Weight and final vowels in the English stress system

Claire Moore-Cantwell

Simon Fraser University

1 Introduction

Metrical theories of stress placement (Liberman and Prince (1977); Hayes (1980, 1982), et. seq.) have historically included a theory of extrametricality - some syllables, nearly always final syllables, are excluded from metrical calculations. They are never stressed, and cannot affect the stress pattern of other syllables in the word. In Optimality Theory (Prince and Smolensky, 1993/2004), the notion of extrametricality does not typically exist as a formal part of metrical stress analyses. Rather, the same job is done by the NonFinality family of constraints. These constraints force the final syllable of a word to either be unstressed, or to be totally unparsed. In this paper, I argue that for the English stress system, NonFinality is not enough. Rather, *both* NonFinality and true extrametricality are required to accurately model the influence that certain final vowels have over an entire word's stress. The constraint NonParse is proposed, which forces certain types of syllables to be parsed outside of the prosodic word. These syllables are therefore rendered invisible to any other stress constraints. Evidence for this constraint comes from the probabilistic interaction between weight generalisations and final vowel effects.

I examine the lexicon of English, using the CMU pronouncing dictionary (Weide, 1994), to characterize the behaviour of actual words of English. Two experiments show that speakers' choice of stress on a novel word is influenced both by the final vowel and the weight of syllables in that word. Speakers' behaviour is then modelled using Maximum Entropy grammar (Goldwater and Johnson, 2003; Hayes et al., 2009), a constraint-based system which predicts probability distributions over candidates.

Words with final syllables containing [i, l̩, ɹ̩] appear to be special in the English stress system. Liberman and Prince (1977) and Hayes (1982) point out that only words with these final segments can host pre-antepenultimate main stress in monomorphemic words like 'cátterpìllar', 'páradìddle', and 'állegòry'. Similar words with a different final nucleus, like *'állegòrin', or *'állegòrow' do not exist. In this paper, I examine the effects of these three final vowels on stress placement more generally. All three exhibit a tendency to push stress leftwards in both long and shorter words, but the effect of final [i] is particularly strong. Words like 'cánnery' are much more common than words like 'canáry'. Words with other final nucleii, like [ə, n̩, m̩, o͡w ...] do not show the same effect. I argue based on both dictionary data and nonce-word experimental data that final [i], and sometimes final [ɹ̩,l̩] as well, can be excluded from the metrical parse in English. The constraint NonParse-i, enforces this.

The evidence that final [i] must be parsed outside the prosodic word, rather than simply left unfooted or unstressed, comes from the interaction between the effects of final vowel, and the effects of syllable weight. English stress in general obeys the Latin Stress Rule (Chomsky and Halle, 1968; Liberman and Prince, 1977; Halle and Vergnaud, 1987; Hayes, 1980, 1982; Kager, 1989; Burzio, 1994):

(1) **Latin Stress Rule for English:** If a word's penultimate syllable is heavy, then that word receives penultimate main stress. If the penultimate syllable is light, then the word receives antepenultimate main stress.

This rule, like most generalisations in the English stress system, has exceptions. Words like *gálaxy* violate the first clause, and words like *vanílla* violate the second clause. However, speakers of English do demonstrate probabilistic knowledge of this rule (Olejarczuk, 2014; Domahs et al., 2014).

While the first clause of the Latin Stress Rule (if penult is heavy → penultimate stress) has very few exceptions in the lexicon, the second clause (if penult is light → antepenultimate stress) has many exceptions. Pater (1994) argues that antepenultimate and penultimate stress compete when the penultimate syllable is light, and that neither is clearly the rule or clearly the exception. In this paper, a search of the CMU pronouncing dictionary of English reveals that the degree of preference for antepenultimate stress in words with light penults varies based on the word's final vowel. Light penult words ending in [i], [l̩], or [ɹ̩] are more likely to take antepenultimate stress than light penult words ending in [n̩/m̩/ŋ̩] or [ə].

Speakers' ability to extend both clauses of the Latin Stress Rule is tested in this paper, using nonce words which speakers must choose a stress pattern for. When the penult is heavy, speakers prefer penultimate stress. Also, when the word ends in [i], speakers prefer antepenultimate stress. These two preferences interact in i-final words with a heavy penult. These words exhibit more antepenultimate stress than words with a heavy penult and a different final vowel, but not as much antepenultimate stress as words with a light penult. This effect cannot be modelled using NonFinality, and instead the constraint NonParse-i is used.

2 The stress generalisations

The stress system of English has a long and rich history as the object of phonologists' interest. Notable comprehensive analyses have been given in Chomsky and Halle (1968); Liberman and Prince (1977); Hayes (1980); Selkirk (1984); Halle and Vergnaud (1987); Kager (1989); Burzio (1994); Alcántara (1998). Throughout these works, a few major themes arise: Main stress nearly always falls on one of the final three syllables of a word, and within that three-syllable window both heavy penultimate syllables and heavy final syllables can attract main stress. The morphological structure of a word is important. Some affixes shift stress to a particular position within the word (e.g. '-ity' demands antepenultimate main stress, so that elástic ∼ elastícity; see Tescner and Whitley, 2004 for a comprehensive discussion of this), but the stress of a base is sometimes preserved in the derived form as a secondary stress (Pater, 2000; Collie, 2008; Zamma, 2012). A word's part of speech can also affect stress placement, so that nouns prefer penultimate stress or antepenultimate stress, while verbs prefer final stress (Kelly and Bock, 1988; Sonderegger and Niyogi, 2013; Sereno and Jongman, 1995). Stress clashes (two stressed syllables in a row) are rare in English. They are rare in words, and avoided in the production of nonwords (Kelly and Bock, 1988), and also can trigger stress retraction at the word and phrase level (Prince, 1983; Tilsen, 2012; Henrich et al., 2014 and many others).

Another property of English stress that is agreed upon throughout the above literature is that lexical exceptions abound. Some generalisations are more exceptionful than others: exceptions to the preference for nouns to take penultimate stress and verbs to take final stress are plentiful, while exceptions to the preference for words with heavy penultimate syllables to take penultimate main stress are vanishingly rare. However, all accounts of the system have used some mechanism for marking exceptions. Chomsky and Halle (1968) and later Burzio (1994) attribute different stress patterns to different underlying forms, for example the difference between the word 'cínema', and 'banána', both nouns with all light syllables, might be that 'banána' is underlyingly /bana<u>nn</u>a/, with a geminate which does not get realized, but which does make the second syllable heavy. Other approaches use diacritics which determine the underlying weight of a syllable (Hayes, 1980), or which specify some syllable as extrametrical on a word-by-word basis (Selkirk, 1984).

In the rest of this section, the lexicon of English in the form of the CMU pronouncing dictionary is examined in order to more precisely characterize the regularities in stress assignment across words. The Latin Stress Rule and the effects of final [i,l,r] are the focus of this examination. Stress on two-syllable words, secondary stress, and the effects of affixation will largely be ignored.

2.1 Setting up the dictionary search

The CMU pronouncing dictionary (Weide, 1994) together with the $\text{SUBTLEX}_{\text{US}}$ corpus (Brysbaert and New, 2009) was used to examine the distribution of stress in long words (3 syllables and longer) of English. The CMU pronouncing dictionary is a dictionary of American English lexical items, and contains about 134,000 entries, phonetically transcribed, and with each vowel annotated for primary, secondary, or no stress. Because it is so exhaustive, it contains a great many entries that are low frequency enough as not to be present in most adult native speaker vocabularies. The CMU pronouncing dictionary also contains in some cases independent entries for forms with inflectional morphology (both 'banana' and 'bananas' for example). English inflectional morphology does not affect a word's stress, so a better model of the stress rules of English would be obtained by looking only at base forms, and excluding inflected forms like 'bananas'. In order to avoid hyper-low-frequency entries and entries with inflectional morphology, a 'cleaned-up' version of CMU was used, namely the input corpus for Hayes's phonotactic probability calculator (Hayes, 2012). This input file contains 18,034 entries, all of which are frequent enough to be in English CELEX (Baayen et al., 1993). This dataset also avoids entries with inflectional morphology, and certain transcription 'errors' are corrected, such as having multiple primary stresses on a single word. A series of scripts was then used to annotate this lexicon further. The first step was to automatically code for morphological complexity.

Morphology: All words were coded for their morphological complexity. This paper is not directly concerned with the effects of affixation, but it is important to know that any observed effects of weight, length, and especially of final vowel, are not merely due to the influence of a particular morpheme. Particular affixes in English can influence the stress pattern of a word, either by attracting stress or by behaving as extrametrical (Chomsky and Halle, 1968; Burzio, 1994; Halle and Vergnaud, 1987). Words were automatically coded for whether or not they were morphologically complex, using the spelling of the word as a proxy. For example, words ending in the orthographic sequence 'tion' were considered to end in the '-tion' affix. The list of suffixes and prefixes in Tescner and Whitley (2004), chapter 2, was used[1]. This method of marking words was surprisingly successful. Table 1 shows the results of a test in which 100 words were sampled from each category (morphologically simple, morphologically complex) for 3 syllable and 4 syllable words. A native English speaker (the author) then checked these randomly sampled words (400 total) and noted the number of incorrect categorizations in each sample. A summary of the accuracy of the categorization method is also reported: the 'F1' statistic. F1 is calculated based on both Precision (of words categorized as complex, how

[1] Some orthographic strings were excluded because too many morphologically simple words contained them by accident. For example, 're-' is a prefix of English, but monomorphemic words like 'real', and 'ready' also contain 're' at the beginning. The affixes excluded for this reason were 'ab-', 'ad-', 're-', '-y' and '-o'.

many actually are?) and Recall (of words that are actually complex, how many are categorized that way?). F1 values range from 0 to 1, with values closer to 1 being better.

Table 1: Categorizations the automatic algorithm made for each sample of 100 words (400 total). F1 score, a measure of goodness of categorization, is given for 3 syllable and for 4 syllable words.

		Categorized as: Simple	Complex	F1 score (goodness of categorization)
3 syllables	Simple	88	12	0.87
	Complex	13	87	
4 syllables	Simple	74	2	0.84
	Complex	26	98	

Weight: Whether the Latin Stress Rule is observed or not depends on whether or not the penultimate syllable is heavy. Chomsky and Halle (1968) and subsequent work on English stress defines 'heavy' penultimate syllables as syllables with either a long vowel, or at least one coda consonant. The long vowel condition is relatively straightforward, but the coda-consonant condition is less so, as we will see below. In the search of the CMU pronouncing dictionary presented below, the monophthongs [ɑ, æ, ə, ɔ, ɛ, ɪ, ʊ, i, u] as well as the syllabic consonants [l̩, ɹ̩, n̩, m̩, ŋ̩] are counted as short vowels, while diphthongs [e͡i, a͡i, o͡ʊ, a͡ʊ, o͡i] were counted as long[2].

Turning to the coda consonant condition: according to the Maximal Onset Principle (Kahn, 1976), consonants between two vowels should be syllabified in such a way that as many consonants as possible are assigned to an onset rather than a coda. For example, in a word like 'tapestry', with three consonants between the final two vowels ([tæpəstɹi]), all three consonants would be assigned to the onset of the final syllable, and the penultimate syllable would be coda-less, and therefore light: [tæ.pə.stɹi]. Any string of consonants which can legally begin a word in English can begin a syllable (for example, [stɹ] begins the word 'string'), but a sequence of consonants which cannot begin a word must be broken up across the coda and the onset. For example, in a word like 'galaxy' ([gæləksi]), the penult is inescapably heavy. [k] must belong to the coda of the penultimate syllable, since [ks] cannot begin a word in English ([gæ.lək.si]).

However, Kahn proposes that the Maximal Onset Principle can be violated in English in the case of stressed syllables - stressed syllables prefer to be heavy, and therefore attract consonants to their coda position. For example, a word like 'digestive', would be syllabified as [da͡i.**ʤɛ́s**.təv], rather than [da͡i.**ʤɛ́**.stəv] - the [s] is parsed as a coda of the stressed syllable which otherwise would be light. Given this possibility, it is

[2] Using the Arpabet, the transcription system for the CMU pronouncing dictionary, (EY, AY, OW, AW, OY) were counted as 'long' while (AA, AE, AH, AO, EH, ER, IH, IY, UH, UW) were counted as 'short'

unclear how to characterize words like 'tapestry' - because the penultimate syllable is unstressed, it is light, but if it were to be stressed then it would be heavy. These words are classified below as having ambiguous weight - either light or heavy depending on context.

A second source of weight-ambiguity arises when words contain syllables closed by sonorants (l, ɹ, n, m, ŋ). If such a syllable is stressed, the sonorant forms a coda, making the syllable heavy (e.g. consúltant [kn̩.**sʌ́l**.tn̩t]), but if the syllable is stressless, the vowel-sonorant sequence is not permitted and the syllable becomes light as the sonorant becomes the nucleus (cònsultátion [kɑn.**sl̩**.t͡ei.ʃn̩]). Unlike the other sonorant consonants, [ɹ̩] can be stressed as a sonorant nucleus (e.g. compérsion [km̩.**pɹ̩́**.ʒn̩]), however it is still disallowed as a coda in a stressless syllable (*[ɛ́.**nəɹ**.ʤi] gets realized as [ɛ́.**nɹ̩**.ʤi]). For this reason, syllables with a sonorant nucleus or a sonorant coda are ambiguous in weight, as long as the following syllable begins with at least one consonant. The penult of 'mystery', [mɪ́.**stɹ̩**.i], would not be heavy even if the stress shifted since the ɹwould become an onset rather than a coda: [mə.**stɛ́**.ɹi].

Table 2: Example words exhibiting different kinds of penultimate syllable weight. 'L' stands for all sonorants which can be syllabic, [l, ɹ, m, n, ŋ]
.

Weight	Penult syllable	Antepenult stress		Penult stress	
H	**VV**	exponent	[ɛ́k.**sp͡oʊ**.nənt]	aroma	[ə.**ɹ͡óʊ**.mə]
	VC(C)	galaxy	[gǽ.**lək**.si]	elixir	[ə.**lɪ́k**.sɹ̩]
H/L	**V.CC**	tapestry	[tǽ.**pə.st**ɹi]	digestive	[d͡âi.**ʤɛ́s**.təv]
	(V)L.C	faculty	[fǽ.**kl̩**.ti]	beholden	[bi.**hól**.dn̩]
L	**V**	settler	[sɛ́.**ɾl̩**.ɹ̩]	appellate	[ə.**pɛ́**.lət]
	V	radio	[ɹ͡éi.**ɾi**.͡oʊ]	bikini	[bɨ.**kí**.ni]
	V	cinema	[sɪ́.**nə**.mə]	banana	[bə.**nǽ**.nə]

In the next section, the effect of the weight of the penult on the stress pattern of the whole word will be examined. Later, effects of the word's final vowel will also be considered, as well as the interaction between the final vowel and the penult weight. Five categories of penult weight will be specifically considered, with examples given above in Table 2: (1) penultimate syllables that are definitely heavy by virtue of having a long vowel (labelled **VV**), (2) penults that are definitely heavy by virtue of having a coda consonant that cannot be resyllabified as an onset (labelled **VC(C)**), (3) penults that are ambiguous because of consonants that can be syllabified as a coda or an onset, depending on stress (**V.CC**), (4) penults that are ambiguous because of having a syllabic sonorant or a sonorant coda (**(V)L.C**), and (5) penults that are definitely light, because they have a short vowel and lack any coda consonants or sonorants that could become a coda if the syllable were stressed (**V**).

2.2 Corpus search results

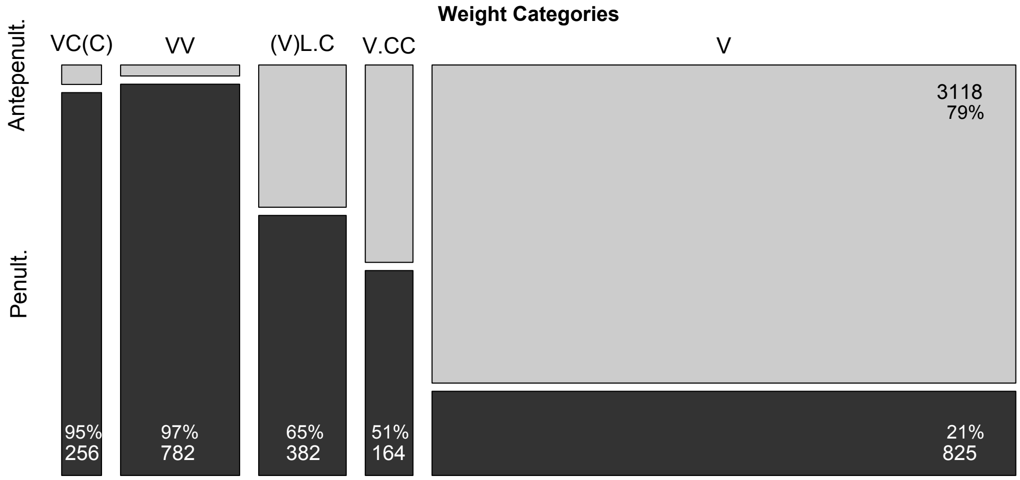
A total of 6531 words were found in CMU that were three syllables long or longer. Main stress was almost always penultimate (2409 words) or antepenultimate (3520 words). A total of 168 words took final stress, 410 took pre-antepenultimate stress, and a mere 24 words had main stress before the preantepenult (all are morphologically complex, like ‘anticipatory’ and ‘professionalism’). The remainder of this section will focus on describing the factors that condition the choice between penultimate and antepenultimate main stress. Unless otherwise noted, morphologically complex and morphologically simple words are both included in the counts - it will turn out that morphological complexity does not matter as much as one might imagine for the trends described here.

Weight: Figure 1 illustrates the effect of the penultimate syllable’s weight on stress placement. Within trisyllabic and longer words with penultimate or antepenultimate stress, 3211 have definitely light penultimate syllables. 1035 have definitely heavy penultimate syllables, with either a diphthong or at least one coda consonant. This leaves 730 ambiguous penults, of two types. The first have monophthongal vowels and no coda consonants, but are followed by a cluster whose first member could be ‘annexed’ as a coda if the penult were stressed. The second type contain a sonorant nucleus which could become a coda if stressed, or a sonorant coda which could become a nucleus if unstressed. Words with definitely heavy penults take penultimate stress around 95% of the time, regardless of whether it is a long vowel (**VV**) or one or more coda consonants (**VC(C)**) that makes the syllable heavy. When the penult is definitely light (**V**), antepenultimate stress is the clear preference: penultimate stress appears just 21% of the time. Penults with ambiguous weight fall in between the definitely light and the definitely heavy penults. Words with short vowels followed by an onset cluster (**V.CC**) take penultimate main stress about half the time, different from the 21% of light-penult words (χ^2=118.47, p< 1x10^{-15}), and also different from the 95-97% of heavy-penult words (χ^2=266.37, p< 1x10^{-15}). Words with sonorant nucleii or codas (**(V)L.C**) take penultimate stress 65% of the time, still different from both light-penult words (χ^2=500.52, p< 1x10^{-16}) and heavy-penult words (χ^2=311.29, p< 1x10^{-16}), and also different from V.CC words (χ^2=15.72, p< 0.0001).

Thus far, the basic conclusions of Pater (1994) seem to be correct: The first clause of the Latin Stress Rule (If a word’s penultimate syllable is heavy, that word receives penultimate main stress) is nearly exceptionless, but the second clause (If the penultimate syllable is light, then the word receives antepenultimate main stress) exhibits a great many exceptions in the lexicon - 21% of all light penult words.

Final Vowels: Figure 2 splits up words by the nucelus of their final syllable. Only a few nucleii appeared with appreciable frequency in the final syllable of words (whether it was open or closed). Syllabic nasals

Figure 1: Effects of the five different types of penult weight on word stress (words three syllables long and longer only)



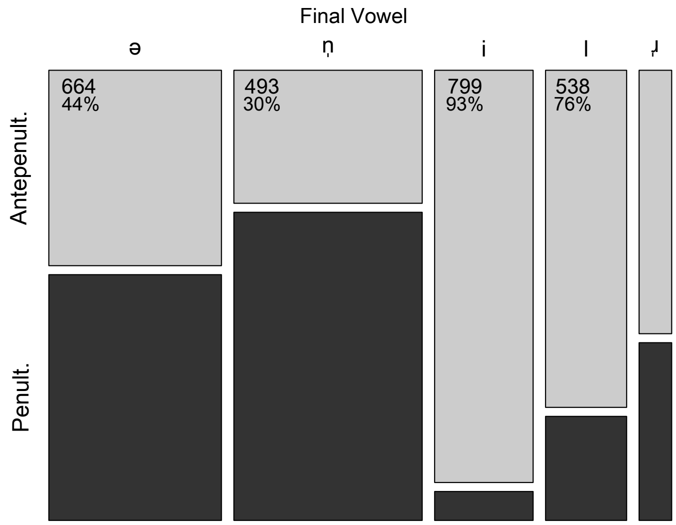
([n̩/m̩/ŋ̩], but overwhelmingly n̩) are the most common (1736 words), followed by [ə][3] (1528), [i] (1080), and [l̩] (772). Additionally, [e͡ɪ] appears in 320 words, mostly with the suffix '-ate'; [ɹ̩] appears in 316 words, [a͡ɪ] appears in 239 words, and [o͡ʊ] appears in 161 words. The remaining vowels of English all appear in the final syllables of less than 100 trisyllabic or longer words.

How likely a particular stress pattern is depends on the word's final vowel. Words ending in schwa or a syllabic nasal prefer penultimate stress, although this preference is relatively weak, and these words exhibit a fairly flat distribution of stresses. Words that end in [i] or a syllabic [l], however, strongly prefer antepenultimate stress.

Many [i/l̩/ɹ̩/n̩]-final words in English are morphologically complex, and specifically contain a suffix which is stressless or which demands antepenultimate stress, e.g. -y (yellow ∼ yellowy), -er (yellow ∼ yellower), -ity (absúrd ∼ absúrdity), -able (avért ∼ avértable), -s/tion (conclúde ∼ conclúsion). Figure 3 illustrates how the relationship between stress and final nucleus changes between morphologically simple vs. morphologically complex words. Words with final [n̩] tend to take penultimate stress in morphologically complex words, but tend to take antepenultimate stress in morphologically simple words. This difference reveals that the general preference for final [n̩] to take penultimate stress is probably because it tends to occur in suffixes that demand penultimate stress, like -s/tion in 'repéat' ∼ 'repetítion'. Final schwa, i, l, and r seem to be

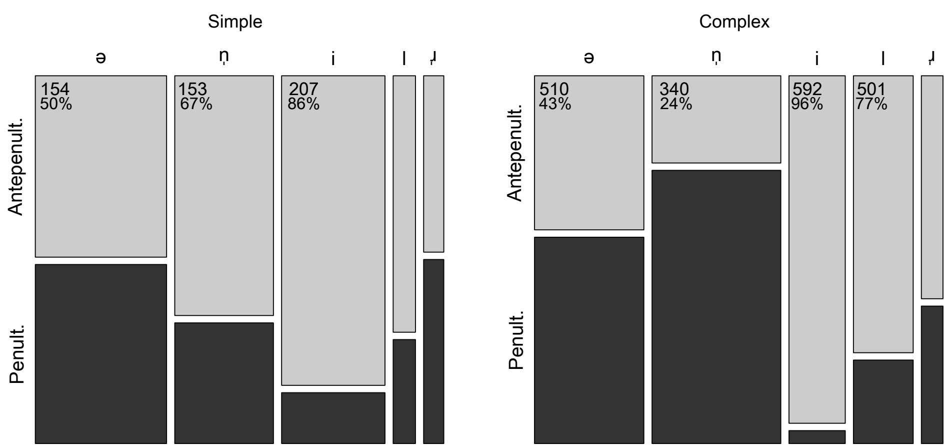
[3]The CMU pronouncing dictionary transcription system uses 'AH' for both [ə] and stressed [ʌ], and additionally uses 'IH' for both [ɪ] and the reduced vowel [ɨ]. For purposes of this search 'AH' and 'IH' in stressless position are both counted as schwas, but in stressed positions are counted as [ʌ] and [ɪ] respectively.

Figure 2: The effect of final nucleus on stress (words three syllables long and longer, with both light and heavy penults)



more stable in their preferences across morphological complexity.

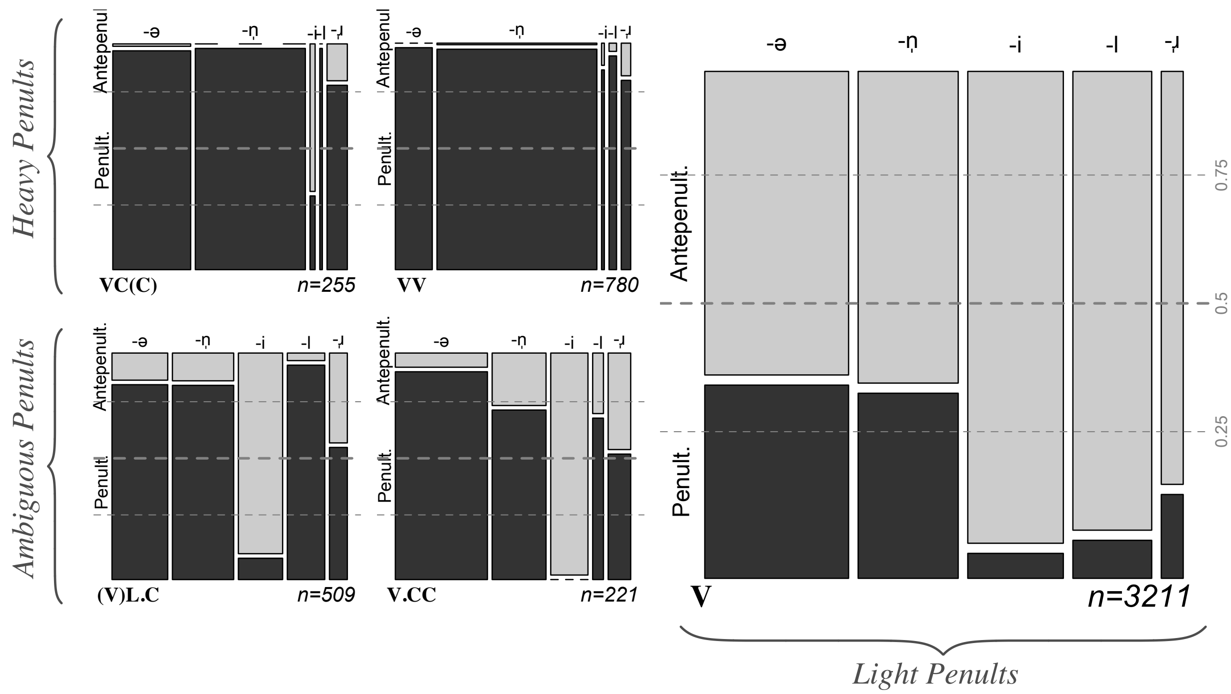
Figure 3: The effects of final nucleus on stress, with words classified as morphologically simple or complex



Interaction between final nucleus and weight:

Having seen that both the weight of the penult and the word's final nucleus affect its chances of taking penultimate or antepenultimate stress, how do these two factors interact? Figure 4 shows the effect of final nucleus within each of the five weight categories discussed above. First, the middling probability of penultimate stress on words with ambiguous penults seen in Figure 1 is an illusion. Once these words are split by final nucleus it is clear that for most vowels these penults prefer to be stressed. Second: Words with final -i prefer antepenultimate stress in every weight class, although there is very little data for words with heavy penults (just 10 i-final words with **VV** penults, and just six with **VC(C)** penults).

Figure 4: Final nucelus and penult weight interact to predict stress patterns.



The fact that non i-final words with ambiguous penults overwhelmingly prefer penultimate stress suggests that the Latin Stress Rule ought to be formulated somewhat differently than it is in (1). In addition to a preference for heavy penults to be stressed, there appears also to be a preference for stressed penults to be heavy (e.g. Stress-to-Weight), such that penults which can only be light tend to be stressed less often than penults which can be heavy if stressed. The behaviour of ambiguous penults will be modelled in section 6.

Summary: The weight of the penultimate syllable overwhelmingly affects stress placement in English. heavy penults, and penults that can be heavy if they are stressed, strongly prefer to host the word's main stress. This tendency has previously been shown to be productive by Olejarczuk (2014) and Domahs et al. (2014). The influence of the word's final nucleus is a novel finding however. While the influence, especially of final [i] in particular, is robust across different types of penult weight and across morphologically complex and morphologically simple words, it is also a bit unexpected. The content of syllabic nucleii does not typically affect stress placement in the world's languages, and when it does it is the content of the *stressed* nucleus, as in Kenstowicz (1997). Because of the typological unexpectedness of the effects of final [i], it is worth checking to see if the tendency is productive. Experiment 1, below, examines English speakers' productions of novel words with and without a final [i]. For simplicity, only final [i] and final [ə] were considered, since their behaviour was the most consistent across the lexicon. Experiment 2 examines the interaction between the effects of final [i] and penult weight in speakers' productions. In this study, words with light penults and heavy penults were included, each with both final [ə] and final [i]. Words with a heavy penult and a final [i] are particularly interesting since they do not occur with much frequency in the lexicon (see Figure 4). Learners of English presumably get only meagre evidence for whether these forms should primarily take penultimate stress, because they have a heavy syllable, or whether they should pattern with ambiguous penults and primarily take antepenultimate stress because they end in an [i].

3 Experiment 1 - effects of final vowel

3.1 Introduction

Both experiments in this paper were modelled after Guion et al. (2003), in which the productivity of certain trends in the English stress system was tested by asking participants to pronounce novel English words. The challenge for a production task for the English stress system is that English orthography is non-transparent, and different participants may interpret one orthographic string in many different ways. Vowels are especially difficult to represent unambiguously in English orthography, which is problematic because the quality of a vowel is one factor which can affect the stress of a word. On the other hand, it is difficult to present a novel word auditorily without giving it some stress pattern. Guion et al. solved this problem by presenting strings of individual syllables, each pronounced as a separate prosodic word, and asking participants to string the syllables together into a word. I copy that methodology here.

3.2 Methods

3.2.1 Participants

The experiment was presented via the world wide web, and participants were recruited through word of mouth, and through Amazon Mechanical Turk. All participants were over 18 years of age, and had an age range of 19 to 61 (mean age: 33). Only IP addresses originating in the United States were accepted.

Participants were asked where they were from, and "when you speak English, where do people think you are from?" If their answer to the second question was a location within the United States, they were assumed to be a native speaker of American English. Participants were paid at a rate of $0.91 for the experiment, which took about 20 minutes. Data was collected from a total of 104 participants, and data from 65 participants was used. The remaining participants were excluded because of problems with the sound recording and native speaker status. The process of excluding participants is described in detail in the results section.

3.2.2 Items

Items were three syllables long, consisting only of light syllables (codaless syllables whose vowel is a monophthong). Both novel words and real words were used. Because participants were asked to pronounce a novel word after listening to three isolated syllables, real words were included in order to encourage participants to make their productions as like real English words as possible - in particular to encourage them to reduce unstressed vowels in their productions. Real words used in the experiment were evenly divided between final vowels and stress patterns, and are shown in Table 3.

Table 3: Real words used in the experiment. Like the nonwords, each was presented split up into three single-syllable prosodic words, for example {æ} {læ} {skʌ} for 'Alaska'.

	Final Vowel	
	i	ə
Antepenultimate stress	colony, recipe	cinema, Canada
Penultimate stress	Alaska, dilemma	bikini, spaghetti

Each item consisted of three auditorily presented individual syllables, and two auditorily presented versions of the full word, with different stress patterns (antepenultimate and penultimate). Participants first heard the syllables, then pronounced the word, then heard the two stress options, and chose between them. An example item is shown in Figure 5.

Figure 5: Example item in four stages. All presentation was auditory.

1)	[bæ] [mæ] [ki]	
2)	*RECORD...*	
3)	What did you say?	
4)	[bǽməki]	[bəmǽki]
	☐	☐

Items were 80% novel words and 20% actual words (32 novel words and 8 actual words). When they were actual words, the two stress choices were (1) the actual word, and (2) a mis-stressed version of the actual word, e.g. [kǽnədə] and [kənǽdə] for 'Canada'. All items (words and nonwords) had the same stressed vowel in each stress version.

For the two stress versions of each item, the stimuli were transcribed into the international phonetic alphabet (IPA) and pronounced in a random order by a male native speaker of American English, in the frame sentence "Say X again." The words were then spliced out of the frame sentence. Eight real words were also included - two words for each stress x vowel quality condition. In each case, the stressed vowels of the correctly stressed and misstressed versions of the word matched.

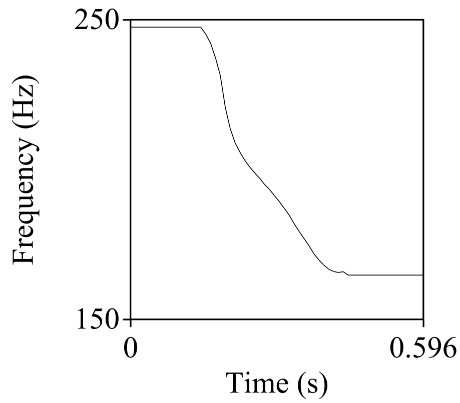
Nonwords were counterbalanced for their final vowel. Two lists were made: in the first list, each nonword was randomly assigned a final vowel so that half were [i] and half were [ə]. In the second list, each item appeared with the final vowel opposite that used in the first list. Participants were assigned one of the two lists at random. The lexical neighbourhood density of each nonword was measured using the Generalized Neighbourhood Model (Bailey and Hahn, 2001). All nonwords used in the experiment had a GNM value of less than 0.01, corresponding to very sparse neighbourhoods.

The isolated syllables were constructed in the following way: A female native speaker of American English (the author) read a list of individual syllables written in IPA. These recordings were then resynthesized in Praat (Boersma and Weenink, 2011) so that each vowel was approximately 400ms long, and faded into silence over the final 100ms. The pitch contour of the syllables was also resynthesized to be identical (a H* pitch accent followed by a H-H% boundary tone, shown in Figure 6). The intensity of the syllables was also normalized.

3.2.3 Procedure

The experiment was presented via the world wide web, using software built on Experigen (Becker and Levine, 2013). When participants arrived at the site, they were first asked to electronically sign a consent form, and then they completed a sound check to test that their microphone and speakers were working. Next, they were instructed that they would hear a sequence of three syllables, and that they should speak the whole word fluently as if it were a real word. They were given an example nonword sequence of syllables and two examples of those syllables strung together into a pseudoword - once with antepenultimate stress and once with penultimate stress. Next, they were given a sample trial which was a real word (they were told in advance that it would be a real word). In each trial of the experiment, they first heard the three syllables, then were asked to speak the word fluently, then they listened to the two stress options for that item and

Figure 6: Pitch contour for individual syllables presented to participants - all syllables were resynthesized to have this contour



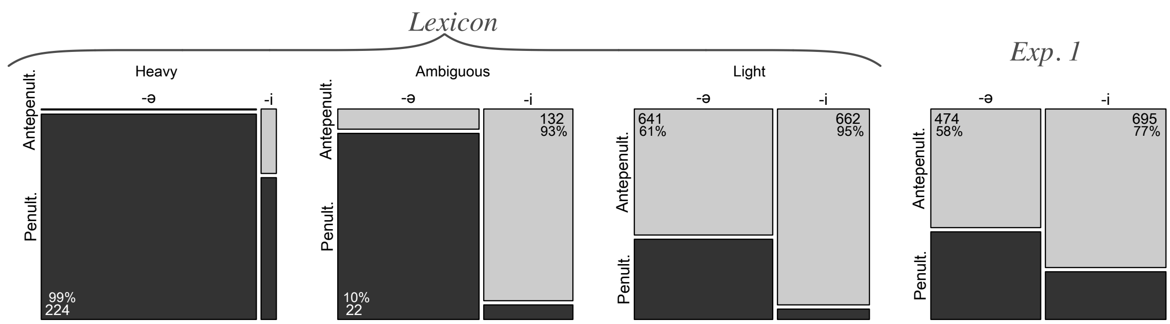
clicked a radio button to choose one. There were 32 nonword items and 8 real word items.

3.3 Results

Participants' success at the production task was assessed in two parts. First, did participants produce the syllables fluently together as a single word, with a single main stress? Second, did their produced stress agree with the stress they reported producing? For each participant, 10 (out of 32) nonword recordings were randomly selected. The author listened to these and annotated whether the production had a single stressed syllable or not, and transcribed the location of the main stress if it had one. Stress was transcribed based on vowel reduction and pitch. If a production had a full vowel in every syllable, or both of the first two syllables, it was counted as 'incorrect'. Also, if the production contained a pitch fall on any syllable but the last, or pauses between the syllables, it was counted as 'incorrect'. A participant was excluded from analysis if more than three of the examined 10 nonwords counted as 'incorrect'. Participants who did not successfully record any sound were also excluded. In total, 22 (out of 104) participants were excluded for these reasons. Two additional participants were excluded because they were not native speakers.

For the 'correct' productions, which followed the criteria of being a single prosodic unit in which at most one syllable bears main stress, participants' accuracy at reporting their own stress pattern was assessed. For 65 participants, their choice of stress pattern in the forced choice task agreed with the author's transcription of their produced stress at least 9 times out of 10. 15 Participants had less than 90% accuracy on the forced choice task, and were therefore excluded from analysis. Both the quality of the productions, and

Figure 7: Counts of stress choices for each final vowel in nonwords



the agreement between participants' productions and the forced choice task were held to relatively stringent standards because the assumption about this data is that it is production data, not perception data. If a participant did not produce the items in a word-like fashion, then their choice of stress patterns in the forced choice task cannot be based on the action of the production system. Likewise, if a participant produced one stress on a form, but chose another after hearing both options, then their choice is not based on the action of their production system, but rather on a combination of the perception of each stress and a high-level judgment.

Figure 7 shows the counts of each type of stress response for each type of final vowel. These counts are responses in the forced choice task, but recall that participants are only included if their choice of stress agreed with their produced stress at least 90% of the time. This means the counts can be thought of as production counts. Additionally, trials in which participants failed to listen to both stress options before responding were excluded.

Overall, participants preferred antepenultimate stress for both ə-final nonwords and i-final nonwords, but this preference was slight in the ə-final case, and relatively strong in the i-final case. This matches the distribution found in the lexicon, where ə-final words exhibit a much weaker preference, tending strongly towards penultimate stress when the penult is ambiguous or heavy, and weakly toward antepenultimate stress when the penult is light.

A mixed effects logistic regression was fitted to this data, with produced stress as the dependent variable, coded as 0 for antepenultimate stress and 1 for penultimate stress. Final vowel was the predictor, and random slopes and intercepts were included for both subjects and items. Negative coefficients indicate a preference for antepenultimate stress. There was a slight preference for antepenultimate stress over penultimate stress over all (Intercept=-0.35, $p=0.049$), and a stronger preference when the final vowel was [i] ($\beta = -1.27$, $p<0.001$).

4 Experiment 2 - effects of penult weight

In order to test for Experiment 2 was similar in methodology to Experiment 1. In this case, items' final vowel was manipulated (half were [i], half [ə]) just as in Experiment 1, but additionally the weight of items' penultimate syllable was manipulated[4].

4.1 Participants

Like Experiment 1, Experiment 2 was presented via the world wide web, and participants were recruited through Amazon Mechanical Turk. All participants were over 18 years of age, and had an age range of 18 to 75 (mean age: 34). Exclusion criteria and pay were the same as for Experiment 1. Data was collected from a total of 101 participants, and data from 38 participants was used.

4.2 Items

The nonword items from Experiment 1, all three syllables long and consisting of only light syllables, constituted the light-penult condition of the experiment. From each of these, a heavy-penult item was constructed by adding a coda to the penultimate syllable. Codas were chosen so that they did not form a legal onset cluster with the onset of the following syllable.

Table 4: Example item in four conditions.

	Penult Weight	
Final Vowel	Light	Heavy
-ə	[pǽ] [lǽ] [kə]	[pǽ] [lǽz] [kə]
-i	[pǽ] [lǽ] [ki]	[pǽ] [lǽz] [ki]

Each item consisted of a written frame sentence[5], followed by the same item structure as in Experiment 1: three auditorily presented individual syllables, and two auditorily presented versions of the full word, with different stress patterns (antepenultimate and penultimate). The same real words used in Experiment 1 were also included.

[4]This experiment also tested for a difference in preferred stress pattern based on a word's part of speech. Studies such as Guion et al. (2003); Domahs et al. (2014); Kelly and Bock (1988) found that speakers prefer initial stress more strongly on two-syllable nouns than on two-syllable verbs. Sonderegger and Niyogi (2013) find evidence for the same pressure in historical records of changes in stress patterns on noun-verb pairs. However, no effect was found for three-syllable words, so this manipulation is not discussed in this paper.

[5]Chosen to manipulate whether the word was understood as a noun or as a verb - no effect of this manipulation was found, and the results are not discussed here.

4.3 Procedure

The procedure was identical to that in Experiment 1, including the number of trials (32 nonwords, 8 real words). Items were counterbalanced so that each participant saw half heavy penults, half light, and half i-final, half ə-final (balanced so that each final vowel x penult weight condition had equal representation).

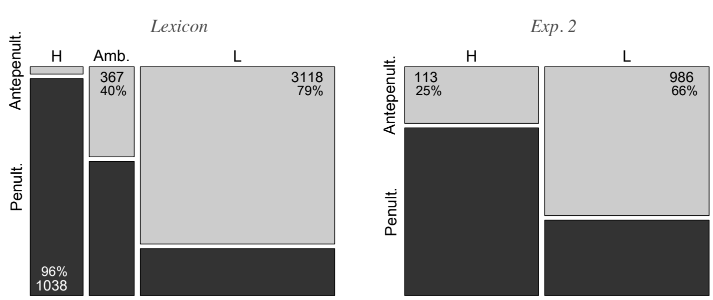
5 Results

As in Experiment 1, participants were included based on their ability to complete the tast correctly, and based on the quality of recordings. Because of a technical problem, a large number of participants - 34 in total, did not record any sound. A total of 38 participants remained after all exclusions.

These 38 participants were all greater than 75% accurate at reporting their produced stress. Participants did not always accurately produce a heavy penultimate syllable when one was present in the three-syllable prompt. On 21% of trials with codas on the penultimate syllable of the prompt, the participant left out the coda, producing a light penult instead. On 6% of trials with a light penult in the prompt, the participant produced a coda on the penult. Because of the relatively high level of mismatch between the prompts and participants' productions, the weight (transcribed by the author) of the actual production is used for analysis, rather than the weight of the penultimate syllable of the prompt.

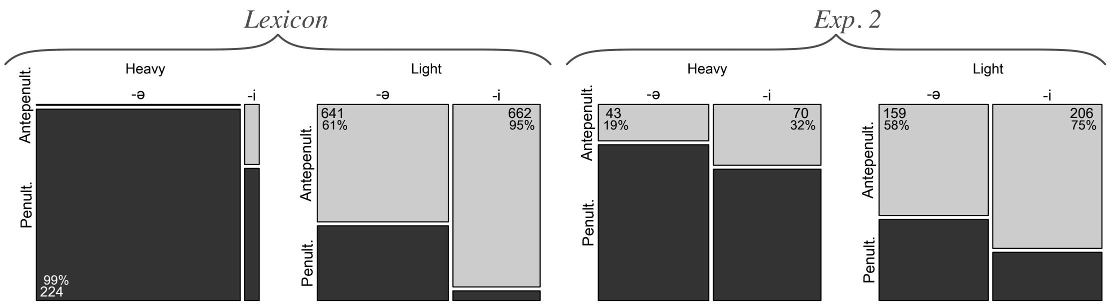
Participants modulated their choice of stress pattern based on the weight of the penultimate syllable. They preferred penultimate stress when the penult was heavy, and antepenultimate stress when the penult was light. In the lexicon 95%-97% of words with a heavy penultimate syllable receive penultimate main stress. Participants under-matched this distribution, producing penultimate stress on words with heavy penults 75% of the time. Figure 8 illustrates this. Note that because participants more often erred towards producing light penults than heavy ones, there are more responses overall in the light penult category.

Figure 8: Participants' stress choices by weight



As in Experiment 1, participants produced antepenultimate stress more when the final vowel was [i] than when the final vowel was [ə]. This effect was independent of the weight of the penultimate syllable - both within heavy items and within light items, i-final items are produced with more antepenultimate stress than ə-final items. This is illustrated in Figure 9.

Figure 9: Participants' choice of stress pattern for different final vowels, broken down by weight category. Regardless of the weight of the penult, participants prefer antepenultimate stress more for i-final items than for ə-final items.



A mixed-effects logistic regression was fitted with Final Vowel and Penult Weight and their interaction as fixed effects, and with random intercepts for participants and for items[6]. When the final vowel was [i], items were significantly less likely to receive penultimate stress than when the final vowel was [ə] (β=-0.96, p<$6x10^{-6}$). When the penultimate syllable was closed by a consonant, items were significantly more likely to receive penultimate stress than when the penultimate syllable was open (β=2.11, p<$2x10^{-16}$). The interaction between the two was not significant (β=0.09, p=0.77).

Table 5: Logistic regression with two factors. Produced stress as a function of penult weight and final vowel

Model: *Produced Stress* ∼ *Final Vowel* + *Penult Weight*

	Estimate	p
Intercept	-0.34	0.119
Final Vowel = i	-0.96	$5.26x10^{-6}$
Penult = H	2.11	<$2x10^{-16}$
Penult = H X Final Vowel = i	0.09	0.78
AIC:	1119	

Participants observed both the trend for i-final words to take antepenultimate stress and the trend for words with heavy penultimate syllables to take penultimate stress. The weight of the penultimate syllable had a stronger effect on stress placement than the final vowel (the magnitude of the coefficient on penult

[6]Random slopes were initially included in the model as well, but the model did not converge.

weight is about twice the magnitude of the coefficient on final vowel), but a heavy penult does not completely trump a final [i]. In fact, the two effects are statistically independent. The quality of the final vowel matters equally in the heavy penult case as in the light penult case.

5.1 Discussion

The results of the two experiments showed that participants have active knowledge of both the trend in the lexicon for heavy penults to be stressed (the Latin Stress Rule), and the trend for i-final words to take antepenultimate stress. There is very little information in the lexicon about how these two trends interact. In particular, very few words with a heavy penult and a final [i] exist in the lexicon. Participants have chosen the option which on the surface seem simplest: that these two trends act independently of each other, both applying equally to every form. This independence is visible in the experiment because of another phenomenon - namely that both trends severely undershoot in the experiments their strength in the lexicon. 95% of i-final words with light penults take antepenultimate stress in the lexicon, while participants produced antepenultimate stress on just 77% of them in Experiment 1, and 75% in Experiment 2. Likewise, words with heavy penults take penultimate stress 95-97% of the time in the lexicon, while participants produced penultimate stress on these items merely 81% of the time for ə-final items, and 68% of the time for i-final items.

The trend for heavy penults to be stressed (the Latin Stress Rule) has been tested with nonce words elsewhere in the literature: Olejarczuk and Kapatsinski (forthcoming) and Domahs et al. (2014) are two recent examples. Olejarczuk and Kapatsinski found, as Experiment 2 did, that participants produced more penultimate stress on heavy-penult words than on light-penult words, but the mean proportion of penultimate stress on heavy-penult words was still very low - less than 50%. Many of the items in this study did use sonorant codas in the penultimate syllable to make it heavy, as in [ma.dal.paz]. I have argued in section 2.2 that rhymes like [al] should be considered ambiguous, since they will be heavy when stressed, but reduced to a syllabic sonorant, and therefore light, when unstressed. This does not completely explain the discrepancy between their results and the results of Experiment 2 however, since at least according to the statistics of CMU, penults with ambiguous weight should mostly still receive penultimate stress unless the word's final vowel is [i]. On the other hand, Domahs et al. (2014) found penultimate stress nearly 100% of the time on items with heavy penults (40% of the time for light-penult items). Both experiments differ in methodology from Experiments 1 and 2, most notably in that the stimuli are presented orthographically in these experiments, rather than auditorily as in Experiments 1 and 2. In Domahs et al. (2014), many different weight patterns were tested in the same experiment, rather than just the two patterns in Experiment 2 and in Olejarczuk and Kapatsinski (forthcoming). All three experiments were production experiments in that participants either read or heard words with no stress pattern specified, and produced a stress pattern for them. It is unclear what effect differences among these methodologies might have on experiment outcomes, but more investigation is in order.

In the next section, I present a model of the effects of penult weight and final vowel in the English stress system. I will argue based on the results of Experiment 2 for the use of a true extrametricality constraint, which completely excludes a final [i] from the metrical parse. While such constraints are not typical in OT-models of stress systems, and NonFinality constraints are typically used instead, I will demonstrate that a NonFinality constraint in this case fails to capture participants' behaviour in the experiments. I introduce the constraint NonParse, which forces a final [i] to be extrametrical, or parsed outside of the prosodic word.

6 Modeling English stress

This section will first describe a model of the quantity-sensitivity in longer words of English: words with a light penultimate syllable weakly tend to take antepenultimate stress, and words with a heavy penultimate syllable strongly tend to take penultimate stress. Next, the effects of the final vowel will be incorporated. Both when the penult is light and when it is heavy, words with final [i] take more antepenultimate stress than words with a final [ə] (or, it is assumed, other vowels). Finally, the behaviour of words with ambiguously weighted penults as described in section 2.2 will be addressed. These words tend to take penultimate stress in the lexicon, but not as strongly as words with truly heavy penults. Also, their behaviour is extremely dependent on the final vowel - when the final vowel is [i], they almost always take antepenultimate, rather than penultimate stress. Words like 'álligàtor' or 'cápillàry', in which a final [i,l̩,ɹ̩] forces stress even farther to the left than the antepenult, will not be modelled in this paper. Their behaviour is mostly predictable from the analysis given below (and crucially also require the NonParse constraint), but difficulties of secondary stress placement crop up that are beyond the present scope.

The model presented here builds on the analyses of primary stress placement given in Pater (2000), and Alcántara (1998), both of which use the constraints Align-R, NonFinality, FootBinarity, and Trochee. To simplify the analyses in this paper, Trochee, the constraint which demands that feet be left-headed, is assumed to have a very high weight, and only candidates which satisfy it are considered.

These constraints are sufficient to describe the weight effects observed in CMU and in Experiment 2, but more is needed to describe the effects of the final vowel. Specifically, I propose a constraint, NonParse-i which requires that a final [i] be parsed outside of the prosodic word. This constraint mimics the effects of Extrametricality rules in analyses like Liberman and Prince (1977); Hayes (1982). In candidates where this constraint is satisfied, Align and Nonfinality assign violations according to the prosodic word edge, which is now inside the word. This causes main stress to surface further left than it would in words without a final [i]. An alternative analysis using a version of Nonfinality specific to i-final words (NonFinality-i), is considered, but is insufficient to account for participants' behaviour in Experiment 2.

Finally, because stress assignment in English is inherently probabilistic, the analysis is incorporated into a Maximum Entropy (MaxEnt, Goldwater and Johnson, 2003) grammatical framework. This framework is used to model the lexical statistics of English as well as participants' behaviour in Experiments 1 and 2.

6.1 The Latin Stress Rule

Main stress in English typically occurs within a three-syllable window at the right edge of the word. This is predicted by the joint action of an alignment constraint demanding that main stress be as close as possible to the right edge of the word (Align-R), and a nonfinality constraint demanding that the final syllable of a word be unfooted (Foot-NonFinality) (Pater, 2000, p. 240, Alcántara, 1998, p. 120-121).

(2) Align-Head-R (Align-R): Assign a violation for every syllable intervening between the right edge of the word and the right edge of the head foot.

(3) Foot-NonFinality (NonFin): Assign a violation if the final syllable of the word is parsed into a foot.

Table 6: Violations assigned by Align-R and NonFin to several candidate stress patterns.

/σσσσ/	NonFin	Align-R
→ a. $\sigma\sigma\sigma(\acute{\sigma})$	1	
→ b. $\sigma\sigma(\acute{\sigma}\sigma)$	1	
→ c. $\sigma\sigma(\acute{\sigma})\sigma$		1
→ d. $\sigma(\acute{\sigma}\sigma)\sigma$		1
e. $\sigma(\acute{\sigma})\sigma\sigma$		2
f. $(\acute{\sigma}\sigma)\sigma\sigma$		2
g. $(\acute{\sigma})\sigma\sigma\sigma$		3

In Table 6, candidates e.-g. violate Align-R too many times and are harmonically bounded by candidates c. and d. Depending on the ranking of NonFin and Align-R, candidates a.-d. are all plausible surface candidates. Parses like a., and c., with a single-syllable foot, are only possible when that foot is heavy. This requirement is enforced by the FtBin constraint.

(4) Foot-Binarity (FtBin): Assign a violation to any foot which is not binary (does not contain exactly two moras).

In Hayes (1995), English is analyzed with moraic trochees, which can consist of a single heavy syllable (H́) or two light syllables (ĹL) but not of a single light (*Ĺ) or a heavy plus a light (*ĹH, *H́L). Foot-Binarity penalizes these disallowed foot shapes. Kager (1999) discusses several examples of other stress patterns for which such a constraint is necessary.

If this constraint is high enough ranked, parses such as a., c., and e. in Table 6 would only be allowed if the syllable bearing the foot is heavy, whereas parses like b. and d. would only be allowed if the two syllables in the foot were both light. Words of shape LHL could only be parsed with a foot on the penultimate syllable: L(H́)L, *(Ĺ)HL, *(ĹH)L. Words with only light syllables, such as those in the experiment, could only have penultimate stress with a right-aligned trochee: L(ĹL), as in b., or antepenultimate stress: (ĹL)L as in d. Penultimate stress would be preferred when Align-R $\gg$ NonFin, and antepenultimate stress would be preferred when NonFin $\gg$ Align-R.

These three constraints, Align-R, NonFin, and FtBin, are already enough to represent the Latin Stress rule. Unlike many languages' stress-weight interactions, English seems to only care about the weight of a syllable when it is final or penultimate. The weight effects of the final syllable are not discussed here, but he interaction of these three constraints explains why the weight of the penult in particular matters: With FtBin highly ranked, an LHL word can only be footed as L(H)L, since all other possible feet are either too big or too small. If the word has only light syllables, or if the word has more than one heavy syllable, the interaction of Align-R and NonFin ensures that the rightmost nonfinal foot will always be selected.

Table 7 illustrates the interaction of these three constraints. If the penultimate syllable is heavy, then a moraic trochee on the penultimate syllable will always be optimal, as the penult is perfectly placed to be aligned as far right as it can without violating NonFinality. If the penultimate syllable is light, then antepenultimate main stress will always be optimal. If the penult and antepenult are both light, then the winning candidate will have a moraic trochee over those two syllables, as in candidate a. below. If the antepenult is heavy, then the winning candidate will be one with a moraic trochee on that heavy antepenult, as in candidate g. below. Even though this syllable is now two syllables away from the right edge of the word, and therefore violates Align-R twice, it is still better than the candidate with a larger-than-binary foot, (HL)H, or the candidate with final stress HL(H).

Table 7: OT grammar illustrating the emergence of the Latin Stress rule from the three constraints Align-R, NonFin, and FtBin.

	FtBin	NonFin	Align-R
/LLL/			
→ a. (LL)L			1
b. (L)LL	1!		2
c. L(LL)		1	
d. L(L)L	1!		1
e. LL(L)	1!	1	
/HLH/			
f. (HL)H	1!		1
→ g. (H)LH			2
h. H(LH)	1!	1	
i. H(L)H	1!		1
j. HL(H)		1!	
/LHL/			
k. (LH)L	1!		1
l. (L)HL	1!		2
m. L(HL)	1!	1	
→ n. L(H)L			1
o. LH(L)	1!	1	
/HHH/			
p. (HH)H	1!		1
q. (H)HH			2!
r. H(HH)	1!	1	
→ s. H(H)H			1
t. HH(H)		1!	

6.2 Syllables with ambiguous weight

As discussed above in section 2.2, syllables can have ambiguous weight in two ways. First, stressed syllables can attract consonants that would otherwise belong to the onset of the following syllable. These consonants then become a coda of the stressed syllable, making it heavy (Kahn, 1976). Second, syllables with a coda liquid or syllabic liquid can alternate between these two states depending on their stress (for example, 'consultant' [kn̩.**sʌ́l**.tn̩t] ∼ 'consultation' [kɑn.**sl̩**.te͡i.ʃn̩]. Both types of syllables have a coda when stressed, and are therefore heavy, but have no coda when stressless, and are therefore light. The dictionary search in section 2.2 showed that these syllables tend to take penultimate stress, except when the final vowel is [i], in which case they pattern with words with light penults in taking mostly antepenultimate stress.

In order to incorporate the behaviour of these syllables into the analysis of Latin Stress, it will be necessary to use the constraint Stress-to-Weight:

(5) Stress-to-Weight (STW): Assign a violation to every stressed syllable which is not heavy.

This constraint does not penalize heavy stressless syllables, but it does penalize light stressed syllables. We will first see how it applies in the case of words like 'digestive', [da͡îi.**ʤɛ́s**.təv], and 'tapestry' [tæ.**pə**.stɹi]. Words like this in the lexicon strongly tend to take penultimate stress - the stress and syllabification of 'tapestry' is an exception. Table 8 shows how STW combined with the previously discussed constraints captures the generalisation that words of this type generally surface with penultimate stress. This is illustrated with the nonce word 'pamastra'. STW outranks NoCoda, the constraint which otherwise forces the s to be part of the onset of the final syllable rather than the coda of the penult. Candidate d., in which the penultimate syllable has absorbed the [s] and become heavy, satisfies FtBin, NonFin, STW, and also Align-R as much as possible. It beats out candidate a., which would be the correct form if there were no onset cluster in the final syllable, because candidate a. violates STW.

Table 8: The constraint Stress-to-weight (STW) outranks NoCoda, causing stressed syllables to attract consonants into their codas.

	FtBin	NonFin	Align-R	Onset	STW	NoCoda
/pamastra/						
a. (pa.ma).stra			1		1!	
b. (pa.mas).tra	1!		1		1	1
c. pa.(ma).stra	1!		1		1	
→ d. pa.(mas).tra			1			1
e. pa.(ma.stra)		1!			1	
f. pa.ma.(stra)	1!	1			1	
g. pa.mas.(tra)	1!	1			1	1
/pamata/						
→ h. (pa.ma).ta			1		1	
i. pa.(mat).a			1	1!		1

The final two candidates of Table 8 show that when an input only has one consonant between the final two vowels, the penultimate syllable cannot absorb that consonant in order to be heavy. That would mean sacrificing the onset of the final syllable, in addition to incurring a violation of NoCoda in the penultimate syllable. Because ONSET outranks STW, the antepenultimately stressed candidate wins.

Let us now turn to the analysis of ambiguous penults where the penultimate syllable consists of a syllabic sonorant, or of a vowel followed by a coda sonorant. When these syllables are in stressed position, they must surface as a vowel plus a sonorant, as in consúltant [kn̩.**sʌ́l**.tn̩t], but when they are in stressless position, they must surface as a syllabic sonorant, as in cònsultátion [kɑn.**sl̩**.te͡i.ʃn̩]. Again, these are syllables which are heavy when stressed, and light when unstressed. Table 9 illustrates the effects of STW on syllables like this. If STW outranks MAX-V and DEP-V as in this example, then words with syllables like this will take penultimate stress as if they had an unambiguously heavy penultimate syllable.

Table 9: The constraint STRESS-TO-WEIGHT (STW) outranks MAX-V and DEP-V, causing stressed syllables to surface with a coda liquid rather than a syllabic one.

		FTBIN	NONFIN	ALIGN-R	STW	MAX-V	DEP-V
/pamalta/							
a.	(pa.mal).ta	1!		1	1		
b.	(pa.ml).ta			1	1!	1	
→ c.	pa.(mal).ta			1			
d.	pa.(ml).ta	1!		1		1	
e.	pa.(mal.ta)	1!	1!				
f.	pa.(ml.ta)		1!		1	1	
g.	pa.mal.(ta)	1!	1		1		
/pamlta/							
a.	(pa.mal).ta	1!		1	1		1
b.	(pa.ml).ta			1	1!		
→ c.	pa.(mal).ta			1			1
d.	pa.(ml).ta	1!		1			
e.	pa.(mal.ta)	1!	1!				1
f.	pa.(ml.ta)		1!		1		
g.	pa.mal.(ta)	1!	1		1		1

For this analysis, it does not matter whether the penultimate syllable underlyingly has a vowel, or has no underlying vowel. Either way, STW will force it to surface as the heavy version, and the entire word to take penultimate stress[7].

[7]The ranking STW $\gg$ DEP-V would wind up forcing any short vowel to be lengthened in stressed syllables. Since we do not observe lengthening of this type in English - the language seems perfectly content with short stressed vowels, as in 'ápple' - this must not be exactly the correct analysis of the liquid alternation either. A complete analysis of the alternation between [kn̩.**sʌ́l**.tn̩t], and [kɑn.**sl̩**.te͡i.ʃn̩] would have to account also for (a) why a syllabic l is allowed in stressless but not stressed position in the first place, (b) whether words like 'consultant ∼ consultation' really have an underlying syllabic l or an underlying vowel, or whether it is different for different words, and (c) if words like 'consultant' have an underlying syllabic l, what determines the quality of the inserted vowel? This complete analysis is beyond the scope of the paper, but presumably the basic idea of the analysis above - that STW exerts influence over whether the liquid surfaces as a coda or as a nucleus - could still be incorporated.

6.3 Making the grammar probabilistic

Stress in English is an inherently probabilistic system. While generalisations such as the Latin Stress Rule certainly explain the stress patterns of many words of English, exceptions abound. The results of Experiments 1 and 2 demonstrate that speakers have access to probabilistic generalisations about their stress system, and are able to extend those to new words. Interestingly, speakers produce somewhat less extreme distributions in the experiments than what is actually observed in the lexicon: While words with heavy penults take penultimate stress 95% of the time in the lexicon, they take penultimate stress just 75% of the time in Experiment 2. Likewise, words with light penults and a final [i] in the lexicon take antepenultimate stress 96% of the time in morphologically complex words, and 86% of the time in morphologically simple words, but they take antepenultimate stress just 77% of the time in Experiment 1, and 75% of the time in the light-penult condition of Experiment 2. Following previous research on speakers' probabilistic knowledge of the patterns in their language (Zuraw, 2000; Hayes et al., 2009; Becker et al., 2011; Zuraw, 2010; Gouskova and Becker, 2013), this paper will assume that the distribution speakers actually produce in the experiment reflects their grammatical knowledge more or less directly.

The model proposed in this section aims first of all to capture the basic trends in the lexicon reported in section 2.2. The preference for heavy penultimate syllables to take main stress, and the preference for words ending in [i] to take antepenultimate stress will be modelled in this section, as well as the behaviour of ambiguously-weighted penultimate syllables. When words with an ambiguously-heavy penult end in a [ə] or a nasal, they pattern with heavy-penult words and strongly prefer penultimate stress. However, when they end in a final [i] specifically, they strongly prefer antepenultimate stress, patterning with light-penult words. Speakers' behaviour on i-final and ə-final words with both heavy and light syllables can be modelled with some precision based on the data from Experiments 1 and 2. However, because speakers' behaviour on ambiguously-weighted penults has not been tested, these trends will be modelled with somewhat less precision.

In order to represent the probabilistic knowledge that English speakers seem to have, I use Maximum Entropy (MaxEnt) grammar (Goldwater and Johnson, 2003; Hayes and Wilson, 2008), a version of Harmonic Grammar (Smolensky and Legendre, 2006; Pater, 2009) which uses the logit transform (also found in logistic and multinomial regressions) to convert Harmony scores into a predicted probability distribution over candidates. MaxEnt grammar, like OT, relies on a universal constraint set, but rather than having the constraints strictly ranked with respect to each other, they are assigned weights. Table 10 shows the same constraints as in Table 7, but now they are assigned weights instead of being ranked. Harmony scores ($\mathcal{H}$)

are calculated by multiplying each weight by each violation and summing, as in (6). Probabilities are then calculated by exponentiating this harmony score and summing over that exponentiated harmony score for all candidates.

(6) MaxEnt Grammar calculations

$$\mathcal{H} = -\sum weight_i * violation_i$$

$$p = \frac{e^{\mathcal{H}}}{\sum e^{\mathcal{H}}}$$

In Table 10, some variation is now introduced: When the penultimate syllable is light, this set of weights predicts antepenultimate stress 81% of the time, and penultimate stress 18% of the time. When the penultimate syllable is heavy, penultimate stress is predicted the majority of the time, but variation emerges based on the detailed weight structure of the word. LHL words take penultimate stress the vast majority of the time, while HHH words vary between penultimate, antepenultimate, and final. This is because in HHH words, there are three configurations (candidates l., n., and o.) which all satisfy FtBin, while in LHL words there is only one. In this configuration, FtBin has a high enough weight to basically rule out candidates which violate it. The penultimate stress case is still preferred in HHH words because it is as right-aligned as possible while still satisfying NonFin. These probability distributions all roughly match the statistics in the CMU pronouncing dictionary, shown in the leftmost column of Table 10.

The final analysis of the English stress system presented in this paper will attempt to match the actual probabilities which participants produced during Experiments 1 and 2, rather than the probabilities in the lexicon. For that, the effects of the final vowel must be taken into account.

6.4 Final [i] vs. final [ə]

The statistics of the lexicon, as well as the results of Experiments 1 and 2 show that in English, long words with final [i] tend to take antepenultimate rather than penultimate stress. The final vowel influences the stress pattern of a word, shifting its probability of taking penultimate stress whether the penult is heavy or light. To capture this basic generalisation, that words with final [i] tend to take stress farther to the left than words with final [ə], I propose the following constraint:

(7) NonParse-i A final [i] should not be parsed into the prosodic word.

This constraint forces a final [i] to be ignored during the assignment of violations from all other metrical

Table 10: The probabilistic preference for penultimate stress on heavy penults, but antepenultimate stress on light penults, modelled with a MaxEnt grammar and the three constraints ALIGN-R, NONFIN, and FTBIN.

CMU %		P	$\mathcal{H}$	FTBIN 5	NONFIN 3.5	ALIGN-R 2	← *weights*
	/LLL/						
74%	→ a. (LL)L	.81	-2			1	
	b. (L)LL	0	-9	1		2	
20%	→ c. L(LL)	.18	-3.5		1		
	d. L(L)L	.01	-7	1		1	
4%	e. LL(L)	0	-8.5	1	1		
	/LHL/						
5%	f. (LH)L	.01	-7	1		1	
	g. (L)HL	0	-9	1		2	
	h. L(HL)	0	-8.5	1	1		
93%	→ i. L(H)L	.99	-2			1	
2%	j. LH(L)	0	-8.5	1	1		
	/HHH/						
	k. (HH)H	.01	-7	1		1	
25%	→ l. (H)HH	.10	-4			2	
	m. H(HH)	0	-8.5	1	1		
62%	→ n. H(H)H	.73	-2			1	
12%	→ o. HH(H)	.16	-3.5		1		

constraints. For example, a parse like {(cana)}ry, where { } indicate the prosodic word edges, would incur zero violations of Align-R and one violation of NonFin. On the other hand, {(cana)ry}, with the final syllable parsed into the prosodic word, would incur no violations of NonFin, and one violation of Align-R, and one violation of NonParse-i. Although the statement of the constraint refers only to the final [i] itself, it is assumed throughout this paper that the entire syllable with the [i] as nucleus would be parsed outside of the prosodic word. Although this is not a crucial assumption for the analysis that follows, it is a simplifying one.

The action of this constraint is illustrated in Table 11. In these tableaux, a new set of weights are used, different from those in Table 10 above. This new set of weights is chosen to more closely match participants' behaviour in Experiments 1 and 2, and will be used in the rest of the paper. With this new set of weights, multiple parses resulting in the same main stress are now possible. For example, candidates k. and l. are both antepenultimate stress, but candidate k. has a trochee on the first two syllables and candidate l. has a subminimal foot on the first syllable. This is because FtBin now has a much lower weight of 2, rather than 5 as it was in Table 10. All three constraints that were necessary to model the probabilities in the lexicon are also necessary to model the experimental results, but their weights are lower, especially FtBin. NonParse-i, on the other hand, has a relatively high weight of 4, which is high enough to almost completely exclude candidates like f.-j., which have a final [i] parsed into the prosodic word.

When the [i] is excluded from metrical structure, NonFin assigns violations differently than it does in the non-i form. This means that candidate l. gets some probability, even though it's corresponding form without the final [i] gets very little. Candidates l. and k. are both antepenultimately stressed candidates, and their summed probability adds up to 0.89, leaving a probability of 0.10 for candidate m., the penultimately stressed candidate. This candidate violates NonFin and FtBin, but not NonParse-i. In the ə-final case (LLL), candidate a. gets the most probability, at 54%, but the penultimately stressed candidate c. gets a healthy amount of probability as well. These probabilities mirror participants' weak but actual preference for antepenultimate over penultimate stress in the experiments.

The NonParse-i constraint is somewhat unusual in OT literature, since in general extrametricality rules (e.g. Hayes, 1982, 1980) were translated into NonFinality constraints in OT (Prince and Smolensky, 1993/2004, section 4.2.2, section 4.3). The NonFinality family of constraints does not do exactly the same work as the extrametricality rules did, but they seem to cover the same empirical ground, in some cases more cleanly (Hyde, 2011, 2007). Two basic types of NonFinality constraints are typically used: a constraint which forbids the main stress of the word from falling on the final syllable ((45) in Prince and Smolensky,

Table 11: NonParse-i forces all-light words with final [i] to take majority antepenultimate stress (89% total), while LLL words that do not end in [i] take only slightly more antepenultimate than penultimate stress - 56% vs. 40%.

	P	$\mathcal{H}$	FtBin 2	NonFin 2	Align-R 1.5	NonParse-i 4
/LLL/						
→ a. { (LL)L }	.54	-1.5			1	
b. { (L)LL }	.02	-5	1		2	
→ c. { L(LL) }	.33	-2		1		
d. { L(L)L }	.07	-3.5	1		1	
e. { LL(L) }	.04	-4	1	1		
/LLi/						
f. { (LL)i }	.02	-5.5			1	1
g. { (L)Li }	0	-9	1		2	1
h. { L(Li) }	.01	-6		1		1
i. { L(L)i }	0	-7.5	1		1	1
j. { LL(i) }	0	-8	1	1		1
→ k. { (LL) } i	.71	-2		1		
→ l. { (L)L } i	.16	-3.5	1		1	
→ m. { L(L) } i	.10	-4	1	1		

1993/2004), and the constraint that has been used so far in this paper, forbidding the final syllable from being parsed into the head foot of the word ((53) in Prince and Smolensky, 1993/2004). The extrametricality rules in popular use before OT kept a final syllable from being considered in any kind of metrical calculations - an even-parity word would be treated like an odd-parity word for example. So far languages requiring this kind of extrametricality have not surfaced. So, NonFinality makes an attractive alternative to extrametricality in OT since it provides a tighter fit to the typology than a family of NonParse constraints presumably would.

However, for the interaction observed between final [i] and weight of the penult in Experiments 1 and 2, NonFinality is insufficient. Table 12 demonstrates this, using a version of NonFinality which specifically penalizes a final [i] that is parsed into a foot. For forms without a final [i], the same weight set results in a set of predicted probabilities which are a pretty good match to participants' behaviour in Experiment 2. All together, penultimately stressed forms receive 12% of the probability, compared with 19% in the experiment. Antepenultimately stressed forms receive a total of 81% of the probability, which perfectly matches participants' 81% in the experiment. However, for i-final forms, this model does not match participants' behaviour. Participants gave forms with a final [i] more antepenultimate stress, and less penultimate stress - they gave penultimate stress 68% of the time. In Table 12, however, LHL words with final [i] get *more* penultimate stress. Because the most probable parse for forms with a heavy penult is L(H)L, which already perfectly satisfies NonFinality, adding an additional NonFinality-i constraint only serves to heap more probability onto this candidate, driving away the small amounts of probability that would otherwise be given to parses like LH(L), and L(HL). No weighting of this set of constraints will result in a *decrease* in probability of penultimate stress for LHL words with final [i], like participants' exhibited in Experiment 2.

Table 12: NonFin-i, which specifically assigns violations to final [i]'s which are parsed into a foot, is insufficient to model participants' behaviour in Experiment 2.

	P	$\mathcal{H}$	FtBin 2	NonFin 2	Align-R 1.5	NonFin-i 4
/LHL/						
→ a. (LH)L	**.10**	-3.5			1	
b. (L)HL	.02	-5	1		2	
c. L(HL)	.06	-4		1		
→ d. L(H)L	**.75**	-1.5	1		1	
e. LH(L)	.06	-4	1	1		
/LHi/						
→ f. (LH)i	**.12**	-3.5			1	
g. (L)Hi	.03	-5	1		2	
h. L(Hi)	0	-8		1		1
→ i. L(H)i	**.86**	-1.5	1		1	
j. LH(i)	0	-8	1	1		1

Table 13, with NonParse-i, provides a much better match for participants' behaviour in Experiment 2 than does table 12. With NonParse-i, candidates f.-h., in which the final [i] is not parsed into the prosodic word get the vast majority of the probability. Among these candidates, the one with penultimate main stress, {L(H)}i, still gets most of the probability, but the parse in candidate g., with a subminimal foot on the leftmost syllable, now gets some probability since it is the only viable candidate that satisfies NonFin. In Table 13, the constraint weights predict a total probability of penultimate stress of 73%. This is not far off participants' behaviour in Experiment 2: they produced 68% penultimate stress.

Next, let us turn to the words in CMU with ambiguously-weighted penultimate syllables. In section 2.2, words like 'digestive' [da͡i.**ʤɛ́s**.təv], and 'tapestry' [tæ.**pə**.stɹi] exhibit an interesting behaviour (Figure 4). When they end in a [ə] or a syllabic nasal, they take mostly penultimate stress. This was modelled above using the constraint Stress-To-Weight (STW), which outranked NoCoda, in the tableaux in Table 8. The first consonant of a complex onset gets appended as a coda in the penultimate syllable, in order to allow that syllable to be a heavy stressed syllable, satisfying STW. On the other hand, when words like this end in a final [i], they exhibit an opposite preference - they take mostly antepenultimate stress. Because

Table 13: NonParse-i forces words with a heavy penult and a final [i] to take more antepenultimate stress than they would normally, by excluding [i] from the metrical parse.

	P	$\mathcal{H}$	FtBin 2	NonFin 2	Align-R 1.5	NonParse-i 4
/LHi/						
→ a. { (LH)i }	0	-7.5			1	1
b. { (L)Hi }	0	-9	1		2	1
→ c. { L(Hi) }	.02	-8		1		1
d. { L(H)i }	0	-5.5	1		1	1
e. { LH(i) }	0	-8	1	1		1
→ f. { (LH) } i	**.10**	-4		1		
g. { (L)H } i	**.16**	-3.5	1		1	
h. { L(H) } i	**.71**	-2	1	1		

the penultimate syllables in these words can become light or heavy without a high cost, they are more able to felicitously take antepenultimate stress when encouraged to because the final [i] is not parsed into the prosodic word.

The probabilistic version of Table 8 is given below in Table 14. The weightings of FtBin, NonFin, and Align-R come from the previous three tables. They are weights that closely predict participants' behaviour in the experiments. NonParse-i is not included in this tableau because the form does not end in [i] - Parse-σ, with nothing militating against it, forces the final schwa to be parsed into the prosodic word. A second NonFin constraint is now included:

(8) Stress-NonFinality (Stress-NonFin): Assign a violation when the final syllable of the word is stressed.

Stress in English strongly disprefers to occur on the final syllable in general. Most exceptions are two-syllable words like 'guitár', or words with a stressed suffix like 'àbsentée'. The NonFin constraint that has been used so far in this paper, which requires final syllables not to be parsed into a foot, partially explains this. Syllables not parsed into a foot can never be stressed. This constraint specifically requiring final syllables not to be stressed is necessary to explain why two-syllable LH words are often stressed on the first syllable despite the preferences of Align-R and FtBin. Words like 'állỳ' make up about 32% of LH words in CMU. For words with ambiguously weighted penults and a final schwa, like 'pamastra', this constraint only serves to make the candidates with final stress, e. and j., even worse. This is a good prediction as three syllable words with final stress in English are exceedingly rare. When the i-final case is considered, Stress-NonFin will become especially useful.

Table 14: Ambiguously-weighted penultimate syllables tend to take penultimate stress, but not as strongly as definitely heavy syllables. NoCoda is not included here - it is presumed to have a very low weight, consistent with the preponderance of codas in English.

	P	$\mathcal{H}$	FtBin 2	NonFin 2	Align-R 1.5	STW 0.5	Stress-NonFin 1
/pamastra/							
→ a. { (pama)stra }	**.26**	-2			1	1	
b. { (pa)ma.stra }	.01	-5.5	1		2	1	
→ c. { pa(ma.stra) }	**.16**	-2.5		1		1	
d. { pa(ma)stra }	.04	-4	1		1	1	
e. { pa.ma(stra) }	.01	-5.5	1	1		1	1
f. { (pa.mas)tra }	.04	-4	1		1	1	
g. { pa(mas.tra) }	.04	-4	1	1			
h. { (pa)mas.tra }	.01	-5.5	1		2	1	
→ i. { pa(mas)tra }	**.44**	-1.5			1		
j. { pa.mas(tra) }	.01	-5.5	1	1		1	1

In Table 14, the candidate with a heavy penult and a heavy foot on it, candidate i., is the most probable, followed by the antepenultimately stressed candidate with a light penult. The heavy penult case is preferred by STW. Candidate c., which is penultimately stressed despite having a light penult, also gets some probability, just as all-light words take penultimate stress with some probability. Together, all penultimate candidates add up to a probability of 68%, while all antepenultimate candidates add up to a probability of 32%.

Table 15 illustrates just the candidates for the form /pamastri/ in which the syllable containing the final [i] is parsed outside the prosodic word. NonParse-i has a weight of 4, as in Tables 11 and 13, which means that candidates for /pamastri/ with the final [i] parsed into the prosodic word all get a probability of very close to zero. With the final [i], and the effect of NonParse-i, the antepenultimately stressed candidate a. now gets the most probability, with the penultimately stressed candidate f. also getting a healthy amount of probability. Both candidate a. and f. violate NonFin, and satisfy Align-R and FtBin. The other candidates all violate FtBin, rendering them lower probability. Candidate a., with an LL foot over the first two syllables, satisfies Stress-NonFin, while candidate b., with an H foot over the penult, satisfies STW. For this form, all penultimate candidates sum up to a probability of 28% while all antepenultimate candidates sum up to 67%.

In the CMU search conducted in section 2.2, words with ambiguously-weighted penults took penultimate stress 90% of the time when ə-final, but just 7% of the time when they were i-final. This MaxEnt model

Table 15: Candidates for the hypothetical word /pamastri/ in which the final [i] is parsed outside the prosodic word. The constraint NonParse-i is not included, but it's weight of 4 is enough to give any candidates that violate it a probability of close to zero.

	P	$\mathcal{H}$	FtBin 2	NonFin 2	Align-R 1.5	STW 0.5	Stress-NonFin 1
/pamastri/							
→ a. { (pa.ma) } stri	**.43**	-2.5		1		1	
b. { (pa)ma } stri	.09	-4	1		1	1	
c. { pa(ma) } stri	.02	-5.5	1	1		1	1
d. { (pa.mas) } tri	.06	-4.5	1	1		1	
e. { (pa)mas } tri	.09	-4	1		1	1	
→ f. { pa(mas) } tri	**.26**	-3		1			1

predicts penultimate stress 68% of the time when the word is ə-final, and 28% of the time when it is i-final. Because the model is designed to match participants' behaviour, rather than the probabilities in the lexicon, for LLL and LHL words, it also mismatches the lexicon here. Neither the preference for antepenultimate stress when the word is i-final, nor the preference for penultimate stress when the word is ə-final are as strong as they are in the data from CMU. However, the basic generalisation, that when these words end in [i] they prefer antepenultimate stress, and when they end in [ə] they prefer penultimate stress, is captured by this model.

This basic generalisation - that words with ambiguous penults prefer penultimate stress when ə-final, but antepenultimate stress when i-final, would be impossible to model using NonFin-i instead of NonParse-i. Table 16 illustrates why. Here, candidates like those in Table 15 are not included. It is assumed that Parse-σ would have enough weight to rule them out. Instead, the same stress candidates compete with each other as did in Table 14 for the input /pamastra/. The constraint NonFin-i forces the final [i] in 'pamastri' to be unparsed, ruling out candidates c, e, g, and j. However, candidate i. is already very good - it satisfies all three NonFinality constraints, as well as FtBin and STW. The addition of NonFin-i only serves to give candidate i, a penultimately stressed candidate, even more probability than it had for /pamastra/ in Table 14. Only NonParse-i can actually force /pamastri/ to prefer antepenultimate stress.

The addition of the two constraints STW and Stress-NonFin also affect the model's behaviour on the LLL and LHL forms discussed in Tables 11 and 13. Table 17 lists the predictions of this final model for the different word shapes discussed in this section. Stress-NonFin actually brings the behaviour of i-final LHL words closer in line with participants' behaviour in Experiment 2 than they were in Table 13, since it penalizes the {L(H)}i parse, preferring {(LH)}i or even {(L)H}i. STW makes the preference for penultimate stress

Table 16: Replacing NonParse-i with NonFin-i results in an incorrect penult preference in words like 'pamastri'

	P	$\mathcal{H}$	FtBin 2	NonFin 2	Align-R 1.5	STW 0.5	Stress-NonFin 1	NonFin-i 4
/pamastri/								
→ a. (pama)stri	**.33**	-2			1	1		
b. (pa)ma.stri	.01	-5.5	1		2	1		
c. pa(ma.stri)	0	-6.5		1		1		1
d. pa(ma)stri	.05	-4	1		1	1		
e. pa.ma(stri)	0	-9.5	1	1		1	1	1
f. (pa.mas)tri	.05	-4	1		1	1		
g. pa(mas.tri)	0	-8	1	1				1
h. (pa)mas.tri	.01	-5.5	1		2	1		
→ i. pa(mas)tri	**.55**	-1.5			1			
j. pa.mas(tri)	0	-9.5	1	1		1	1	1

on ə-final LHL words even stronger, actually making that preference a bit stronger than what participants exhibited in Experiment 2. Finally, Stress-NonFin makes the preference for antepenultimate stress on i-final LLL words even stronger than what it was in Table 11, or in the experiments. This is because it penalizes parses like {L(L)}i, preferring {(LL)}i or {(L)L}i.

Table 17: The final MaxEnt model's predictions for the various word types discussed in this section. Model predictions are in bold, and CMU percentages as well as percentages from the two production studies (Experiments 1 and 2) are shown for comparison.

			% Antepenultimate				% Final	
	Word Type	Final vowel	**Model**	CMU	Exp. 1	Exp. 2	**Model**	CMU
	LLL	ə	**57**	68	58	58		
		i	**94**	96	77	76		
	LHL	ə	**8**	1		19		
		i	**35**	-		32		
Light Penult	LLH		**64**	85			**23**	4
	HLL		**43**	62			**2**	5
	HLH		**50**	90			**31**	5
Heavy Penult	LHH		**7**	17			**16**	7
	HHL		**25**	5			**1**	1
	HHH		**22**	25			**13**	12
Ambiguous	pamastra	ə	**32**	10				
	pamastri	i	**72**	93				

For the remaining word shapes, this final model matches the behaviour in the lexicon in that most word shapes with a light penult take more antepenultimate stress than penultimate, and all word shapes with a heavy penult take more penultimate than antepenultimate stress. The model incorrectly predicts that HLL words should take slightly more penultimate stress than antepenultimate stress (54% vs. 43%). Beyond this, some word shapes match well, like HHH and LLH, and others match quite badly, like HLL. While weights could be assigned to this set of constraints that would result in a more perfect match to the lexical frequencies of these different forms, this set of weights mismatches primarily because it is designed to match the experimental results first. More experimental work is required to ascertain English speakers' knowledge of the stress preferences of forms like HLL, HLH, HHH, etc. Once that work is done, a full model of English speakers' weight preferences can be created.

7 Conclusion

This paper presents both dictionary evidence and experimental evidence that the quality of a word's final vowel matters in assigning main stress in English. Specifically, a final [i] pushes main stress leftwards - in three syllable words this means stress strongly tends to be antepenultimate. The final vowel's influence can best be captured via the constraint NonParse-i, which forces a final [i] to be parsed outside of the prosodic word. Because it is parsed outside the prosodic word, English's NonFinality constraints then conspire to ensure that the penultimate syllable is not stressed either - pushing stress back to the antepenultimate syllable.

NonParse-i is an atypical stress constraint in OT-style models. Although it essentially enforces an extrametricality requirement (Liberman and Prince, 1977; Hayes, 1980, 1982), OT analyses typically do not include extrametricality directly, but rather rely on constraints from the NonFinality family to force stress off final syllables, or to force them to be unfooted. The effects of the constraint NonParse-i were compared in this paper to the effects of a constraint NonFinality-i, which specifically demands that a final [i] not be footed. The NonFinality-i constraint was unable to account for the facts of Experiment 2, however. In Experiment 2, participants demonstrated a preference for i-final words to take antepenultimate stress, even when the word's penult was heavy. This pattern is impossible to account for using NonFinality-i, because when the penult is heavy, the optimal parse is already for the final syllable to be unfooted. A constraint requiring a final [i] to be unfooted can't really change the metrical structure between i-final vs. ə-final words.

In order to come to this conclusion about the usefulness of extrametricality, in the form of the NonParse constraint, it was necessary to examine the interaction between final vowels and weight, and it was also necessary to take seriously the probabilistic nature of the English stress system. In particular, Experiment

2 demonstrated that English speakers prefer penultimate stress over antepenultimate stress when the penult is heavy, but the *degree* to which they do differs when the final vowel is [i] vs. when it is [ə]. The interaction between weight and final vowel can also be seen clearly in the words with ambiguously weighted penults, like ‘tapestry’, or ‘digestive’. These words take penultimate stress when they end in [ə], but antepenultimate stress when they end in [i]. A model with both NonFinality and NonParse-i explains why: when the final vowel is [i] and it is parsed outside of the prosodic word, then words like ‘tapestry’ are able to surface with a light penultimate syllable and antepenultimate stress. They satisfy both FtBin and NonFinality. As with words with a truly heavy penult, an analysis with NonFinality-i instead of NonParse-i only serves to heap even more probability onto the penultimately stressed candidates.

One final issue remains: What is so special about a final [i]? Why should it be final [i], and not [ə] that drives stress leftward? A few possibilities present themselves, though all come with their own problems. In Section 1, I mentioned words like ‘álligàtor’, ‘cápillàry’, and ‘pártic̀iple’, suggesting following (Hayes, 1982; Liberman and Prince, 1977) that final [i,l̩,r̩] all pattern together, to the exclusion of final [ə,n̩]. However, the results of the search of CMU in section 2.2 did not present a clear picture as to whether these three should be treated as a uniform class or not. In light-penult words final [i,l̩,r̩] all exhibit a stronger preference for antepenultimate stress than do [ə,n̩]. In words with truly heavy penults there are very few words that end with [i,l̩,r̩], but in words with ambiguously heavy penults (Table 4), their behaviour is inconsistent. [l̩]-final words with ambiguously heavy penults, like ‘orchéstral’, or ‘ínterval’ tend to take penultimate stress fairly strongly, while [r̩]-final ones sit in the middle, and [i]-final ones strongly prefer antepenultimate stress. This consistency across types of words in the behaviour of final [i] was why it was chosen as the focus of the two experiments.

If final [i,l̩,r̩] *do* pattern together in driving stress leftwards, then Chomsky and Halle (1968) offer a possible explanation. They claim that [i,l̩,r̩] can be underlying consonants in English, so that words that end in them seem to have one fewer syllable. For example, ‘cannery’ would underlyingly be /kænɑɹj/, with the [i] being stored as a glide. Then, syllabification and stress would proceed to produce {(kǽ.nɑɹj)}, and finally a rule would transform the remaining glide back into a vowel after stress has been assigned. An obvious conclusion then is that [i,l̩,r̩] are special nucleii because they can be changed to consonants with only the change of their syllabicity. No other feature changes are necessary. Other vowels, including [ə] do not have this privilege. The nasals do not fit the pattern however - because nasals also can change from a consonant to a vowel and back with just the change of syllabicity, final syllabic nasals should pattern with [i,l̩,r̩], but they do not.

Another possible explanation is that a final [i], but not a final [ə], actually bears secondary or tertiary stress, since it is a full vowel. This would mean words like 'círema', but not words like 'banána' would contain a stress lapse. Meanwhile, 'cánnerỳ' would be perfectly alternating, and 'canárỳ' would contain a stress clash. If this is the right explanation, it would work for final [r] as well, since a syllabic [r] can be stressed (bird [bɹ̩]). The other common final nucleii, [ə,l̩,n̩] cannot be stressed and so would not be subject to the same pressure. This does not fully explain the behaviour of final [l̩], since it does pattern with [i,r̩] in light-penult words. Additionally, the assumption that every final full vowel bears stress brings with it its own problems, such as how to account for varying stress judgments between words like 'cánnery' and words like 'mánatèe' and 'filigrèe'.

Finally, one possibility is simply that learners of English, and all languages, are exceptionally good a noticing more or less arbitrary patterns in their lexicons and incorporating them into their grammar, at least probabilistically. Other work on speakers' probabilistic knowledge of their languages has demonstrated similar effects (e.g. Hayes et al., 2009; Gouskova and Becker, 2013). The statistical trend for i-final words to take stress farther left than they would otherwise at least partially comes from the morphological system of English. Many derivational suffixes, such as '-ity', shift stress, or demand that the derived word have a certain stress pattern. For example, '-ity' shifts stress to the antepenult, as in 'eléctric' ∼ 'èlectrícity'. Other suffixes shift stress to the penult, such as '-ic' ('ácid' ∼ 'acídic'). It happens to be the case that all derivational suffixes that end in a final [i] either shift the stress to the antepenult (like '-ity', '-ology', '-cracy', '-pathy') or are stress-preserving like '-ly' in 'cáreful' ∼ 'cárefully', or '-y' in 'wíllow' ∼ 'wíllowy'. Stress preserving suffixes have the effect of shifting stress leftwards because they do not change the syllable that stress is on, but they do add a syllable to the right edge of the word - so 'wíllowy' and 'cárefully' have antepenultimate stress. On the other hand, suffixes containing [ə] fall into both the antepenult-shifting category, like '-ious', '-ica' ('hármony' ∼ 'harmónica'), and the penult-shifting category, like '-ic', '-ive'. There are also stress preserving ə-final affixes, like '-ness', '-less', '-ish'. It is important to note that the trend for i-final words to take antepenultimate stress is robust even in monomorphemic words, as is the difference between ə-final and i-final words. However, since most long words are multimorphemic, the learning data available to a child acquiring the stress system of English would be mostly multimorphemic. Furthermore, children start learning the details of their native language's stress patterns before they are even one year old (Jusczyk et al., 1993), well before they have acquired the derivational morphology of the language. It is possible that children pick up on the trend for final [i] to push stress leftwards before they can parse out the difference between words like 'recipe' with no suffix and words like 'willowy' with a suffix. If so, they might never bother to correct their grammar once they do learn the morphology.

To conclude, this paper presents a probabilistic analysis of a portion of the English stress system, including the Latin Stress Rule as it exists in the lexicon of English, and in speakers' productions of nonce words (Experiment 2). The effects of final vowel in the English stress system are also modelled, as well as (crucially) the interaction between weight effects and final vowel effects. Data from actual English speakers was crucial in showing that the final-vowel effect was productive (Experiment 1), and for understanding the interaction between final vowel and weight, since the lexicon did not provide enough evidence (Experiment 2). The use of a probabilistic grammatical model to model these experimental results was crucial in landing upon the best analysis of the final vowel effects, which was that certain final vowels are excluded from metrical considerations because they are parsed outside of the prosodic word.

References

Jonathan B. Alcántara. *The architecture of the English lexicon*. PhD thesis, Cornell University, 1998.

R Harald Baayen, Richard Piepenbrock, and Rijn van H. The CELEX lexical data base on CD-ROM. 1993.

Todd M. Bailey and Ulrike Hahn. Determinants of wordlikeness: Phonotactics or lexical neighborhoods? *Journal of Memory and Language*, 44:568–591, 2001. doi: 10.1006/jmla.2000.2756.

Michael Becker and Jonathan Levine. Experigen - an online experiment platform, 2013. URL `http://becker.phonologist.org/experigen`.

Michael Becker, Nihan Ketrez, and Andrew Nevins. The surfeit of the stimulus: Analytic biases filter lexical statistics in Turkish laryngeal alternations. *Language*, 87(1):84–125, March 2011. doi: 10.1353/lan.2011.0016.

Paul Boersma and David Weenink. Praat: doing phonetics by computer [computer program]. In *http://www.praat.org/*, 2011.

Marc Brysbaert and Boris New. Moving beyond Kučera and Francis: A critical evaluation of current word frequency norms and the introduction of a new and improved word frequency measure for American English. *Behavior Research Methods*, 41(4):977–990, November 2009. doi: 10.3758/BRM.41.4.977.

Luigi Burzio. *Principles of English stress*, volume 72 of *Cambridge studies in Linguistics*. Cambridge University Press, Melbourne, Australia, 1994.

Noam Chomsky and Morris Halle. *The sound pattern of English.* Harper and Row, New York, Evanston, and London, 1968.

Sarah Collie. English stress preservation: the case for 'fake cyclicity'. *English Language and Linguistics*, 12 (3):505–532, 2008.

Ulrike Domahs, Ingo Plag, and Rebecca Carroll. Word stress assignment in german, english and dutch: Quantity-sensitivity and extrametricality revisited. *The Journal of Comparative Germanic Linguistics*, 17:59–96, 2014. doi: 10.1007/s10828-014-9063-9.

Sharon Goldwater and Mark Johnson. Learning OT constraint rankings using a maximum entropy model. In Jennifer Spenader, Anders Eriksson, and Osten Dahl, editors, *Proceedings of the Stockholm Workshop on Variation within Optimality Theory*, pages 111–120, 2003.

Maria Gouskova and Michael Becker. Nonce words show that Russian yer alternations are governed by the grammar. *Natural Language and Linguistic Theory*, 31(3):735–765, 2013.

Susan G. Guion, J.J. Clark, Tetsuo Harada, and Ratree P. Wayland. Factors affecting stress placement for English nonwords include syllabic structure, lexical class, and stress patterns of phonologically similar words. *Language and Speech*, 46(4):403–427, December 2003. doi: 10.1177/00238309030460040301.

Morris Halle and Jean-Roger Vergnaud. *An essay on stress.* The MIT press, Cambridge, Massachusetts, 1987.

Bruce Hayes. Extrametricality and English stress. *Linguistic Inquiry*, 13(2):227–276, 1982.

Bruce Hayes. *Metrical stress theory: principles and case studies.* University of Chicago press, 1995.

Bruce Hayes. Blick - a phonotactic probability calculator. Project page at www.linguistics.ucla.edu/people/hayes/BLICK/, 2012. URL `http://www.linguistics.ucla.edu/people/hayes/BLICK/CMUDictionaryEditedWithCELEXFrequenciesPublicVersio`

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

Bruce Hayes, Kie Zuraw, Péter Siptár, and Zsuzsa Londe. Natural and unnatural constraints in Hungarian vowel harmony. *Language*, 85(4):822–863, December 2009. doi: 10.1353/lan.0.0169.

Bruce Philip Hayes. *A metrical theory of stress rules.* PhD thesis, Massachusetts Institute of Technology, 1980.

Karen Henrich, Kai Alter, Richard Wiese, and Ulrike Domahs. The relevance of rhythmical alternation in language processing: an ERP study on English compounds. *Brain and Language*, 136:19–30, 2014.

Brett Hyde. Non-finality and weight sensitivity. *Phonology*, 24(2):287–334, 2007.

Brett Hyde. Extrametricality and Non-Finality. In *The Blackwell companion to phonology*, pages 1–25, 2011.

Peter W. Jusczyk, Anne Cutler, and Nancy J. Redanz. Infants' preference for the predominant stress patterns of English words. *Child development*, 64(3):675–687, 1993. doi: 10.1111/j.1467-8624.1993.tb02935.x.

René Kager. *A metrical theory of stress and destressing in English and Dutch.* Foris, Dordrecht - Holland/Providence RI - USA, 1989.

Rene Kager. *Optimality Theory.* Cambridge Textbooks in Linguistics. Cambridge University Press, 1999.

Daniel Kahn. *Syllable-based generalizations in English phonology*, volume 156. Indiana University Linguistics Club Bloomington, 1976.

Michael H. Kelly and J. Kathryn Bock. Stress in time. *Journal of Experimental Psychology: Human Perception and Performance*, 14(3):389–403, 1988.

Michael Kenstowicz. Quality-sensitive stress. *Rivista di Linguistica*, 9(1):157–187, 1997.

Mark Liberman and Alan Prince. On stress and linguistic rhythm. *Linguistic Inquiry*, 8(2):249–336, 1977.

Paul Olejarczuk. The productivity and stability of competing generalizations in stress assignment. In *Laboratory Phonology*, Tachikawa, Japan, 2014.

Paul Olejarczuk and Vsevolod Kapatsinski. The metrical parse is coarse-grained: phonotactic generalizations in stress assignment. University of Oregon, forthcoming.

Joe Pater. Against the underlying specification of an'exceptional'english stress pattern. *Toronto Working Papers in Linguistics*, 13, 1994.

Joe Pater. Non-uniformity in english secondary stress: the role of ranked and lexically specific constraints. *Phonology*, 17(2):237–274, 2000. doi: 10.1017/S0952675700003900.

Joe Pater. Weighted constraints in generative linguistics. *Cognitive Science*, 33:999–1035, 2009.

Alan Prince. Relating to the grid. *Linguistic Inquiry*, 14:19–100, 1983.

Alan Prince and Paul Smolensky. *Optimality Theory: constraint interaction in generative grammar*. Blackwell, Malden, MA, 1993/2004.

Elisabeth Selkirk. *Syntax and Phonology: The relation between sound and structure*. MIT Press, Cambridge, Massachusetts, 1984.

Joan A Sereno and Allard Jongman. Acoustic correlates of grammatical class. *Language and Speech*, 3(1): 57–76, 1995.

Paul Smolensky and Geraldine Legendre. *The harmonic mind: from neural computation to optimality-theoretic grammar*. MIT Press, Cambridge, Massachusetts, 2006.

Morgan Sonderegger and Partha Niyogi. Variation and change in English noun/verb pair stress: Data, dynamical systems models, and their interaction. In A.C.L. Yu, editor, *Origins of sound patterns: approaches to phonologization*, pages 262–284. Oxford University Press, 2013.

Richard V. Tescner and M. Stanley Whitley. *Pronouncing English: A stress-based approach with CD-ROM*. Georgetown University Press, Washington, DC, 2004.

Sam Tilsen. Utterance preparation and stress clash: Planning prosodic alternations. *Fuchs, S., Perrier, P., Weirich, M. & Pape, D.(Eds.) Speech production and perception: Planning and dynamics*, 2012.

Robert L. Weide. CMU pronouncing dictionary, 1994. URL `http://www.speech.cs.cmu.edu/cgi-bin/cmudict`.

Hideki Zamma. *Patterns and categories in English suffixation and stress placement: a theoretical and quantitative study*. PhD thesis, University of Tsukuba, 2012.

Kie Zuraw. *Patterned Exceptions in Phonology*. PhD thesis, University of California, Los Angeles, 2000.

Kie Zuraw. A model of lexical variation and the grammar with application to tagalog nasal substitution. *Natural Language and Linguistic Theory*, 28(2):417–472, 2010.