarth et al., 2018; Schmitz, 2022; Stein, 2023; Strycharczuk, 2019).

Most published studies concern Dutch or English suffixes, but those that address prefixes show comparable patterns to suffixes (Ben Hedia & Plag, 2017; Hay, 2007; Smith et al., 2012; Zuraw et al., 2021). Yet the most detailed use controlled materials spoken in standard accents.

We do not know how systematically prefix-nonprefix distinctions are preserved across speaking styles—especially the wide range in conversational speech—nor how they are realized in less well-studied varieties of English. The distinction exhibits many of the factors associated with variation in syllable stress, while also interacting with lexical stress and utterance prosody, and being influenced by historical changes (see e.g., Raffelsiefen, 1999). So is prefixedness powerful enough to systematically appear in conversational speech as a property in its own right? That is, does it manifest in a way that transcends other powerful effects, such as word frequency and stress, which vary along the same acoustic parameters? Furthermore, examining the distinction in different dialects may ultimately help to elucidate how syllable reduction (the ‘other face’ of syllable stress) manifests in different phonological/phonetic systems. This in turn should contribute to better models of speech production, perception and comprehension.

The focus of this paper is thus on understanding better the acoustic-phonetic correlates of the prefix/nonprefix distinction, by exploring circumstances in which they are (or are not) preserved. Specifically, we examine corpora of conversational speech, in four very different varieties of English: New Zealand, Glasgow (Scotland), Liverpool and environs (NW England), and Ohio (USA).

Amongst prefixes, we focus only on *dis-* and *mis-*, for two main reasons. First, unlike many other monosyllabic putative prefixes—*un-*, *in-*, and *pre-*, for example—*dis-* and *mis-* have a rich yet relatively reliably-segmentable internal acoustic structure, potentially affording greater insight into phonetic processes underlying syllable reduction—one motivation for this study. As summarized below, variation in this internal acoustic structure due to prefix status was comprehensively described for one speech style in Standard Southern British English (SSBE) by Smith et al. (2012), but much about its generality remains to be learned. Most other studies measured only durations and/or single segments such as relative darkness of /l/ (Strycharczuk, 2019); these cannot offer the same level of insight into reduction processes available to speakers, and their implications for models of speech production.

Second, prefixes are of great interest because of their potential role in speech perception. Because prefixes occur at the beginning of a word, they could focus attention onto a restricted cohort of possible following words, thus facilitating prediction of meaning (e.g., Gaskell & Marslen-Wilson, 1997; Warren & Marslen-Wilson, 1987). A phonetic difference between prefixed *discolour* and unprefix *discover*, for example, could help to identify the word earlier than would be the case if the first phonetic difference is not heard until towards the end of the second syllable. In an eye tracking experiment, Clayards et al. (2021) showed that listeners do use the fine phonetic detail of the first syllable, in real time, to predict the identity of the rest of the word, confirming that they are sensitive to whether the first syllable is a prefix or not while they are hearing it. But while listener adaptation to distribution of stimulus properties was demonstrated, the stimuli still came from a single speaker and speech style. We do not know if the perceptual cues under these controlled conditions are robust enough to be preserved within all the variety of normal conversational speech and multiple talkers.

As noted above, the triphones that form the focus of the present study are *dis-* and *mis-*, such as in the prefixed forms *discolour*, *displace*, *disambiguate*, *mistimes*, *mismanage*, in contrast with unprefixed *discover*, *distress*, *discern*, *mistakes*, *mysterious*, *misdemeanour*, etc. Amongst the several acoustic distinctions between the prefixed and unprefixed forms that Smith et al. (2012) found, prefixes have a relatively longer vowel and shorter [s] than nonprefixes, making for a lower (more equal) ratio of durations of aperiodicity to periodicity in the prefix's syllable rhyme. This is referred to as the *s:i ratio*, or *[s]:[ɪ]*. Prefixes also had higher F2 frequency in [ɪ]. Center of Gravity measures showed that spectral differences in the [s] were relatable to the word's utterance prosody as nuclear or postnuclear, but not to prefix status. There were also robust differences in the VOT of voiceless stops in the onset of the words' second syllables (all the words' fourth phonemes were voiceless stops), with longer VOT after a prefix than a nonprefix, but as discussed below, VOT was not examined in this study.

The prefix/nonprefix distinction can thus be described as marked in SSBE by multiple acoustic parameters that covary to form two distinctive patterns. In Smith et al.'s (2012) data, no single parameter carried the distinction by itself, which is why we call these patterns the triphone's internal acoustic structure. When the triphone is a prefix, its vowel takes up a greater proportion of the triphone's overall duration. This presumably contributes to the percept that, though both types of syllable bear weak stress, the prefixed form conveys a somewhat heavier beat than the unprefixed form (Smith et al. (2012) and references therein). This heavier beat need not be due to a longer syllable. In Smith et al.'s (2012) data, triphone + closure + VOT was longer in prefixes than nonprefixes, but both triphone and rhyme total durations (i.e., excluding following stop closure and VOT) were remarkably similar regardless of the triphones' prefix status.

Smith et al. (2012) obtained their data from phonetically controlled 'scripted dialogues' written to naturally produce desired prosodic structures when fluently read by pairs of talkers. In some respects this can be considered good, clean data, much more like everyday conversational speech than most read laboratory speech. But it is not spontaneous speech. Moreover, in addition to following strict criteria on prefix status as agreed in the literature, their study deliberately used a highly restricted range of material in order to achieve maximum phonetic control: 10 word pairs, all second syllables taking primary lexical stress and beginning with a voiceless stop, for example *mistimes* (prefixed) versus *mistakes* (unprefixed). There are no comparable data for words in which the triphone's [s] is followed by a vowel (as in *misalign*), or where it can never form an onset (*misread*).

While the grammatical distinction can be represented as binary, matters are more complex from the point of view of an individual saying the relevant words, or using the phonetic detail they offer to facilitate perception. One theoretical and practical problem in studies of morphological variation of phonetic forms is that while morphological status is often manifest phonetically in terms of relative degree of syllable reduction, other influences on pronunciation

such as word frequency and rate of speech are also manifested as reduction. Distinguishing the various influences can be difficult.

For example, segmentable prefixes, which do not normally receive primary stress, nonetheless tend to be more ‘canonically’ articulated than their nonaffixed phonemic equivalents (Smith et al., 2012). But they also tend to be less frequent than the nonaffixed forms and less frequent words are often longer and/or more canonically articulated than more frequent words in connected speech (e.g., Bybee, 2006; Gahl, 2008). Consequently, phonetic reduction due to word frequency is difficult to disentangle from the same patterns due to morphological structure. Smith et al. (2012) discussed these issues with respect to the prefixes *dis-* and *mis-*, acknowledging the influences of frequency and rate of speech on reduction, but concluded that the regular differences they found throughout those syllables and in the onsets of a stressed second syllable in their phonetically controlled words are nonetheless better accounted for by morphological structure than by word frequency. This conclusion is supported in independent analyses by Plag and Ben Hedia (2018). Zuraw et al. (2021) found longer VOTs associated with lower word frequency and higher stem frequency but did not assess effects on the triphone. Their speech was collected in very artificial conditions, with words for their frequency analyses selected for variability (instability) in degree of aspiration. Nonetheless, their data nicely confirm expected interacting effects of word and stem frequencies.

Another factor making it challenging to assess phonetic effects of morphological structure itself is that, while some candidate words can be clearly categorized as prefixed or not (e.g., *discolour* and *discover*), many are more ambiguous, and there is both systematic and idiosyncratic variation in the degree to which a prefix is transparent. Such ambiguity can stem partly from the fact that at any given time, particular words are in various stages of progress from prefixed to unprefixed, or from unprefixed to prefixed, as discussed by Bauer (1983) and Raffelsiefen (1999) and summarized by Smith et al. (2012, sec. 4.2.3). For example, *disaster* originally meant *ill-starred* but now is clearly unprefixed, to the extent that the orthographic ‘s’ is realized as /z/ in the varieties of English we know of—whereas the prefix of *disarmament* is so clear that it can even be followed by a glottal stop.

Degree of perceived affixedness of particular words can vary across populations and contexts, and there is both systematic and idiosyncratic variation in the degree to which a prefix is transparent. Systematic variation is evident from dictionaries, which document variation between standard dialects: for example, *pre-* in *premature* can retain its phonetic transparency as a prefix in American English, /pri:/, but not in UK English, where it is /prɛ/. Idiosyncratic variation can be heard in day-to-day situations, so for example, in some words with a voiceless stop in the onset of syllable 2, for example *disposition*, the VOT, criterial of prefix status in unambiguous words, can vary considerably in the same individual’s speech—also reported by Zuraw et al. (2021). All of this raises the question of whether gradient morphological status leads to gradient phonetic marking. Are morphologically ‘ambiguous’ words also phonetically ambiguous?

Finally, it is not clear whether the differences found by Smith et al. (2012) are dialect-specific. To examine this question, we examined words beginning with *dis-* or *mis-* in corpora of conversational speech from the four English dialects listed earlier: New Zealand (NZ), Glasgow (GLA), Liverpool and environs (LIV), and Ohio (US). These were chosen partly because corpora were available with comparable segmentation, and, importantly, because they sound very different from SSBE and from each other, sometimes to the point of mutual unintelligibility (Adank et al., 2009; Garrett et al., 2005; Sharma et al., 2022; Wells, 1982). We start from the strong position that all dialects are likely to make a phonetic distinction between prefixes and nonprefixes. Thus, for each dialect, we look for differences in the internal acoustic structure of these phonemically identical triphones that reflect their prefix status.

The known differences in segmental quality between these dialects do not mean that they cannot make the same quality of distinction due to prefix status. Each might make the distinction in the same way as SSBE, a different way, or not at all. If it makes the distinction in a different way, then it could be because the articulation fosters some particular pattern typical of syllable reduction in that dialect, and/or because the pattern used is not articulatorily favoured, but fits with something else in the dialect's phonological system. These are side issues in this paper. It is the existence of and influences on prefix-related patterns that we are primarily interested in. What seems most likely to predict actual differences is how each dialect distinguishes more from less stressed syllables of this type.

In sum, the results presented by Smith et al. (2012) leave open a number of questions. A corpus study from natural conversational speech was therefore undertaken to answer the following four research questions. 1) To what extent is the prefixed-unprefixed distinction found in carefully controlled speech by Smith et al. (2012) also observed in spontaneous speech and in a wider range of phonological contexts? 2) Is the distinction observable across different dialects of English? 3) Can it be definitively distinguished from lexical frequency? 4) Is it best described as a binary phonetic distinction, or is there evidence of morphological gradience which is reflected in gradient phonetic realization?

To reduce repetition and to keep related analyses together for easier reference, this paper is structured in an atypical way, without an explicit Methods section. Rather, Sections 2 through 6 present the analytical structure (Corpora, Word Selection, Predictor Variables, Dependent Variables, and Statistical Approach). Section 7 presents the main analysis. This analysis identifies five Components (also called PCs) in a Principal Components analysis (PCA)—sets of acoustic variables that covary in the data set. Section 7 also offers a graphical overview of PC loadings for the acoustic variables, in each of the five main Components, together with the interpretative and analytic conventions we use. In Section 8, we analyse the main acoustic dimensions each Component relates to, and fit regression models to investigate the relationship between the acoustic Component and prefixedness. Section 9 provides a general discussion. Appendix A links to the data and analysis code.

2. Corpora

Words beginning with /dis/ or /mis/ were examined from corpora of New Zealand (NZ), Glasgow (GLA), Liverpool and environs (LIV), and Ohio (US) conversational English. As noted above, these varieties contrast greatly with SSBE and each other in phonemic realisations, and sometimes behave in different ways with respect to multiple and often covarying parameters involved. Segmentation used forced alignment (using DARPA criteria) from phonemic transcriptions: ESPS Aligner with subsequent hand-correction for Buckeye, and HTK (see Young et al., 2006; Fromont & Watson, 2016) for the other three.

The NZ corpora are drawn from a number of sources. The Intermediate Archive are speakers born between 1900–1930 (Gordon et al., 2007). Several corpora contribute recordings of contemporary General NZE: the Canterbury Corpus, the QuakeBox, Darfield and Southland corpora (Clark et al., 2016; D’Arcy, 2012; Gordon et al., 2007; Villarreal et al., 2021). Finally, we include data from the MAONZE corpus, which includes the English of speakers in NZ who are bilingual in NZ English and Māori (King et al., 2011). The Glasgow corpora comprise separate vernacular and standard Glaswegian talkers born between the 1890s and 1990s (Stuart-Smith et al., 2017). The Liverpool corpora comprise three separate regions, Liverpool city itself, and two nearby smaller towns, Skelmersdale and St. Helen’s. There is again considerable time depth, with speakers born between 1890 and 1994 (Watson & Clark, 2017). The US data comes from the Buckeye Corpus (Pitt & Fosler-Lussier, 2007), whose speakers were born between 1960 and 1970. Initial modelling attempted to investigate patterns of usage within different subcorpora and across time, but models were unstable due to low speaker numbers. We therefore report results that collapse within each broad dialect area, while acknowledging that there may be structured variation within each corpus.

Table 1 shows the numbers of tokens analysed (after removing outliers, as described in the next section). Tables S3.1–S3.4 (Supplementary Materials) show tokens organized by particular combinations of variables.

Corpus	Counts
US	152
NZ	579
LIV	78
GLA	193

Table 1: Distribution of Tokens across Corpora.

3. Preliminary word selection and classification

All words from the corpora that met the following criteria were included. 1) Polysyllabic and beginning with the orthographic letters *dis* or *mis*. 2) The first syllable does not carry the main lexical stress, since prefixes do not (this excludes words like *discus*, *distant*, *mischief* and *mishaps*). When necessary, this criterion was applied after listening to the particular utterance (cf. *discount*: the triphone has primary stress when used as a noun). 3) In citation form the orthographic ‘s’ must normally be pronounced as [s] not [z], which excludes words like *disease* and *disaster*.

The word types were then classified as prefixed or not, according to the semantic criteria used by Smith et al. (2012), which in turn followed those of Wurm (1997): 1) when the first syllable (the candidate prefix) is removed, what is left is a free-standing word (which could thus form the stem of the prefixed word); 2) the prefixed word has a semantically-transparent relationship with its stem, for example *displease*, *please*; 3) the meaning of the candidate prefix is consistent with other uses of that prefix. Thus the prefix *dis* “expresses negation...reversal or absence of an action or state”, and the prefix *mis* means “wrongly, badly, or unsuitably” (Oxford English Dictionary Online, 2008). For items whose prefix status was unclear, the Oxford Dictionary Online (2013) and Merriam-Webster Online were consulted, since each indicates prefix status, albeit with different systems.

These procedures resulted in a total of 1119 candidate tokens. Because all but one of the corpora are automatically aligned, and have not been manually checked, there is considerable scope for measurements of individual tokens to be affected by alignment errors. Prior to statistical analysis, we therefore inspected distributions for all acoustic measures, and removed tokens that were conspicuously outlying on any measure. This was done using cutoffs chosen manually, based on the shapes of the distributions. Specific thresholds are in the RMarkdown code (linked to in Appendix A). The final data set comprised 1002 tokens: 388 prefixed tokens (111 types) and 614 unprefixed tokens (93 types). These remaining tokens are not outlying on any durational or acoustic measure. Thus, although some alignment errors may remain in the dataset, we have some confidence that our results are not an artefact of alignment error.

4. Predictor Variables

We are aiming to understand variation in the acoustics of words beginning with *mis* and *dis*. As noted, we expect multiple influences on acoustic patterns. Accordingly, as well as the basic binary morphological predictor variable of *PrefixStatus*, we identified five other independent variables that we expected to influence the dependent measures in addition to any influence of the word’s prefix status. These are listed in **Table 2**. They comprised a measure of word frequency, three phonological variables, and a rating score representing how ‘prefixed’ each word ‘feels’ to respondents. See Stein (2023) for discussion of the role of these and other predictor variables.

Word frequency (*LogWordFreq*, shortened to *Frequency* except when confusable with spectral frequency) was measured as $\log_{10}(\text{word frequency} + 1)$ from the CELEX lexical database (which represents counts from the Cobuild corpus). The three phonological variables were: whether the second syllable in the word (immediately following the *dis-* or the *mis-*) bears primary lexical stress (*Stress*); whether the word begins with *mis* or *dis* (*MisDis*); and a two-level factor (*BoundaryType*) encoding the segmental phonological context in terms of the phonotactic syllabification possibilities of the first syllable's coda. *BoundaryType* distinguishes cases in which the fourth phoneme, that is, the phoneme following the [s], may potentially form part of the coda of the first syllable (*CodaPossible* e.g., *discover*, *distrust*), from those where it cannot, either because it is a vowel so cannot be in a coda (e.g., *misalign*), or because it is a consonant that cannot form a coda cluster with /s/ (*CodaImpossible* e.g., *misread*, *dislike*). *CodaPossible* forms therefore potentially allow a phonological analysis in terms of two-consonant (branching) codas. *CodaImpossible* forms do not—they restrict the coda of the first syllable to /s/. This division was chosen to reflect phonetic consequences of the presence or absence of a phonological syllable boundary due to phonotactic influences. More complete phonotactic classifications were rejected because they resulted in data sets that were strongly imbalanced with respect to important variables like stress, or were too small to allow reliable statistical analyses.

Measure	Description
PrefixStatus	A binary code for whether the word is prefixed or unprefixed, according to the criteria in Section 3.
LogWordFreq (abbr. Frequency)	$\log_{10}(\text{word frequency} + 1)$ in the CELEX lexical database.
Stress	Whether the syllable immediately following <i>dis/mis</i> bears primary lexical stress (early) or not (late). It is very strongly related to syllable count, so both cannot be used to predict the acoustic measures.
BoundaryType	A binary factor, distinguishing two syllabification possibilities. For <i>CodaPossible</i> forms, the 4th phoneme (that following the [s]) can potentially form part of the coda of the first syllable (e.g., <i>distaste</i>). For <i>CodaImpossible</i> forms, the 4th phoneme cannot form part of the coda of the first syllable, either because it is a vowel (e.g., <i>misalign</i>), or because it cannot follow [s] as part of a coda cluster (e.g., <i>mislay</i>).
MisDis	A binary factor distinguishing <i>mis</i> words and <i>dis</i> words.
Rating	Ratings of “felt” degree of prefixedness. See Section 4.1.

Table 2: Predictor variables.

There is considerable variation in the number of items per category examined: prefix status, stress, boundary type and MisDis (Tables S2–S4). This variation to some extent dictated the questions that could be examined statistically.

4.1. Subjective ratings of ‘prefixedness’

Rating is used as a gradient measure of degree of prefixedness. To address the issue of gradience of these word forms, objective measures can and have been previously used (e.g., Hay, 2007; Plag & Ben Hedia, 2018; Stein, 2023; Wurm, 1997), but examining gradience in this way is often complex and open to conflicting interpretations. We wanted a simple global measure that would reflect native speakers’ intuitive feelings about the extent to which each word was experienced as being prefixed or unprefixed. Accordingly, we used an online rating task and solicited the opinions of volunteers in NZ, Great Britain, USA, Canada and Australia via social media and a few professional mailing lists.

The questionnaire was run in Qualtrics (<https://www.qualtrics.com/uk/>). Participants rated how much the word seems to them to contain a prefix. Full instructions are given in Appendix B. To keep the rating task to about 20 minutes, 14 words were removed. These 14 were inflectional [s] variants, chosen for exclusion provided that the base word was represented in the ranking task, and that neither the base nor the inflected variant are ambiguous as to part of speech. Pilot work indicated that the addition of suffixes does affect ratings, but the above constraints on removing words in effect made minimal changes to the word’s meaning and maintained the same number of syllables for each word type. Excluded words were: *disabilities, disagreements, disappears, disappointments, discoveries, discussions, disruptions, disrupts, distinctions, distractions, distributors, misdemeanours, misfortunes, mispronunciations*. In later analyses, these words were assigned the mean rating given to their noninflected counterparts. This left 214 words, which are listed in Appendix C.

Each of the remaining 214 words was presented individually on the computer screen, centred above a numerical scale ranging in 11 steps of 10 from 0 on the left to 100 on the right. There was a horizontal slider below these numbers. Respondents used their mouse to slide a short vertical bar on the slider to the desired numerical value. Spatial configuration was not counterbalanced: 0 was always designated ‘Completely unprefixed’, and 100 always ‘Completely prefixed’. At the start of each trial, the slider was centered at 50, which was labelled ‘Neutral’ on the numerical scale. A response was required for each word. A horizontal bar near the bottom of the screen showed how much of the task had been completed. The task was self-paced: the participant clicked on a ‘Next’ button when ready to proceed to the next item. After rating all 214 words, respondents completed a short demographic questionnaire. This did not allow respondents to be identified, but did allow their responses to be sorted by regional accent of English and age.

After completing the questionnaire, the participant was informed on the screen that pressing the final ‘Submit’ button constituted consent for their data to be used (anonymously) for research purposes. Ethical approval was granted by the University of Canterbury, Christchurch, and by the Faculty of Music, University of Cambridge, UK.

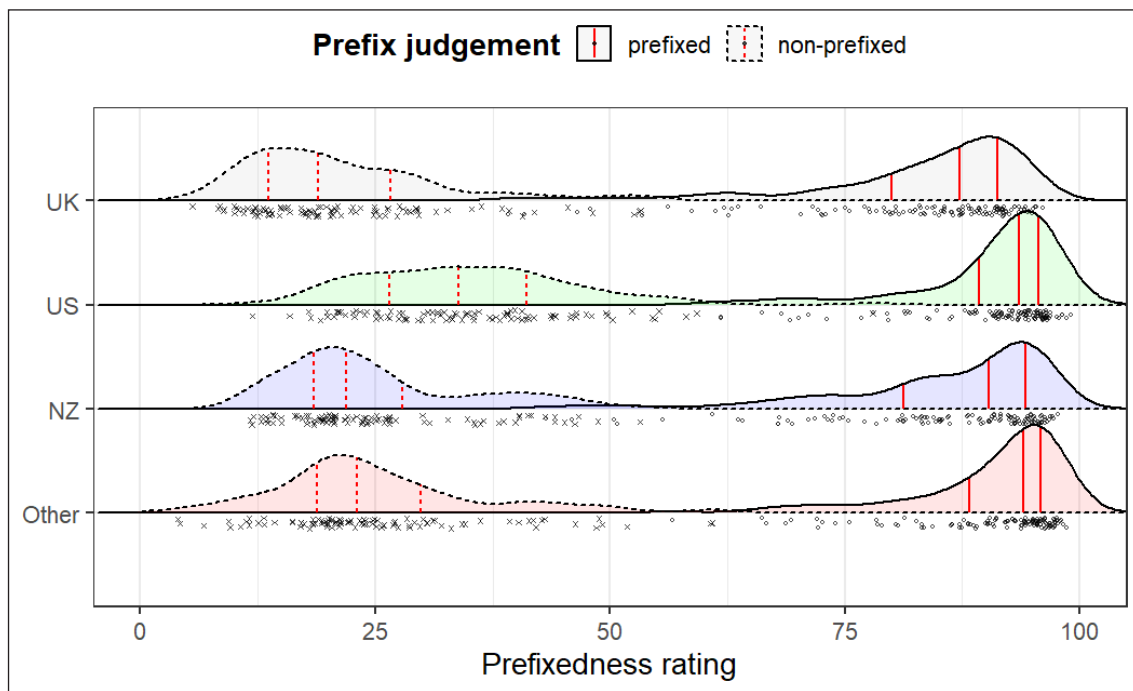


Figure 1: Density plots of mean word prefixedness Rating scores, by geographical group. Solid lines: ratings of prefixed forms. Dashed lines: unprefixed forms. Vertical lines under each curve show quartiles. Individual datapoints below x axes: circles = classed in Appendix C as prefixed; x = classed as unprefixed.

Ratings were analysed from the 78 respondents deemed suitable: native speakers of English, defined as having English as their main language since infancy (with regional dialect immaterial). Mean degree of ‘felt prefixedness’ was calculated for each word within each geographical subgroup. **Figure 1** shows density plots of the mean responses per word, with one panel and colour for each region. Responses to words classed as prefixed in Appendix C are shown with solid lines and to those classed as unprefixed are shown with dashed lines. For all regions, the distribution is roughly bimodal. There are many judgments close to 100% for words felt to be prefixed, whereas there seems to be less agreement about unprefixed words: there are relatively fewer judgments close to 0%, and the spread of the lower (more unprefixed) judgments is wider than that of the higher (prefixed) judgments. There also seems to be some regional variation in terms of how prefixed the words seem overall. In particular, the judgments from the United States skew more strongly toward prefixedness than those of the other regions, with a wide spread of ratings amongst unprefixed forms and a somewhat narrow spread amongst prefixed forms.

Despite the fact that the shape of the distribution of responses varies slightly across different regions, the rankings of the words were remarkably aligned. **Figure 2** shows these same data as bar charts (undifferentiated by formal prefixed/unprefixed classification) separately for each geographical subgroup, together with Pearson r and Spearman r_s correlations across words between geographical subgroups. The correlation coefficients between different geographical groups are all positive and highly significant, with r_s ranging between .9 and .94, indicating significant consistency between each geographical subgroup's mean ratings. This indicates that individuals are able to rank the words in terms of how prefixed they feel, and that the ranking is extremely consistent, even when compared across speakers of different dialects of English.

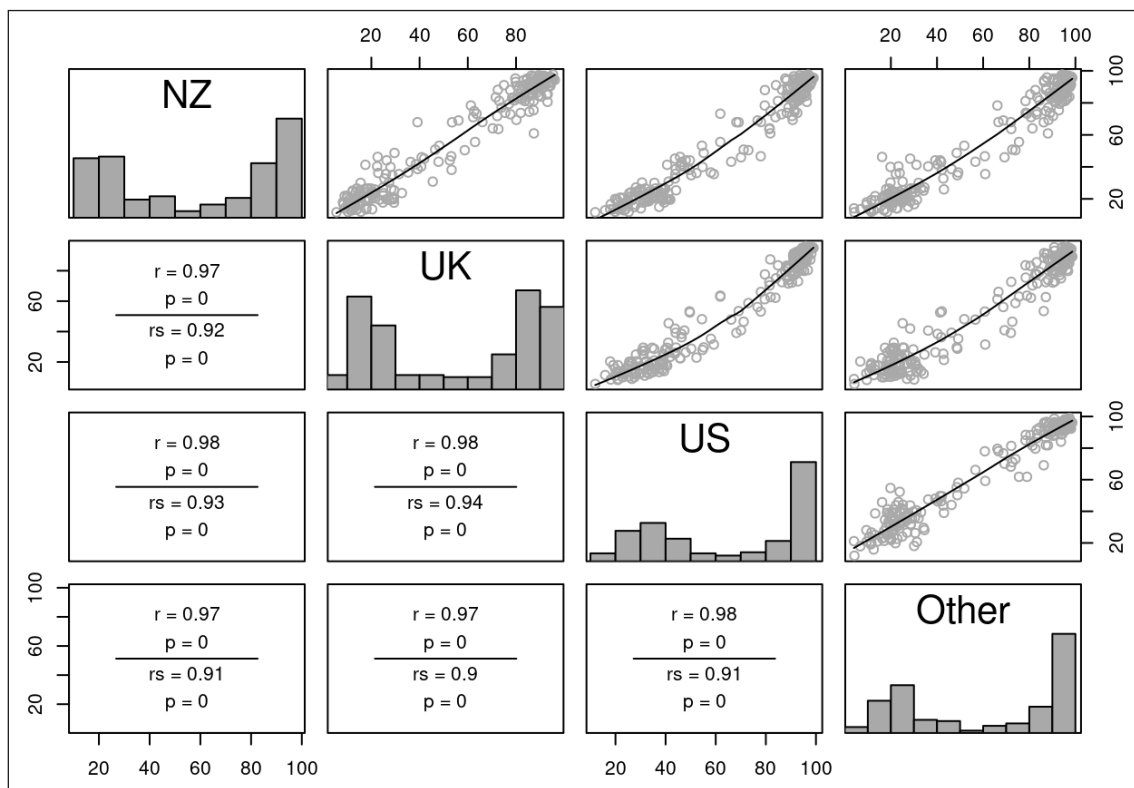


Figure 2: Barplots (diagonal): distributions of words' mean ratings of 'felt prefixedness', grouped by respondents' country of origin: 14 NZ respondents, 15 UK, 34 US, and 11 Other (comprising Canada (8), Australia (3), India (1)). Correlations between words' mean ratings according to country of respondent are shown by scatterplots (right of diagonal) with r and r_s coefficients and associated p values left of diagonal.

For the purposes of predicting acoustic realization, it is clear that there is no need to use regionally specific ratings, as the ratings are very highly correlated across regions. The prefixedness ratings were therefore pooled across geographical groups to yield a single rated value for each word. The mean rating for each word is shown in Appendix C, together with its initial classification as prefixed or unprefixed. The prefixed/unprefixed classification assigned

by our formal criteria corresponded almost perfectly with the mean ratings of felt prefixedness. Only one word that we formally classified as unprefixes, *misnomer*, gained a higher mean rating than words we classified as prefixed, and the overlap was small: the unprefixes *mistaken* (53%) and *misnomer* (65%) were separated by just five words we formally classed as prefixed, all of which are, like *misnomer*, semantically opaque to many native speakers: *dismantled*, *disconcerting*, *discharging*, *misdeemeanour*, *discharge*. All subsequent analyses of gradience use this single rating scale.

We used simple tests to assess the roles of five variables: four of the predictor variables defined above: Frequency, word Stress (whether the syllable immediately following *dis-mis* bears primary lexical stress), BoundaryType (CodaPossible/CodaImpossible, further subdivided due to large differences in cell numbers (see Table S3.4) into whether phoneme 4 was a Consonant or Vowel), and whether the word begins with *mis* or *dis* (MisDis). The fifth variable, word length, contrasted two versus three or more syllables. Length could be included because our analyses of influences on ratings do not require that our ratings are independent of one another, making its aforementioned interdependency with word Stress less important.

We first examined the distribution of these five properties *between* the two morphological categories. All differ significantly between prefixed and unprefixes word types. Prefixed forms are significantly less frequent than unprefixes forms (Unpaired Wilcoxon test, $W = 2801.5$, $p < .0001$). Prefixed forms are also significantly more likely to have the primary stress on the third syllable or later (51% vs. 11%, chi-square = 35.99, $p < .0001$), to be three syllables or longer (49% of prefixed forms vs. 23% of unprefixes forms, chi-square = 8.323, $p < .01$), and to have a vowel following the [s] (47% vs. 6%, chi-square = 38.619, $p < .0001$). Frequency, early/late stress and number of syllables are presumably correlated: the addition of a prefix often (though not inevitably) results in a less common word, and inevitably increases the syllable count and shifts the main lexical stress one syllable later. However, there seems to be no a priori reason for stems to begin with vowels more often than the second syllable of unprefixes forms.

We then investigated whether these five properties also influenced ratings *within* rather than between each morphological category. Word frequency approaches significance within prefixed forms (Spearman's $r_s = -.16$, $p = .09$), with lower frequency words likely to be rated as more prefixed. It is not a significant predictor of variation within unprefixes forms. Wilcoxon tests show that words with later lexical primary stress are more likely to be given higher ratings, within both prefixed forms ($p < .001$) and unprefixes forms ($p < .01$). Longer words are more likely to be given higher ratings within prefixed forms ($p < .001$), but not within unprefixes forms. And Consonant versus Vowel as the fourth phoneme is a significant predictor of ratings for prefixed forms ($p < .02$), with forms with a vowel following the [s] attracting higher ratings. Very few unprefixes forms have a vowel following the [s], so this factor cannot account for variation in the ratings among unprefixes forms.

Overall, a number of linguistic properties seem to be more characteristic of prefixed forms than of unprefixed forms. These properties also drive variation *within* these categories in terms of native speakers' judgments of how prefixed each word form feels. Within each category, words appear to feel maximally prefixed when they most resemble other prefixed forms phonetically and phonologically, and minimally prefixed when they most resemble unprefixed forms. That is, native speakers' judgments are at least somewhat gradient, in a way that reflects the phonological distributions in the language.

Of course, the factors analysed here are unlikely to be the sole ones driving this gradience. Semantic factors surely play an important role in affecting gradience, despite the fact that they are not formally analysed here. Semantic criteria were of course central to the formal binary classification, but there were nonetheless cases that were difficult to decide. The anomalous *misnomer* and *dismantled* discussed above are good examples of these, and their semantic indeterminacy is reflected in the subjective ratings. The mean word ratings derived from this task were a predictor variable—Ratings—in subsequent regression analyses.

5. Dependent Variables

The dependent variables are all acoustic. They broadly followed those of Smith et al. (2012): absolute and relative segmental durations, formant frequencies and their difference, and the spectral balance of [s]. For practical reasons, the present study introduced minor measurement differences for durations and formant frequencies. We also included a wider range of measures allowing a more nuanced assessment of the spectral balance of [s]. This section lists the parameters, together with their definitions.

Table 3 shows the 14 dependent acoustic measures, and briefly describes each. Measures relating to [s], which may be less familiar, are illustrated in **Figure 3** below. The following sections give reasons for these measurement decisions, and technical details of the methods.

Measure	Description
tri-dur	Duration of the full <i>mis</i> or <i>dis</i> triphone.
o-dur	Duration of the first syllable onset, [m] or [d].
i-dur	Duration of the vowel.
s-dur	Duration of [s].
s:i-ratio	Ratio of s-dur to i-dur.
F1	Frequency of F1 (Hz) at midpoint of vowel.

(Contd.)

Measure	Description
F2	Frequency of F2 (Hz) at midpoint of vowel.
F2-F1	Difference between F2 and F1 frequencies (greater in more peripheral /l/).
s-freqM	Frequency of the highest-amplitude peak in the mid-frequency spectrum (3–7 kHz). See Figure 3 . Probably closely correlated with the ‘lower-frequency cutoff’ described from spectrograms (the abrupt increase in amplitude at about 4 kHz). Compare also Chodroff and Wilson (2022).
s-freqH	Frequency of the peak amplitude in the high-frequency range (> 7 kHz). Recommended by C. Shadle (Pers. Comm, 2018). See Figure 3 .
s-freqMH	Overall peak frequency, indicating which of s-freqM and s-freqH has higher amplitude.
s-ampDiff	Peak amplitude in the mid-frequency band (3–7 kHz) minus the minimum amplitude in the low-frequency band (0.55–3 kHz): the vertical red line in Figure 3 . Designed to capture the degree of perceived sibilance, or [s] ‘goodness’. Always positive for [s], it quantifies the difference in amplitude of the low-frequency antiresonance and the lowest front-cavity resonance and is greater for smaller constriction areas. It can be thought of as mainly a resonance characteristic affected by degree of constriction.
s-levelDiff	Relative spectral balance in the mid- and high-frequency ranges, taking into account the whole spectral shape in each range. The difference between the sound levels (areas under curves) in the mid- and high-frequency bands, identified in Figure 3 as <i>areaM</i> and <i>areaH</i> respectively. Specifically, $areaM - areaH$, normalized by dividing each by its frequency range. This addresses nuances in high frequency energy. It can be thought of as mainly a noise-source characteristic with two causes. When the constriction area is small, greater air particle velocity increases turbulence noise in the vicinity of the constriction, especially at high frequencies; and when the jaw moves higher, high frequencies in the source spectrum may also be enhanced because the airstream may hit the teeth more effectively. The smaller s-levelDiff’s value (i.e., the more negative, or the closer to zero if positive), the more energy at frequencies > 7 kHz.
s-var	Variance of the multitaper [s] (the second spectral moment, or square of the standard deviation of the spectral spread around the spectral centre of gravity) above 550 Hz. Potentially affected by multiple interacting factors. E.g., when the spectrum has a high-amplitude, narrow-bandwidth mid-frequency prominence, s-var is normally low (most energy concentrated in the mid-frequency region). When energy is higher in high-frequency than in mid-frequency regions, s-var is likely to be large (due to likelihood of wide bandwidths in such spectral shapes).

Table 3: Acoustic Measures.

5.1. Durations

Durations were measured only for the first three phonemes, /mis/ or /dis/. Smith et al. (2012) also measured the closure duration and VOT of the voiceless stop that was the fourth phoneme in all their words. But since in our data the fourth phoneme encompassed a wide range of vowels and consonants, and the vast majority of the tokens do not contain a following stop, closure duration and VOT are less useful in this study. We refer to the joint duration of the first three phonemes as *triphone duration*.

Duration measurements extracted were guided by Smith et al. (2012)'s findings: the absolute duration of the [m], [d], [ɪ], [s] and of the three segments together (i.e., the whole triphone). From these, relative durations of [s]:[ɪ] were calculated.

Some segmentation from HTK's forced alignment is not optimal. To test whether this was likely to affect results, a small sample of the corpus was hand-corrected. The sample was 196 words in which the critical *dis-* or *mis-* is followed by a voiceless stop, for example *discarded*. Of these, 157 came from the NZE corpus, and 39 from the LIV corpora. (This sample was chosen because the original intention was to include measures of VOT in words where the fourth phoneme was a voiceless stop, and HTK segmentation does not distinguish stop closure from VOT. However, this plan was abandoned because there were too few such words in the corpus.) Hand correction reduced the variance in the data and brought mean durations between the first syllables of prefixed and unprefixed words slightly closer. However, in our initial investigations, hand correction made no difference to statistical results. Therefore, uncorrected HTK segmentation was used for all corpora except Buckeye.

5.2. Vowel formant frequencies

Smith et al. (2012) hand-corrected frequencies of the first three formants. For within-category comparisons of (unrounded) high front vowels, F3 frequency contributes little that cannot be inferred from the difference between F1 and F2 ($F2 - F1$), and automatic measures of F3 frequency tend to be unreliable. Consequently, it was decided not to try to measure F3.

Frequency measurements for F1 and F2 were made at the vowel midpoint, using Praat's formant tracker (version 6.0.28, 'To Formant (burg)...') with 5 formants estimated below maxima of 5 kHz for men and 5.5 kHz for women, a 50 ms Gaussian window, pre-emphasis of +6 dB/octave, and a time step of 2.5 ms. From these midpoint frequencies, the frequency separation, $F2 - F1$, was calculated, to give a single measure of vowel quality for /ɪ/.

5.3. Spectral properties of [s]

Standard spectral measures of [s] (e.g., the centre of gravity (first spectral moment), and/or frequency of a spectral peak), calculated across the whole spectral range, can characterize differences between places of articulation. But we sought spectral measures that would be able

to capture nuances within the spectral shape of [s] that might stem from stress- and/or duration-related production differences relevant to the prefix/nonprefix distinction. These could affect relative amplitudes and bandwidths in mid versus high frequencies, which could in turn affect perceived [s] loudness or quality.

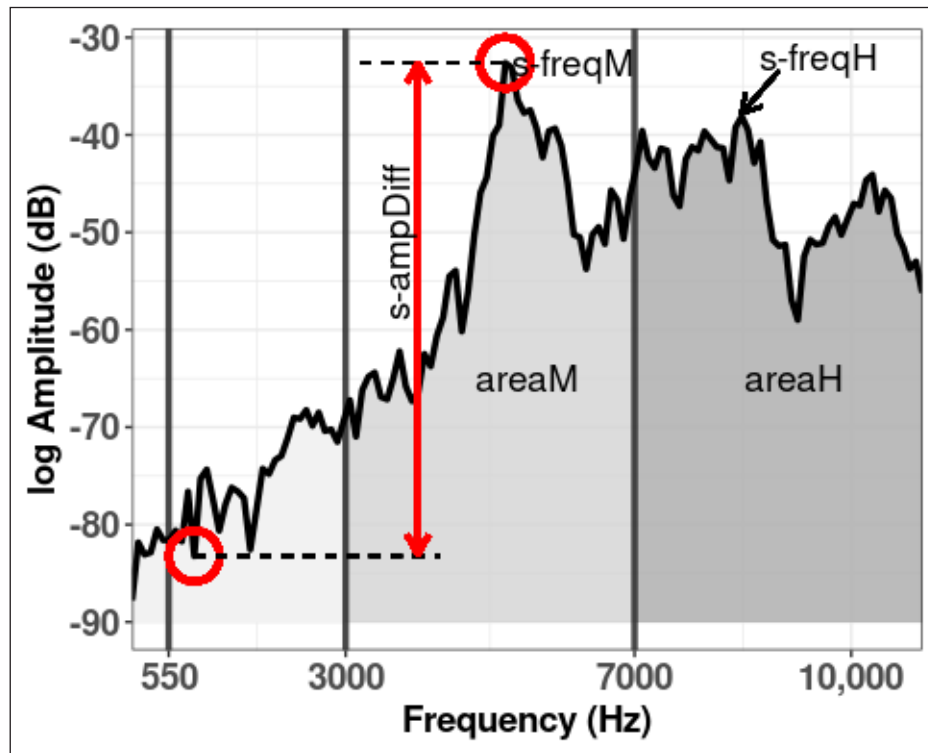


Figure 3: Long-term average spectrum (LTAS) of an [s] spoken by the first author, illustrating the measures used for [s] as listed in Table 3 and described in the text. The three shaded areas under the spectral envelope distinguish the frequency regions within which particular parameter values are measured: 0.55–3 kHz, 3–7 kHz, and above 7 kHz. The circled point in the lowest frequency range has the lowest amplitude in that range, and contributes to the measure *s-ampDiff* (vertical red line), whose upper limit is *s-freqM* in the mid-frequency range. See text for other measures, after Koenig et al. (2013). This and subsequent example LTAS spectra were calculated in wavesurfer over the mid portion of the [s], with a 256-point FFT, Hamming window, step size 128 points, 0.96 pre-emphasis.

Figure 3 illustrates Koenig’s (2013) measures, using the coding terms adopted in this paper (see also Table 3). These terms closely mirror those of Koenig but are typically shorter and without sub- or superscripts, to simplify graphs. The three frequency bands distinguished are shown with different shadings of grey under the spectral envelope. There were two measures of amplitude differences. One is the maximum amplitude difference between the two lowest frequency bands, *s-ampDiff*, applied to unique frequency-amplitude points shown by the vertical red line between

the points within the red circles. The other is the relative amplitude over broad frequency ranges in the mid- and high-frequency bands, *s-levelDiff*, calculated as $areaM - areaH$ in **Figure 3**, each normalized by being divided by the frequency range over which it was calculated. That is, all *areaM* and *areaH* means were divided by 4 kHz (3–7 kHz, and 7–11 kHz respectively) except for Buckeye, whose Nyquist frequency, at only 8 kHz, meant the mean of *areaH* was (effectively) divided by 1 kHz. This ‘normalization’ gave results in dB close to those reported by Koenig et al. (2013), and allowed us to explore very high-frequency spectral shapes when initial sampling rates allowed. Peak frequencies within the mid- and high-frequency ranges were also compared, *s-freqM* and *s-freqH*. The higher-amplitude of these two was designated *s-freqMH*.

To obtain these [s] spectral measures, sound files were first downsampled as necessary to 22 kHz (e.g., Quakebox, MAONZE, and Glasgow were all sampled at > 44 kHz). Buckeye, whose sampling rate was 16 kHz, was unchanged. Then, for each [s], multitaper spectra were calculated over a 25 ms Hamming window centred in the middle of the [s] segment, with a time bandwidth of 4.

Using an R implementation kindly adapted by Patrick Reidy from Reidy (2015), methods followed Koenig’s (2013) exactly except in three respects. First, as noted above, the mid- and high-frequency areas were normalized by dividing by the frequency range contributing to them. Second, following C. Shadle’s recommendation (pers. comm. 2018), we introduced two more parameters, *s-freqH* and *s-freqMH*, to compare peak frequencies above 7 kHz and in the mid-frequency range, 3–7 kHz. Lastly, while Koenig et al. (2013) measured spectra at the beginning, middle and end of their fricatives, we took spectra only from the middle of the [s]. This was for two reasons. The first reason was phonetic: Koenig et al. (2013) observed that the greatest differences due to articulatory differences tended to be in the middle of the fricative, which they interpreted as due to maximum constrictions being achieved at this point. Narrow constrictions for an articulation like [s] increase amplitudes of frication at high frequencies by increasing air particle velocity, which affects the shape of the source spectrum, and also by increasingly decoupling the front and back cavities, which reduces low-frequency spectral energy contributed by back-cavity resonances. The second reason was pragmatic: forced alignment of segments in our database tends to produce segmentation errors at the edges, which would be undesirable in these nuanced measures. A dynamic approach to measuring [s] might differentiate prefix status in *dis-mis-*, but might be better addressed in a more controlled study.

5.4. Overall Predictions

Our overall prediction is that all dialects will make a phonetic distinction between prefixes and phonemically identical nonprefixes, and be subject to the same nonacoustic influences, including any effects of frequency. The dialects differ in how they realize the phonemes involved, and need not necessarily use the same pattern of acoustic parameter values to mark prefixedness. Broadly,

however, we expect the internal acoustic structure of nonprefixes to reflect the pattern typical of more reduced syllables for each dialect, compared with that of prefixes. But since prefixes are also relatively weak syllables, how exactly this will manifest is uncertain because not enough is known about this type of variation amongst weak *dis* and *mis* syllables in each dialect. For two reasons, we did not make specific predictions about the effect of prefixedness on the acoustic dependent variables, either separately or once the PCA had identified dimensions of acoustic covariation in the dataset. First, there are no relevant data for all the new dimensions for [s]. More importantly, because we are dealing with multiple interacting parameters and a wide range of influences on articulation, there are usually arguments that support at least two conflicting patterns. Some of these influences are discussed when we explain observed patterns.

6. Statistical approach

The acoustic measurements were normalized by z-scoring within each (sub)corpus, separately for males and females. The normalized values were subjected to a Principal Components Analysis (PCA) to condense the data into a manageable number of model parameters that would adequately address the covariance between acoustic measures. Then, for each PC, we conducted mixed-effects linear regression analyses, using R's lme4 library (Bates et al., 2015), to determine whether degree of prefixedness was predictive of the Component structure, and whether this varied across corpora. See Supplementary Materials Section 6: Statistical approach, for more details of these standard methods.

As explained in Section 7, our principal components analysis leads us to interpret five PCs. For each, we fit four separate regression models, predicting the component value. Two models address all the data. The Binary model uses a binary PrefixStatus code. The other (Rating) uses the more gradient 'prefixedness' rating derived from the subjective rating task (Section 4.1). The other two models use the rating data to assess evidence for gradient effects within prefixed and unprefixed forms separately.

The data are too sparse to test for all possible interactions. For the binary and rating models fit to all the data, we test for theoretically motivated interactions, as follows. First, we test for an interaction between the prefixedness measure and LogWordFrequency, because distinguishing prefixedness from lexical frequency is one of the research questions. Since the effects of either prefixedness or frequency could vary across dialect, we further test for interactions between corpus and prefixedness, and corpus and frequency. We also test for an interaction between prefixedness and stress pattern, as the realization of a syllable adjacent to a stressed syllable may well differ from one that is not adjacent to a main stress, and this could differ across prefixed and unprefixed forms. Similarly, we test for an interaction between prefixedness and boundary type, as any phonetic differences that are due to differences in syllabification would not affect all BoundaryTypes equally.

In initial exploration we also considered information about the length of the word. A word's stress pattern and length in syllables are inherently correlated, making them problematic to assess in the same statistical model. In particular, two-syllable words in these data inevitably contain early stress in our classification. As these two predictors offer very similar information, we retained lexical stress in the model. It is important to keep in mind that this also controls for word length to a certain degree.

For the binary models, we start with the fully-specified mixed-effects regression model: $PC \sim \text{PrefixStatus} * \text{LogWordFrequency} + \text{PrefixStatus} * \text{Corpus} + \text{LogWordFrequency} * \text{Corpus} + \text{PrefixStatus} * \text{Stress} + \text{PrefixStatus} * \text{BoundaryType} + \text{MisDis} + (1|\text{Speaker})$. For gradient models, the structure is the same, but Rating (gradient prefixedness) replaces PrefixStatus. We then use a backwards selection process, pruning each of the two-way interactions in turn. If an ANOVA comparison between a model containing one of the two-way interactions and a model lacking it is significant ($p < .05$) the more complex model is retained. If the comparison is not significant, we continue to work backwards, until all the two-way interactions have been assessed. Models selected through this process are then inspected and, following Vittinghoff (2005), interactions leading to a variance inflation factor (vif) greater than 10 are also excluded. All main effects are retained regardless of significance, and so the simplest model we report contains main effects of PrefixStatus/Rating, Corpus, Frequency, Stress, BoundaryType, and MisDis, regardless of significance.

All models contain speaker as random intercept. Random slopes were explored early in the model fitting procedure, but tended to lead to convergence problems, so were abandoned. Our data are not well suited to fitting slopes because they are sparsely sampled across word forms and (particularly) speakers. The mean number of observations per speaker is 2.2, and the mean number of observations per word form is 5.3. Random intercepts for word were also explored. They tend to explain very little variance, and trigger singularity warnings, but do not change the overall results. They are omitted from the models reported here. That is, reported models are free of singularity and convergence issues.

As noted at the start of this section, to investigate whether gradient effects exist within the categories of prefix and nonprefix, we assessed effects of Rating separately within prefixed and unprefixed forms. We can interpret a significant effect of Rating *within* category as confirming that Rating has improved predictive power over and above the binary PrefixStatus code.

7. Overall results of Principal Components Analyses

We conducted a Principal Components Analysis on our 14 variables. **Figure 4** plots the variance explained by each principal component in the resulting analysis (in red). Our first five PCs collectively explain 74% of the variance (20%, 17%, 14%, 13%, 10%). To determine which to interpret, we followed Wilson Black et al. (2022): permute the data 1000 times and plot

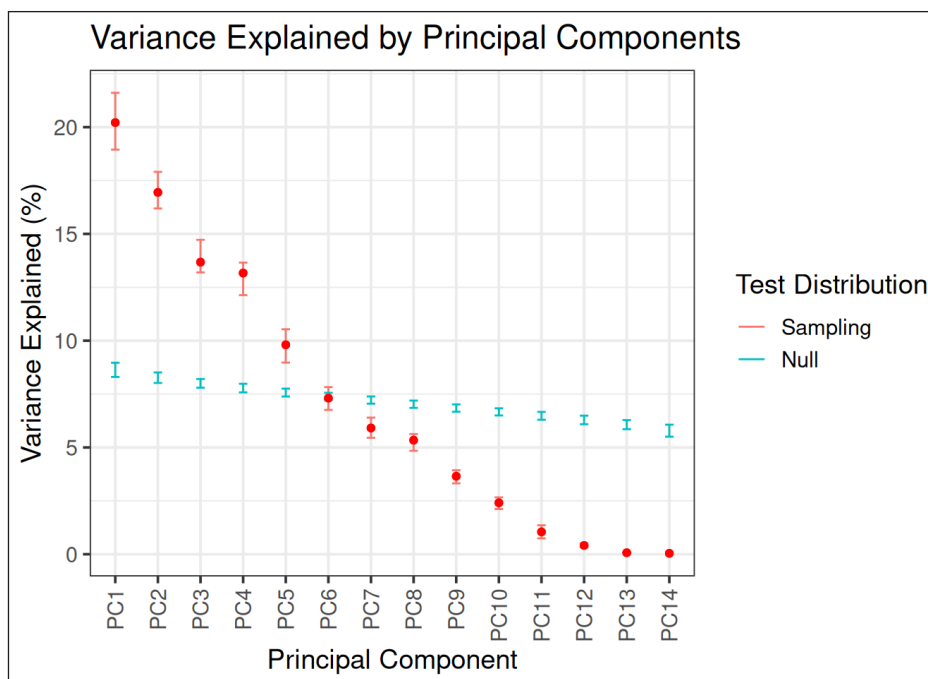


Figure 4: Variance explained by Principal Components. Observed loadings are shown as red 95% confidence intervals about the mean (circles). Null distribution is shown in blue.

the distribution of these random data (**Figure 4**, blue symbols); PC loadings lying above the random distribution are regarded as contributing significantly to the analysis (with due regard to confidence intervals—see below). This confirmed that PCs 1–5 merited analysis.

To interpret our PCs, we again followed Wilson Black et al. (2022, p. 9)’s procedure to “avoid looking at loadings which are random noise” and “to avoid loadings which are unstable” in permuting and bootstrapping to generate a 90% null distribution and 95% confidence bands for the index loadings for each of the PCs. An index loading “multiplies the variance explained by a PC and each of its loadings” (Wilson Black et al., 2022). **Figure 5** shows these distributions for each of the first five PCs.

Each plot has both a bootstrapped null distribution with a 90% confidence band and 95% confidence intervals for the actual observed loadings. The actual index loadings from the PCA appear as either a black plus or minus, indicating the sign of the loading. If the sign sits above the blue null distribution, we can say that that factor contributes significantly to the PC in this data set. The distributions around the plus or minus symbols show bootstrapped confidence intervals, indicating the stability of the loading in different subsamples of the dataset. If the confidence interval for an actual loading falls outside the corresponding interval for the null distribution, then the actual loading reliably falls outside the 90% confidence interval of the null distribution, and can be regarded as very reliable. However, Wilson Black et al. (2022) recommend that more

caution is used if the confidence interval falls inside the null distribution, and that “we should be very cautious about those whose confidence bands stretch down towards zero.” (Wilson Black et al., 2022 – Supplementary materials).¹ In sum, if the sign is above the null distribution, the index loading is significantly above chance in the full dataset. If the confidence interval is also above the null distribution, this significance is robust across subsamples of the dataset.

PCs 1–5 include loadings on acoustic variables that fall reliably above the null distribution. We thus have five major patterns in the data to work with. As noted in Supplementary Materials: Statistical Approach, we have unequal numbers of acoustic parameters for each of the three phonetic segments (six measures for [s], five for duration, three for vowel formant frequencies). Moreover, they are likely to covary for various reasons that are not necessarily related to prefixedness. Consequently, we cannot use the five PCs’ ordering to infer the degree to which these PCs might relate to prefixedness. That is, PC1 and PC5 may or may not be the most and least important, respectively.

The panels in **Figure 5** show the loadings of the 14 acoustic variables measured, for each of the five main PCs. Each panel is headed with the PC number, together with a shorthand phrase that describes a simple phonetic interpretation of the bundle of acoustic features that are most loaded on that PC. These terms are intended as broad-brush mnemonics. PC1: “[s] duration and peakiness”, PC2: “[s] spectral balance”, PC3: “vowel duration”, PC4: “F2-F1 spacing in vowel” and PC5: “[s] low-frequency cutoff.” The specific way in which the shorthand manifests at extreme positive and negative values of the component appears as an annotation on each panel, as described further below.

Variables whose covariation drives a particular PC have high absolute loadings on that PC. A strong positive loading of a particular variable (indicated with a +) indicates that positive values on that PC are associated with higher (more positive) values of that variable. Likewise, a strongly negative loading (indicated by a –) indicates that high absolute values on the PC are associated with smaller/lower values of that variable. So when two variables share extreme loadings in the same direction, they are highly positively correlated. When two variables each have extreme loadings but in opposite directions, they are negatively correlated.

Within each PC, actual signs returned by the PCA (as opposed to the relationships between the signs) are arbitrary, and carry no interpretation. To make the results easier to understand, and as noted in the legend to **Figure 5**, we somewhat preempt the results of the following sections, and manually impose a polarity on each PC. The PCs, as pictured in **Figure 5**, and as modelled in our regression models, always have the dimension of the PC most associated with prefixes oriented to be positive. This dimension is named, and its annotation highlighted in red, in each of the panels.

¹ The bootstrap procedure for PC3 and PC4 includes a filtering procedure, to accommodate their overlapping confidence intervals in degree of confidence explained. See Wilson Black et al. (2022) for details.

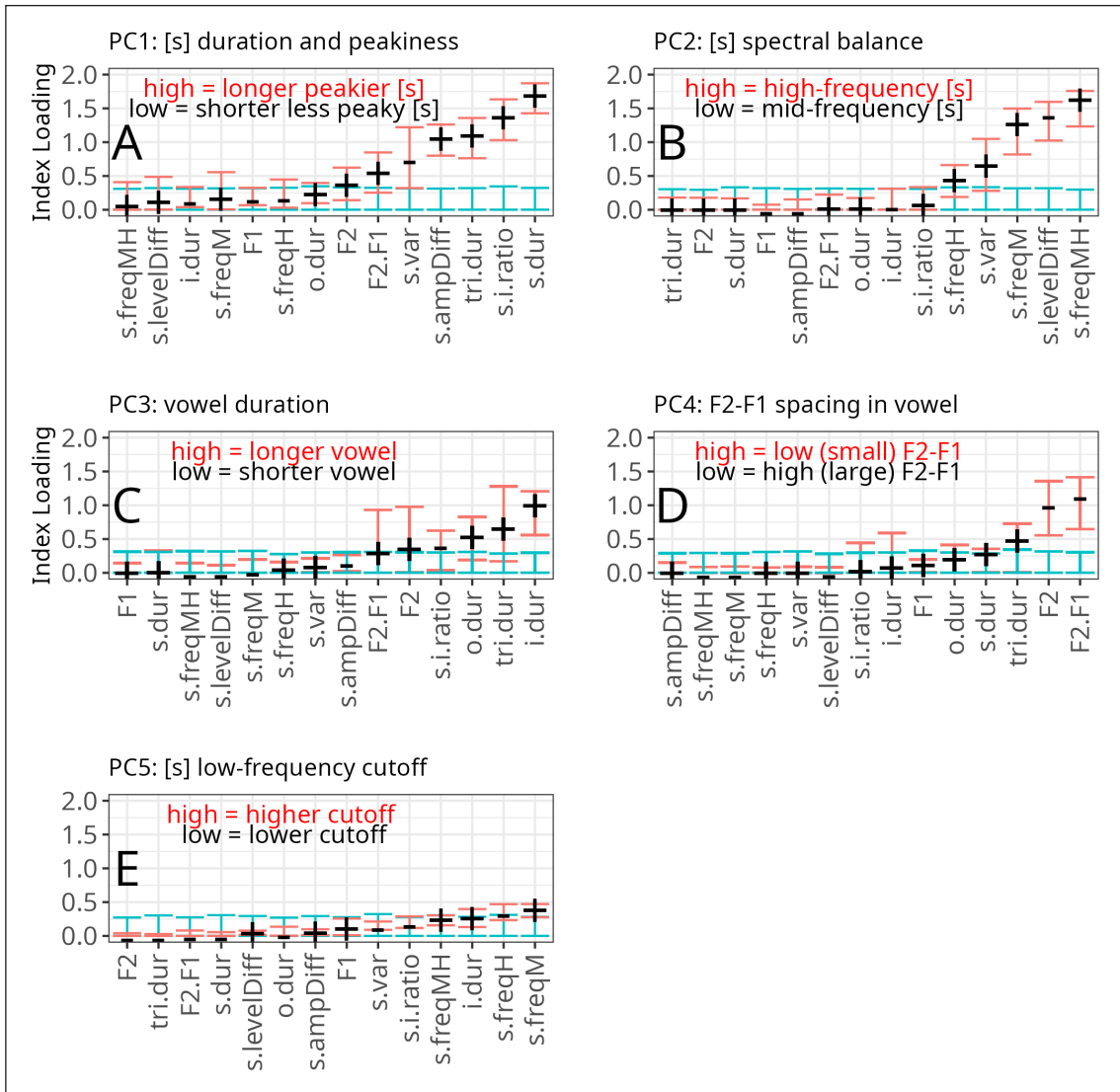


Figure 5: Principal Components Index loadings. Observed loadings are shown as red 95% confidence intervals about the mean, with black sign indicating the absolute mean loading and whether it is positive or negative. Null distribution is shown in blue. Each PC is labelled with phrases summarising the nature and direction of the pattern. The red label shows the direction most associated with prefixes in the analysis: for memorability, it is always shown as positive (see text).

For example, for PC1 the most strongly loaded variables are (from right to left) s-duration, s:i ratio, triphone duration and s-ampDiff. These all have positive loadings on PC1. This indicates that, in the results to be presented, the strongest trend with respect to prefixes is for them to have high values of PC1, that is, high values of the above-listed variables. Nonprefixes, on the other hand, are more likely to have low values of these variables.

In the following sections, we take each PC in turn. We describe the acoustic pattern that the PC captures. We then describe the results of regression models, which assess the relationship between that PC and our binary PrefixStatus and gradient Rating measures.

8. Results and interpretations of Components 1 to 5

8.1. PC1: “[s] duration and peakiness”

8.1.1. Acoustic interpretation of PC1

PC1 captures 20% of the variance. Our shorthand for the direction of PC1 is indicated by “longer peakier [s]” and “shorter less peaky [s]”. **Figure 5A** shows that the acoustic variables driving PC1 are related both to duration of [s] and to its spectral shape. Stronger PC1 loadings are associated with positive values for s-dur and s:i ratio, meaning that tokens with high positive loadings on PC1 have an [s] that is long both in absolute duration, and relative to the duration of the [i], compared with other triphones in the dataset.

S-ampDiff also has a positive loading on PC1, and s-var a negative loading, albeit with a large confidence interval, which very slightly overlaps the null distribution. This indicates that, a very small proportion of the time, a subsample of the full dataset does not result in a significant contribution of s-var to this PC. S-ampDiff captures the peak amplitude in the mid-frequency band (3–7 kHz) minus the minimum amplitude in the low-frequency band. The observed inverse relationship between s-ampDiff and s-var suggests that, when s-ampDiff is large, most of the spectral energy is typically concentrated in a comparatively narrow bandwidth (low variance), representing a relatively prominent spectral peak in the frequency range 3–7 kHz.

We refer to low (negative) PC1 [s], then, as being less peaky. High (positive) PC1 [s] is more peaky. Example spectra from our data are shown in **Figure 6**, with the same three frequency bands as those in **Figure 3**. The ‘peaky’ high-PC1 example in the upper panel has a notable mid-frequency peak that the low-PC1 example in the bottom panel lacks.

The raw data confirm the tight relationship between [s] duration and s-ampDiff: longer [s] typically has a mid-frequency spectral peak. Koenig et al. (2013) noted that this spectral peak tends to be more prominent when [s] is longer. In describing the acoustic conditions that produce it, they relate it to a smaller cross-sectional constriction area, especially at [s] midpoint. They noted it creates louder and more sibilant percepts, where sibilance meant “goodness” of [s]-like quality.

Finally, we note a significant, but less stable, contribution of F2-F1. In the full dataset, when the [s] is longer, there is greater separation between F2 and F1. In the bootstrap analysis, however, this is not reliably present across multiple subsamples of the data.

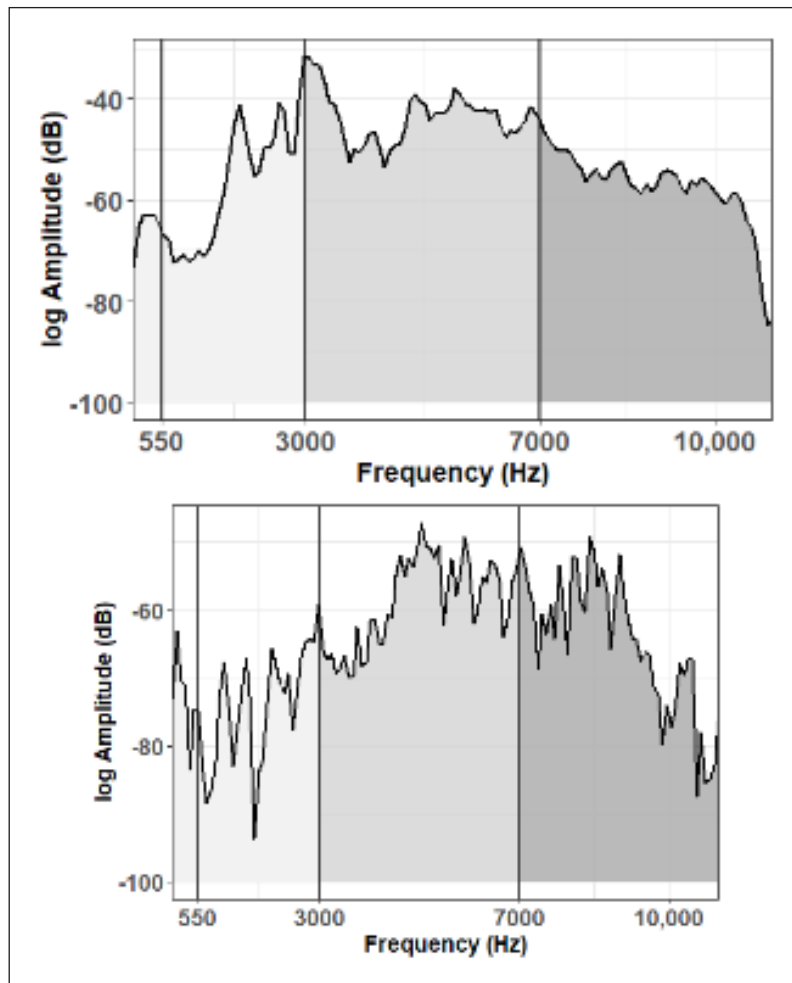


Figure 6: Example LTAS for high (positive) PC1 (top) and low (negative) PC1 (bottom). Neither example has extreme loadings for other PCs. The high-PC1 example (top) shows a prominent narrow-band peak in the 3–7 kHz range. The lower panel shows high-amplitude energy distributed more evenly over a wide frequency range. Top: “discharging”, Liverpool male. Bottom: “disadvantage”, New Zealand male.

In sum, high PC1 is associated with a longer peakier [s]. Low PC1 is associated with a shorter [s] that is less peaky. While triphone duration and the s:i ratio have high positive loadings on PC1, this is due to [s] duration. Neither vowel duration (i-dur) nor onset duration (o-dur) contribute significantly to PC1, but there is some association with vowel formants.

8.1.2. Relationship between PC1 and prefixedness

We fit a series of models to predict tokens’ PC1 values as a function of prefixedness, as outlined in Section 6. The selected regression model for the analysis of binary prefix status on PC1 is shown in Table S10.1.1. The coefficients are plotted in **Figure 7**. The phonological control variables of

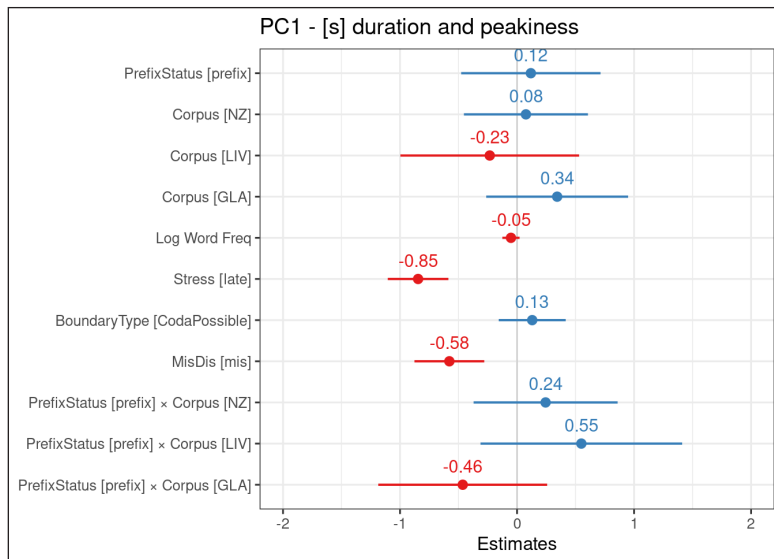


Figure 7: Coefficients and 95% confidence intervals (CIs) from binary model of PC1. Formula: $PC1 \sim Corpus * PrefixStatus + LogWordFreq + Stress + BoundaryType + MisDis + (1|Speaker)$. Significant predictors are Stress ($p < .0001$), MisDis ($p < .001$) and PrefixStatus * Corpus ($p < .04$). Numbers centered above the circles show each coefficient's value: positive in blue, negative in red.

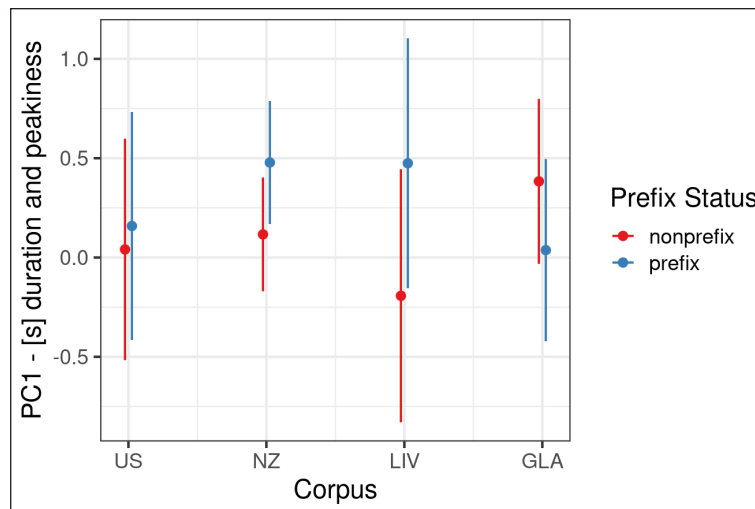


Figure 8: PC1: Model estimates and 95% CIs for binary PrefixStatus × Corpus. Higher values reflect longer, more peaky [s].

MisDis and Stress are highly predictive. The lowest PC1 (shortest, least peaky [s]), is associated with *mis* forms and forms where the base word contains a late stress. With these factors controlled, there is a significant PrefixStatus × Corpus interaction ($p < .04$), shown in **Figure 8**. Post hoc tests using estimated marginal means reveal that the difference between prefixes and nonprefixes

only reaches significance within the NZ Corpus ($p < .05$). Post hoc significance levels were not corrected for multiple comparisons for two reasons: we expected differences between dialects; and the increased risk of Type II errors with small sample sizes seems unlikely to materially improve the reliability of our conclusions (Barnett et al., 2022).

When prefixedness Rating was used as a predictor rather than the binary classification of PrefixStatus, it did not reach significance, either overall or when tested individually within prefixes and nonprefixes. Stress and MisDis remained strongly significant ($t = -6.488$ and -3.804 respectively), but no other main effects or interactions were significant (Tables S10.2.1–S10.2.3).

In sum, the overall pattern suggests that prefix status behaves in a weakly binary fashion for the parameters accounted for by PC1. Although Glasgow's [s] tends to be shorter in prefixes than in nonprefixes, and the other corpora trend the other way, the only statistically significant difference between prefixes and nonprefixes is for longer [s] in prefixes within the New Zealand corpus.

8.1.3. PC1: Discussion

PC1 is strongly related to [s] duration and spectral shape, but the corpora show only weak differentiation due to prefix status. The mid-frequency peak in the [s] spectrum is almost certainly due to the longer [s] durations allowing a narrower constriction to develop and be maintained. It could also result from greater pulmonic pressure increasing velocity through the constriction, which would be expected to be associated with greater stress (Stevens, 1998). So only the variation in [s] duration needs accounting for.

Three of the four dialects suggest a main effect in the opposite direction from that reported by Smith et al. (2012), who found a shorter [s] and lower s:i ratio, for SSBE prefixes.

The difference could be because SSBE is not represented in the current dataset, but seems at least as likely to be because the present data include many different word types in uncontrolled prosodic contexts, whereas Smith et al. (2012) used almost-minimal pairs in highly controlled prosodic contexts and a single speech style. So the present data seem likely to reflect “the general case” for three of our four varieties, taking into account a number of stronger influences as discussed below.

Why is variation in [s] duration so dominant in the data (as indicated by the PCA), yet only differentiates prefix status in the NZ corpus? One reason is the number of [s]-related variables with significant loadings on PC1. As Section 6 notes, a relatively large number of correlated dependent variables inevitably accounts for greater observed variance. Additionally, however, inspection of the raw data suggests that the variation may reflect well accepted influences on segment and syllable duration, most of which can be summarized under the umbrella-term ‘syllable reduction’.

We offer three reasons why syllable reduction may affect the [s] particularly strongly, despite literature showing that English vowels are generally more compressible than consonants (e.g., Klatt, 1976). First, all our triphones are weak, whereas the literature showing greater vowel compressibility typically examines stressed syllables. Weak syllables can undergo great ‘segmental reorganisation’ when severely reduced in natural speech, in English usually at the expense of the vowels. (Ernestus (2014) offers a brief review, and Rathcke & Smith (2015) a more comprehensive treatment in the context of rhythmic typologies.) Second, while the short, weak [ɪ] can be produced with just one or two glottal pulses or even none, there are limits on how long it can be while still being heard as /ɪ/ as in *fit*. If [ɪ] is prolonged relative to its context, most native English listeners will hear it as the phonologically long vowel /i/, as in SSBE *feet*. Thirdly, [s] is unlikely to sound like [s] if it is very short. It is usually notably longer than 50 ms (i.e., it cannot be shortened as much as [ɪ] can), yet it is easily lengthened. Thus, while it is possible to lengthen the triphone (especially its [s]), constraints on how short the triphone can be are largely dictated by how short its [s] can be. It follows, then, that [s] is likely to be the acoustically dominant segment of these unstressed or weakly stressed *dis-mis-* triphones, as well as the segment whose duration varies most under any condition that favours longer triphones.

In sum, variation in [s] duration may largely reflect degrees of syllable reduction. It is a major source of variation in these data because segment durations are susceptible to many influences, including lexical form, frequency, and rate and style of speech (e.g., Smith et al., 2012, sec. 4.2). Weak syllables are particularly variable in that their internal acoustic structure can undergo major reorganisation. Our regression analyses suggest that, at least in NZ, prefix status is also an influence on [s] durational variation, albeit a minor one.

8.2. PC2: “[s] spectral balance”

8.2.1. Acoustic interpretation of PC2

Figure 5B shows the loadings for PC2, which accounts for 17% of the variance. This is straightforwardly related to variation in the spectral balance of the [s]: whether there is relatively more energy in the high or mid frequencies. High PC2 relates to robustly negative values of *s-levelDiff*, and positive values of *s-freqM* and *s-freqMH*. There is a weaker but still significant positive effect of *s-var* in the full dataset, with confidence intervals that only slightly overlap the null distribution. This pattern reflects an [s] spectrum whose balance is tilted towards higher amplitudes in the high frequencies, with relatively broad spectral prominences. Conversely, low values on PC2 reflect an [s] with high-amplitude energy in the mid-frequency range, and relatively less above 7 kHz.

Figure 9 gives example spectra from a high PC2 (top) and a low PC2 (bottom) [s]. The high-PC2 example has more energy in the high frequencies, whereas the low-PC2 example has more energy in the mid frequencies. The relatively broad spectral prominence, especially in

the high-PC2 case, may simply reflect the fact that a formant's bandwidth typically widens significantly as its centre frequency increases. It may also reflect that s-var varies with PC2 and is negatively correlated with the strongly significant s-levelDiff. A high value of s-var and low s-levelDiff implies a relatively broad spectral prominence in the high frequencies.

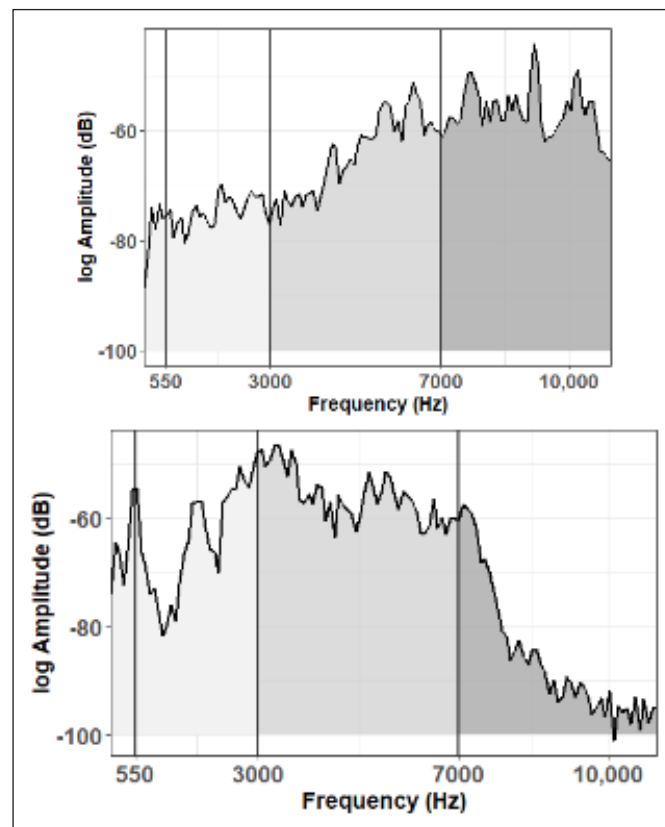


Figure 9: Example LTAS [s] spectra for high PC2 (top) and low PC2 (bottom). Neither example has extreme values for other PCs. The high-PC2 spectrum shows more energy in the high frequencies. The low-PC2 spectrum shows more mid-frequency energy. Top: “disappointing”, New Zealand female. Bottom: “disappear”, vernacular Glaswegian male.

8.2.2. Relationship between PC2 and prefixedness

The selected regression model of binary prefixedness is shown in Table S11.1.1 and the coefficients are plotted in **Figure 10**. BoundaryType affects [s] spectral balance, with the spectrum skewed towards low frequencies (low PC2) for CodaPossible forms ($p < .01$). Figure S11.1.1 shows that PrefixStatus \times Stress ($p < .05$) is due solely to unprefixed forms with late stress having lower PC2 than the other three conditions. That is, the nonprefix [s] spectrum skews away from the high frequencies, but only when the second syllable in the word is weak—when a word has more than two syllables and its main lexical stress falls on syllable 3 or later

(unprefixed *disappointing*, *distillation*). When the second syllable carries the main lexical stress, then prefixed and unprefixed forms have essentially identical, slightly positive, loadings on PC2 (prefixed *discomfort*, unprefixed *distinguished*). This is confirmed by post hoc tests with estimated marginal means: with late Stress, prefixes have higher values of PC2 than nonprefixes (more spectral energy in the high frequencies [$p < .026$]) whereas prefixes and nonprefixes with early Stress have similarly high values of PC2 ($p = .83$); within nonprefixes, late Stress lowers PC2 values ($p < .007$) whereas prefixes are not affected by Stress ($p > .76$).

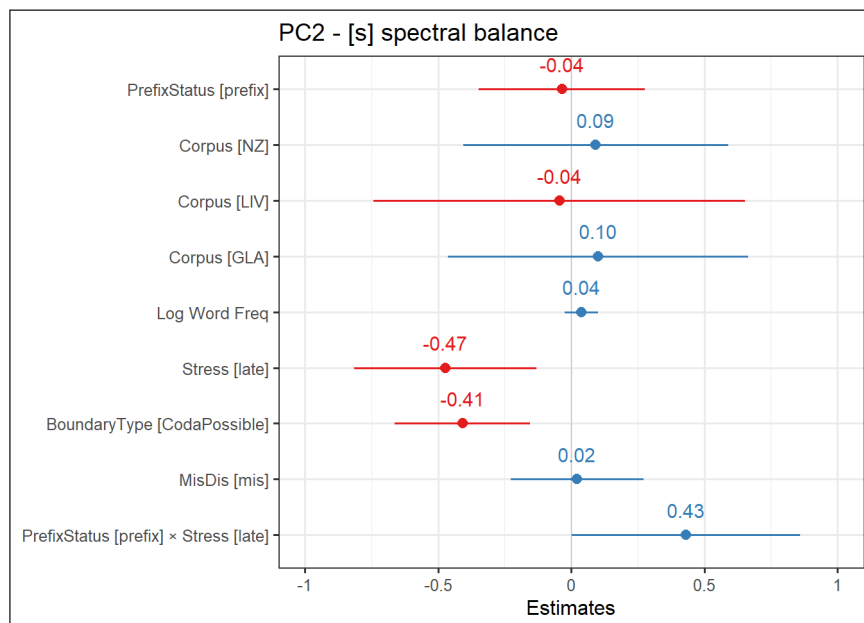


Figure 10: Coefficients and 95% CIs from binary model of PC2. Formula: $PC2 \sim Corpus + LogWordFreq + PrefixStatus * Stress + BoundaryType + MisDis + (1|Speaker)$. Significant predictors are *BoundaryType* ($p < .01$) and *PrefixStatus* × *Stress* ($p < .05$). Numbers centered above the circles show each coefficient's value: positive in blue, negative in red.

When prefixedness is viewed as a continuum, *Rating* does not reach significance either as a main effect or in interaction with *Stress*. This is true when tested over the whole dataset, and also when tested separately within prefixes and nonprefixes (Tables S11.2.1–S11.2.3).

8.2.3. PC2: Discussion

PC2, then, affects prefix status only insofar as nonprefixes that are not adjacent to the main lexical stress have lower-frequency spectral energy. What could cause a change in spectral tilt towards mid frequencies, that is consistent with the articulation of both unprefixed, late-stress words and words whose fourth phoneme is a voiceless stop? These patterns may reflect

differences in constriction area affecting air flow, which presumably would be subject to other articulatory influences too. Specific candidate influences are cross-sectional area at the major oral constriction, pulmonic airflow, and the degree of tapering of the tongue immediately behind the oral constriction. Supplementary Materials Section 11.3 develops these arguments.

8.3. PC3: “vowel duration”

8.3.1. Acoustic interpretation of PC3

PC3 captures 14% of the variation. **Figure 5C** shows that when PC3 is high, loadings are strongly positive on vowel duration (i-dur). High values of PC3 represent a longer vowel. Related duration measures are less robustly loaded on PC3, with confidence intervals that overlap the null distribution. Effects for triphone and onset durations are significant, but absent in some proportion of the bootstrapped samples. The effect of the s:i ratio is significant, but with its lower confidence interval approaching zero—a case in which Wilson Black suggests caution. It is clear that PC3 captures a robust effect of vowel duration. Vowel duration is itself somewhat correlated with duration effects elsewhere in the word, but those effects are not necessarily in the same place across all words in the sample.

8.3.2. Relationship between PC3 and prefixedness

The binary regression model of PC3 is shown in Table S12.1.1. The coefficients are plotted in **Figure 11**. Consistent with a wide range of previous research, PC3 is significantly predicted by word Frequency, with higher frequency words having shorter vowels. There is a more strongly significant effect of PrefixStatus \times Corpus, shown in **Figure 12**. Post hoc testing with estimated marginal means reveals that prefixed forms have higher PC3 (longer vowels) than unprefixed forms for US ($p < .0005$), NZ ($p < .05$) and Liverpool ($p < .01$), but not Glasgow ($p = .65$). It also reveals that the difference between the corpora is carried by differences within the prefixed set. While there are no differences in PC3 across corpora for nonprefixes, within prefixes the US has a significantly higher PC3 than both NZ ($p < .005$) and Glasgow ($p < .01$). **Figure 13** shows the PrefixStatus \times BoundaryType interaction: prefixed forms have higher PC3 (longer vowels) for both BoundaryTypes. The difference is more robust for CodaPossible cases, when the fourth phoneme is /p t k/. (Posthoc estimated marginal means tests: CodaPossible $p < .001$, CodaImpossible $p < .01$).

The ratings model also contains a significant Rating \times Corpus interaction (Table S12.2.1, Figure S12.2.1). The pattern is much the same as in the binary model. However when tested separately within the prefixes and nonprefixes, Rating does not reach significance within either subset (Tables S12.2.2–S12.2.3). There is thus no evidence that there is a separate gradient effect that explains variation within either the prefixed or the unprefixed words.

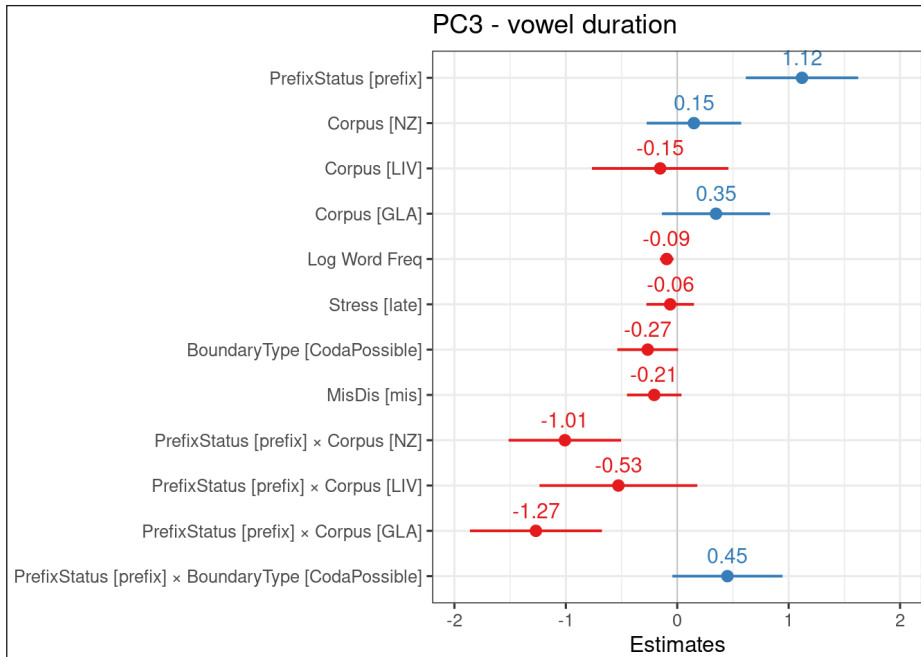


Figure 11: Coefficients and 95% CIs from binary model of PC3. Formula: $PC3 \sim Corpus * PrefixStatus + LogWordFreq + Stress + PrefixStatus * BoundaryType + MisDis + (1|Speaker)$. Significant predictors are LogWordFreq ($p < .01$), PrefixStatus * BoundaryType ($p < .05$) and PrefixStatus * Corpus ($p < .001$). Numbers centered above the circles show each coefficient's value: positive in blue, negative in red.

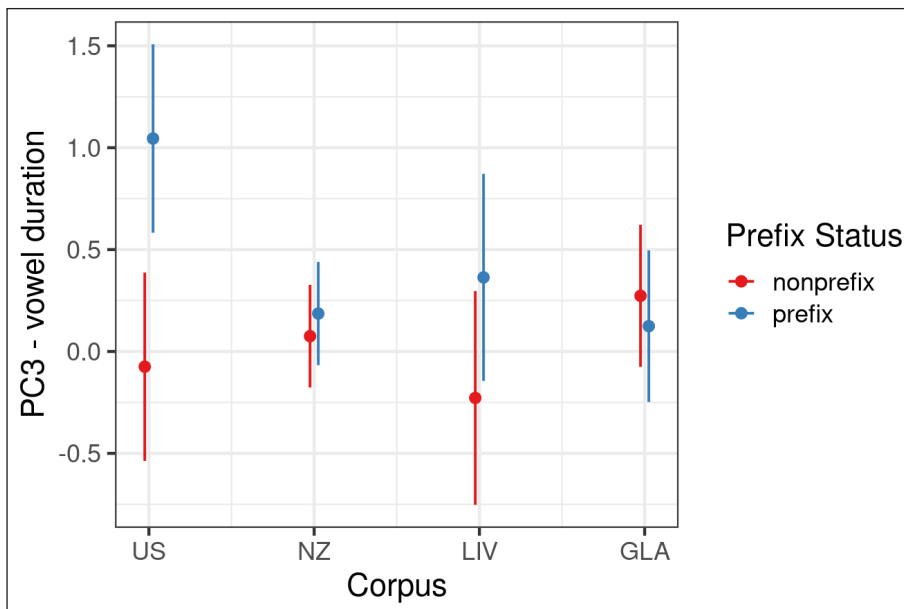


Figure 12: PC3: Model estimates and 95% CIs for binary PrefixStatus × Corpus interaction. Color: prefix status. Higher values reflect a longer vowel.

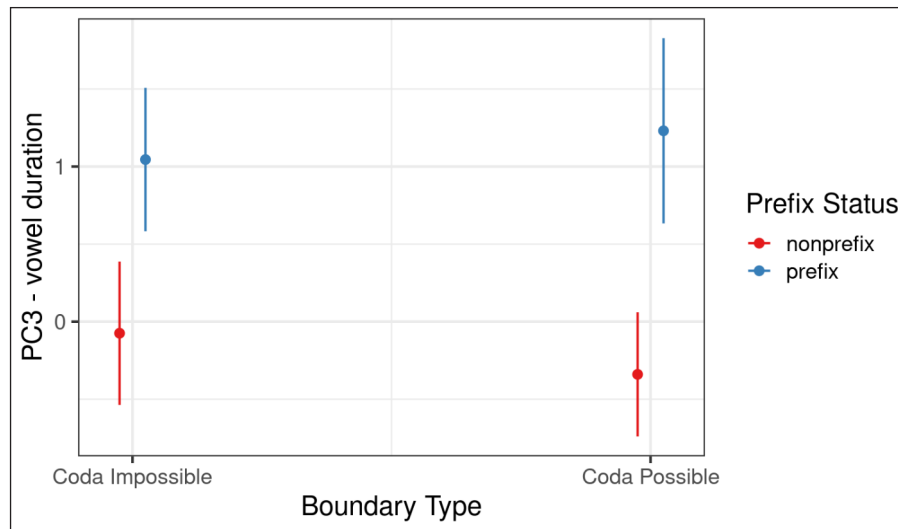


Figure 13: PC3: Model estimates and 95% CIs for binary PrefixStatus \times BoundaryType. Left: CodaImpossible. Right: CodaPossible. Color: prefix status. Higher values reflect a longer vowel.

8.3.3. PC3: Discussion

PC3 mainly reflects vowel duration. Although there is the expected positive covariation of vowel duration with triphone duration, onset duration, F2 frequency and F2-F1 spacing, none of these other measures are robust across the bootstrapped samples.

The vowel in prefixes is longer in all dialects except Glasgow (**Figure 12**). Smith et al. (2012) also found a small (6 ms) but reliable difference favouring longer prefixed vowels in SSBE. Across dialects, the prefix-nonprefix difference is more robust in CodaPossible forms (**Figure 13**). In these, the fourth phoneme is a stop, so could be classed as part of the coda of the first syllable, as well as the onset of the second. Vowel duration appears much more affected than [s] duration in PC1. The implication is that (absolute) vowel duration is a strong indicator of prefixedness.

8.4. PC4: “F2-F1 spacing in the vowel”

8.4.1. Acoustic interpretation of PC4

PC4 captures 13% of the variation. The loadings for PC4 are shown in **Figure 5D**. High values of PC4 are strongly associated with smaller F2-F1 spacing, caused mainly by lower F2 frequency. This suggests that high PC4 values reflect a more centralized and/or open vowel. This could result from various combinations of lower jaw, wider lip aperture, and less fronted tongue constriction. All other measures have null distributions extending toward zero.

8.4.2. Observed relationship between PC4 and prefixedness

The binary regression model of PC4 is shown in Table S13.1.1, and the coefficients are plotted in **Figure 14**.

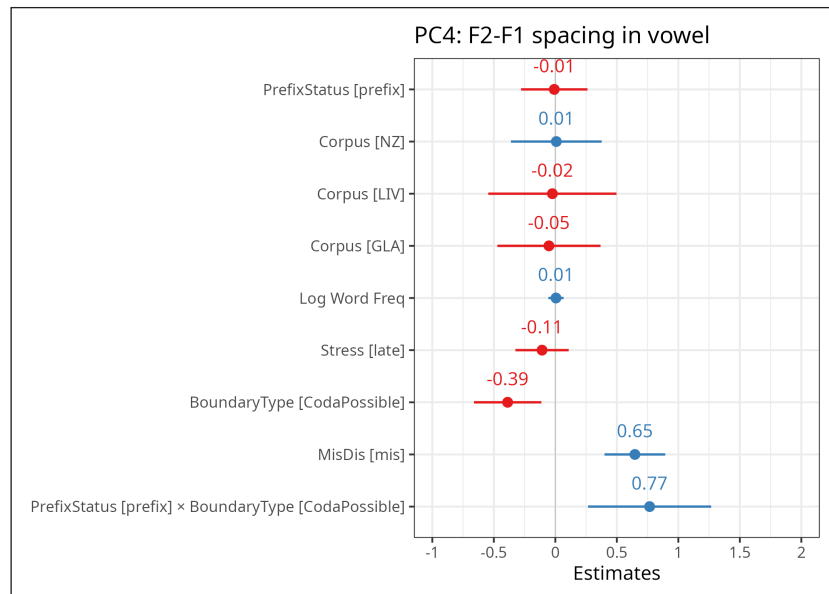


Figure 14: Coefficients and 95% CIs from binary model of PC4. Formula: $PC4 \sim Corpus + PrefixStatus + LogWordFreq + Stress + PrefixStatus * BoundaryType + MisDis + (1|Speaker)$. Significant predictors are *BoundaryType* ($p < .01$), *PrefixStatus * BoundaryType* ($p < .01$) and *MisDis* ($p < .001$). Numbers centered above the circles show each coefficient's value: positive in blue, negative in red.

Words beginning with *mis-* have significantly higher PC4 values. This is presumably due to different coarticulatory constraints imposed by the /m/ or /d/ onset, which could affect formant spacing in the vowel. However, the prefix effect is not an artefact of the strong *mis-dis-* effect: There are no substantive differences in parameter values between models with *mis-* included versus excluded. When *mis-* is excluded, the three estimated parameters of greatest interest are virtually identical to those in Table S13.1.1. (*PrefixStatus*: 0.025, $SE = 0.137$, $t = 0.180$. *BoundaryType*: 0.384, $SE = 0.139$, $t = -2.765$. *PrefixStatus × BoundaryType*: 0.774, $SE = 0.265$, $t = 2.927$.)

PrefixStatus interacts significantly with *BoundaryType* (Figure 15). Post hoc estimated marginal means tests confirm that PC4 differentiates *PrefixStatus* in *CodaPossible* words ($p < .002$), but not in *CodaImpossible* words ($p = .89$) which form a majority of tokens within the prefixed category (345 vs. 43, Table S3.3). In other words, the formant frequencies suggest that, for triphones followed by /p t k/, the vowel is more open and/or centralized in prefixes, and more close or fronted in nonprefixes.

To check whether it is specifically a following phonologically voiceless stop, rather than a following consonant of any sort, that creates the prefix effect on vowel formant spacing, we tested models that distinguished all vowels from all consonants following the triphone. These showed no difference between boundary type, confirming the current analysis of PC4: prefixes and nonprefixes were distinguished only when phoneme 4 is a stop consonant.

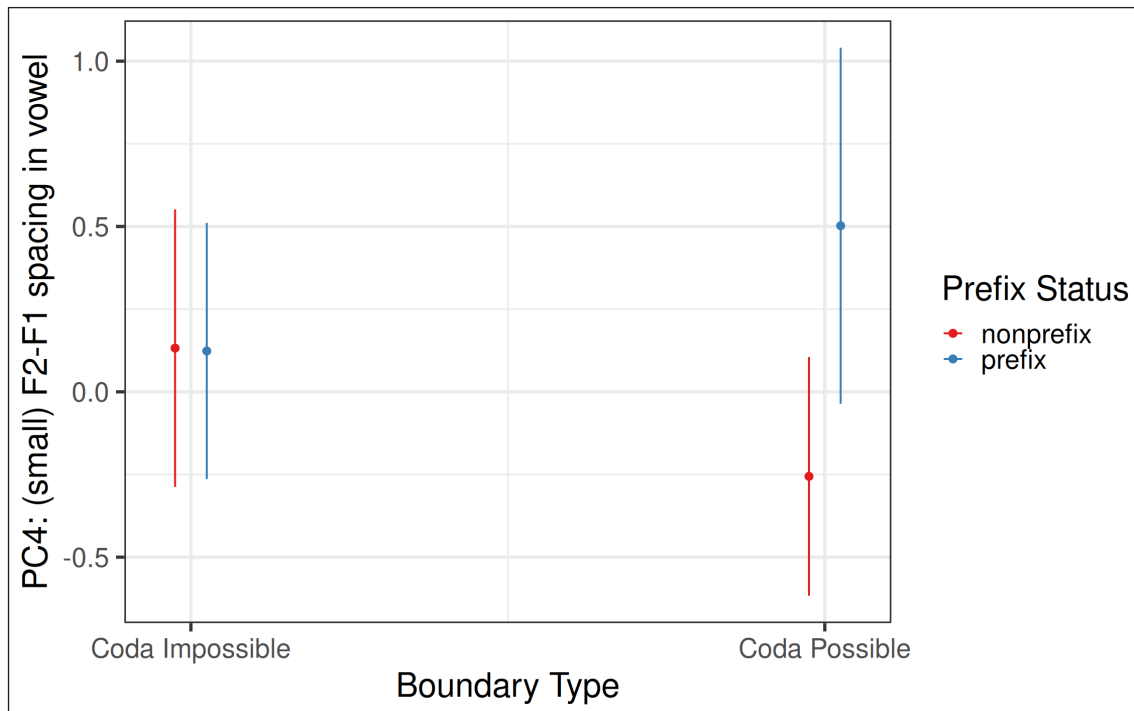


Figure 15: PC4: Model estimates and 95% CIs for binary PrefixStatus \times BoundaryType interaction. Left: CodaImpossible. Right: CodaPossible. Color: prefix status. Higher values reflect smaller F2-F1 frequency spacing.

The Ratings model shows no overall significant gradient effect of prefixedness on PC4 (Table S13.2.1). However, when prefixes and nonprefixes are tested separately, Rating is significant within the prefixed forms, but not within unprefixed forms (Tables S13.2.2.–S13.2.3). **Figure 16** shows that prefixed forms that are rated as more prefixed have higher PC4 (Rating estimate = 0.0267, $SE = 0.008$, $t = 3.3$). There is no significant interaction with BoundaryType. However, most prefixed forms are CodaImpossible (345) as opposed to CodaPossible (43).

8.4.3. PC4: Discussion

PC4 captures a pattern in which prefixes have less separation between F2 and F1. How do we reconcile this with PC1, in which there is a significant loading of F2-F1, and which seems to capture a pattern in which prefixes have more separation between F2 and F1? This is the only variable which is significantly loaded onto different PCs in opposite directions, so requires a little unpacking.

Importantly, while PC4 captures a pattern that exists within F2-F1, this does not mean that it is identical to raw F2-F1. (Pearson's correlation between PC4 and the z-scored F2-F1 is $-.81$, $p < .0001$.) Rather, PC4 captures that aspect of F2-F1 separation that is not captured by other components. PC1 reflects that longer /s/ (and longer segments in general) are associated with

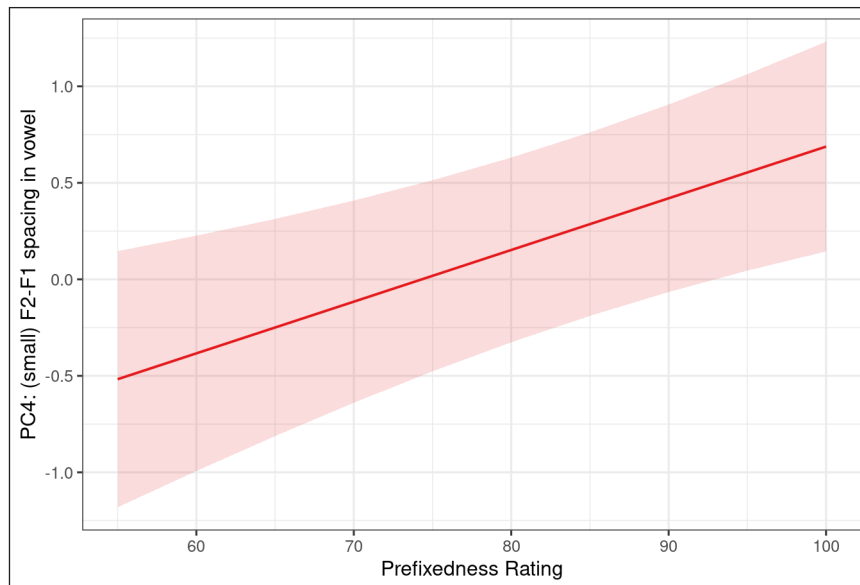


Figure 16: PC4: Significant effect of Rating within prefixes only.

more F2-F1 separation. This is as we might expect through phonetic peripheralisation. And PC1 is associated with prefixes, at least for NZ. PC4, on the other hand, captures the fact that there is substantial extra variation in formant separation *that is not correlated with or explained by differences in [s] duration*. To solidify the intuition that prefixes can have both *more formant separation than nonprefixes*, and *less formant separation than expected given the duration patterns*, **Figure 17** shows the z-scored formant values, for the three tertiles of [s]-duration values. (Note the x-axis shows raw F2-F1 [z-scored Hz], not PC4, which is negatively associated with F2-F1).

A number of things are apparent. First, distributions in the top panel sit to the right of the others. This reflects the effect of PC1: tokens with a long [s] tend to have more formant separation. Second, prefixes are much more likely to be associated with the long [s] distribution (41% of prefixes) than the short distribution (17% of prefixes). This means that, across all data, prefixes have more formant separation. Third, within long-duration tokens, prefixes have more formant separation than nonprefixes. They are also likely to be longer—though **Figure 17** does not show that because duration is treated as categorical.

However, in PC4, the association between duration and formant separation is controlled (by PC1). PC4 therefore represents *unexpected formant separation* that is not associated with duration increases. Our regression model tells us that this is associated with nonprefixes. Indeed, **Figure 17** shows that tokens with a very short [s] (bottom panel) clearly differentiate prefixes from nonprefixes: most prefixes have less formant separation than most nonprefixes, though a few prefixes have greater F2-F1 separation. For mid-duration tokens, prefixes have a broader range of formant separation than nonprefixes—some with more separated formants, others with less.

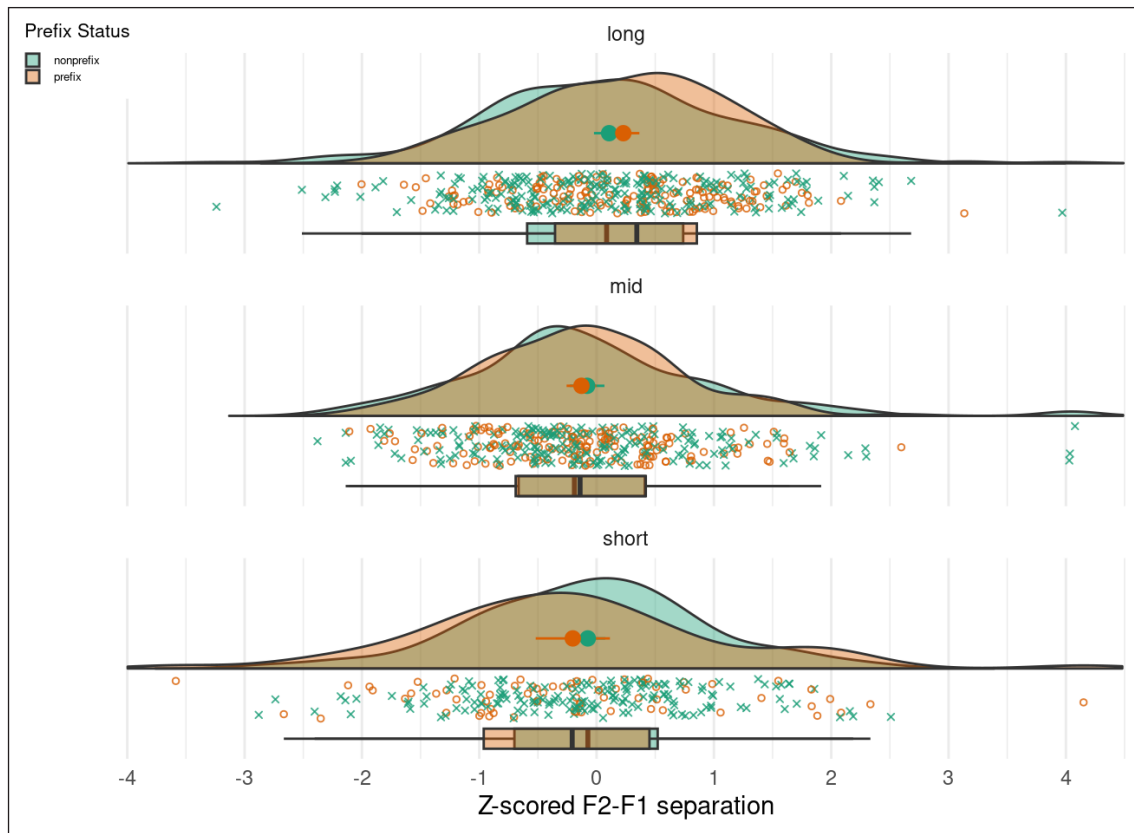


Figure 17: Z-scored formant separation (x-axis), for different tertiles of z-scored [s] duration: long (top), mid, and short (bottom) durations. These are z-scored raw values ($\text{Hz} \times \text{ms}$), not principal components.

This seems to reveal a second pattern of production for prefixed words. Most often, they are associated with a stronger ‘beat’. But they also show less formant separation than we would expect given the degree of ‘beatness’, and this is reflected in the association of triphone duration with PC4 (**Figure 5**), longer triphones being associated with reduced formant separation.

Our binary model tells us that the pattern of reduced formant separation in prefixes is prevalent in CodaPossible prefixed words (where phoneme 4 is a stop e.g., *mislacement*, *mischaracterization*, *discontented*). Moreover, words rated ‘more prefixed’ are more likely to be CodaPossible. The *mis-dis-* in these words may be most easily recognized as a unit—semantically, and inasmuch as the vowel may differ maximally between prefix and nonprefix due to phoneme 4 being syllabified with the triphone in nonprefixed words but not in prefixed ones.

What, then, would that syllabification difference imply for realisation of the vowel? More canonical vowels are expected in prefixed words. What counts as more canonical is complex in this dataset of multiple dialects. Our prefixed CodaPossible tokens come mainly from NZ (65%) and Glasgow (14%) speakers, whose canonical KIT vowel we know (from the literature and our

own impressionistic listening) approximates schwa. Wells (1982, vols 1 & 3) states that NZ [ɪ] is lowered and centralised and does not functionally contrast with schwa, while Glasgow /ɪ/ can be as open as [ʌ] or [ɛ]. Since only direction of change is important to the current argument, we term all these sounds *schwa*. When canonical KIT vowels are realised as close to [ə], then, effectively destressing the *dis-mis-* syllables may shift their phonetic quality closer to [ɪ]—i.e., less-stressed triphones will have greater F2-F1 formant spacing in NZ and Glasgow (Figure S13.3.1). The five CodaPossible prefixed forms in our US data are consistent with a trend in the opposite direction, as would be expected since Ohio KIT approximates IPA [ɪ] and its high vowels are typically centralised when unstressed.

In short, although small numbers mean that this conclusion can only be tentative, the observed patterns in PC4 are consistent with an interpretation in which, for all dialects, the change in vowel quality from prefix to nonprefix (within CodaPossible) represents a relative “relaxation”, or less extreme articulation: less centralized than expected for NZ and Glasgow, and trending towards more centralized than expected for Ohio.

What the PC4 pattern highlights is that, while prefixes are phonetically distinct from nonprefixes, it does not follow that there is only one way to be distinct. There are clearly multiple forces operating together to create the observed patterns. Supplementary Materials Sections 13.3–4 give details.

8.5. PC5: “[s] low-frequency cutoff”

8.5.1. Acoustic interpretation of PC5

PC5 captures 10% of the variation. The only significant loading, using Wilson Black’s criteria, is *s-freqM* (Figure 5E), the frequency of the highest peak in the 3–7 kHz spectral range (Table 3 and Figure 3). Following Koenig et al. (2013), this is interpreted as the frequency of the lowest front-cavity resonance during the [s], seen as closely correlated with the ‘low-frequency cutoff’ associated with [s] in spectrograms—the frequency above which high-amplitude energy dominates the spectral envelope, typically around 4 kHz for adults’ [s]. We use ‘low-frequency cutoff’ for shorthand, although a spectrogram’s ‘low-frequency cutoff’ and *s-freqM* are not synonymous: for a standard alveolar [s], *s-freqM* is likely to be higher than 4 kHz (cf. Figure 3). The positive loading means that the cutoff frequency is higher for high values of PC5.

8.5.2. Observed relationship between PC5 and prefixedness

The binary regression model of PC5 is shown in Table S14.1.1, and the coefficient estimates are plotted in Figure 18. PrefixStatus is strongly significant, as are word Frequency and Stress. There are no significant interactions. The cutoff frequency is higher when the [s] is in a prefix compared with a nonprefix ($p < .02$), in higher frequency words ($p < .01$), and when Stress is late i.e., an unstressed syllable follows the [s] ($p < .05$).

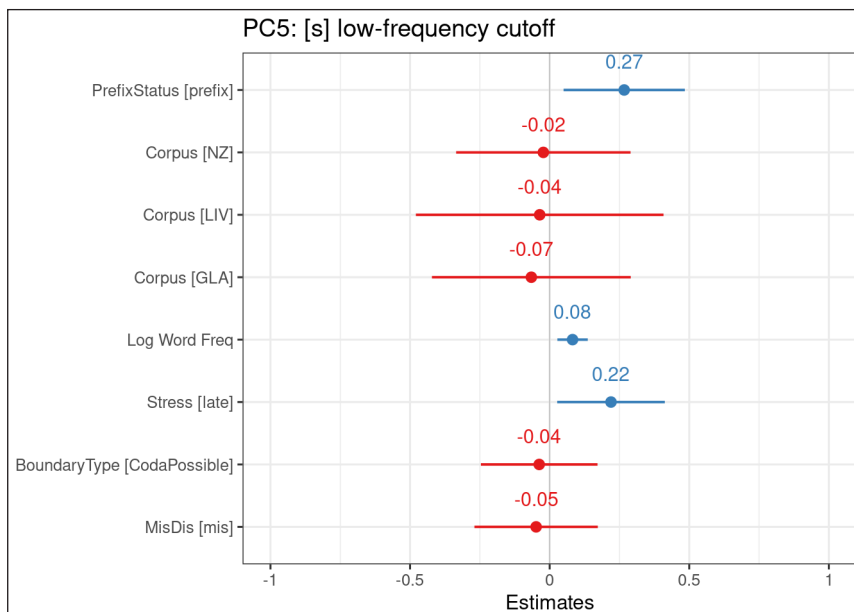


Figure 18: Coefficients and 95% CIs from binary model of PC5. Formula: $PC5 \sim Corpus + PrefixStatus + LogWordFreq + Stress + BoundaryType + MisDis + (1|Speaker)$. Significant predictors are Prefixedness ($p < .02$), LogWordFreq ($p < .01$) and Stress ($p < .05$). Numbers centered above the circles show each coefficient’s value: positive in blue, negative in red.

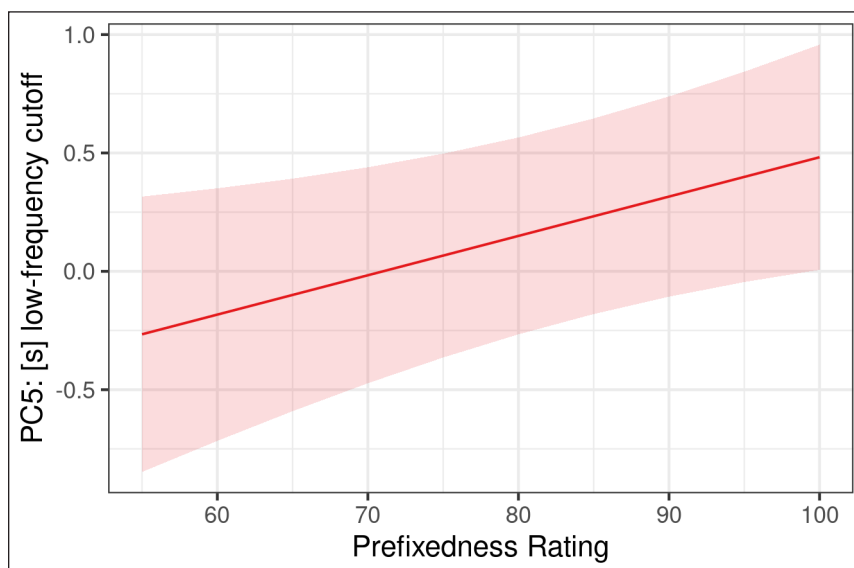


Figure 19: PC5: Gradient effect of prefixedness rating, within prefixes only.

The Ratings model has a significant effect of prefixedness ($p < .01$, Table S14.2.1, Figure S6). Considered separately, there is a significant gradient effect within prefixes ($p < .05$, Table S14.2.2), but not within nonprefixes ($p = .86$, Table S14.2.3). The stronger the ‘felt prefixedness’ of a word, the higher its low-frequency cutoff in [s] (Figure 19).

8.5.3. PC5: Discussion

Though PC5 accounts for only 10% of the variance, it is robustly affected by prefixedness, as well as Stress and word Frequency as independent main effects in the binary model. S-freqM is associated with the frequency of F5 in adjoining vowels, so is strongly influenced by lip rounding and gender (Stevens, 1998; Koenig et al., 2013). Our analyses will be affected by lip rounding, but not by gender because we normalized measurements within gender.

PC5 is unlikely to simply reflect a skew in the amplitude envelope towards high frequencies, because no other measure associated with spectral skew towards high frequencies is significant (s-freqH approaches significance according to Wilson Black's criteria, but it is negatively loaded). Furthermore, PC2 reflects skews in spectral balance, independently of PC5. We return to this and related points below.

8.6. Summary of binary prefix effects across corpora

All five PCs showed a significant binary effect of PrefixStatus, either in isolation or in interaction with another factor. For PC1 and PC3, PrefixStatus interacted with Corpus: for these PCs, post hoc tests reported above show that prefixedness reaches significance within some corpora, but

		Significant within-corpus effects from emmeans post hoc testing			
		NZ	US	LIV	GLA
PC1 [s] duration and peakiness	PrefixStatus × Corpus				
PC2 [s] spectral balance	PrefixStatus × Stress				
PC3 vowel duration	PrefixStatus × Corpus; PrefixStatus × BoundaryType				
PC4 F2-F1 spacing in vowel	PrefixStatus × Boundary Type				
PC5 [s] low frequency cutoff	PrefixStatus				

Table 4: Summary of significant effects of prefixedness for each PC, and the significant within-corpus effects revealed by post hoc testing (in green).

not others. Of relevance is whether each corpus marks the prefix phonetically. For example, PC2 and PC5 do not have a significant overall effect of Corpus, suggesting that the same general trend is likely to be present in all corpora. But all corpora containing the same trend is somewhat different from the effect being robust enough to reach significance *within* each corpus: the overall effect could be carried by some corpora while others are not significant. We tested this by adding a Prefix \times Corpus interaction into those models for which that interaction is not significant (PC2,4,5), and conducting post hoc tests to assess for a prefixedness effect within each corpus. **Table 4** summarises the results across PCs and Corpora, both in the overall binary model, and in the post hoc within-corpus testing (details in Supplementary Materials Section 15).

We see different patterns of significance across the corpora. It makes sense that the two larger corpora have more significant effects. The smaller number of effects for Glasgow and Liverpool should not be taken as evidence that they mark the distinction more weakly, especially since only 2/5 of the PCs have a significant Corpus \times PrefixStatus interaction. **Table 4** does suggest, however, that the patterns may be somewhat different across the corpora, and that each corpus shows a significant effect for at least one PC.

9. Discussion

9.1. Acoustic parameters: PCs 1–5

The first five PCs capture all the main dimensions of acoustic covariation in the data: PC1, PC2 and PC5 for [s], and PC3 and PC4 for the vowel. PCs 1, 2 and 5 all involve differences in [s] spectral shape, but in different ways. PC1 mainly reflects longer [s] for prefixes, significantly so only in the NZ corpus. We interpret PC1's spectral difference—degree of mid-frequency peakedness—as a secondary effect of the much stronger influence of [s] duration. PC2 reflects overall balance in relative amplitude between mid versus high frequencies. It is affected by prefix status only insofar as nonprefixes that are not adjacent to the main lexical stress have lower-frequency spectral energy. The spectrum is also skewed towards mid frequencies when phoneme 4 is a stop, regardless of prefix status. PC2 is discussed further with PC5 below, and Supplementary Materials Section 11.3 discusses articulatory causes of PC2's overall pattern. PC5 reflects the frequency at which the spectral envelope of [s] gains its highest amplitude within the range 3–7 kHz—a range which covers most vocal tract lengths for this property. So it too captures an aspect of spectral shape, but, unlike PC1, not one that reflects degree of peakedness.

PC3 captures vowel duration. In the binary model, prefixes have longer vowels than nonprefixes in all corpora except Glasgow. The difference is especially large for the US. Prefixes also have longer vowels for both boundary types, with a more robust effect for CodaPossible forms than CodaImpossible forms. That is, the difference is greater for forms whose fourth phoneme is /p t k/. PC3 Ratings show the same PrefixStatus \times Corpus interaction as the binary

model, but there is no evidence of a separate gradient effect that explains variation within either the prefixed or the unprefixed words: tested separately within prefixes and nonprefixes, neither ratings model reaches significance.

PC4 captures spectral variation in vowel quality, as does, to a lesser extent, PC1. With the association between duration and F2-F1 separation controlled (in PC1), prefixedness and formant spacing are negatively associated. In the CodaPossible set, but not the CodaImpossible set, prefixes have smaller F2-F1 spacing than nonprefixes. The directions of change in PC4 differ because the dialects' KIT vowels differ in quality, and these nonprefixes carry a weaker beat than prefixes. Though both prefixed and unprefixed *mis-dis-* are weak syllables, nonprefixes are more reduced, and so have a less canonical KIT vowel than prefixes. For NZ and Glasgow this means a shift towards [ɪ] with increasing syllable reduction. For Ohio speakers, it means a shift towards [ə]. Though only, as noted, within CodaPossible words in our data. We suggest PC3 and PC4 combine to give CodaPossible prefixes a heavier beat than nonprefixes.

We also see a gradient effect within prefixes, with higher levels of felt 'prefixedness' having high PC4 (and thus smaller unexplained F2-F1 differences). More data are required to fully disentangle these effects. What can be said, however, is that the present data support interpretations in the literature (cited in the Introduction) that prefixed cases like *discolour* differ in their syllabification from unprefixed cases like *discover*, in ways that simple phonotactic rules cannot explain: an analysis reflecting morphological structure is required.

Turning now to the broad picture, prefixes are marked by a tendency for NZ, US and Liverpool [s] to be slightly longer and peakier (only significantly so for NZ), and a relatively long vowel, whose F2-F1 spacing represents a more canonical KIT vowel for the specific dialect, but only when phoneme 4 is a voiceless stop. Vowel duration is the most robust differentiator of prefix status.

PCs 1, 3 and 4 show an effect of binary PrefixStatus, although always in interaction with one or more other predictor variable. PrefixStatus interacts with Corpus for PC1 and PC3, and with BoundaryType for PC3 and PC4. Within these generalizations, there is dialectal variation as already described.

PC5 stands out as the only component showing an effect of binary PrefixStatus independent of other predictor variables. PC5 also has one of the two significant gradient effects in our Rating data, the other being in PC4: the stronger the 'felt prefixedness' of a word, the more canonical the vowel for its dialect (PC4) and the higher its low-frequency cutoff in [s] (PC5). These two properties could both reflect speakers' intuitive sense that greater syllable stress accompanies stronger prefixedness.

PC5's pattern is compatible with that for PC2. Though independent of each other, both reflect effects on [s]'s spectral shape. Moreover, Stress is implicated in both. In PC5, Late Stress (in which syllable 2 is unstressed [*disillúision*]) raises s-freqM, as does prefixedness. In PC2, Late Stress skews spectral amplitudes of nonprefixes towards lower frequencies. Thus these nuanced

spectral measures of [s] potentially capture the degree of stress the segment carries, broadly defined as encompassing prefixedness, lexical stress, and presumably utterance prosody, along with related influences like word frequency. Together PC5 and PC2 may contribute to subtle gradations observable in English utterance and lexical stress: prefixes vary more than nonprefixes in comparable environments, including nuclear and postnuclear positions (Smith et al., 2012), with consecutive unstressed syllables tending to be somewhat differentiated from one another. More broadly, these tentative observations suggest how multiple parameters combine to determine subtle gradations in degree of syllable stress, from fully stressed to highly reduced. Confirmation would require work with very large corpora, or highly controlled parameters in standard experiments.

These results are largely compatible with the findings of Smith et al. (2012), but go beyond them not just by offering more nuanced observations about [s], but also allowing us to examine differences across dialects and BoundaryType, to seek gradient as well as binary distinctions, and to identify any independent influences of Stress and word Frequency, as well as patterns of covariation amongst the acoustic parameters.

9.2. General Discussion

We set out to answer four interlinked research questions: (1) To what extent is the prefixed-unprefixed distinction described by Smith et al. (2012) marked in spontaneous speech and different phonological contexts? (2) Is it observable across different dialects of English? (3) Can it be definitively distinguished from lexical frequency? (4) Is it best described as a binary phonetic distinction, or is there evidence of morphological gradience which is reflected in gradient phonetic realization?

For question (1), the distinction is clearly present in spontaneous speech. While the results are broadly consistent with Smith's (2012) findings for SSBE, there are also differences. These likely stem from combinations of three influences: the methodology, allowing us to decouple different aspects of covariation and consider them independently from each other, albeit with much less control over segmental, prosodic and other contextual variation; the words analysed, which encompass a much broader range of following phonological environments; and perhaps the absence of SSBE amongst the varieties analysed here. (SSBE is phonetically very different from all four dialects analysed here, though in the relevant acoustical aspects it behaves rather like our US corpus.)

Prefixes contain longer vowels (except for Glasgow) and longer [s] (though only significantly so for NZ). The vowel effect is more robust than the [s] effect, the net result being consistent with Smith et al.'s (2012) findings that prefixed forms have a smaller s:i ratio, albeit with the dialect-specific differences noted in Section 9.1. PC4 offers evidence that in spontaneous speech, prefixes can differ from nonprefixes in multiple, sometimes competing ways.

For question (2), all dialects mark prefix status acoustically in spontaneous speech. Duration (of vowel and [s] independently), formant frequencies in the vowel, and attributes of the spectral envelope of [s] can all contribute to the difference. Not all dialects use all these properties (see **Table 4**), but all use some, though sometimes only when other, more influential factors are statistically controlled, especially BoundaryType, Stress, and MisDis.

We do not wish to read too much into the specific differences observed in this dataset. The recordings are not completely comparable in terms of speech style, quantity of observations, or the specific words produced. To be sure of the exact nature of the dialectal differences, more controlled experiments are needed to replicate the observed differences, or a repeated analysis on other corpora. What does seem clear, though, is that not all the dialects behave in exactly the same way. This was expected given the number of acoustic properties involved and other differences between the four dialects. Nonetheless, for all of them, the vowel is longer in prefixes than nonprefixes, except for Glasgow, for which vowel duration is constrained by other morphophonological factors (Rathcke & Stuart-Smith, 2016).

The answer to question (3) is that the prefixedness effect is not an artefact of word frequency. Word frequency clearly affects the production of such affixes and nonaffixes: it has been reported in the literature (e.g., Pluymaekers et al., 2005), and significantly influenced vowel duration (PC3) in our own data. However, it did not interact with prefix status even for vowel duration. Importantly, for all PCs we controlled for frequency effects, and still found an effect of prefixedness over and above that of lexical frequency. This conclusion confirms Smith et al.'s (2012) more tentative observations.

For question (4), work on the cognitive representation of morphological structure has found evidence of gradient representation, with words containing the same affix seemingly affixed to different degrees. Factors that contribute to perceived degree of affixedness include, among other factors, semantic transparency (Gonnerman & Andersen, 2000; Wurm, 1997), the (relative) lexical frequency of the word form and the base (Hay, 2003), and the phonological transparency of the boundary (Gonnerman & Andersen, 2000). See Hay and Baayen (2005), Raffelsiefen (1999) and Stein (2023) for reviews, and Smith et al. (2012) Section 4.2 for brief further comments, including on historical change. Rather than quantify such factors separately, we attempted to get ratings reflecting an aggregate measure of subjective frequency that captured the combined effect of such factors.

As **Figure 1** shows, participants responded bimodally to this rating task. Even though they were given a continuous scale and encouraged in the instructions to think in terms of gradience, they sorted words into two clear categories of 'prefixed' and 'nonprefixed'. However, there was considerable variation within each category and they are not completely nonoverlapping (see *misnomer*, Table 12 [Appendix C] and its discussion in Section 4.1). So some words seem subjectively highly prefixed, while other prefixed words feel much more ambiguously prefixed.

Indeed, we classed *misnomer* as unprefixes with some misgivings, which were confirmed by the collected ratings. The correlation plots in **Figure 2** demonstrate substantial agreement across dialect areas in ranking the words, including the ambiguous words with intermediate ratings. This is not just an effect that emerges in the aggregate. Individuals used the scale gradiently. Just 11% of the individual ratings given were 0, 32% were 100, and the rest fell somewhere between the ends of the scale. No participant exclusively used the ends of the scale. Speakers, then, do have a sense of gradience within these categories. This perceived degree of membership of the category was associated with the same factors that significantly distinguished the two categories: lexical frequency, stress pattern, word length, and phonological context.

The statistical analysis of our acoustic-phonetic data likewise supports a largely binary prefixed-unprefixed distinction in phonetic realization. For all five PCs, a binary analysis shows some effect, and while subjective ratings of gradient prefixedness can sometimes also predict the acoustics well, for PCs 1–3 we found no evidence that the rating continuum adds significant explanatory power beyond that of the binary classification.

However, PC4 and PC5 are significantly higher in prefixed forms rated as more prefixed. This provides evidence of phonetic gradience in production within the prefixed category. Words can seem prefixed to different degrees and this affects their production. In sum, the weight of the evidence is that the prefix-nonprefix distinction is heavily bimodal, and reflected largely categorically in the acoustics, but there are also gradient phonetic effects in vowel quality and low-frequency cutoff of [s] within the prefixed category.

While ‘rating’ does not significantly predict acoustic realizations better than prefix status for PCs 1–3, it is important to consider that our models of the acoustics control for many of the factors that statistically influenced these subjective ratings (stress, frequency, following environment). In our preliminary exploratory modelling, which did not contain all these controls, we did in fact find multiple cases where rating could predict realization within either prefixed or unprefixes forms. In many ways, the combined linguistic factors that we are controlling for *are* in large part the gradience (although other factors are involved too, most notably semantic transparency). The consistency of raters’ behaviour does suggest that gradience within the categories has some cognitive reality. It seems that the sense of affixedness and the linguistic properties are interrelated and likely work together in affecting the acoustic pattern.

Our data illustrate a wide set of available acoustic properties that can be harnessed to indicate prefix status, and how they work together, with different dialects using them in different combinations. If, as we and others have suggested, a broad overarching principle drives these results, then this raises interesting questions regarding how the dialect-specific differences we have found to reflect the prefix distinction can be generalized to how each dialect marks rhythmic prominence, and its opposite, syllable reduction. For example, speakers of Ohio English might vary vowel duration more substantially with the prominence of a syllable than speakers

of Glaswegian English. For varieties that vary the vowel duration, if the overarching principle is governed by rules of syllable prominence and reduction, then we might expect the prefix status of the initial triphone to be less clear under conditions that favour extremely short or extremely long vowels: sometimes prefixes need to be spoken very fast, causing highly-reduced vowels, while at other times the talker may need to emphasize the first syllable of an unprefix word, resulting in a heavier syllable and longer vowel. The implication is that prefix status may be most clearly conveyed when vowel durations have intermediate (non-extreme) values. Further work into the wider dialectal differences, then, may shed considerable light into the processes driving these acoustic consequences of prefixedness. If there are not comparable differences in the wider phonological systems of these dialects, then that would locate the observed differences very locally in the *dis-mis-* prefix, and suggest that the differences are instead located in phonetically-rich dialect-specific representations of the prefixes themselves.

What we have observed here goes beyond standard descriptions of differences between dialects. The bread and butter of the field of language variation and change is differences across dialects and speakers in terms of language usage. This might be grammatical variation (e.g., Tagliamonte & Baayen, 2012), phonological variation (e.g., Labov, 2001), phonetic variation (e.g., Kiesling, 1998), or—very occasionally—morphological variation (e.g., Nevalainen et al., 2011). Looking at the interface between these—as in, for example, how morphological structure affects interacting covariant phonetic cues—can potentially lead us away from a focus on superficial differences in isolated variables, toward an understanding of differences in the systemic organization of the wider communicative system.

Bybee (1985, pp. 89–90) cites Pagliuca's (1976) unpublished study of the prefix *pre*, which the Shorter Oxford English Dictionary listed as having four different vowel qualities. She discusses the complex acoustic consequences for pronunciation dependent on frequency, transparency of meaning between putative prefix and its base, and the variety of meanings a word (or its base) can have: “there is a strong relationship between the vowel quality and the frequency of the word, and the vowel quality and the semantic predictability of the word.” (p. 89). And “the higher the frequency of a derived word, the more likely it is to occur in a variety of contexts, including some in which its related base word, and the semantic notions expressed by it, do not occur.” (p. 90). Together with our own results and those of others (e.g., Engemann & Plag, 2021), this illustrates that the pronunciation of prefixed words is likely to be subject to the multiple influences that affect all word production, ranging across phonological and prosodic word structure, phonotactics and neighborhood density, sequential and paradigmatic probability, and frequency of occurrence (of potentially the prefix, the stem and the whole word), as well as situation-specific influences stemming from the pragmatic and functional load on the prefix in the particular discourse.

Individual variation amongst talkers provides a further dimension of complexity, even within a single variety of the same language. Using ultrasound and acoustic measures, Strycharczuk and Scobbie (2016) found evidence of gradient differences between 20 SSBE speakers in realisation of vowels and /l/ in word sets like bimorphemic *fooling* versus monomorphemic *hula* (and also *fool*). While all speakers distinguished the bi- and monomorphemic forms, there was a fairly uniform distribution across speakers in the degree to which they made the distinction: some made a largely categorical difference, others made very little distinction, and most talkers were intermediate, resulting in gradiency overall, but not necessarily for everyone.

It is not surprising, therefore, that one can find gradient and sometimes contradictory effects on particular prefixes. But it does seem that the prefix status of our critical syllables exerts its own effect, which speakers of different dialects of English respect in one way or another, usually closely related to the way somewhat stronger rhythmic (yet still weak) beats are realized. This sensitivity to prefixedness is problematic for any model predicting that implementation should be blind to the morphology, and adds to a growing body of research which provides evidence of acoustic consequences of morphological structure (see Plag & Ben Hedia, 2018; Schmitz, 2022; Tang & Bennett, 2018 and references therein). A number of papers suggest that there is an impetus to avoid reduction of a morpheme to facilitate communication of affix meaning (Arnon & Cohen Priva, 2013; Hanique & Ernestus, 2012; Pluymaekers et al., 2010; Rose, 2017), and this interpretation accounts well for prefixes. Prefixes convey an important meaning. Many, including *dis* and *mis*, transform the meaning of the lexical stem to something close to its opposite. We know that listeners use the acoustic properties signalling prefixedness in real time to distinguish between lexical competitors (Clayards et al., 2021; Blazej & Cohen-Goldberg, 2015; see also Hawkins, 2011). Ensuring that the additional meaning is intelligible, and/or that listeners' attention is drawn to the prefix, may explain why prefixes are typically more rhythmically/metrically prominent than nonprefixes.

How might these factors be conceptualised? Clayards et al. (2021) concluded that while listeners used the phonetic detail of both types of initial triphone to predict word identity, the rhythmic emphasis that comes with a true prefix probably allows faster and more accurate prediction of meaning than the absence of such a focus found in nonprefixes. They relate this observation to the centrality of metrical rhythm in a general, probabilistic and context-dependent theory of speech perception, drawing evidence from neuroscientific work that shows that the listening brain constructs metrical beats from an incoming rhythmic signal (which can be multimodal), and focusses attention onto them, thereby facilitating prediction and allowing meaning to be efficiently accessed. When such a metrical beat is lacking, attentional processes shift to more continuous monitoring. These systems can operate in tandem, with the rhythmicity of the stimulus determining their relative balance (Schroeder & Lakatos, 2008). Presumably,

talkers normally adapt their speech to accommodate both types of listening process, such that the intelligibility of the utterance as a whole falls within acceptable limits for the dialect and communicative situation.

A theoretical prosodic account and communicative accounts such as those discussed by Clayards et al. (2021) need not be mutually exclusive. They have in common multiple interacting influences on behaviour. Advances in analytical methods and a wider variety of databases are beginning to indicate how these multiple influences may work within and across languages. See for example Strycharczuk (2019) for a number of European languages, Stein and Plag (2022) and Stein (2023) for English, and Tang and Bennett (2018) for Kaqchikel Mayan. A new type of statistical modelling Linear Discriminative Learning, LDL (Baayen, Chuang, Shafaei-Bajestan and Blevins, 2019; Heitmeier et al., 2021; Stein & Plag, 2021), uses the highly abstract concept of ‘lexome’ to map form to meaning in networks that preserve the integrity of words, whether morphologically complex or not (content lexomes), while recognising structural patterns that signal, for example, morphological relationships (derivational lexomes), and inflectional functions such as number and tense (inflectional lexomes). Thus form and meaning work together in an integrated system. What matters in the case of prefix status is that there is a structure, or form, some of whose details depend on whether the last segment abuts a prosodic boundary with the rest of the word, and this structure forms an identifiable pattern. Such patterns interact with other structural and statistical influences that may enhance or obscure the prefix-related pattern. In an implicit learning model like LDL, these patterns can be learned and will generalise to the rest of the language to form a largely (though not completely) coherent morphophonological system with predictable semantic consequences (e.g., Kuperman et al. (2007) and references cited above). They can be reflected elsewhere too, such as in the content-word versus function-word distinction (Bell et al., 2009).

One issue arising from our own work concerns whether the patterns of syllable reduction observed in the present study generalize to other syllables with similar phonetic structure but without potential prefix status, for example *wisteria*, *hysterical*, *hysteresis*. We hypothesize that they should, and will be subject to the same influences such as lexical stress, wider prosodic context and word frequency/predictability. Other challenges concern perceptual processing. Which combinations of acoustic properties convey the heavier beat said to characterize prefixes? For dialects like the US/Ohio, is the s:i ratio critical regardless of overall rhyme duration, as the SSBE data of Smith et al. (2012) suggest it might be, or does a relatively long triphone always convey prefixedness regardless of its internal acoustic structure? For CodaPossible *dis-* and *mis-* words, can a long VOT override acoustic-perceptual cues that would otherwise be heard as an unprefix word?

10. Concluding remarks

We have considered 14 different acoustic measures, and how they work together to mark prefixedness across four broad varieties of English. Our analysis focuses on clusters of acoustic measures (PCs) identified via a Principal Components Analysis. Three of these related primarily to properties of the [s], one to vowel duration, and one primarily to the relative spacing of formant frequencies in the vowel, that is to vowel quality. A series of regression models shows that each of these five PCs is used to differentiate prefixed from unprefixed forms in at least one variety of English.

Morphology affects phonetic detail. The exact details of how it does so depend on a variety of influences, including regional dialect. Future work should disentangle whether these dialect-specific details relate to broader dialectal differences associated, for example, with the phonetic realisation of stress and prominence, or whether they reflect norms specific to *mis-dis-* prefixes. In either case, these differences open new directions for work on variation between dialects of English. There is much to be gained from moving beyond the study of isolated variables, towards looking at complex patterns of covariation of phonetic cues (in our case, the ‘internal acoustic structure’ of relevant triphones), and how each dialect and speaker uses these as a coordinated system to implement grammar in context.

Appendix A: Supplementary materials, code and reproducibility

<https://osf.io/hq8kr/> gives access to Supplementary Materials as a 5 MB html file downloadable from the first page. For R code and data, see Files section.

Appendix B: Instructions for rating task

How prefixed is this word?: We are interested in prefixed words. A prefix is a meaningful part at the beginning of a word, such as ‘sub’ in ‘substandard’, or ‘inter’ in ‘intergenerational’. Other words might begin with the same sounds, but not be prefixed. The ‘sub’ in ‘sublime’ for example is not a prefix. ‘Substandard’ and ‘sublime’ are clear cases of prefixed and unprefixed words. But many cases are not so clear cut, and may feel somewhat in between. For example, ‘submarine’, might feel somewhat prefixed to you, but not as clearly prefixed as ‘substandard’. Our experiment is about words that begin with the letters ‘dis’ and ‘mis’. We would like to put these words on a scale, representing how much the word seems to have a prefix. When you consider each word, how much do you tend to think of it as obviously containing the prefix ‘mis’ or ‘dis’, or obviously not containing a prefix? For each word, please use the slider to mark the degree of prefixedness, from 0% (completely unprefixed) to 100% (completely prefixed).

There are just over 200 words. There are no right or wrong answers. We’re simply after your first intuition, so just work quickly and give your best guess. This shouldn’t take you more than around 20 mins.

Appendix C: Mean Word Prefixedness Ratings

Prefix Status	Word	Rating	Rating Rank	Stress	Boundary Type
prefix	misinformation	97.36	213	late	CodaImpossible
prefix	disempowered	96.77	212	late	CodaImpossible
prefix	mispronouncing	96.49	211	late	CodaPossible
prefix	mispronunciation	96.27	210	late	CodaPossible
prefix	disorganized	96.24	209	early	CodaImpossible
prefix	misunderstandings	96.03	208	late	CodaImpossible
prefix	disadvantage	95.65	207	late	CodaImpossible
prefix	disloyal	95.64	206	early	CodaImpossible

(Contd.)

Prefix Status	Word	Rating	Rating Rank	Stress	Boundary Type
prefix	mischaracterisation	95.55	205	late	CodaPossible
prefix	disempowerment	95.44	204	late	CodaImpossible
prefix	disagreed	95.33	203	late	CodaImpossible
prefix	disorderliness	95.31	202	early	CodaImpossible
prefix	misbehave	95.24	201	late	CodaImpossible
prefix	disadvantages	95.14	200	late	CodaImpossible
prefix	disassembling	95.13	199	late	CodaImpossible
prefix	disapproving	95.04	198	late	CodaImpossible
prefix	disrespect	95.04	198	late	CodaImpossible
prefix	dislike	95.01	196	early	CodaImpossible
prefix	displeased	94.94	195	early	CodaPossible
prefix	disagreeable	94.87	194	late	CodaImpossible
prefix	disinterested	94.83	193	early	CodaImpossible
prefix	disagreement	94.82	192	late	CodaImpossible
prefix	disassemble	94.82	192	late	CodaImpossible
prefix	disadvantaged	94.78	190	late	CodaImpossible
prefix	disbelief	94.68	189	late	CodaImpossible
prefix	misbehaving	94.67	188	late	CodaImpossible
prefix	disengaged	94.62	187	late	CodaImpossible
prefix	disliking	94.55	186	early	CodaImpossible
prefix	disbelieving	94.21	185	late	CodaImpossible
prefix	dishonest	94.04	184	early	CodaImpossible
prefix	disconnected	94.00	183	late	CodaPossible
prefix	discomfort	93.97	182	early	CodaPossible

(Contd.)

Prefix Status	Word	Rating	Rating Rank	Stress	Boundary Type
prefix	dislikes	93.90	181	early	CodaImpossible
prefix	disembarking	93.78	180	late	CodaImpossible
prefix	disconnect	93.72	179	late	CodaPossible
prefix	misbehaved	93.69	178	late	CodaImpossible
prefix	disliked	93.58	177	early	CodaImpossible
prefix	disinherited	93.40	176	late	CodaImpossible
prefix	disobeying	93.38	175	late	CodaImpossible
prefix	disapproval	93.33	174	late	CodaImpossible
prefix	disengage	93.22	173	late	CodaImpossible
prefix	disembark	93.06	172	late	CodaImpossible
prefix	disqualifying	93.05	171	early	CodaPossible
prefix	disassociated	93.00	170	late	CodaImpossible
prefix	disqualified	92.97	169	early	CodaPossible
prefix	misguided	92.88	168	early	CodaImpossible
prefix	disproportionate	92.87	167	late	CodaPossible
prefix	disadvantaging	92.82	166	late	CodaImpossible
prefix	discontinued	92.77	165	late	CodaPossible
prefix	disembarked	92.65	164	late	CodaImpossible
prefix	discontented	92.55	163	late	CodaPossible
prefix	disfunctional	92.45	162	early	CodaImpossible
prefix	misbehaviour	92.40	161	late	CodaImpossible
prefix	disbelieved	92.31	160	late	CodaImpossible
prefix	disqualification	92.31	160	late	CodaPossible
prefix	distasteful	92.27	158	early	CodaPossible

(Contd.)

Prefix Status	Word	Rating	Rating Rank	Stress	Boundary Type
prefix	disenchanted	92.26	157	late	CodaImpossible
prefix	disregarded	92.23	156	late	CodaImpossible
prefix	disorderly	92.04	155	early	CodaImpossible
prefix	disengagement	91.95	154	late	CodaImpossible
prefix	dismounted	91.83	153	early	CodaImpossible
prefix	disagree	91.81	152	late	CodaImpossible
prefix	disorientation	91.76	151	late	CodaImpossible
prefix	misplaced	91.74	150	early	CodaPossible
prefix	disheartened	91.69	149	early	CodaImpossible
prefix	discolour	91.40	148	early	CodaPossible
prefix	misadventure	91.24	147	late	CodaImpossible
prefix	misconception	90.85	146	late	CodaPossible
prefix	disheartening	90.65	145	early	CodaImpossible
prefix	disorientated	90.56	144	early	CodaImpossible
prefix	disoriented	90.53	143	early	CodaImpossible
prefix	disorientating	90.14	142	early	CodaImpossible
prefix	misplacement	89.56	141	early	CodaPossible
prefix	disorienting	89.46	140	early	CodaImpossible
prefix	misfortune	89.21	139	early	CodaImpossible
prefix	disappearing	89.19	138	late	CodaImpossible
prefix	disheartens	89.10	137	early	CodaImpossible
prefix	disused	88.87	136	early	CodaImpossible
prefix	disregard	88.33	135	late	CodaImpossible
prefix	disappeared	88.22	134	late	CodaImpossible

(Contd.)

Prefix Status	Word	Rating	Rating Rank	Stress	Boundary Type
prefix	disowns	88.13	133	early	CodaImpossible
prefix	dislodged	87.79	132	early	CodaImpossible
prefix	disenfranchised	87.55	131	late	CodaImpossible
prefix	misleading	87.45	130	early	CodaImpossible
prefix	miscarry	87.24	129	early	CodaPossible
prefix	disabled	87.10	128	early	CodaImpossible
prefix	disillusioned	87.00	127	late	CodaImpossible
prefix	disrepair	86.90	126	late	CodaImpossible
prefix	disable	86.87	125	early	CodaImpossible
prefix	disillusionment	86.05	124	late	CodaImpossible
prefix	disorder	85.81	123	early	CodaImpossible
prefix	disappear	85.68	122	late	CodaImpossible
prefix	disinfectant	84.97	121	late	CodaImpossible
prefix	dislodging	84.83	120	early	CodaImpossible
prefix	displacement	84.58	119	early	CodaPossible
prefix	disbanded	84.33	118	early	CodaImpossible
prefix	disability	84.23	117	late	CodaImpossible
prefix	mislaid	83.50	116	early	CodaImpossible
prefix	disintegrated	82.47	115	early	CodaImpossible
prefix	disjointed	82.23	114	early	CodaImpossible
prefix	disintegrate	81.87	113	early	CodaImpossible
prefix	displaced	81.56	112	early	CodaPossible
prefix	disorders	79.46	111	early	CodaImpossible
prefix	disband	78.47	110	early	CodaImpossible

(Contd.)

Prefix Status	Word	Rating	Rating Rank	Stress	Boundary Type
prefix	mislay	77.41	109	early	CodaImpossible
prefix	disarray	76.50	108	late	CodaImpossible
prefix	disfigurement	74.82	107	early	CodaImpossible
prefix	disgraced	73.65	106	early	CodaImpossible
prefix	discouraged	71.58	105	early	CodaPossible
prefix	disgrace	70.86	104	early	CodaImpossible
prefix	discharged	67.96	103	early	CodaImpossible
prefix	disgorged	66.87	102	early	CodaImpossible
prefix	discomfiture	66.00	101	early	CodaPossible
nonPrefix	misnomer	65.35	100	early	CodaImpossible
prefix	discharge	63.00	99	early	CodaImpossible
prefix	misdemeanour	62.99	98	late	CodaImpossible
prefix	discharging	62.64	97	early	CodaImpossible
prefix	disconcerting	59.90	96	late	CodaPossible
prefix	dismantled	55.97	95	early	CodaImpossible
nonPrefix	mistaken	53.18	94	early	CodaPossible
nonPrefix	disappointed	49.47	93	late	CodaImpossible
nonPrefix	disappointing	49.17	92	late	CodaImpossible
nonPrefix	discovered	48.68	91	early	CodaPossible
nonPrefix	mistaking	48.60	90	early	CodaPossible
nonPrefix	disappointment	47.72	89	late	CodaImpossible
nonPrefix	discover	44.28	88	early	CodaPossible
nonPrefix	disclosed	44.14	87	early	CodaImpossible
nonPrefix	mistakenly	43.87	86	early	CodaPossible

(Contd.)

Prefix Status	Word	Rating	Rating Rank	Stress	Boundary Type
nonPrefix	discovering	43.69	85	early	CodaPossible
nonPrefix	disappoints	43.18	84	late	CodaImpossible
nonPrefix	disclose	42.42	83	early	CodaImpossible
nonPrefix	mistake	42.26	82	early	CodaPossible
nonPrefix	discriminate	41.41	81	early	CodaPossible
nonPrefix	discriminated	39.85	80	late	CodaPossible
nonPrefix	mistakes	39.06	79	early	CodaPossible
nonPrefix	disposition	38.95	78	late	CodaPossible
nonPrefix	discriminative	37.15	77	early	CodaPossible
nonPrefix	discrimination	36.96	76	late	CodaPossible
nonPrefix	dismissed	36.22	75	early	CodaImpossible
nonPrefix	discovery	36.06	74	early	CodaPossible
nonPrefix	disrupting	34.96	73	early	CodaImpossible
nonPrefix	disrupted	33.47	72	early	CodaImpossible
nonPrefix	disgusting	33.19	71	early	CodaImpossible
nonPrefix	distressed	32.94	70	early	CodaPossible
nonPrefix	dismissive	32.72	69	early	CodaImpossible
nonPrefix	disruptive	31.87	68	early	CodaImpossible
nonPrefix	distraction	31.24	67	early	CodaPossible
nonPrefix	dismayed	30.82	66	early	CodaImpossible
nonPrefix	distressing	30.71	65	early	CodaPossible
nonPrefix	disgustingly	30.56	64	early	CodaImpossible
nonPrefix	disposing	30.49	63	early	CodaPossible
nonPrefix	disparaging	30.13	62	early	CodaPossible

(Contd.)

Prefix Status	Word	Rating	Rating Rank	Stress	Boundary Type
nonPrefix	disgust	30.12	61	early	CodaImpossible
nonPrefix	dismiss	29.86	60	early	CodaImpossible
nonPrefix	distress	29.78	59	early	CodaPossible
nonPrefix	disgusted	29.76	58	early	CodaImpossible
nonPrefix	distorted	29.26	57	early	CodaPossible
nonPrefix	disruptors	29.17	56	early	CodaImpossible
nonPrefix	distortion	29.13	55	early	CodaPossible
nonPrefix	disrupt	28.91	54	early	CodaImpossible
nonPrefix	disposed	28.56	53	early	CodaPossible
nonPrefix	distracting	28.41	52	early	CodaPossible
nonPrefix	distracts	28.38	51	early	CodaPossible
nonPrefix	disperses	28.19	50	early	CodaPossible
nonPrefix	disruption	28.05	49	early	CodaImpossible
nonPrefix	dispatched	27.95	48	early	CodaPossible
nonPrefix	dismissal	27.68	47	early	CodaImpossible
nonPrefix	disposable	27.35	46	early	CodaPossible
nonPrefix	distort	27.29	45	early	CodaPossible
nonPrefix	disturbance	27.21	44	early	CodaPossible
nonPrefix	dispose	27.08	43	early	CodaPossible
nonPrefix	disputing	26.85	42	early	CodaPossible
nonPrefix	disperse	26.67	41	early	CodaPossible
nonPrefix	disputed	26.65	40	early	CodaPossible
nonPrefix	distraught	26.50	39	early	CodaPossible
nonPrefix	distract	26.09	38	early	CodaPossible

(Contd.)

Prefix Status	Word	Rating	Rating Rank	Stress	Boundary Type
nonPrefix	distracted	26.06	37	early	CodaPossible
nonPrefix	displayed	25.95	36	early	CodaPossible
nonPrefix	disturbed	25.88	35	early	CodaPossible
nonPrefix	dispersed	25.63	34	early	CodaPossible
nonPrefix	disturbances	25.36	33	early	CodaPossible
nonPrefix	dispute	25.06	32	early	CodaPossible
nonPrefix	dispatching	24.77	31	early	CodaPossible
nonPrefix	disturber	24.56	30	early	CodaPossible
nonPrefix	distribute	24.46	29	early	CodaPossible
nonPrefix	disturbing	23.22	28	early	CodaPossible
nonPrefix	distributed	22.78	27	early	CodaPossible
nonPrefix	distributing	22.08	26	early	CodaPossible
nonPrefix	disposal	21.99	25	early	CodaPossible
nonPrefix	disturbs	21.95	24	early	CodaPossible
nonPrefix	dispatch	21.46	23	early	CodaPossible
nonPrefix	disputes	21.19	22	early	CodaPossible
nonPrefix	dispatches	21.13	21	early	CodaPossible
nonPrefix	displays	20.81	20	early	CodaPossible
nonPrefix	discussed	20.76	19	early	CodaPossible
nonPrefix	distinguish	20.58	18	early	CodaPossible
nonPrefix	distillation	19.90	17	late	CodaPossible
nonPrefix	discuss	19.79	16	early	CodaPossible
nonPrefix	discussion	19.38	15	early	CodaPossible
nonPrefix	distinguished	19.01	14	early	CodaPossible

(Contd.)

Prefix Status	Word	Rating	Rating Rank	Stress	Boundary Type
nonPrefix	distributor	18.97	13	early	CodaPossible
nonPrefix	discretion	18.73	12	early	CodaPossible
nonPrefix	distill	18.63	11	early	CodaPossible
nonPrefix	distinction	18.47	10	early	CodaPossible
nonPrefix	distinctly	18.33	9	early	CodaPossible
nonPrefix	distinctive	18.21	8	early	CodaPossible
nonPrefix	display	17.40	7	early	CodaPossible
nonPrefix	discussing	17.09	6	early	CodaPossible
nonPrefix	dispensary	16.51	5	early	CodaPossible
nonPrefix	distinct	15.95	4	early	CodaPossible
nonPrefix	discreetly	15.36	3	early	CodaPossible
nonPrefix	distillery	14.13	2	early	CodaPossible
nonPrefix	disciplinary	9.68	1	late	CodaImpossible

Table 12: Word Prefixedness Ratings and related information.

Acknowledgements

This work has been supported by a University of Canterbury Erskine Fellowship to the first author, a Rutherford Discovery Fellowship to the second author, and University of Canterbury Summer Scholarship funding. The ONZE data was collected by Rosemary Goodyear, Lesley Evans, members of the NZ English class of the Linguistics Department, University of Canterbury, and members of the ONZE team. The NZ Darfield Data was collected by Alex D'Arcy. The MAONZE data is provided courtesy of the MAONZE team: Jeanette King, Catherine Watson, Margaret Maclagan, Peter Keegan, and Ray Harlow. We are grateful to Kevin Watson and Lynn Clark for their work on the OLIVE corpus, and for making it available to us, and to NZILBB and UC CEISMIC for their work in creating and sharing the QuakeBox corpus. The work done by all members of the multiple corpus teams in preparing the data, making transcripts, and obtaining background information is also gratefully acknowledged. We are grateful to Nick Hight for his work on implementing the online ratings survey, and to our colleagues at NZILBB for feedback on various stages of this work. Finally, we thank the Editors and two anonymous reviewers for their rigorous standards and their enthusiasm about the paper's content.

Competing Interests

The authors have no competing interests to declare.

References

- Adank, P., Evans, B. G., Stuart-Smith, J., & Scott, S. K. (2009). Comprehension of familiar and unfamiliar native accents under adverse listening conditions. *Journal of Experimental Psychology: Human Perception and Performance*, 35(2), 520–529. <https://doi.org/10.1037/a0013552>
- Arnon, I., & Cohen Priva, U. (2013). More than words: The effect of multi-word frequency and constituency on phonetic duration. *Language and Speech*, 56(3), 349–371. <https://doi.org/10.1177/0023830913484891>
- Aronoff, M., & Sridhar, S. N. (1983). Morphological levels in English and Kannada or atarizing Reagan. In J. F. Richardson, M. Marks, & A. Chukerman (Eds.), *Papers from the parasession on the interplay of phonology, morphology, and syntax*. Chicago Linguistic Society (pp. 3–16).
- Baayen, R. H., Chuang, Y.-Y., Shafaei-Bajestan, E., & Blevins, J. P. (2019). The discriminative lexicon: A unified computational model for the lexicon and lexical processing in comprehension and production grounded not in (de)composition but in linear discriminative learning. *Complexity*, 2019, 4895891. <https://doi.org/10.1155/2019/4895891>
- Barnett, M. J., Doroudgar, S., Khosraviani, V., & Ip, E. J. (2022). Multiple comparisons: To compare or not to compare, that is the question. *Research in Social and Administrative Pharmacy*, 18(2), 2331–2334. <https://doi.org/10.1016/j.sapharm.2021.07.006>
- Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67(1), 1–48. <https://doi.org/10.18637/jss.v067.i01>
- Bauer, L. (1983). *English word formation*. Cambridge University Press.
- Bell, A., Brenier, J. M., Gregory, M., Girand, C., & Jurafsky, D. (2009). Predictability effects on durations of content and function words in conversational English. *Journal of Memory and Language*, 60(1), 92–111. <https://doi.org/10.1016/j.jml.2008.06.003>
- Ben Hedia, S., & Plag, I. (2017). Gemination and degemination in English prefixation: Phonetic evidence for morphological organization. *Journal of Phonetics*, 62, 34–49. <https://doi.org/10.1016/j.wocn.2017.02.002>
- Blazej, L. J., & Cohen-Goldberg, A. M. (2015). Can we hear morphological complexity before words are complex? *Journal of Experimental Psychology: Human Perception and Performance*, 41(1), 50–68. <https://doi.org/10.1037/a0038509>
- Bybee, J. (1985). *Morphology: A study of the relation between meaning and form*. John Benjamins.
- Bybee, J. (2006). From usage to grammar: The mind's response to repetition. *Language*, 711–733.
- Chodroff, E., & Wilson, C. (2022). Uniformity in phonetic realization: Evidence from sibilant place of articulation in American English. *Language*, 98(2), 250–289. <https://doi.org/10.1353/LAN.2022.0007>
- Clark, L., MacGougan, H., Hay, J., & Walsh, L. (2016). “Kia ora. This is my earthquake story.” Multiple applications of a sociolinguistic corpus. *Ampersand*, 3, 13–20.

- Clayards, M., Gaskell, M. G., & Hawkins, S. (2021). Phonetic detail is used to predict a word's morphological composition. *Journal of Phonetics*, 87(3), 101055. <https://doi.org/10.1016/j.wocn.2021.101055>
- Cohn, A., & McCarthy, J. J. (1998). Alignment and parallelism in Indonesian phonology. *Working Papers of the Cornell Phonetics Laboratory*, 12, 53–137.
- D'Arcy, A. (2012). The diachrony of quotation: Evidence from New Zealand English. *Language Variation and Change*, 24(3), 343–369.
- Engemann, M., & Plag, I. (2021). Phonetic reduction and paradigm uniformity effects in spontaneous speech. *The Mental Lexicon*, 16(1), 165–198. <https://doi.org/10.1075/ml.20023.eng>
- Ernestus, M. (2014). Acoustic reduction and the roles of abstractions and exemplars in speech processing. *Lingua*, 142(3), 27–41. <https://doi.org/10.1016/j.lingua.2012.12.006>
- Fromont, R., & Watson, K. (2016). Factors influencing automatic segmental alignment of sociophonetic corpora. *Corpora*, 11(3), 401–431.
- Gahl, S. (2008). Time and thyme are not homophones: The effect of lemma frequency on word durations in spontaneous speech. *Language*, 84(3), 474–496.
- Garrett, P., Williams, A., & Evans, B. (2005). Attitudinal data from New Zealand, Australia, the USA and UK about each other's Englishes: Recent changes or consequences of methodologies? *Multilingua*, 24(3), 211–235. <https://doi.org/10.1515/MULT.2005.24.3.211>
- Gaskell, M. G., & Marslen-Wilson, W. (1997). Integrating form and meaning: A distributed model of speech perception. *Language and Cognitive Processes*, 12, 613–656.
- Gonnerman, L., & Andersen, E. (2000). Graded semantic and phonological similarity effects in processing morphologically complex words. In S. Bendjaballah, W. U. Dressler, O. E. Pfeiffer, & M. D. Voeikova (Eds.), *Morphology 2000: Selected papers from the 9th Morphology Meeting, Vienna, 24–28 February 2000* (pp. 137–148). <https://doi.org/10.1075/cilt.218.12gon>
- Gordon, E., Maclagan, M., & Hay, J. (2007). The ONZE corpus. In J. C. Beal, K. P. Corrigan, & H. L. Moisl (Eds.), *Creating and digitizing language corpora: Volume 2: Diachronic databases* (pp. 82–104). Palgrave Macmillan UK. https://doi.org/10.1057/9780230223202_4
- Hanique, I., & Ernestus, M. (2012). The role of morphology in acoustic reduction. *Lingue e Linguaggio*, 11(2), 147–164. <https://doi.org/10.1418/38783>
- Hawkins, S. (2011). Does phonetic detail guide situation-specific speech recognition? Keynote address. In *Proceedings of the 17th International Congress of Phonetic Sciences*. http://www.icphs2011.hk/ICPHS_CongressProceedings.htm
- Hay, J. (2003). *Causes and consequences of word structure*. Routledge.
- Hay, J. (2007). The phonetics of “un.” In J. Munat (Ed.), *Lexical creativity, texts and contexts* (pp. 39–57). John Benjamins.
- Hay, J., & Baayen, R. H. (2005). Shifting paradigms: Gradient structure in morphology. *Trends in Cognitive Sciences*, 9(7), 342–348.
- Heitmeier, M., Chuang, Y.-Y., & Baayen, R. H. (2021). Modeling morphology with linear discriminative learning: Considerations and design choices. *Frontiers in Psychology*, 12. <https://doi.org/10.3389/fpsyg.2021.720713>

- Kiesling, S. F. (1998). Men's identities and sociolinguistic variation: The case of fraternity men. *Journal of Sociolinguistics*, 2(1), 69–99.
- King, J., Maclagan, M., Harlow, R., Keegan, P., & Watson, C. (2011). The MAONZE project: Changing uses of an indigenous language database. *Corpus Linguistics and Linguistic Theory*, 7(1), 37–57.
- Klatt, D. H. (1976). Linguistic uses of segmental duration in English: Acoustic and perceptual evidence. *Journal of the Acoustical Society of America*, 59(5), 1208–1221.
- Koenig, L. L., Shadle, C. H., Preston, J. L., & Mooshammer, C. R. (2013). Toward improved spectral measures of /s/: Results from adolescents. *Journal of Speech, Language and Hearing Research (Online)*, 56(4), 1175–1189. [https://doi.org/10.1044/1092-4388\(2012/12-0038\)](https://doi.org/10.1044/1092-4388(2012/12-0038))
- Kuperman, V., Pluymaekers, M., Ernestus, M., & Baayen, R. H. (2007). Morphological predictability and acoustic duration of interfixes in Dutch compounds. *Journal of the Acoustical Society of America*, 121(4), 2261–2271. <https://doi.org/10.1121/1.2537393>
- Labov, W. (2001). *Principles of linguistic change, volume 2: Social factors*. Wiley Online Library.
- Local, J. K. (1995). Syllabification and rhythm in non-segmental phonology. In J. Windsor-Lewis (Ed.), *Studies in general and English phonetics: Essays in honour of Professor J. D. O'Connor*. (pp. 350–366). Routledge.
- McCarthy, J. J., & Prince, A. (1993). *Prosodic morphology I: Constraint interaction and satisfaction*. Ms. University of Massachusetts & Rutgers University.
- Nevalainen, T., Raumolin-Brunberg, H., & Mannila, H. (2011). The diffusion of language change in real time: Progressive and conservative individuals and the time depth of change. *Language Variation and Change*, 23(1), 1–43. <https://doi.org/10.1017/S0954394510000207>
- Ogden, R., Hawkins, S., House, J., Huckvale, M., Local, J. K., Carter, P., Dankovicová, J., & Heid, S. (2000). ProSynth: An integrated prosodic approach to device-independent, natural-sounding speech synthesis. *Computer Speech and Language*, 14, 177–210.
- Pitt, D., M. A., & Fosler-Lussier, E. (2007). *Buckeye Corpus of Conversational Speech (2nd release)*. Department of Psychology, Ohio State University. <https://buckeyecorpus.osu.edu/>
- Plag, I., & Ben Hedia, S. (2018). The phonetics of newly derived words: Testing the effect of morphological segmentability on affix duration. In S. Arndt-Lappe, A. Braun, C. Moulin, & E. Winter-Froemel (Eds.), *Expanding the lexicon: Linguistic innovation, morphological productivity, and the role of discourse-related factors* (pp. 93–116). de Gruyter Mouton. <https://doi.org/10.1515/9783110501933-095>
- Plag, I., Homann, J., & Kunter, G. (2017). Homophony and morphology: The acoustics of word final S in English. *Journal of Linguistics*, 53(1), 181–216. <https://doi.org/10.1017/S0022226715000183>
- Pluymaekers, M., Ernestus, M., & Baayen, R. H. (2005). Lexical frequency and acoustic reduction in spoken Dutch. *Journal of the Acoustical Society of America*, 118(4), 2561–2569. <https://doi.org/10.1121/1.2011150>
- Pluymaekers, M., Ernestus, M., Baayen, H., & Booij, G. (2010). Morphological effects on fine phonetic detail: The case of Dutch –igheid. In C. Fougeron, B. Kühnert, M. d'Imperio, & N. Vallée

(Eds.), *Laboratory phonology 10: Variability, phonetic detail and phonological representation* (pp. 511–531). Mouton de Gruyter.

Raffelsiefen, R. (1999). Diagnostics for prosodic words revisited: the case of historically prefixed words in English. In T. A. Hall & U. Kleinhenz (Eds.) *Studies of the phonological word*. (pp. 133–201). Benjamins. (*Current Issues in Linguistic Theory* 174).

Rathcke, T., & Smith, R. H. (2015). Speech timing and linguistic rhythm: On the acoustic bases of rhythm typologies. *Journal of the Acoustical Society of America*, 137(5), 2834–2845. <https://doi.org/10.1121/1.4919322>

Rathcke, T. V., & Stuart-Smith, J. H. (2016). On the Tail of the Scottish Vowel Length Rule in Glasgow. *Language and Speech*, 59(3), 404–430. <https://doi.org/10.1177/0023830915611428>

Reidy, P. F. (2015). A comparison of spectral estimation methods for the analysis of sibilant fricatives. *Journal of the Acoustical Society of America*, 137(4), EL249. <https://doi.org/10.1121/1.4915064>

Rose, D. E. (2017). *Predicting plurality: An examination of the effects of morphological predictability on the learning and realization of bound morphemes* [Doctoral dissertation, University of Canterbury]. University of Canterbury Repository. <https://doi.org/10.26021/4881>

Schmitz, D. (2022). Production, perception, and comprehension of subphonemic detail: Word-final /s/ in English. *Studies in Laboratory Phonology 11*. Language Science Press. https://library.oapen.org/bitstream/id/8e1417cc-b2ed-464b-b20d-3681fd90b2ec/external_content.pdf

Schroeder, C., & Lakatos, P. (2008). Low-frequency neural oscillations as instruments of sensory selection. *Trends in Neurosciences*, 32, 9–18. <https://doi.org/10.1016/j.tins.2008.09.012>

Selkirk, E. O. (1982). The syllable. In H. van der Hulst & N. Smith (Eds.), *The structure of phonological representations*. (Part 2, pp. 337–384). Foris; De Gruyter. <https://doi.org/10.1515/9783112423325-010>

Seyfarth, S., Garellek, M., Gillingham, G., Ackerman, F., & Malouf, R. (2018). Acoustic differences in morphologically-distinct homophones. *Language, Cognition and Neuroscience*, 33(1), 32–49. <https://doi.org/10.1080/23273798.2017.1359634>

Sharma, D., Levon, E., & Ye, Y. (2022). *50 years of British accent bias: Stability and lifespan change in attitudes to accents*. *English World-Wide*, 43(2), 135–166. <https://doi.org/10.1075/EWW.20010.SHA>

Siegel, D. (1974). *Topics in English morphology* [Doctoral dissertation, Massachusetts Institute of Technology]. DSpace@MIT. <http://hdl.handle.net/1721.1/13022>

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

Stein, S. D. (2023). *The phonetics of derived words in English: Tracing morphology in speech production* (Vol. 585). de Gruyter.

Stein, S. D., & Plag, I. (2021). Morpho-phonetic effects in speech production: Modeling the acoustic duration of English derived words with linear discriminative learning. *Frontiers in Psychology*, 12, 678712.

- Stein, S. D., & Plag, I. (2022). How relative frequency and prosodic structure affect the acoustic duration of English derivatives. *Laboratory Phonology*, 13. <https://doi.org/10.16995/labphon.6445>
- Stevens, K. N. (1998). *Acoustic phonetics*. MIT Press.
- Strycharczuk, P. (2019). *Phonetic detail and phonetic gradience in morphological processes*. Oxford University Press. <https://doi.org/10.1093/acrefore/9780199384655.013.616>
- Strycharczuk, P., & Scobbie, J. M. (2016). Gradual or abrupt? The phonetic path to morphologisation. *Journal of Phonetics*, 59, 76–91. <https://doi.org/10.1016/j.wocn.2016.09.003>
- Stuart-Smith, J., José, B., Rathcke, T., Macdonald, R., & Lawson, E. (2017). Changing sounds in a changing city: An acoustic phonetic investigation of real-time change over a century of Glaswegian. In C. Montgomery & E. Moore (Eds.), *Language and a sense of place: Studies in language and region* (pp. 38–65). Cambridge University Press.
- Sugahara, M., & Turk, A. (2009). Durational correlates of sublexical constituent structure. *Phonology*, 26, 477–524.
- Szpyra, J. (1989). *The phonology–morphology interface: Cycles, levels and words*. Routledge.
- Tagliamonte, S. A., & Baayen, R. H. (2012). Models, forests, and trees of York English: Was/were variation as a case study for statistical practice. *Language Variation and Change*, 24(2), 135–178.
- Tang, K., & Bennett, R. (2018). Contextual predictability influences word and morpheme duration in a morphologically complex language (Kaqchikel Mayan). *The Journal of the Acoustical Society of America*, 144(2), 997–1017. <https://doi.org/10.1121/1.5046095>
- Villarreal, D., Clark, L., Hay, J., & Watson, K. (2021). Gender separation and the speech community: Rhoticity in early 20th century Southland New Zealand English. *Language Variation and Change*, 33(2), 245–266.
- Vittinghoff, E. (2005) *Regression methods in biostatistics: Linear, logistic, survival, and repeated measures models*. Springer Verlag.
- Warren, P., & Marslen-Wilson, W. (1987). Continuous uptake of acoustic cues in spoken word recognition. *Perception & Psychophysics*, 41, 262–275.
- Watson, K., & Clark, L. (2017). The origins of Liverpool English. *Listening to the Past: Audio Records of Accents of English*, 114.
- Wells, J. C. (1982) *Accents of English*. Cambridge University Press.
- Wilson Black, J., Brand, J., Hay, J., & Clark, L. (2022). Using principal component analysis to explore co-variation of vowels. *Language and Linguistics Compass*, 17(1), e12479.
- Wurm, L. H. (1997). Auditory processing of prefixed English words is both continuous and decompositional. *Journal of Memory and Language*, 37, 438–461.
- Young, S., Evermann, G., Gales, M., Hain, T., Kershaw, D., Liu, X., Moore, G., Odell, J., Ollason, D., & Povey, D. (2006). *The HTK book* (Vol. 3). Cambridge University Engineering Department.
- Zuraw, K., Lin, I., Yang, M., & Peperkamp, S. (2021). Competition between whole-word and decomposed representations of English prefixed words. *Morphology*, 31, 201–237. <https://doi.org/10.1007/s11525-020-09354-6>

