How Prefixation, Onset Length, and Coda Length Determine Lexical Stress Assignment when Reading Aloud

Kayla Hensley
Washington University in St. Louis

WASHINGTON UNIVERSITY IN ST. LOUIS
Department of Brain and Psychological Sciences

How Prefixation, Onset Length and Coda Length Determine Lexical Stress Assignment when Reading Aloud.

by
Kayla Hensley

A thesis presented to
Washington University in St. Louis
in partial fulfillment of the
requirements for the degree
of Master of Arts

December 2023
St. Louis, Missouri

Table of Contents

List of Tables

Acknowledgments

I would first like to acknowledge Dr. Treiman's support and feedback throughout the entire process of writing this thesis. It would not have been possible without her. I would also like to thank Dr. Van Engen and Dr. Sommers for their time and support. Additionally, I would like to thank Dr. Kessler and the other members of the Reading and Language Lab for their feedback and advice over this past year, the research assistants for running participants, and the undergraduates who participated in my studies.

I want to thank my parents, sister, and friends (and all their pets) for offering their love and encouragement during this process.

Kayla Hensley

Washington University in St. Louis

December 2023

ABSTRACT OF THE THESIS

How prefixation, onset length and coda length determine

lexical stress assignment when reading aloud.

by

Kayla Hensley

Master of Arts in Brain and Psychological Sciences

Washington University in St. Louis, 2023

Professor Rebecca Treiman, Chair

Lexical stress, or which syllable is emphasized in a multisyllabic word, can provide important cues for speech segmentation, lexical class, and reading acquisition. Despite this, much of the literature on reading has focused on single syllable words, neglecting the role that stress can play. In this project, I examine the role of prefixation, onset and coda length in assigning stress in disyllabic words when reading aloud. Study 1 is a corpus analysis of English words with grade level information to replicate the model developed by Treiman et al. 2020 and check for age differences. There were no significant differences on account of age, nor did age predict any factors that predicted stress. Studies 2 and 3 examined the role of coda and onset length when any impact of prefixation was removed by having subjects read sentences containing novel nonwords aloud in Study 2 and assign stress to these nonwords in a survey. Coda length has a significant impact in both studies such that a greater coda length leads to less first syllable stress assignment. Onset length produced mixed results across the two studies with an increased onset length increasing first syllable stress in Study 3 but having no effect in Study

2. Studies 4 and 5 extended the investigation into onset length by varying onset length along with the presence of prefixation. Study 4 employed the same method as Study 2 with reading aloud while Study 5 mirrored Study 3's procedure of self-report. Study 4 found significant effects of both onset and prefixation as well as an interaction. Onset length again increased first syllable stress as it increased. Going from a prefixed word to a word with no prefix also increased stress. The interaction showed that prefixation strengthened the effect of onset length on stress assignment. Study 5 found similar significant effects of both variables but no interaction. These findings suggest that previous work investigating the impact of onset length on lexical stress assignment may have been confounded by prefixation.

Chapter 1: Introduction

Lexical stress is the emphasized syllable in a multisyllabic word. It is also the reason it is possible to understand the sentence, "The parents per***mit*** the children to get a ***per***mit for their lemonade stand." The noun and verb forms of the word "permit" do not differ in their spelling or pronunciation, only in which syllable is stressed. While words like "permit" that differ only in stress assignment are relatively rare, stress is important for many things in language. Additionally, stress differs across language (Yu & Andruski, 2010), and this thesis will focus on issues concerning stress assignment in English. Stress can provide cues for determining lexical class: figuring out if a word is a noun, verb, adjective, or something else, (Guion et al., 2003). It also assists with speech segmentation, or determining where one word starts and another word ends in spoken language (Mattys & Samuel, 1997). Despite playing these important roles, stress is often overlooked in studies that examine how phonology is mapped onto orthography, as prior work in reading aloud has primarily focused on monosyllabic words. This paper will focus on the connections between lexical characteristics of words and pronunciation when reading aloud with disyllabic words. Disyllables are the focus here because the rules for stress become more complex with longer words as secondary stress can come into play, and it is important to get a foundational model for disyllabic words before expanding it to longer multisyllabic words.

1.1 Models of Reading

While reading occurs in the context of sentences and paragraphs, it is impossible to understand them without first understanding the individual words that make up those parts. Being able to decode individual words quickly and efficiently is essential to being able to fully understand the text being read. Two early categories of models of reading words include the dual-route model (Coltheart et al., 2001) and connectionist or statistical models (Seidenberg & McClelland, 1989). Dual-route models posit that reading is based on two possible paths, or routes. The first route is a sublexical route which applies rules that match letters to sounds. The second route is a lexical route that stores the phonological form of an item in memory; this is particularly important for words which constitute exceptions to rules. Following rules for spelling-to-sound correspondences and storing exceptions to those rules accounts for some interesting findings in reading. One such example is the frequency by regularity interaction (Hino & Lupker, 2000). Words that follow the rules of the language prompt faster and more accurate decisions in a lexical decision task, but this relationship is moderated by the frequency of the words. More frequent irregular words still have relatively quick reaction times.

However, dual-route models cannot account for everything. Specifically, exceptions—particularly common exception words—can create groups that have their own patterns that resemble rules. Rule-based models would not be able to account for these patterns in the language or their implementation by untrained individuals in behavioral work (Treiman et al., 2003), but connectionist models can account for these patterns. One of the most famous connectionist models was proposed by Seidenberg and McClelland (1989) and is known as the triangle model. These models are graded rather than all-or-nothing like rule-based models, and they are based on connections formed through experience. In this way, the connections between spelling and sounds (and meaning) are formed by taking statistics about the environment. Though both of these types of models can tell us a lot about reading, they are unsuited to capturing stress assignment as their properties were developed using single-syllable words.

Though these principles could apply to multisyllabic words, early models of reading neglected to account for lexical stress and its potential impact on learning to read.

There have been attempts to build models of reading that do consider stress when reading aloud. One example of such a model was Rastle and Coltheart (2000). This model builds on work by Fudge (1984) which holds that readers first determine if a word contains a prefix. If it does, it will have second-syllable stress, so long as it is possible to pronounce that syllable as a monosyllabic word. This model follows from the dual-route category of reading models. It has one route that follows a rule: if the word contains a prefix, stress the second syllable. Then there is a secondary mechanism for words that do not fit into the rule-based category: if the second syllable cannot be pronounced as a monosyllabic word, it must be remembered as an exception that contains first-syllable stress. Treiman et al. (2020) found that behavioral data with nonwords did not support this model, as nonwords with a plausible monosyllabic word in the second syllable and a prefix in the first did not always have second syllable stress. There was, however, an impact of prefixation on stress assignment, so the model still provides some insight into predicting stress assignment.

Another type of model is a statistical model built from a corpus of words. One such model is Mousikou et al (2017) which predicts lexical stress based on orthographic weight. Orthographic weight is a measure they developed that creates a ratio of the first syllable and the second syllable based on how many letters they have. This then predicts whether a word will have first or second syllable stress. This model predicted stress correctly for 93.5% of words in the Unisyn Corpus (Centre for Speech Technology Research, n.d.) in Treiman et al.'s (2020) paper. However, in the same paper, Treiman et al. (2020) examined how the model accounted for behavioral data and found that it did not do as good a job at predicting stress than separately including coda, onset, and medial cluster identity into a model, perhaps because coda length has such a strong influence on stress assignment but Mousikou's model weights it evenly with onset length and medial identity.

1.2 What Makes Stress Difficult to Study

One reason for neglecting stress as an area of research is the lack of consensus on how exactly to measure it. Linguists disagree on what features are best to use as the acoustic correlates of lexical stress as there are many features that correlate with stress. Thus far, much of linguistic work concerns itself with the prediction of stress placement rather than creating a precise definition for stress. Some of the most used features include volume, change in fundamental frequency, or even syllable duration (Gordon & Roettger, 2017). The disagreement over the constituents of stress means that developing a standard methodology for behavioral studies of stress is difficult, and there are currently frequent inconsistencies among the scoring methods used.

Two common categories of methodology exist: self-scoring and trained judge scoring. Each of these categories can be broken down into more specific considerations that vary along a spectrum. On the side of self-scoring, participants generally read an item aloud or to themselves and indicate which syllable they think should carry stress (Kelly, 2004). Some such studies involve additional training for participants such as feedback from instructors and peers (Sippel & Martin, 2022) which does produce improvement on the perception of lexical stress. Alternatively, the participants could listen to multiple pronunciations of the target item with alternating stress patterns and report which one sounds the most correct (Fear et al., 1995). On one hand, a self-scoring method requires low effort from the researcher, and it measures the meta-linguistic knowledge on stress for the participants in addition to stress placement. On the other hand, evidence suggests that untrained individuals who are native speakers and readers of English have a hard time identifying stress.

One study by O’Brien (2019) showed that untrained individuals had little success in generating rules for stress assignment. They had Native English and French speakers who were studying German read German sentences and indicate which syllable they stressed and where the stress was supposed to fall in a word. They then asked the participants to generate a rule for where stress should be assigned in each word. The researchers found that the ability to verbalize where the subjects had placed stress was significantly predictive of accurate stress assignment. Further there was little accuracy generating rules for stress assignment across even after accounting for different fluency levels with German. This issue with stress identification could be due to ambiguity in the lexical stress of the item to be scored (Benrabah, 1997) or a general problem focusing on stress (Yu & Andruski, 2010).

These issues surrounding a lack of knowledge about stress found in self-report scoring would be less of a problem for trained judges scoring participant productions. In these designs, the participants simply read aloud to be recorded or scored on the spot. Then, trained judges score their productions to be analyzed by the researchers. However, there is a lot of variation in exactly how much training the judges receive, how they are scoring stress, and the reliability of their assessments. Additionally, these judges are still subject to bias because of the ambiguity of lexical stress. If a production is ambiguous and the judge knows the hypotheses of the experiment, they could—consciously or unconsciously—score in a way that will support the study’s hypotheses.

Outside of these concerns about both self-report and trained judge scoring, one must consider if these two methods are measuring the same things. I would argue that a self-report method taps into the meta-linguistic knowledge that participants possess about lexical stress. It shows what they think should be stressed. Trained judge scoring, on the other hand, measures what participants are actually doing when they encounter new multisyllabic words. My project will examine these possibilities by using two types of study designs for each set of stimuli. One will be a self-report design and the other will involve trained judge scoring. If there are differences in the results, it could be that this is due to the different scoring methods measuring different concepts. The same pattern of results would suggest that they are both measuring lexical stress production, albeit in different ways.

1.3 Linguists' View of Stress

Despite these issues measuring stress, linguists and psycholinguists have long been interested predicting stress patterns for different languages, and their efforts have produced varying results. Asserting that the weight of the syllable, or the number and duration of elements in the rime (the vowel and coda of a syllable), is more important for predicting stress assignment then the morpheme (Booij, 1983), linguists have stated that the rightmost syllables are the most informative for predicting stress in English (Halle & Vergnaud, 1987; Fudge, 2015). In other words, what happens at the end of a syllable is the most useful to figure out where the stress will occur when the word is produced. Linguistic theory also acknowledges that different lexical classes carry different stress patterns in English (Chomsky & Halle, 1968).

Psycholinguistics behavioral research has confirmed these factors as important for predicting stress assignment and postulated new factors on the morphemic level like prefixation (Rastle & Coltheart, 2000) and the word level like coda length (Smith & Baker, 1976) and onset length (Kelly, 2004). Coda length is defined here as the number of phonemes occurring after the nucleus, or sonorous part (typically a vowel), of the final syllable of a word. More phonemes are interpreted as a longer coda. For example, a word like ‹revolt› would have a coda length of two because it has two phonemes in the end of the word. Onset length is the same but for the phoneme(s) that precede the nucleus of the syllable. ‹Revolt› would have an onset length of one. For the present research, the onset length for the first syllable in a word will be considered, and length will be measured in letters which map onto a single phoneme. Coda and onset length can be equal to zero. Coda length could be seen as a mechanism that accounts for the process observed by linguists that the ends of words impact stress assignment. Onset length is more contentious as it would not be expected to impact stress assignment heavily using linguistic theory since it has to do with the beginning of the syllable and not the rime. This project aims to determine if onset length does, in fact, play a role in stress assignment or not by eliminating a possible confound in prefixation.

1.4 Corpus Analysis Approach to Stress

One way that researchers have approached lexical stress assignment is by building statistical models based on different corpora of words. This method allows researchers to determine what patterns are actually present in the language of interest. Further, this can give insight into if those patterns in the language are matched by the behavior of untrained individuals. A corpus analysis can be conducted for any language, not just English. One such example is Jouravlev and Lupker’s (2014) analysis of Russian stress consistency. Using this approach, the model inputs different variables to predict the likelihood of stress on a certain syllable for a given corpus of words. This can then be compared to the base likelihood of first syllable stress occurring in disyllabic words in English. There have been few attempts to capture predictors into a statistical model of reading based on a corpus of words that can account for multisyllabic words and capture lexical stress. Often the success of these models is judged against the success one would find if they guessed first syllable stress for all English words which would be correct around 88% of the time (Treiman et al. 2020). One recent corpus-based statistical model to predict lexical stress in a corpus, developed by Treiman et al. (2020), predicted stress using factors based on previous behavioral research. The model combines the effects of lexical class, onset length, coda length, prefixation, and medial consonant count to predict the likelihood of first syllable stress in each word with around 94% accuracy.

Arciuli et al. (2010) began to examine if the predictors for stress changed based on age. In a corpus analysis which used a corpus that contained grade level information (Zeno, 1995), they split words into two chunks: the first consonant/consonant cluster and vowel and the rime of the second syllable. For example, ‹zebra› would be split into ‹ze› and ‹a›, with the ‹br› in the middle not being considered in the analysis. They then looked at if these beginnings and endings of words were differentially predictive based on the grade levels of the text. They found that, while both were significantly predictive for stress placement, word endings overall were more accurate at predicting stress than word beginnings. In the earlier example of ‹zebra› this means that the beginning, ‹ze›, was predictive of first syllable stress, but that the ending ‹a› had an even bigger effect. Additionally, they found that the proportion of words in the youngest grade's assigned corpora contained a higher proportion of orthographically-useful information about stress position than the word lists assigned to older grade's corpora. In other words, the most predictive parts of words, so the ‹a› in ‹zebra› showed up in a higher proportion in words from lower grade levels.

Study 1 endeavors to see if the predictors used in the model developed by Treiman et al (2020) can be predicted by or change because of age through a corpus analysis of Zeno's (1995) Word Frequency Guide. In this way, there is more specificity then Arciuli et al (2010) about exactly what parts of a word a child is using to predict lexical stress. Additionally, the medial portion of the word is considered by looking at if there is a geminate cluster or not. If there is still an impact of age, then it would be more convincing that the specific predictors of lexical stress vary by age.

1.5 Behavioral Approach to Studying Stress

Another way to look at stress is to design a behavioral study in which participants are presented with novel multisyllabic nonwords to read. By doing both behavioral work, and a corpus analysis it is possible to see if the behavior of participants matches the subtle patterns found in corpus analyses. For example, it may be true that onset is a predictor of stress in the English lexicon, but that untrained speakers of English do not use it as a cue for determining the stress of a novel word. Alternatively, untrained individuals might overuse other cues like prefixation. A study by Ktori et al. (2016) had dyslexic subjects assign stress to words that either had a prefix with second-syllable stress, had a prefix with first syllable stress, or had no prefix and first syllable stress. They found that the subjects made the most errors when the words had a prefix but were meant to have first syllable stress.

Kelly (2004) explored whether the number of letters in the onset of the first syllable of a word impacted lexical stress. They had participants look at nonword items that had either one or two consonants at the beginning of the item. The items were identical otherwise. They then read the word and indicated whether the "stress or accent" was on the first or second syllable (Kelly, 2004, p. 238). Kelly found that words with more letters in the onset produced more first-syllable stress judgements.

This result was also found in the Treiman et al (2020) studies. Studies 1 and 2 of that paper had participants read aloud sentences which contained novel nonword items that varied on their coda and onset lengths. Their productions were scored for stress by a trained judge, and the researchers found that longer coda reduced first syllable stress and longer onset increased first syllable stress. Study 4 examined the same question as a self-report task with isolated nonwords and found the same pattern of results. Study 5 examined data from previous studies and added in medial consonant count as variable. They found the same pattern of results with respect to coda and onset length and found that more consonants in the middle of the word predicted more first syllable stress.

These results seem promising for the idea that onset length predicts stress; however, both Kelly's (2004) and Treiman et al.'s (2020) stimuli raise some doubts about if they truly accounted for prefixation when developing stimuli. The presence of a prefix reduces first syllable stress (Ktori et al., 2018), and most prefixes have less than two letters in the onset of the syllable. For example, of the prefixes in the Rastle and Coltheart model, only 6% start with more than one consonant letter (Rastle & Coltheart, 2000). Given the high frequency of unstressed prefix syllables, it is common to pronounce them using a schwa (/ə/ in IPA[1]). In English, this phoneme can be spelled using almost any vowel which means that words with the first consonant of a prefix like ‹be› could also be spelled like ‹bu›, ‹ba›, or ‹bo› and they would all have the pronunciation /bə/. For studying stress assignment, this means that stimuli designed to avoid orthographic prefixes (like ‹be›) might still contain sound-alike prefixes, or syllables that can be pronounced as prefixes, in shorter onset conditions. However, longer onset conditions would be less likely to have such a confound. The impact of onset length may really be due to changing

[1] Note on symbols: ‹› are used for orthographic representation, and // are used for phonetic representations.

from an item that contains a syllable that sounds like a prefix to an item with no sound-alike syllables.

Both papers which found an impact of onset length (Kelly 2004; Treiman et al., 2020) outlined how they excluded words with prefixes, but there were stimuli that contained a sound-alike prefix in the first syllable, even if it was not orthographically a prefix. In Kelly's (2004) study nonword pairs were developed to include both pairs with and without prefixes. There was no significant effect of prefix in the Kelly study. However, Kelly included items that began with ‹bo-› and ‹de-› as items that did not include prefixes. For example, ‹bolay› was an item in this study that was not marked as containing a prefix but could be pronounced in a way that includes a prefix. Additionally, out of 48 pairs of words, only 11 contained prefixes which would make it difficult to find any impact of prefixation with such a small sample. Treiman et al. (2020) were careful to exclude any orthographic prefixes, but there were still items that contained syllables that could be pronounced like prefixes. For example, in Study 2, the first syllable in ‹bupel› could be pronounced /bə/ and the first syllable in ‹dunnep› could be pronounced /də/. Even though they do not contain any prefixes, these items may register phonologically to participants as the orthographic prefixes ‹be-› and ‹de-›, respectively.

Looking at specific factors which predict lexical stress, this project will start by reaffirming the role of coda and onset length in predicting lexical stress via a corpus analysis and behavioral studies. As mentioned earlier, coda length has been consistently named as a predictor across both the linguistic and psycholinguistic literatures such that a longer coda produces less first syllable stress. More letters in the coda pulls the stress toward the consonant cluster at the end of the word. Study 1 is a corpus analysis that will investigate a new corpus with grade level information to affirm the impact of coda and first-syllable onset length in English. Studies 2 and

3 will behaviorally explore the impacts of onset and coda length on stress assignment while reading novel words aloud while eliminating any possibility for prefixes, either orthographically or as a sound-alike. These studies provide a start at getting at if onset length shows an impact on stress assignment when prefixation is taken out of the equation. If there is not an impact of onset length on stress assignment with no prefixation, then that would support the idea that prefixation is a confound for first syllable onset length. Additionally, studies 2 and 3 will compare two measures of stress, a self-report survey and a trained judge scoring to determine how different methods of measurement might change the pattern of results in a study of stress.

Studies 4 and 5 will compare similar issues methodologically while also continuing to tease apart the question of how much the impact of onset length can be accounted for using prefixation. If there is a difference in items based on prefixation that is greater than the difference because of changing onset length with no prefix in either item, it could be inferred that prefixation is acting as a confound for onset length in previous studies.

Chapter 2: Study 1

Study 1 is a corpus analysis that looks at how different orthographic predictors of lexical stress in disyllabic words might vary by age. The factors considered were lexical class, prefixation, and coda, onset, and medial consonant cluster length, replicating the Treiman et al. (2020) paper. I wanted to see if incorporating all these factors into a single model does a better job of predicting stress than always assuming first syllable stress (which would be correct 88% of the time, as they found in the 2020 paper.) Adding onto those findings, I wanted to see if the grade level of the words predicted either their subsequent lexical stress when combined with this model, or the predictive value of the different factors mentioned earlier.

2.1 Methods

I created a lexicon of words that contained grade level data by selecting words from the Zeno Word Frequency Guide (Zeno et al., 1995), excluding hyphenated and foreign words. Zeno Word Frequency Guide is a corpus of over 17 million words that are rated on readability and grade level for 13 different grades: 1st through 12th grade and a beyond high school level. The words used in the present study contained 0-3 consonants at the onset and coda of the word and a medial consonant that was either one consonant or the same consonant repeated (a geminate). The words were also required to have exactly one vowel before and after the medial consonants. Then I joined that corpus with the Unisyn corpus (Centre for Speech Technology Research, n.d.) to get IPA data. Using this word list, I eliminated words that falsely appeared as disyllabic by requiring that words have one intervocalic consonant. Words that only differed in their stress placement (e.g., “record”) were also excluded from the final corpus. The final corpus contained 940 distinct words.

The Unisyn data (Centre for Speech Technology Research, n.d.) also provided information about stress assignment, and onset, medial and coda consonant counts. I coded the part of speech for the words using noun, verb, adjective, and miscellaneous for the categories. The miscellaneous words did not fit into any of the other categories described or fit into more than one of them. Words were also coded for if they contained prefixation as defined in the Rastle and Coltheart model (2000). This model is based on whether the spelling appears to be a prefix, regardless of if it is morphologically a prefix. This means that their model starts with the orthography of the word and then translates it to the phonology to determine if something is a prefix.

2.2 Results

I looked at the following predictors of first syllable stress: number of consonants in the onset, middle, and coda of a word, presence of a prefix, and type of speech (see Table 1 for regression results). All the following results were for words coded as a fifth-grade level in the Zeno corpus, though the results are similar at all grade levels. Words that have more consonants in the onset were more likely to have first syllable stress. The proportion of words with first syllable stress was .63, .9, and .93 for 0, 1, and 2 consonants respectively. Words that have medial geminate consonants were more likely to have first syllable stress (.93 compared to .84 for no geminate medial). Consonants in the coda had the opposite effect of onset with more consonants in the coda leading to less likelihood of first syllable stress. The proportion of words with first syllable stress was .96 (0 consonants), .94 (1 consonant), .52 (2 consonants), and .5 (3 consonants). Prefixation made a word less likely to have first syllable stress. The proportion of words with a prefix that had first syllable stress was .48 compared to .95 for those that did not have a prefix. Lastly, the proportion of words with first syllable stress was .96 for nouns, .95 for adjectives, .26 for verbs, and .79 for words that did not fit into one of these other classes.

Table 1

Results of Regression Analysis for Words of Different Grade Levels from the Zeno Corpus

Term	*Grade 2*	*Grade 5*	*Grade 9*	*Post High School*
Intercept	1.47 (.62)*	1.44 (.57)**	1.42 (.57)*	1.31 (.58)*
Onset Count	.92 (.28)***	1.08 (.26)***	1.08 (.26)***	1.09 (.26)***
Coda Count	-1.31 (.24)***	-1.29 (.22)***	-1.28 (.22)***	-1.23 (.22)***
Medial Consonant Count	1.01 (.36)**	0.96 (.33)**	.95 (.33)**	1.02 (.34)**
R&C Prefix	-1.73 (.31)***	-1.76 (.28)***	-1.74 (.28)***	-1.80 (.28)***
Noun vs. Misc.	1.44 (.32)***	1.32 (.30)***	1.34 (.30)***	1.38 (.31)***
Adj. vs. Misc.	2.06 (1.08)	1.55 (.67)*	1.55 (.67)*	1.55 (.67)*
Verb vs. Misc.	-2.17 (.52)***	-2.13 (.44)***	-2.13 (.44)***	-2.15 (.44)***

I replicated Treiman et al.'s (2020) corpus analysis using the data from the 13^{th} (or post-high school) reading level using the same variables: number of consonants in the onset, middle, and coda of a word, presence of a prefix, and type of speech. The model predicted the stress pattern correctly around 94% of the time for both word lists. The successful replication indicates that the Zeno corpus of words is satisfactorily like the Unisyn lexicon of words used in the 2020 study. This prediction rate outperforms assuming first syllable stress which would be correct 88% of the time.

Adding grade level as a variable to the regression model produced no significant effect on predicting first syllable stress. Additionally, running separate regressions on the second, fifth, ninth, and postsecondary grades showed no significant differences in the different prediction factors mentioned above (as seen in Table 1). At each individual grade level, the model also predicted stress correctly 94% of the time.

2.3 Discussion

The lack of influence by grade level was somewhat unexpected given previous work by Arciuli et al (2010) which found that their predictors of stress, though those factors were different from the exact ones examined here, changed as a factor of grade level. This may be explained by the differences in how the words were split into parts in that paper. This analysis was based on the work of Treiman et al. (2020) which analyzed the onset length, coda length, and medial cluster presence. In Arciuli et al. (2010), the medial consonant/consonant cluster was omitted from the analyses. Given that the corpus analysis in this study and Treiman et al. (2020) found significant effects of the medial consonant cluster, this difference in how the words were treated may account for some of the difference in findings.

The other findings, though, were expected, and I will use them to build the stimuli for the subsequent studies. First, I will look at the question of the impacts of coda and onset length when prefixation is removed from the possible predictors in Studies 2 and 3.

Chapter 3: Study 2

This first behavioral study tests whether the corpus patterns from Study 1 are reflected by the behavior of individuals when faced with a new word. A statistical learning account would predict that individuals can implement even subtle patterns in the corpus. A more rule-based account would assert that individuals pick up on general rules but miss more subtle patterns. Specifically, this first behavioral study examines if the findings that coda and onset length decrease and increase first-syllable stress—respectively—are present in the pronunciations of subjects as they are reading sentences with nonword items. Additionally, it will address the question about scoring. Study 2 uses a trained scorer to measure lexical stress assignment, and Study 3 uses meta-linguistic self-reports on the same stimuli. This combination of studies means that I will be able to investigate if these different methods are measuring the same concept related to lexical stress based on how well the results align.

3.1 Methods

3.1.1 Participants

I recruited 42 undergraduate students from Washington University in St. Louis to participate in this study. Subjects in this and the other experiments received course credit for their participation. The mean age of participants was 20 years old (range 18–22 years). Twenty-six of participants identified as female, 15 as male, and one as nonbinary. All the participants in this and the following studies reported no hearing, vision, reading, or language disorders and spoke English as a native language. No participant took part in more than one of the studies reported in this project.

3.1.2 Stimuli

I designed 36 quadruplets of disyllabic nonwords which varied on the number of letters (1 or 2) in the onset (beginning) and coda (ending) of an item. One example of a quadruplet would be ‹cubbaf›, ‹crubbaf›, ‹cubbaft›, and ‹crubbaft›. The first letter in each onset corresponded to a stop or fricative. In the condition with two letters in the onset, the letters ‹r›, ‹w›, or ‹l› were added after the stop. For example, ‹cubbaf› in the single onset condition became ‹crubbaf› in the double onset condition. Any beginning to a word that was either an orthographic prefix or a sound-alike prefix was not used in the stimuli. For example, ‹de-› was not used in the stimuli because it is an orthographic prefix, and ‹du-› was not used because it could be a sound-alike prefix for the prefix ‹de-›. Following the one or two letters in the onset, a vowel (‹a›, ‹e›, ‹i›, ‹o›, or ‹u›) was selected followed by a consonant cluster. All the items had a geminate medial consonant cluster to reduce the variation in the vowel pronunciation of the first syllable as geminate medial consonant clusters produce overwhelmingly short vowel pronunciations (Treiman et al., 2023). After this, there was another vowel. The coda of each item was made up of endings that occur in real English words. If one of the quadruplet endings qualified as a suffix, all the endings in a quadruplet were also made to be suffixes. For example, ‹kippen› and ‹kippent› both end in suffixes. Each participant was presented with one item from each quadruplet (a total of 36 different nonwords), counterbalanced so they saw an even number of each item type. This allows us to be sure that there is no influence of previously-seen words from the same quadruplet on the pronunciation of each new stimuli.

Training items were also disyllabic nonwords, but they were markedly simpler. They contained a CVCVC pattern where C is a consonant and V is a vowel. The consonants at the beginning and endings of items could not appear in those positions in the experimental stimuli.

For example, no items used in the experimental stimuli began with the letter ‹z›, so one of the training words was ‹zedav›. All the experimental and training stimuli are included in the Appendix.

The words appeared in simple phrases, referred to as sentence frames from this point forward. Sentence frames were used to emulate the natural prosody of reading aloud and, therefore, ensure a more natural stress production. A sentence frame could be "The girls ‹zedav› today." The sentence frames contained one of four subjects ("girls", "boys", "cats", or "dogs") preceded by a determiner. The nonword appeared after the subject in the verb position of the sentence and was followed by one of four adverbs ("inside", "outside", "today", and "away") with second syllable stress. Using plural subjects meant that the stimuli would not change because the common conjugation of a verb for plural subjects does not generally change form. The sentence frames all followed a pattern that would further encourage second-syllable stress using rhythmic alternation. Each even-numbered syllable surrounding the nonword in the sentence was stressed which makes the nonword more likely to be second syllable stressed (Kelly, 1988) We took these steps because avoiding ceiling effects of first-syllable stress pronunciation was a priority for this research. In the Treiman et al. (2020) paper, the sentence frames varied on whether the items were presented as verbs or nouns, and the proportion of first syllable stress remained between .80 and .96. By making all the sentence frames verbs, this should allow for more variation in the stress placement. The same frames were used for experimental and training items.

3.1.3 Procedure

This study was carried out over Zoom and participants' audio was recorded. Participants completed four training trials in which they were presented with a nonword in isolation and then in a sentence and were asked to read aloud what they saw on the screen. Each screen was presented for six seconds. The experimenter demonstrated for the first two trials, using first-syllable stress for the first trial and second-syllable stress for the second so that participants would not take cues for stress frequency from the experimenter. Participants were instructed to take their time to sound out the nonword when it was presented alone, but to read it fluently when it was presented in the sentence.

The participants did the experimental trials without any input from the experimenter. These trials followed the same procedure as the training trials. After completing the experimental trials, participants filled out a short demographic survey. The study took approximately 20 minutes to complete.

3.1.4 Scoring

The researcher scored the utterances within the word frame for the final stress scoring. Utterances were scored for audio first, looking at if the audio recording was clear enough to score. Then the recording was scored for plausibility. Plausibility here refers to a pronunciation that includes phonemes that can be matched to all the letters in the proper order. An implausible response for the word ‹kebbolg› would be if an individual produced /'kɛblɑg/ because they changed the order of the phonemes included. Then participants were scored on if there was stress in the nonword. Pausing in the middle of a word would be scored as not having stress, for example. Lastly, the participants were scored on if their productions were stressed on the first or second syllable.

A second judge scored 20% of responses (*n* = 324). These were half of the responses from a group of 18 randomly selected participants. The judges agreed on the audio quality 99.7%. Of the trials for which they agreed that the audio was of good enough quality to score the stress, they agreed on the plausibility 99.4% of the time. The two judges agreed on if the remaining trials had stress 99.4% of the time. Of the trials that the judges agreed on the audio, plausibility, and stress presence, they agreed on the location of the stress (first syllable or second) 92.8% of the time.

3.2 Results

Thirty-eight trials (2% of the total trials) were excluded for issues relating to audio, plausibility, or stress. Of the remaining trials, 58% had first syllable stress. Given how much lower this is than the expected 88% first-syllable stress in English disyllabic words, the efforts to avoid ceiling first-syllable stress were successful. Additionally, there were no words that began with a vowel and no consonants in the first-syllable onset which would also lower first-syllable stress. Table 2 shows the percentage of first syllable stress grouped by the item type. Item type is differentiated by how many letters are in the coda and onset for each word.

Table 2

Percentage of First-Syllable Stress by Item Type in Study 2

Onset Length	Coda Length	Example Item	Mean Percentage of First-Syllable Stress	Subject Standard Deviation
One	One	‹cubbaf›	73	24
One	Two	‹cubbaft›	38	25
Two	One	‹crubbaf›	74	21
Two	Two	‹crubbaft›	44	29

A generalized logistic mixed-model regression was run using the lme4 package in R (Bates et al., 2015) with the factors being onset length, coda length, and the interaction between onset and coda length. Quadruplet and subject were included as random effects. Onset length did not produce a significant effect (*b* = – 0.218, *SE* = 0.432, *p* = 0.613). There was a strong significant effect of coda length (*b* = – 2.367, *SE* = 0.430, *p* < .001) such that 2 letters in the coda meant that first-syllable stress was much less likely than a single letter in the coda. In other words, more letters in the coda pulled the stress to the second syllable of the word. There was not a significant interaction between onset and coda length (*b* = 0.267, *SE* = 0.268, *p* = 0.318). These results suggest that there is no influence of onset length in the absence of prefixation, but there is a strong influence of coda length on lexical stress.

3.3 Discussion

The strong effect of coda length on first syllable stress assignment was expected in the results, but it was surprising that onset length did not have an effect as the corpus analysis and past literature psycholinguistic has shown such an effect (Kelley, 2004; Treiman et al., 2020). One account for this finding is that onset length has a small effect, and the number of trials to be analyzed for the study was relatively small with only 36 trials for each participant. It could also be that linguists are correct when they state that the beginnings of syllables (and this would include onsets of words) do not impact the tendency of a syllable to hold weight and therefore stress (Halle & Vergnaud, 1987).

Alternatively, as I raised in the introduction, it is possible that there is another variable that is accounting for the previously-found impact of onset length on lexical stress assignment: prefixation. In my corpus analysis, 48% of the words with a prefix had first syllable stress while 95% words without prefixes had first syllable stress. This difference was significant across grade levels with the college results being most relevant for this study as it worked with a college student sample (*b* = -1.80, *SE* = 0.28, *p* < .001) This means that it could appear as though onset length is the driving force for moving stress assignment, or it could be that going from a word with a prefix to a word that likely does not have a prefix accounts for the change in stress assignment. Since this study eliminated all orthographic and sound-alike prefixes, the effect of onset length was also diminished. I conduct Study 4 to further explore these different possibilities, but before that I do a version of Study 2 as a metalinguistic task.

Chapter 4: Study 3

This study serves as a compliment to Study 2. It compares self-reports of lexical stress with the trained scorings of the pronunciations of Study 2 with the same stimuli. This addresses the question of if individuals' understanding of stress matches what they are doing when encountering new words.

4.1 Methods

4.1.1 Participants

Ninety-six participants were recruited from the same pool as Study 2. Of the 96 participants who began the study, five participants were excluded because they failed to complete the training trials and move on to the experimental trials, and one was excluded for completing the study twice. An additional 31 participants were excluded for failing to correctly identify stress in 75% of the English filler words. A total of 59 participants' data was analyzed.

The included participants' mean age was 20 (range=18 – 20). Forty-one participants identified as female and 18 as male.

4.1.2 Stimuli

The nonwords included in the study were from the same set of quadruplets as in Study 2 plus 20 filler words as a check that the participants understood the concept of stress. The filler words were two syllable English words. Half of the words had first-syllable stress and half had second-syllable stress. The words were equated for length across stress conditions with both having an average of 5.9 letters per word. Each participant saw all 20 filler words and 36 nonwords for a total of 56 items. Only 55 items were analyzed, however, because one quadruplet was repeated in the presentation to a quarter of the participants and removed from further analyses.

The training stimuli consisted of 18 real English words, half first-syllable stress and half second-syllable stress.

4.1.3 Procedure

This study was completed on Qualtrics in a place and time of the participants' choosing. Participants began by completing a short demographic survey. They then completed the 18 training trials in which they read a real English word and selected which syllable (1 or 2) was stressed. The participant received feedback on their answers. For example, if they incorrectly specified "battle" as second-syllable stress, they would see a screen with the words "No, that is first-syllable stress." The subject was required to select the correct answer for every item before being able to move on to the experimental trials. After completing the training trials, participants randomly saw one of four lists. As in Study 2, participants only saw one item from each quadruplet with an even spread of the four different word types. Following the same procedure as the training trials, participants read the nonword or filler word and indicated which syllable (1 or 2) they thought was stressed in the word. The survey took an average of 15 minutes to complete.

4.2 Results

Among the 59 participants who made it past the exclusion criteria, they scored an average of 93% on the filler words. Participants assigned items first syllable stress 59% of the time. Table 3 shows the mean percentage of first syllable stress and standard deviation for subjects by item type.

Table 3

Percentage of First-Syllable Stress by Item Type in Study 3

Onset Length	Coda Length	Example Item	Mean Percentage of First-Syllable Stress	Subject Standard Deviation
One	One	‹cubbaf›	67	26
One	One	‹cubbaft›	52	25
Two	One	‹crubbaf›	76	22
Two	One	‹crubbaft›	58	22

I ran a logistic mixed model regression using the same variables (onset length, coda length, and the interaction between onset and coda length) and random effects (participant and quadruplet) as Study 2. There were significant effects of coda (b = –0.65, SE = 0.33, p = 0.048) and onset (b = 0.70, SE = 0.34, p = 0.041). As in Study 2, the likelihood of first-syllable stress assignment decreased with greater coda length. In Study 3, unlike in Study 2 where there was no influence of onset length, greater onset length increased the likelihood of first-syllable stress. There was no interaction between coda and onset length (b = –0.20, SE = 0.21, p = 0.35).

4.3 Discussion

These results were less surprising than those of Study 2 in following the psycholinguistic literature, but they brought up their own considerations about how best to go about scoring stress assignment. They were less surprising in that both coda and onset length had a significant effect on first syllable stress assignment. This follows what is expected from the corpus analysis that I conducted and previous research (Kelly 2004; Treiman et al. 2020).

The more interesting finding in this study is how the results differ from Study 2. The two studies were designed to show the same results in two different ways, yet they had a different pattern of results. One possibility for this is that there was a larger sample size in Study 3, so it was able to pick up on the effect onset length simply because there was more statistical power. However, the difference between results in studies 2 and 3 could also have to do with the differences in methods or scoring. In Study 3, participants were instructed to self-report their stress for the nonword items after reading aloud. Nearly 1/3 of the participants failed to accurately report the stress assignment for the real English words which demonstrates how difficult it is for untrained individuals to identify stress. This would follow with the results from the O'Brien (2019) study which found that untrained individuals struggled to generate rules surrounding stress assignment. The two different results in the present studies further show how difficult it is to score stress, so future studies involving lexical stress should be cautious in how they score the productions or ask participants to self-report stress. These differences could also reflect that the metalinguistic knowledge about stress, which is tapped in self-report studies, is not the same as the behavior when encountering novel words, which would be better reflected by trained judges. Given the differences in the patterns of results, these different scoring methods could be tapping two different aspects of stress. Next, I continue the investigation into whether prefixation may account for a portion of the effect of onset length by directly testing words with and without prefixes.

Chapter 5: Study 4

Given the findings of Study 2 and Study 3, this study looks more precisely at the relationship between onset length and prefixation in stress assignment. I directly compare pairs of items that contain no prefixes or clusters that could sound like prefixes with items that contain a prefix in the single-onset item of a pair and no prefix in the double-onset item. By varying the onset length of stimuli with and without the influence of prefixation, I will be able to see if prefixation can account for, or at least influence, the impacts of onset length on lexical stress.

5.1 Methods

5.1.1 Participants

Participants were recruited from the same population as previous studies with a mean age of 19.82 (range = 18 – 23 years). Twenty-nine participants self-identified as female, 10 identified as male, and one as non-binary. All participants reported no vision, hearing, speech, or reading disorders and spoke English as a native language.

5.1.2 Stimuli

Training items and word frames were the same as those used in Study 2.

I designed 46 quadruplets of disyllabic nonwords (listed in Appendix) that included two pairs of words. For example, one quadruplet was, ‹pegav›, ‹plegav›, ‹begav›, ‹blegav›. Each pair contained an item with a single letter in the onset of one item and two letters in the onset of the other. Additionally, the single onset item in one pair contained a prefix while the other pair did not contain any prefixes or combinations of letters that could sound like prefixes. In the example given, the prefix pair would be ‹begav› and ‹blegav›. The first syllable vowel and second syllable of the words were held constant within the quadruplets. The second syllable could contain a single medial consonant followed by a single vowel and another single consonant in the coda. Each participant was presented with two items from each quadruplet for a total of 92 nonwords. The presented items were counterbalanced so that each type of item was presented equally across participants.

The prefixes included, ‹be-›, ‹de-›, ‹re-›, ‹co-›, ‹mis-›, and ‹sub-›. When possible, the no-prefix, single onset condition simply changed the voicing of the onset consonant. For example, ‹be-› became ‹pe-› in the no-prefix condition. For ‹re-› and ‹sub-›, there were multiple possibilities for the no-prefix condition, but all these consonants were stops or fricatives. ‹Mis-› became ‹fis-› for the no-prefix condition. The resulting quadruplet in this case was, ‹fisord›, ‹flisord›, ‹misord›, ‹smisord›. The double-onset clusters were formed by adding ‹r›, ‹l›, or ‹m› after the consonant in the single onset condition for both the prefix and no-prefix conditions. The prefix ‹mis-› was a special case which formed the double-onset cluster by adding ‹s› before ‹m› rather than anything after ‹m›. The prefix ‹re-› was also a special case and it was transformed to a double-onset cluster by adding a stop or fricative to the beginning of the word. For example, ‹renov› became ‹grenov›.

5.1.3 Procedure

The procedure was identical to the procedure in Study 2 except that there were more trials. The study took approximately 30 minutes to complete.

5.1.4 Scoring

Two judges who were blinded to the experiment hypotheses but trained in lexical stress scored Zoom recordings of the productions. They used the same scoring guide as the scorer in Study 2, and their agreement was calculated for each measure successively. The judges disagreed on whether the audio was usable 0.2% of the time (*n* = 6). Among the trials for which they agreed on the audio usability, there were disagreements on plausibility of pronunciations 2.2% of the time (*n* = 61). Disagreements on the presence of stress in a pronunciation occurred 0.3% of the time (*n* = 3) after the plausibility and audio usability disagreements were removed. When audio, plausibility, and stress presence were all in agreement, the scorers agreed on the stressed syllable 92.96% of the time. A dataset containing only the trials in which the scorers agreed on every measure was created (*n* = 3330) as an aggregate score of the stress assignment. Analyses were also run on the data from each scorer individually.

5.2 Results

A total of 4.13% of trials (*n* = 152) were excluded due to issues with audio, plausibility, or stress by either scorer. Of the remaining trials, Scorer 1 recorded 35.6% first syllable stress and Scorer 2 recorded 33.7% first syllable stress.

A generalized logistic regression was run using onset length, prefixation, and the interaction between onset length and prefixation as factors with subject and quadruplet as random effects for Scorer 1. There was a small but significant effect of onset length (b = 0.322, SE = 0.132, p = 0.01) such that moving from one letter in the onset of an item to two increased the presence of first syllable stress. Prefixation was also significant (b = -0.928, SE = 0.301, p < 0.01). This result shows that changing from an item with no prefix to an item with a prefix decreases the likelihood of first syllable stress. The interaction between onset length and prefixation was significant (b = 0.534, SE = 0.188, p < 0.01) which can be interpreted as the effect of onset length being more dramatic when there is change in prefixation. A post hoc analysis was run to check for an effect of onset without any influence of prefixation and a small but significant effect was found (b = 0.329, SE = 0.132, p = .01). This means that increasing onset length does result in an increase in first syllable stress above and beyond any influence of prefixation for Scorer 1.

The same analyses were run for Scorer 2 and produced very similar results. Table 4 shows the mean first syllable stress for both of the scorers side by side. The effect of onset length was significant in the same direction ($b = 0.445$, $SE = 0.133$, $p < 0.001$) as was the effect of prefixation ($b = -1.07$, $SE = 0.306$, $p < 0.001$). Similarly, the interaction was significant ($b = 0.620$, $SE = 0.190$, $p = 0.001$). As with Scorer 1, a post hoc analysis was run to check for the influence of onset length in isolation. There was a significant effect ($b = 0.434$, $SE = 0.133$, $p = 0.001$) showing that Scorer 2 also reported an increase in first syllable stress along with an increase in onset length independent of prefixation. Both scorers found that increasing onset length increases first syllable stress, adding a prefix decreases first syllable stress, and—importantly—increasing onset while changing from a word with a prefix to a word without a prefix produces a bigger effect than increasing onset alone.

Table 4

Percentage of First-Syllable Stress by Item Type in Study 4 for Scorer 1 and Scorer 2

Onset Length	Prefix in Single Onset Item in Pair?	Example Item	Mean Percentage of First-Syllable Stress for Scorer 1	Mean Percentage of First-Syllable Stress for Scorer 2
One	No	‹pegav›	34	32
Two	No	‹plegav›	39	37
One	Yes	‹begav›	29	26
Two	Yes	‹blegav›	40	40

Of the trials where the scorers agreed on the audio quality, plausibility, presence of stress, and which syllable contained the stress, the percentage of first syllable stress was 33.5%. The final analyses reported are from this set of data. The stress for each item type in this aggregated data is reported in Table 4 with the item types being differentiated by the presence or absence of a prefix and the number of letters in the onset.

Table 5

Percentage of First-Syllable Stress by Item Type in Study 4 when Scorers Agreed on All Measures.

Onset Length	Prefix in Single Onset Item in Pair?	Example Item	Mean Percentage of First-Syllable Stress	Subject Standard Deviation
One	No	‹pegav›	31	47
Two	No	‹plegav›	37	48
One	Yes	‹begav›	26	44
Two	Yes	‹blegav›	39	49

I ran another generalized logistic mixed-model regression with the same factors as the regressions for the individual scorers. Onset length produced a significant effect (b = 0.409, SE = 0.144, p = 0.004) such that a longer onset produced greater amounts of first syllable stress. There was also a significant effect of prefixation (b = – 1.079, SE = 0.332, p = .001) such that prefixation produced less first syllable stress than when there was no prefix. Most importantly, there was a significant interaction between onset and prefixation (b = 0.652, SE = 0.207, p = 0.002). These results suggest that the impact of onset length that increases first syllable stress is strengthened by the presence of a prefix in the single onset condition. In other words, the change from a single onset with a prefix to a double onset with no prefix is more likely to shift stress than simply shifting from a single onset to a double onset with no change in prefixation.

A post hoc analyses was run to look at the impact of onset length in isolation. The trials in conditions where there was a prefix in the single onset condition were removed from the data, and a generalized logistic regression was run using onset length as the only factor. Subject and quadruplet were used as random effects. There was a small, but significant effect of onset length, even when prefixation was removed ($b = 0.407$, $SE = 0.145$, $p = .005$). This shows that prefixation does not account for the impact of onset length on stress assignment. However, it does create a larger difference when there is a prefix present in the single onset condition than when there is not.

5.3 Discussion

These results support the hypothesis that the impact of onset length on stress assignment is intertwined with prefixation; however, prefixation does not completely account for the impact of onset length as there was still a significant effect when prefixation was removed from the data. This suggests that onset has an effect above and beyond prefixation, but it is not a very big effect, and that much of the effect seen in prior research and my corpus analysis can be accounted for by prefixation. In terms of Kelly (2004) and Treiman et al (2020), this means that the significance of onset in those studies could very well be confounded by the presence of phonological prefixes in their stimuli. Future work should take these findings into account when developing stimuli to look at stress assignment. Moreover, the results of this study show that the number of consonants in the onset do in fact affect stress assignment when reading aloud beyond the impact of sound-alike prefixes.

Further looking at the scoring, there can be greater confidence placed in these results as the scorers were blinded to the study hypotheses. This means that when a stress assignment could be interpreted as ambiguous, the scorers would not unconsciously select stress to suit the hypotheses. Additionally, being able to run analyses on only the data in which the scorers completely agreed imbues the conclusions with greater confidence than previous studies of stress assignment. Study 5 looks at the same stimuli with a metalinguistic task to compare the findings with the scorers' results.

Chapter 6: Study 5

Like how Study 3 was a complement to Study 2, this study serves as a metalinguistic complement for Study 4. There were differences in the findings of the metalinguistic measure in Study 3 and the scorer findings in Study 2, so it would be reasonable to expect that these might also differ. If they do, then that points to a lack of conscious awareness of the factors that influence lexical stress in untrained individuals.

6.1 Methods

6.1.1 Participants

Eighty undergraduate students were recruited from the same sample as previous studies. A total of 24 participants were excluded for failing to correctly identify stress in the real English filler words. A total of 55 participants' data was uses for the analyses.

Of the participants whose data was included in the final analyses, participants' mean age was 20.13 and 38 self-identified as female, 13 as male, and 1 as non-binary. All participants reported no reading, hearing, speech, or vision disorders and spoke English as their native language.

6.1.2 Stimuli

The stimuli included in the study were drawn from the same 46 quadruplets as in Study 4 and the same 20 filler words from Study 3. Each participant saw 92 nonwords and 20 filler words for a total of 112 experimental items. As in Study 4, each participant saw two items from each quadruplet counterbalanced so that there was an equal distribution of the different types of items. The training stimuli was the same as in Study 3.

6.1.3 Procedure

The procedure was the same as in Study 3 except that it contained more items. The survey took approximately 30 minutes to complete.

6.2 Results

Among the 55 participants who met the inclusion criteria, they scored an average of 90.55% on the filler words. Participants assigned items first syllable stress 52.5% of the time. Table 5 shows the mean percentage of first syllable stress and standard deviation for subjects by item type.

A final logistic mixed model regression was run using onset length, prefixation, and the interaction between onset length and prefixation as factors with subject and quadruplet as random effects. There was a significant effect of onset length (b = 0.37, SE = 0.09, $p < .001$). This would suggest that an increase in onset length increases the likelihood of subjects to self-report a first-syllable stress assignment. There was a marginal effect of prefixation (b = - 0.37, SE = 0.2, p = .066) and no significant interaction between prefixation and onset length (b = 0.21, SE = 0.13, p = 0.11) suggesting that there may not be a strong influence of prefixation in self-reports of stress.

Table 6

Percentage of First-Syllable Stress by Item Type in Study 5

Onset Length	Prefix in the single onset pair?	Example Item	Mean Percentage of First-Syllable Stress	Subject Standard Deviation
One	No	‹pegav›	49	18
Two	No	‹plegav›	57	15
One	Yes	‹begav›	46	16
Two	Yes	‹blegav›	57	18

6.3 Discussion

Comparable to how Study 3 differed from Study 2, these results differed from Study 4. Beyond that, these results are not what would be expected given past literature and the present research. Additionally, there continues to be a large amount (30% in this study) of participants excluded from the data due to poor filler performance. These differences suggest that the method for measuring stress assignment could determine the results of a given study. It seems that individuals report that onset length should be important as it was significant in both self-report studies. However, there is a somewhat smaller effect of onset length on the actual productions when presented with a nonword. The consistent variations between the different study designs indicate a disconnect between what happens with stress when reading novel words aloud and what untrained individuals guess will happen when they encounter a new word. This may be related to the metalinguistic knowledge afforded to an untrained individual by the writing system used in their language. Since there is no indicator for stress in English, untrained speakers of English are less sensitive to stress than speakers of a language that does notate stress, like Greek.

Chapter 7: General Discussion

The goal of this project was to examine some of the intricacies of lexical stress assignment in disyllabic words. I was first interested in if predictors found in a corpus analysis of lexical stress could vary by age. I then wanted to explore if the impacts of coda and/or onset length were still impactful when the presence of prefixes was eliminated—especially the presence of clusters that sounded like prefixes. Further, I wanted to know if the impact of onset length specifically could be predicted by prefixation. Additionally, I aimed to determine the impact that using different methods of scoring could have on the results of a study, or if they even measure the same concept.

7.1 Conclusions on Predictors of Stress in English Corpora

My corpus analysis showed that coda length, onset length, and prefixation (along with lexical class and the presence of a word-medial geminate cluster) all impact lexical stress assignment, and that these factors do not vary as a function of age. These results were inconsistent with the past literature from Arciuli et al. (2010), but that could be at least partially explained by differences in the way the words were analyzed. Arciuli et al. (2010) eliminated any possible influence of medial cluster identity when they broke apart the words to analyze while this analysis keeps the influence of the medial cluster identity. Given that the medial cluster identity was a significant predictor for both the corpus analysis here, and in Treiman et al. (2020), this omission of the middle parts of the words may have exaggerated the effect of grade level on prediction.

Overall, the corpus results show subtle patterns that reflect a lot of predictors for stress assignment in English including onset, coda, and prefixation. It is generally agreed upon that these patterns exist; a more interesting question is if they are reflected in the behavior of individuals who are naïve to their existence.

7.2 Conclusions on Models of Reading and Stress

As mentioned in the introduction, the two basic categories of modeling reading are a statistical/connectionist model and a dual-route model of reading. A statistical model would be supported by individuals picking up on patterns in the language, even untaught and subtle patterns. Conversely, a dual-route model would be supported by evidence that individuals learn the general rules of a language but fail to notice more subtle patterns. On the topic of onset length, there is a small but significant effect of onset length on lexical stress assignment such that a longer onset increases the likelihood of first-syllable stress found in studies 4 and 5. However, prefixation accounts for some portion of this effect as Study 4 found an interaction between onset and prefixation. This follows from linguistic work which identifies prefixes as stress repellent (Fudge, 2015). The Ktori et al. (2016) with dyslexics assigning stress is another example of prefixation being found to reduce first syllable stress in a behavioral study, albeit with a different subject population. Additionally, there was no effect of onset length found in Study 2, which does not follow from the patterns found in the corpus analysis. This suggests that a dual-route model of reading may be more supported by that present studies, at least for stress assignment, then a statistical model.

This may be accounted for by how the writing system of English fails to explicitly represent stress. There is evidence that the writing system of a language can impact its users' metalinguistic awareness of features in the language. Read et al. (1986), found that adults who were only familiar with the Chinese character writing system had greater difficulty manipulating the phonemes in a word than adults of a similar literacy level who were familiar with both the

Chinese character system and an alphabetic writing system. They asserted that this was because a character writing system did not break down the words into sounds in the way that an alphabetic system would—roughly—represent each sound with a letter. This shows that the orthography of a language affects one's understanding of the spoken form of their language. A similar mechanism might translate to stress in English, as stress is not explicitly noted in the writing. Other languages (like Greek) explicitly note the stress in a word, often using an accent mark. This question was investigated by Jimenez and Garcia (1995) when they found that Spanish-speaking children (as Spanish is a language that explicitly notates stress in the writing system more than English does) outperformed English-speaking children on a pronunciation task which varied the stress of the words. Metalinguistic awareness—which is already low for English stress assignment—could play a large role in the success of that assignment for novel words. Going back to O'Brien's (2019) findings about subjects' success at verbalizing their own stress assignment predicting their accuracy of stress assignment, it seems that metalinguistic awareness is a factor that should be considered when modeling stress.

7.3 Conclusions on Methodological Concerns with Stress

The last aim of the study gets at if lexical stress can be studied, and the importance of the choices made by researchers. The findings of this project suggest that a trained judge scoring pronunciations of nonwords measure something different than a metalinguistic report of where the stress should fall in those same nonwords. Studies 1 and 2 and studies 3 and 4 both had different patterns of results. Further, there were high rates of error in the filler words for the self-report. Taken together, these suggest that untrained individuals do not have a good understanding of what predicts lexical stress when they come across novel word, nor do they know where the stress occurs in familiar words. The metalinguistic self-report measures this understanding, or perhaps this misunderstanding, while the trained judges measure the instinctive production of lexical stress when encountering a new word.

One example of these measures getting at different concepts, were discrepancies in the Arciuli et al (2010) paper as they also ran a behavioral study to check for age differences in stress assignment. In this study, they had children read isolated nonwords that contained beginnings and endings that they identified as indicating either first or second syllable stress. For example, the item, ‹bedoon› would have both a beginning and an ending that suggested second-syllable stress, while ‹beppet› has a beginning that produces second-syllable stress and an ending that produces first-syllable stress. They found that children of all ages were sensitive to both beginnings and endings' contributions to stress, but that older children were more sensitive to the influence of word endings. This does not agree with their corpus study findings that the words in the younger-grade levels provide more orthographic information about stress assignment than words in higher grade levels. Even though, the differences in the corpora suggested that there is more information for younger children, the behavioral results showed better performance for older children. Another example of this discrepancy between corpus and behavioral results comes from the Treiman et al. (2020) paper. In the corpus analysis, medial cluster identity is a significant predictor of stress assignment, but it did not have an influence on the actual behavior of participants. Results like these suggest that individuals are not necessarily sensitive to the subtle patterns present in the corpora to which they are exposed. It would benefit future research to examine why this discrepancy exists and how to make these patterns more explicit.

7.4 Limitations

It should be noted that the present research is limited to disyllabic words, and the rules for English stress become more complicated as words get longer and secondary (and tertiary, etc.) stress must be considered. That is not to say that the general principles of stress assignment and the predictors would not hold for longer words. Rather, as the predicting factors increased for disyllabic words, there will be more factors and more ways the predicting factors can interact as word length increases.

Additionally, the sentence frames in the behavioral studies were, perhaps, overly simple. As the behavioral studies were originally designed to be able to run with children as young as second grade, the frames needed to be simple enough that young children could read them. These frames only used four possible subjects and four possible adverbs shuffled around to form only 16 total sentences. This meant that the sentence frames were seen multiple times even in Study 1 which had the fewest trials. This design may have limited the natural fluency when reading aloud for skilled readers because the sentence frames became rehearsed. The rehearsal of the word frames means that participants could, in a sense, memorize the prosody, or rhythms, of the sentences as well. This could serve as a confound with the results in the study if the participants were reciting the whole sentences as memorized chunks rather than reflecting the influence of word level characteristics.

7.5 Future Directions

Future work should continue the line of questioning about the role of prefixation in stress assignment by next directly investigating whether clusters that sound like prefixes, like those included in Kelly (2004) and Treiman et al (2020) but avoided in these studies, can influence stress assignment. This would allow for a more complete answer to the question: is it important for stress assignment that a cluster be both orthographically a prefix and phonologically a prefix or is the phonology of a prefix enough to influence how likely that syllable is to carry the stress of a word. In other words, does a syllable need to be spelled as a prefix and pronounced as a prefix for it to impact stress? Or is enough that a syllable sounds like a prefix to influence its stress? The present work suggests that it may be the case that the phonology is enough, but a direct test would be best to emphatically state that phonology of prefixes is most important for stress assignment. This test could be done by creating stimuli that contain no prefixes, real prefixes, and syllables that sound like prefixes and seeing if there is an influence of the soundalike prefixes as well as the real prefixes. Then participants would read them aloud and be scored in a method like studies 2 and 4. If there are not differences between the sound-alike prefixes and the lexical prefixes, that would show that the phonology of prefixes is the most important for determining stress. If there is a difference, it would follow that the orthographical knowledge is important for stress assignment.

Another line of future work should study how predictors of stress vary as word length increases. The word in this paper sets up predictors about disyllabic words, and those principles can extend to longer multisyllabic words, but there are undoubtedly other variables that come into play when the word length increase. I suspect that medial consonant clusters will become increasingly important in the determination of stress. It would also be interesting to see if the factors that determine primary versus secondary stress are the same as those that determine stress when there is only one possible stress pattern. This could be looked at through corpus analysis of English words to find general patterns and followed up with behavioral data to see how individuals intuitively assign stress.

Additionally, though the patterns of stress predictors are the same in the corpora for all grade levels, it is important to clarify if there are differences in the behavior of children over time. As there were differences in the Arciuli et al (2010) study between the corpus and behavioral studies, it is possible that there are age-based behavioral differences in stress assignment even if there are not differences in the words that children are exposed to. The present studies could be given to children of different age groups. Then, in a similar method to the corpus analysis of Study 1, the researcher would look if there was different predictive success because of grade level. The researchers could also look at if different factors were more or less important for different grades.

Overall, studying lexical stress is already a complicated process, and this project increases the knowledge about the nuance involved when undertaking a research question involving stress. Onset, coda length, and prefixation contribute to lexical stress assignment. Greater coda length leads to less first-syllable stress while greater onset length leads to more first-syllable stress, though this effect is smaller. Prefixation affects stress in that going from a prefix to a word without a prefix increases the likelihood of first-syllable stress, so prefixes reduce stress. This also accounts for some of the effect of onset length found in prior models of stress assignment like Treiman et al. (2020). These factors do not seem to change as a factor of age, but the scoring method is important. Self-report methods act as a metalinguistic view to show participants' misconceptions about stress while trained judges can identify stress assignment when faced with novel nonwords.

References

Arciuli, J., Monaghan, P., & Seva, N. (2010). Learning to assign lexical stress during reading aloud: Corpus, behavioral, and computational investigations. *Journal of Memory and Language*, *63*(2), 180-196. https://doi.org/10.1016/j.jml.2010.03.005

Baker, R. G., & Smith, P. T. (1976). A psycholinguistic study of English stress assignment rules. *Language and Speech*, *19*(1), 9-27. https://doi.org/10.1177/002383097601900102

Bates, D., Mächler, M., Bolker, B. M., & Walker, S. C. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Benrabah, M. (1997). Word-stress–a source of unintelligibility in English. *International Review of Applied Linguistics in Language Teaching 35* (3), 157–165. https://doi.org/10.1515/iral.1997.35.3.157

Booij, G. E. (1983). Principles and parameters in prosodic phonology. *Linguistics 21*(1), 249-280. https://doi.org/10.1515/ling.1983.21.1.249

Centre for Speech Technology Research (n.d.). Unisyn lexicon release, version 1.3. http://www.cstr.ed.ac.uk/projects/unisyn/.

Chomsky, N., & Halle, M. (1968). *The sound pattern of English.* New York: Harper and Row.

Coltheart, M., Rastle, K., Perry, C., Langdon, R., & Ziegler, J. (2001). DRC: a dual route cascaded model of visual word recognition and reading aloud. *Psychological review*, *108*(1), 204.

Fear, B. D., Cutler, A., & Butterfield, S. (1995). The strong/weak syllable distinction in English. *The Journal of the Acoustical Society of America*, *97*(3), 1893-1904.

Fudge, E. (1984). *English word-stress*. Routledge.

Gordon, M., & Roettger, T. (2017). Acoustic correlates of word stress: A cross-linguistic survey. *Linguistics Vanguard*, *3*(1), 1-11.

Guion, S. G., Clark, J. J., Harada, T., & Wayland, R. P. (2003). Factors affecting stress placement for English nonwords include syllabic structure, lexical class, and stress patterns of phonologically similar words. *Language and Speech*, *46*(4), 403-426. https://doi.org/10.1177/00238309030460040301

Halle, M., & Vergnaud, J. R. (1987). *An essay on stress*. Cambridge, MA: MIT press.

Hino, Y., & Lupker, S. J. (2000). Effects of word frequency and spelling–to–sound regularity in naming with and without preceding lexical decision. *Journal of Experimental Psychology: Human Perception and Performance*, *26*(1), 166. https://doi.org/10.1037/0096-1523.26.1.166

Jiménez González, J. E., & García, C. R. H. (1995). Effects of word linguistic properties on phonological awareness in Spanish children. *Journal of Educational Psychology*, *87*(2), 193. https://doi.org/10.1037/0022-0663.87.2.193

Jouravlev, O., & Lupker, S. J. (2014). Stress consistency and stress regularity effects in Russian. *Language, Cognition and Neuroscience*, *29*(5), 605-619. https://doi.org/10.1080/01690965.2013.813562

Kelly, M. H. (1988). Rhythmic alternation and lexical stress differences in English. *Cognition*, *30*(2), 107-137. https://doi.org/10.1016/0010-0277(88)90037-6

Kelly, M. H. (2004). Word onset patterns and lexical stress in English. *Journal of Memory and Language*, *50*(3), 231-244. https://doi.org/10.1016/j.jml.2003.12.002

Ktori, M., Tree, J. J., Mousikou, P., Coltheart, M., & Rastle, K. (2016). Prefixes repel stress in reading aloud: Evidence from surface dyslexia. *Cortex*, *74*, 191-205. https://doi.org/10.1016/j.cortex.2015.10.009

Ktori, M., Mousikou, P., & Rastle, K. (2018). Cues to stress assignment in reading aloud. *Journal of Experimental Psychology: General*, *147*(1), 36-61. https://doi.org/10.1037/xge0000380

Mattys, S. L., & Samuel, A. G. (1997). How lexical stress affects speech segmentation and interactivity: Evidence from the migration paradigm. *Journal of Memory and Language*, *36*(1), 87-116. https://doi.org/10.1006/jmla.1996.2472

Mousikou, P., Sadat, J., Lucas, R., & Rastle, K. (2017). Moving beyond the monosyllable in models of skilled reading: Mega-study of disyllabic nonword reading. *Journal of Memory and Language*, *93*, 169-192. https://doi.org/10.1016/j.jml.2016.09.003

O'brien, M. G. (2019). Attending to second language lexical stress: exploring the roles of metalinguistic awareness and self-assessment. *Language awareness*, *28*(4), 310-328. https://doi.org/10.1080/09658416.2019.1625912

Rastle, K., & Coltheart, M. (2000). Lexical and nonlexical print-to-sound translation of disyllabic words and nonwords. *Journal of Memory and Language, 42*, 342–364. https://doi.org/10.1006/jmla.1999.2687.

Read, C., Zhang, Y.-F., Nie, H.-Y., & Ding, B.-Q. (1986). The ability to manipulate speech sounds depends on knowing alphabetic writing. Cognition, 24(1/2), 31–44. https://doi.org/10.1016/0010-0277(86)90003-X

Seidenberg, M. S., & McClelland, J. L. (1989). A distributed, developmental model of word recognition and naming. *Psychological review*, *96*(4), 523. https://doi.org/10.1037/0033-295X.96.4.523

Sippel, L., & Martin, I. A. (2022). Immediate and long-term improvement in lexical stress perception: the role of teacher and peer feedback. *International Review of Applied Linguistics in Language Teaching*. https://doi.org/10.1515/iral-2021-0175

Treiman, R., Kessler, B., & Bick, S. (2003). Influence of consonantal context on the pronunciation of vowels: A comparison of human readers and computational models. *Cognition*, *88*(1), 49-78. https://doi.org/10.1016/S0010-0277(03)00003-9

Treiman, R., Kessler, B., & Hensley, K. (2023). Number and syllabification of following consonants influence use of long versus short vowels in English disyllables. *Journal of Memory and Language*, *129*, 104399. https://doi.org/10.1016/j.jml.2022.104399

Treiman, R., Rosales, N., Cusner, L., & Kessler, B. (2020). Cues to stress in English spelling. *Journal of Memory and Language*, *112*, 104089. https://doi.org/10.1016/j.jml.2020.104089

Yu, V. Y., & Andruski, J. E. (2010). A cross-language study of perception of lexical stress in English. *Journal of psycholinguistic research*, *39*, 323-344. https://doi.org/10.1007/s10936-009-9142-2

Zeno, S., Ivens, S. H., Millard, R. T., & Duvvuri, R. (1995). *The educator's word frequency guide*. Touchstone Applied Science Associates.

Appendix

Studies 1 & 2 Nonword Experimental Items: cubbaf, cubbaft, crubbaf, crubbaft; tidduc, tidduct, twidduc, twidduct; fannel, fanneld, frannel, franneld; feppor, fepporb, fleppor, flepporb; guttis, guttist, gruttis, gruttist; piffam, piffamp, pliffam, pliffamp; pozzef, pozzeft, plozzef, plozzeft; tammof, tammoft, trammof, trammoft; paffis, paffist, plaffis, plaffist; kibbaf, kibbaft, klibbaf, klibbaft; foppuc, foppuct, floppuc, floppuct; kaddom, kaddomp, kladdom, kladdomp; gummil, gummilp, glummil, glummilp; fassop, fassopt, frassop, frassopt; pozzuf, pozzuft, plozzuf, plozzuft; febbor, febborg, flebbor, flebborg; puddac, puddact, pluddac, pluddact; kippen, kippent, krippen, krippent; kozzup, kozzupt, krozzup, krozzupt; kaggor, kaggord, klaggor, klaggord; tullan, tullant, trullan, trullant; fassul, fassuld, frassul, frassuld; cussil, cussilt, crussil, crussilt; kezzap, kezzapt, krezzap, krezzapt; guppor, gupporm, gluppor, glupporm; tagguf, tagguft, tragguf, tragguft; fommep, fommept, flommep, flommept; tozzim, tozzimp, trozzim, trozzimp; pubban, pubbant, plubban, plubbant; pizzoc, pizzoct, plizzoc, plizzoct; godduf, godduft, glodduf, glodduft; kebbol, kebbolg, krebbol, krebbolg; tuggle, tuggelt, truggel, truggelt; fissac, fissact, frissac, frissact; follum, follump, frollum, frollump; teffap, teffapt, treffap, treffapt

Studies 3 & 4 Nonword Experimental Items: pegav, plegav, begav, blegav; tezok, trezok, dezok, drezok; senug, stenug, renug, frenug; govik, grovik, covik, crovik; fisord, flisord, misord, smisord; fuboz, fluboz, suboz, sluboz; pefin, plefin, befin, brefin; tepum, trepum, depum, drepum; sezab, slezab, rezab, grezab; gofaw, glofaw, cofaw, clofaw; fisost, frisost, misost, prisost; fubaz, frubaz, subaz, smubaz; pegub, plegub, begub, blegub; tefaw, trefaw, defaw, drefaw; fezoy, frezoy, rezoy, grezoy; gobiv, grobiv, cobiv, crobiv; fisuf, flisuf, misuf, plisuf; fubork, flubork, subork, slubork; pemaz, plemaz, bemaz, blemaz; telig, trelig, delig, drelig;

gepab, grepab, repab, frepab; gomar, glomar, comar, clomar; tubost, trubost, subost, stubost; penord, plenord, benord, blenord; tevack, trevack, devack, drevack; gevam, grevam, revam, frevam; fisib, flisib, misib, smisib; gopog, gropog, copog, cropog; fubil, frubil, subil, slubil; fisork frisork, misork prisork; pevorn, plevorn, bevorn, blevorn; tegob, tregob, degob, dregob; fenov, flenov, renov, grenov; fisog, frisog, misog, prisog; gomab, glomab, comab, clomab; tubap, trubap, subap, slubap; pegip, plegip, begip, blegip; tekaz, trekaz, dekaz, drekaz; petoz, pletoz, retoz, fretoz; fisot, frisot, misot, prisot; govug, grovug, covug; crovug, fubek, flubek, subek, slubek

Studies 1 & 3 Training Items: zedav, rudit, jecud, hetib

Studies 2 & 4 Training Items: susan, suzanne, ella, elaine, jerome, jersey, balloon, battle, supper, supply, commute, comment, pretend, ignore, pretty, illness, revolt, ramble

Studies 2 & 4 Filler Words: alley, anger, comet, dazzle, inward, otter, pattern, pirate, prudent, sisters, allow, around, command, deny, invent, occur, patrol, polite, predict, sustain