KEY, KRISTIN M. M.A. Sublexical representations in adult skilled spellers. (2022)
Directed by Dr. Robert Wiley. 81 pp.

Despite being an important aspect of written language, our understanding of the processes underlying spelling remains greatly limited. Spelling new words is not done via a series of random guesses. One prominent theory proposes that spelling occurs through a probabilistic system that is able to track what combinations of letters are likely to occur in the English language. This system is known as the sublexical system, which processes orthographic and phonological information at some level below that of a whole word. Processed words can then be assimilated into our orthographic long-term memory (O-TLM). I wish to further investigate the processes at work in the sublexical system, and how it interacts with the lexical system. Over a series of three experiments, I investigate four hypothesized representations of sublexical spelling to see which are best supported by our findings. I will investigate which of these mechanisms best explain pseudoword spelling responses, how they relate to real word spelling, and what influence they may have on learning.

SUBLEXICAL REPRESENTATIONS IN ADULT SKILLED SPELLERS

by

Kristin M. Key

A Thesis
Submitted to
the Faculty of The Graduate School at
The University of North Carolina at Greensboro
in Partial Fulfillment
of the Requirements for the Degree
Master of Arts

Greensboro

2022

Approved by

____
Dr. Robert Wiley
Committee Chair

DEDICATION

For Scott, whose passion should continue to be an inspiration to us all, and who was undoubtedly the kindest person I ever met. We miss you.

APPROVAL PAGE

This thesis written by Kristin M. Key has been approved by the following committee of the Faculty of The Graduate School at The University of North Carolina at Greensboro.

Committee Chair ____________________________
Dr. Robert Wiley

Committee Members ____________________________
Dr. Michael Kane

____
Dr. Peter Delaney

February 22, 2022
Date of Acceptance by Committee

October 5, 2022
Date of Final Oral Examination

ACKNOWLEDGEMENTS

I would like to extend my deepest gratitude to my advisor Dr. Robert Wiley, without whom I would have never completed this research. His guidance and endless patience were a truly a blessing even during the most difficult of times.

I would also like to thank my committee members Dr. Michael Kane and Dr. Peter Delaney for their wonderful insights, and for making my defense a collaborative and enjoyable experience.

TABLE OF CONTENTS

LIST OF TABLES

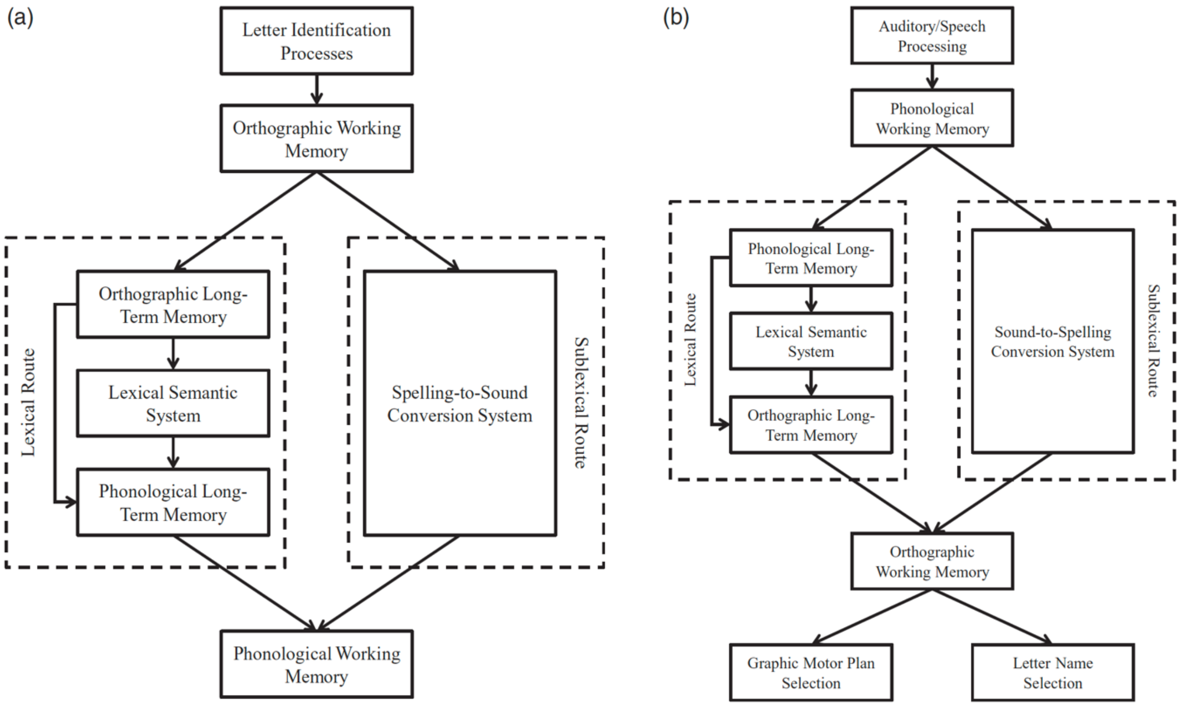
LIST OF FIGURES

CHAPTER I: THE COGNITIVE ARCHITECTURE OF SPELLING

I. Introduction

How do we go from hearing a novel word to generating a spelling? Dual route models are the dominant account, and state that the two routes - the lexical and the sublexical - operate in parallel and share information in order to generate the most accurate spellings of given words (Coltheart et al., 2001; Rapp et al., 2001). The lexical route involves using stored knowledge in order to provide the proper spelling and pronunciation of given words. This stored knowledge involves phonological, orthographic, and semantic representations held in long term memory (Rapp et al., 2002). In simple terms, this is the process used by skilled readers during written language tasks when they already know the presented words. The sublexical route, on the other hand, is thought to operate with smaller parts of words, such as phonemes and graphemes and the relationship between the two. This route is thought to contribute to parsing what is heard or read when novel words are encountered and gives a best estimate of how that word should be spelled/pronounced (Rapp et al., 2002). However, as seen in *Figure 1*, whereas the lexical route consists of multiple, well-studied components, the sublexical route is simply referred to as a monolithic “spelling-to-sound conversion system.” There is general disagreement about what the representations within the sublexical route are, how they interact with lexical processes, and their role in learning. This research will investigate four proposed mechanisms that may be at work within the sublexical system: a probabilistic Phoneme-Grapheme (P-G) Mapping system, a Phonological Neighborhood system, an Oncleus, and a Rime system.

Figure 1. ***Model from Hepner et al. (2017)***



Note: (a) the reading model and (b) the spelling model.

Before detailing the proposed mechanisms internal to the sublexical route, let us describe how the heard word "yacht" would be analyzed for spelling to dictation by both routes, in order to briefly elaborate on the basic dual-route model. When processing the heard word "yacht," the sublexical route uses lower-order representations to map the phonemes, /j/, /ɑ/, and /t/ to corresponding graphemes. The most plausible graphemes might be Y-A-T, or Y-O-T. However, the lexical route functions simultaneously using higher-order representation of whole word forms to retrieve the spelling from orthographic long-term memory (O-LTM), where the string of graphemes Y-ACH-T is stored. The retrieved information of the encoded spelling overrides the sublexical route in this case, as it is providing known, concrete evidence of the irregular spelling of the word. According to the lexical quality hypothesis (Perfetti & Hart, 2002), individuals have differing qualities of phonological, semantic, and orthographic representations, meaning that some individuals will be more or less likely to rely on stored information. If it is the case that there is no information in O-LTM, or it is poorly consolidated such that the recalled information is incomplete, the spelling provided by the sublexical route may be used. Let us take for example spelling behaviors in dysgraphic patients, such as in the case of an individual who spelled "bouquet" → BOUKET (Rapp et al., 2002). Here, the K is a plausible spelling likely provided by sublexical conversion, while the T is a fairly nonobvious spelling and was likely provided by long-term lexical knowledge. This example of disordered spelling provides evidence that the lexical and sublexical routes do in fact interact. Ideally, the lexical takes precedence in cases of known word and the sublexical takes precedence in cases of poorly known or novel words to provide the most accurate spelling possible.

There is relatively little consensus about the representations used internally in the sublexical route, or the specific mechanisms used to process novel word stimuli (Rapp et al., 2001). One proposal, which will be referred to as P-G Mapping, argues that spelling is achieved by a one-to-one mapping of phonemes onto the most probable graphemes based on the entire English lexicon. Another, the Phonological Neighborhood account, asserts that phoneme-grapheme mapping probabilities are influenced by the local phonological neighborhood of the spoken word. Finally, another account contends that units larger than P-G's, specifically the oncleus (onset + nucleus) and rime (nucleus + coda), contribute to selection of the mappings between sound and spelling.

The primary focus will be on investigating these four proposed mechanisms by accounting for individual variations in graphemic assignment (i.e., spelling to dictation). I will compare the multiple proposed accounts to determine which, if any, show evidence of being utilized during pseudoword spelling tasks. The pages that follow will describe 1) the proposed mechanisms, laying out evidence provided for each by previous literature; 2) the series of experiments designed to examine the proposed mechanisms; and finally 3) the results of said experiments and what they indicate.

II. Sublexical Mechanisms

Here I provide a brief overview of proposed mechanisms utilized by the sublexical route: P-G Mappings, Phonological Neighborhoods, Oncleuses, and Rimes. To examine these mechanisms, a series of experiments was designed tasking people with generating spellings for heard pseudowords. Their patterns of behavior on this task were measured, along with how those patterns of behavior relate to real word spelling and new word learning. These experiments are necessary because while it is generally understood that spelling a word from auditory input involves using information provided by phonemic parsing to generate graphemic assignments (Meng et al., 1994), there is wide disagreement regarding how this graphemic assignment occurs. To further elaborate what this means, it is important to understand that phonemic parsing involves how a heard word is broken apart into its individual sound units, e.g., “when” will be parsed as /w/, /ɛ /, /n/. When attempting to spell a new word, this parsed phonemic information is used by the sublexical route for graphemic assignment, which entails the assignment of specific letter combinations to heard phonemes in order to represent them in a written format. The phonemes for “when” would be graphemically mapped as /w/ → WH, / ɛ / → E, /n/ → N. However, this P-G mapping of /wɛn/ is only one of a variety of possible mechanisms which may be utilized by the sublexical route.

A potential factor of note for the sublexical route is consistency: the extent to which similar sounding words are also spelled similarly, and vice versa. Consistency of feedforward (spelling-to-sound, such as in oral reading) and feedback (sound-to-spelling, such as in spelling to dictation) information is extremely important during spelling and reading tasks. In general, consistency means that if two words are spelled similarly, you should pronounce them similarly (feedforward), and if two words are pronounced similarly, they should be spelled similarly (feedback). The relationship between feedforward and feedback consistency is not a symmetric one; it is entirely possible that a language can be consistent in one direction but inconsistent in another (Marjou, 2019). However, English tends to be generally inconsistent in both directions, and not all phonemes and graphemes within English are inconsistent in the same manner. For example, the grapheme RH can only be pronounced in one way (/r/), but there are many ways to spell /r/ (R, RR, RH, WR). In the same vein, the phoneme /ð/ can only be spelled as TH, but TH can be pronounced in numerous ways (/ð/, /θ/, /t/). Perfect consistency would generate 1-to-1 mappings between sound and spelling, while perfect inconsistency would indicate no relationship at all between the sound and the spelling (neither of which happen in the English language). It should be noted that these analyses are not going to be looking at feedforward (spelling-to-sound) information, as this research is solely examining what happens when participants hear a pseudoword for the first time and attempt to generate a spelling response. People may, either overtly or covertly, alter their spellings due to feedforward information after completing their P-G conversion, but that is beyond the scope of this current study.

While it is clear that consistency matters in regard to spelling tasks, how consistency is defined is still up for debate. For example, there is a widespread idea that “friends and enemies” have major implications for spelling (e.g., Jared et al., 1990). Here, a friend indicates that a set of words are all pronounced and spelled with the same rhyme (e.g., TINT and LINT are both friends of MINT), whereas an enemy indicates that the pronunciation or rhyme are not the same (PINT is a phonological enemy of MINT, and MINE is a graphemic enemy of MINT).

Researchers (e.g., Pexman et al., 1999) have determined that having more friends has facilitatory effects for spelling behaviors. Words with higher ratios of friends compared to enemies demonstrate lower naming latencies and lower response error rates (Bolger et al., 2007; Stemberger, 2004). Enemies, on the other hand, can hinder spelling tasks as the disagreement between spellings or pronunciations causes slower processing and less accurate responses due to higher variability (Bolger et al., 2007).

Friends and enemies have primarily been defined in one of two ways. The first involves phonological neighbors, which are words that are only one phoneme different (e.g., MINT and MINCE). A word's phonological neighborhood is thus the set of all such words. Some researchers state that words in a phonological neighborhood can either be orthographically similar (friends) or orthographically dissimilar (enemies; e.g., Pecher et al., 2011). For example, the word MINT has the orthographically similar phonological neighbor MIST but also the dissimilar neighbor MISSED. It merits mentioning that friends and enemies are also defined in the other direction, meaning that orthographic neighbors can be either phonologically similar or dissimilar. Some researchers have even asserted the importance of the "phonographic" language network, which are words that are both phonologic and orthographic neighbors (Siew and Vitevitch, 2019). However, for the purposes of this research I am specifically interested in friends and enemies as they relate to spelling, and therefore focus exclusively on phonological neighbors.

Most typically, friends and enemies are defined based on the rime of the syllables within a word. For example, the rime of MINT is -INT, with the consistent mapping -/ɪnt/ (e.g., LINT, TINT, HINT) and the inconsistent mapping -/aɪnt/ (PINT). Both the phonological neighbor and rime definitions of friends and enemies argue for a case of consistency, but differ in what representations are relevant to determining consistency. Below I will further explore the possible units of sublexical representation that may be utilized during spelling tasks.

II.1 Phoneme-Grapheme Mappings

In an effort to examine how to effectively teach phonics and spelling, Hanna et al. (1966) sought to use a computerized algorithm to sequence P-G correspondences in English. They compiled a list of 17,310 English words that were meant to be representative of the English language and documented each P-G mapping that appeared within these words. Each mapping was then examined for its frequency of appearance relative to its position within a syllable. Finally, they had a computer algorithm attempt to spell words based on the provided frequencies to examine whether or not English is alphabetic in nature. The corpus Hanna et al. generated provides a probabilistic database of English P-G mappings, which can theoretically be used to explore the likelihood of a given spelling. The database probabilities are sensitive not only to the mapping itself, but also to its placement within its respective syllable.

Although there have been attempts to adapt these tables for reading (grapheme to phoneme, G-P) probabilities (Berndt et al., 1987), and to update the presentation of these tables (Fry, 2004), there has not been a definitive reproduction of these now somewhat outdated tables. I argue that instead of using an updated or revised version of the Hannah mappings, a completely revised version of these G-P and P-G mappings is necessary. This is because there are a few notable issues in the original work that hinder the databases' overall usability. There were serious discrepancies in how words were parsed, with no systematic approach to addressing syllable breaks. For example, despite sharing rimes the word SATYR was parsed as SAT // YR, while the word MARTYR was parsed as MAR // TYR. There was also the issue of many words having outdated pronunciations, as the English language has changed over time. For example, Hanna et al. (1966) mapped the phoneme they labeled "I3" to the bolded graphemes in MOV**IE**, G**UI**TAR, and YESTERD**AY**. Contemporary sources for American English pronunciations map these as three different vowels: MOV**IE** → /i/, G**UI**TAR → /ɪ/, and YESTERD**AY** → /eɪ/. Finally, there were multiple low-frequency mappings that were not included in the original work by Hanna and colleagues, such as /aʊ/ → AU for S**AU**DI.

To correct the issue of irregular syllabification, I turned to the maximum onset principle (MOP) as described by Kahn (2015). In simple terms, this principle seeks to place consonants in the onset of new syllables whenever possible, so long as it is phototactically legal. I adhere to the MOP, with a few notable exceptions (see *Appendix E*). Any words mapped with outdated pronunciations have been remapped to reflect contemporary pronunciation using the CMU Pronouncing Dictionary (1995), and low frequency mappings that were not originally a part of Hanna et al.'s (1966) database have been added.

This mapping system operates based on the P-G probabilities provided by approximately 7000 words English words, which are intended to be representative of the frequencies of each P-G mapping in the English language relative to their position within the word and syllable. For this system there were four designated positions: word initial, syllable initial, medial, and final. Hanna et al. (1966) did not include word initial as a possible position, and instead considered the onset of a word to simply be a syllable initial. However, I chose to separate the first syllable initial from 2nd-or-later initial as it became evident that some mappings simply do not occur at the start of a word, but do occur at the start of later syllables. For example, /p/ → PP never occurs at the beginning of a word in English, but can occur in a later syllable initial (e.g., A-PPEAL). Other systems are even stricter in regard to syllable initials; Chee et al. (2020) considers the onset of each syllable as having its own probability, such that the R's in READ, MIS-READ, and CI-GA-RETTE are all treated differently from one another. However, as there is currently no strong evidence that each syllable’s onset should be treated differently beyond the word initial, I elected to keep the system as simple as possible and did not consider onsets beyond the word initial as having their own specialized probabilities.

As an example of a P-G probability used by this system, we find that the mapping /ɑ/ → O has the probabilities 0.408 if it’s the initial phoneme of a word (e.g., ON), 0.182 if it’s the initial phoneme of a syllable (e.g., CHA-OS), 0.458 if it’s the medial phoneme (e.g., GOT), and 0.731 if it’s the final phoneme (e.g., BLO-SSOM). This indicates that the mapping is highly likely to occur when a syllable ends in /ɑ/, but infrequently when a 2nd or later syllable begins in /ɑ/.

Using the P-G Mapping method, the pseudoword /raɪ nə/ is most probably spelled as RINA, because the maximally probable mappings for each segment are: /r/ → R 92%, /aɪ/ → I 59.52%, /n/ → N 85%, and /ə/ → A 36.85%. In the most simplistic terms, the P-G Mapping account conceptualizes the sublexical route as a probabilistic system that draws from overall knowledge of the English language. Now, I will explore how this output might differ when influenced by the lexical route, via Phonological Neighbors.

II.2 Phonological Neighborhood

Research has indicated that phonological neighbors may have an impact on word recognition and spelling (e.g., Tainturier et al., 2013; Yates and Slattery, 2019). Some researchers have found that phonological neighbors can facilitate spelling tasks (Stemberger, 2004), while others demonstrate evidence that increasing neighborhood sizes can actually hinder spelling speed and accuracy (Karim and Diaz, 2020). Frequency of neighbors has also been found to have an influence, with neighbors that are frequently used in English having more influence over spelling behaviors than those that are infrequent (Oldfield and Wingfield, 1965; Sadat et al., 2014).

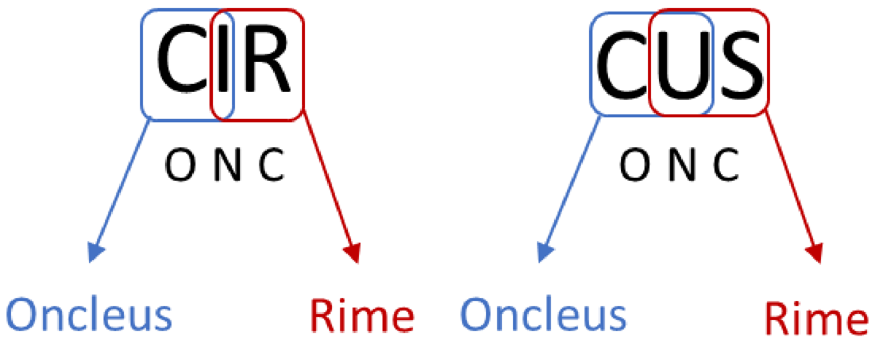
Both Tainturier et al. (2013) and Patterson & Folk (2014) proposed that the lexical and sublexical routes interact via neighborhoods influencing the P-G mapping probabilities. According to this account, the activation of neighbors in the lexical system causes a corresponding dynamic reweighting of P-G mappings based on what is prevalent in that neighborhood. However, it is important to note that previously this reweighting has been demonstrated using priming, in which participants are presented with a word and later provided a phonological neighbor. I aim to demonstrate P-G reweighting without relying on priming methods.

For an example of how the phonological neighborhood account operates, let us examine the pseudoword /raɪ nə/, which has the phonological neighbors RHINO, CHINA, and MYNA/MYNAH. Despite the low probability of the P-G mapping /r/ → RH (0.9%), the mapping will still occur frequently due to spreading activation from the neighbor **<u>RH</u>**INO to the mapping /r/ → RH, raising the probability that mapping will be selected. The phoneme /aɪ/ would be more likely to map onto I (RH**<u>I</u>**NO, CH**<u>I</u>**NA) than it would otherwise be, and the /ə/ would more likely be A (CHIN**<u>A</u>**, MYN**<u>A</u>**). Therefore, the most likely spelling for /raɪ nə/ based on this Phonological Neighborhood account is RHINA.

II.3 Oncleus and Rimes

Before diving into the Oncleus-Coda and Onset-Rime accounts, let us explain the onset, nucleus, and coda. The nucleus is the vowel sound in a given syllable, with the onset being the initial consonant(s) and the coda being the final consonant(s). In any given English syllable, there must always be a nucleus but there is not necessarily a consonant in the onset or coda position. As seen in *Figure 2*, the onset and nucleus combine to create a higher-order representation referred to as the oncleus, while the nucleus and coda combined are referred to as the rime[1].

Figure 2. Oncleus and Rime



Note: A model demonstrating the oncleus and rime for the word "circus".

Burt and Blackwell (2008) proposed the idea that spelling errors occur not on a whole word, morpheme, or phoneme-grapheme basis, but rather at the rime level of the word. According to this theory, the words "sink, mink, [and] link" (pg. 78, Burt and Blackwell, 2008) all share rimes and are orthographically consistent with one another (they are "friends"), making them more likely to be spelled accurately. Meanwhile, the words "bear, care, [and] hair" (pg. 78, Burt and Blackwell, 2008) all share rhymes but are inconsistent (they are "enemies") because they have different rimes, which may contribute to spelling errors among those rimes. One prediction made by this framework is that spellings for pseudowords will adhere to existing rimes in English, and particularly so for consistent rimes. For example, the pseudoword /snoʊb/ should be spelled SNOBE because all rhyming words are consistently spelled with -OBE in

[1] The "rime" refers to the graphemes used to represent the phonological form, i.e., the "rhyme".

English (e.g., GLOBE, ROBE, LOBE), whereas the spelling SNOAB would not be predicted, as no existing words in English have the rime -OAB.

The rime account is arguably the most prominent sub-syllabic mechanism put forth in the literature, but this does not rule out the possibility of other sub-syllabic representations contributing to pseudoword spelling (Chee et al., 2020). The oncleus is of particular interest as it tends to be highly salient in reading tasks, due to the primacy of the beginning of words (Chee et al., 2020; Zeven and Seidenberg, 2005). For an example of how the oncleus may influence spelling, let us take a look at the phoneme /k/. The probability of /k/ → C in the initial position of a word is ≈ 88% (Hanna et al., 1996). However, if the oncleus consists of /ki/, C becomes extremely unlikely because CI and CE almost never occur in the initial position to represent that oncleus. They would instead typically represent the oncleus /sɛ/ or /sɪ/ as in “cell” or “cinch”. KI or KE is instead pushed forward as the most likely mapping, follow by CH + Vowel or QU + Vowel.

The pseudoword /raɪ nə/ has the oncleuses /raɪ/ and /nə/, and the rimes /aɪ/ and /ə/. Based on those representations, the most likely spellings are RINA (oncleus-based) and RYNA (rime-based).

II.4 Summary

The P-G Mapping account predicts that, upon hearing a word, individuals access knowledge derived from all words within their lexicon regarding which graphemes are most likely to correlate to each phoneme (based on where they are within the syllable). The Phonological Neighborhood account predicts that stored knowledge in the lexical system activates those words that are closest to the one you have heard. Activation of these words in the lexical route entails activation of their P-G mappings as well, and this input in turn skews the P-

G probability distribution away from that which matches the entire lexicon towards one that is specific to the activated neighborhood. The Oncleus and Rimes account operates on the basis of units larger than individual phonemes but smaller than entire words. These can generate different spellings, as probabilities are shifted based on the additional constraints imposed by these larger units.

Each of these accounts provide specific predictions about spelling behavior. In the following sections, I outline how I intend to operationalize those predictions in order to test the relative strength of evidence for each account.

III. Generating the Hypotheses

Above, I described four possible accounts of sublexical spelling. The following sections elaborate on testable predictions that are generated by these accounts. Specifically, it is known there is variability in pseudoword spelling across spellers, and so I will examine how well the mechanisms account for that variability. It is also known that real word spelling varies across spellers, and I wish to see if it is possible to systematically relate this to pseudoword behavior. Finally, it is known that people vary in their ability to learn new words, which I again seek to relate to pseudoword spelling behaviors. The sections below will further elaborate on individual differences in pseudoword spelling.

III.1 Individual Differences

How do we know that there are individual differences in how people spell? Wouldn't it make more sense if researchers and teachers simply assumed everyone approached written language the same way? Coltheart and Ulicheva (2018) analyzed data collected by Pritchard et al. (2012) about pronunciation differences of read pseudowords in adult skilled readers, indicating that there is a great deal of variability in phonetic assignment even among experienced populations. Pseudoword studies such as this are often conducted with young children who are learning written language for the first time. However, even in adulthood individuals are likely to encounter words they have never heard or seen before. This is especially true in the case of proper names, which can vary widely from person to person. I want to understand how such lifelong learning is achieved, hence the need for pseudoword studies in adults as well as children.

Figure 3. ***Model from Coltheart and Ulicheva (2018).***

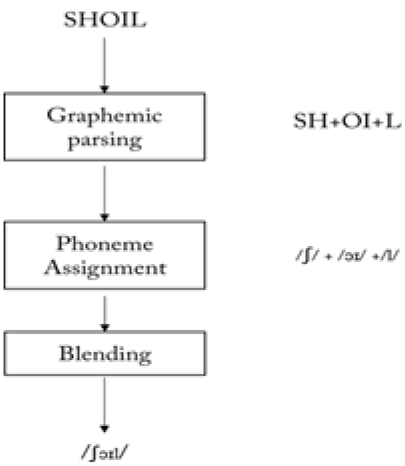
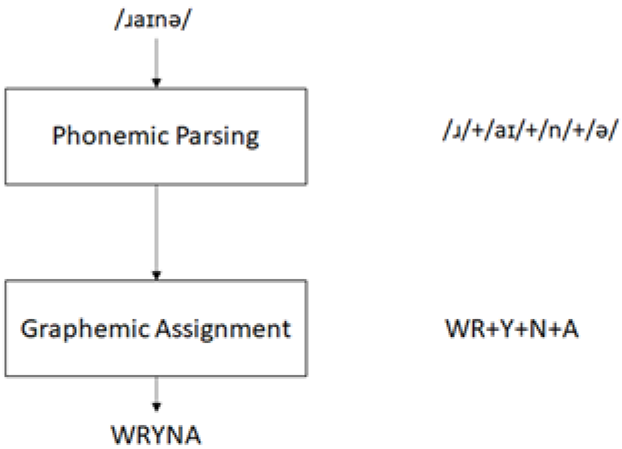


Figure 3 demonstrates the three major components hypothesized to be used by the sublexical route in reading-to-pronunciation tasks. Graphemic parsing is the first step of reading a newly presented pseudoword. It entails breaking the entire word, SHOIL, apart into its respective graphemes, SH + OI + L. The next step is phonemic assignment, in which a phoneme is assigned to each grapheme. The final step is blending, in which all phonemes are combined into a pronounceable word.

Coltheart and Ulicheva's research revealed a wide degree of across-item variation in graphemic parsing, with a single word receiving as many as 24 unique pronunciations. Entropy (*H*), or within-participant variability in graphemic parsing and phoneme assignment, was also calculated. Higher entropy indicates that a single grapheme is being assigned multiple phonemes across pronunciations. Each participant read at least one grapheme with an entropy of 0 (i.e., they consistently read the grapheme the same way), with averages ranging from 0.25 (relatively low) to 0.71 (relatively high). Interestingly, some participants tended to pick one phoneme for one grapheme and stick with it (e.g., always read IE as in "pie"), while others tended to vary widely (read IE as in "pie" *or* "chief" *or* "friend" *or* "mischief", etc.). This variation harkens back to the aforementioned lexical quality hypothesis by Perfetti and Hart (2002), in that participants have differing qualities of lexical information and therefore may be relying more or less on their sublexical system. Coltheart and Ulicheva's findings raise the question: can the same variation be found in spelling tasks? Instead of phonemic assignment, I seek to focus on graphemic assignment (see *Figure 4*) to see if I also obtain high rates of variability with a pseudoword spelling task as Coltheart and Ulicheva demonstrated with a pseudoword reading task. The four proposed sublexical mechanisms allow me to examine whether I can account for which spellings are most prevalent. In other words, can I explain why certain spellings are popular using either P-G Mappings, Phonological Neighbors, Oncleus-Coda or Onset-Rime measures? Furthermore, I investigate whether individual differences in pseudoword spelling behaviors are significantly associated with individual differences in both real word spelling behaviors and the ability to learn new words. Do "good spellers" tend to adhere to certain probabilities, such as those provided by the Phonological Neighborhood or Oncleus-Coda, etc.?

Figure 4. Phoneme-to-Grapheme Route



III.2 Learning

Lexical and sublexical interactions, while important for spelling, may also impact participants' behaviors when learning new words. Share's self-teaching hypothesis asserts that the sublexical route is responsible for phonological recoding, which allows for the acquisition of new items in the orthographic lexicon (Share, 1994). In essence, once a reader has developed an understanding of a sufficient number of phoneme-to-grapheme correspondences, they are able to parse newly encountered words and use their knowledge of other correspondences to learn new ones. Increasing orthographic knowledge eventually alerts the student to constraints within the language, such as context-dependent P-Gs and letter position within the word. This process of "lexicalization" (i.e., consolidating word knowledge into O-LTM) eventually leads to learners depending increasingly upon their lexical knowledge when reading and spelling, so much so that their sublexical system may become obsolete. Of course, there are others who argue that language learning is not a dual-route system at all, and depends wholly on lexical knowledge. Many proponents of the rime/oncleus mechanism, such as Burt and Blackwell (2008), question the extent to which sublexical processes are at all active in literate adults. Their argument is that spelling is rime-based, and therefore the spelling of new words operates based on knowledge of other words' rimes—that is, pseudoword spelling is achieved by an analogical process that creates new spellings that adhere to the English lexicon. This account, however, is challenged to explain neologisms or other novel spellings that do not adhere to existing rimes but nonetheless occur with some regularity. For example, the pseudoword /snoʊb/ should be spelled only with -OBE, along the lines of ROBE and GLOBE; however, other spellings are in fact attested, such as SNOAB.

No matter the case, it is clear that these possible sublexical mechanisms leave open the possibility for individual differences in lexical knowledge, aligning with Perfetti and Hart's Lexical Quality Hypothesis (2002). Of course, there is no guarantee that a neurotypical individual's sublexical system is particularly strong either; perhaps some people fail to accumulate a good probabilistic representation of the English lexicon, or have difficulty accessing it. By extension, both lexical and sublexical processing are extremely important for both for diagnosis and treatment of language deficits in neurodivergent patients. It is common practice to use pseudowords as a diagnostic tool for possible cases of dyslexia/dysgraphia (Tilanus et al., 2013), with the goal of exploring sublexical processing specifically. Irregular real words are frequently used to examine lexical processing, as these words contain unusual spellings/pronunciations that require specific knowledge. Numerous case studies (e.g., Rapp, 2015) have classified individuals based on their accuracy with different types of real words (regular vs irregular, high vs low frequency), as well as what sort of errors they make (e.g., for MARTYR, a substitution error would be MART**I**R, while a transposition error would be MAR**Y**T_R). Using these pseudo and irregular real words allows researchers to more precisely detect where deficits may lie in a patient's lexical or sublexical route, which can lead to targeted treatment interventions and outcomes. Evidence of lexical and sublexical interaction could prove to be extremely beneficial for developing future diagnostics and interventions.

If sublexical processing is an integral part of learning new words, then I expect that some degree of participants' ability to learn faster and/or retain new words longer than others can be attributed to their sublexical route. Therefore, learning behaviors should provide further evidence for the four proposed routes of sublexical processing. I anticipate that strong sublexical spellers will be more likely to make correct "first guesses" during a new word learning task, which should contribute to higher rates of successful recall. Additionally, a strong lexical knowledge base should contribute to successful new word learning given that it indicates past success with integrating new words into O-LTM.

IV. Summary

The lexical and sublexical systems are both integral parts of our ability to pronounce and spell words, with the sublexical in particular being key to processing and decoding newly encountered words. Yet there is little agreement as to what mechanisms or representations are at work within this system. Evidence from Coltheart and Ulicheva (2018) indicates that whatever mechanism(s) is/are at work may vary across participants, as a high degree of pronunciation variability was present both across and within participants. This variation is likely due to differing qualities in lexical knowledge, and therefore more or less dependence upon the sublexical system (Perfetti and Hart, 2002).

Multiple mechanisms have been proposed regarding how people spell newly encountered words, ranging from relatively simple probability-based P-G mappings, to middle order units of oncleus and rimes, to higher order phonological neighborhood influences. There is significant debate over whether one or all of these accurately captures the mechanism used for spelling new words.

How learning is achieved through the sublexical system is also debated. Share (1994) asserts that the sublexical system is part of a self-teaching mechanism that comes online to enhance lexical knowledge after a minimum amount of P-G mapping knowledge is acquired. Others, like Burt and Blackwell (2008), question the extent to which the sublexical system is active in literate adults, or whether all learning is solely based on analogizing to lexical representations. No matter the case, though, there appears to be a degree of variability across individuals due to what processes and representations are operating within the sublexical route. I am particularly interested in examining the effects of these individual differences in regard to how people complete novel spelling tasks and assimilate new word knowledge. In the next section, I will detail how I plan to conduct a series of experiments to obtain evidence relevant to the four hypothesized sublexical mechanisms.

V. Goals and Hypotheses

Our goals are to explore how well each of the four proposed mechanisms, P-G Mapping, Phonological Neighborhood, Oncleus-Coda, and Onset-Rime, 1) explain variability in pseudoword spelling across people, 2) relate to real word spelling behaviors, and 3) account for individual differences in new word learning. I believe that different people may use different mechanisms depending on their relative lexical and sublexical strengths. I expect that those with relatively weak lexical systems will tend towards using simple probabilistic P-G Mappings, while those with strong lexical systems will show influence from Phonological Neighborhoods and Oncleuses/Rimes.

The remainder of this paper will describe in detail two experiments conducted to explore the aforementioned hypotheses. Experiment 1 predominantly demonstrated that it is possible to detect variability in pseudoword spelling across participants and examined how well each of the four proposed mechanisms accounted for this variability. It also examined the relationship between these four mechanisms and real word spelling behavior. Experiment 2 examined the influence of the four mechanisms on new word learning.

CHAPTER II: EXPERIMENT 1

The goal was to find evidence for the four proposed sublexical mechanisms. I did so in two ways within Experiment 1. First, I investigated variability in adult skilled spellers similar to the variation Coltheart et al. (2018) found in adult skilled readers. I specifically examined this variability in terms of four proposals: P-G Mappings, Phonological Neighborhoods, Oncleus-Coda, and Onset-Rime. Second, I investigated whether there is a systematic relationship between real word spelling and pseudoword spelling behaviors. Specifically, I assessed whether more accurate real word spellers err on the side of a particular mechanism(s), while less accurate spellers tend towards the probabilities given by another mechanism(s). For example, are those who tend to use high probability mappings according to phonological neighbors better real word spellers than those who adhere to rime probabilities are worse real word spellers?

Coltheart and Ulicheva's (2018) research exclusively used monosyllabic pseudowords, but I aim to expand their findings by using a mix of both pseudowords and low frequency, irregular real words presented in a mono- or disyllabic format. This specific type of real words was chosen because they are unlikely to be consistently known across all participants due to low frequency, guaranteeing some amount of spelling errors and thereby differentiating participants in terms of their lexical knowledge (i.e., O-LTM). In this way, I examined if there is any correlation between mechanisms being used for pseudoword spelling and better real word spelling.

VI. Methods

VI.1 Materials

Monosyllabic pseudowords were selected from the ARC Nonword Database (2002) from Macquarie University. There were no restrictions, such as word length, placed on these words.

As this database did not have multi-syllabic words, two-syllable pseudowords were taken from a paper by Mousikou et al. (2017) involving disyllabic non-word reading. After selection of pseudowords, low frequency irregular words were found using Northwestern University's CLEARPOND database (Marian et al., 2019), with each real word being matched through CLEARPOND by number of syllables and phonetic neighbors to a respective nonword. A total of thirty pseudowords and thirty low-frequency irregular real words were selected, ranging from words with minimal variability in spelling (e.g., two spellings across all participants) to words with high variability in spelling (e.g., eight spellings across all participants). Pseudowords were selected based both on response diversity given and the likelihood of auditory errors (e.g., hearing "sprumps" as "sprumth"). Response diversity means that if every single participant spelled a given word the exact same way, then it was too regular and would not provide valuable entropy information. These stimuli are presented in Appendices A (real words) and B (pseudowords).

The cut-off for low frequency was 10 appearances per million words. The cut-off for irregular was at least one PG mapping < 16% probability (based on Hanna et al., 1966). Phonological and orthographic measures of the stimuli are summarized in *Table 1*, and the stimuli are presented in Appendix A, with the most common pseudoword spellings in Appendix B. All 60 words were prerecorded to mitigate any variability in pronunciation that may influence subsequent spelling. The survey was created in Qualtrics and distributed through Amazon's Mechanical Turk. Javascript was embedded into the Qualtrics file to disable spellcheck.

Table 1. Phonological and Orthographic Properties of the Stimuli

	Length (letters)		Length (phonemes)		Syllables		Frequency		PG Probability		PTAN	
	M	*s*	M	*s*	M	*s*	M	*s*	M	*s*	M	*s*
Irregular words	6.6	*1.35*	4.57	*1.07*	1.6	*0.5*	0.56	*0.42*	53%	*17%*	*4.95*	*3.81*
Pseudowords	--	--	4.6	*0.89*	1.43	*0.5*	0	*0*	--	--	--	--

Note: Phonological and orthographic properties of the stimuli. For the irregular words, the most common spellings CZAR and SCHLEP are used here (although alternate spellings, e.g., TSAR, were also accepted as correct). Frequency refers to appearances per million obtained from CLEARPOND (Marian et al., 2019). PTAN = number of phonological neighbors.

VI.2 Participants

A sample of 100 participants initially filled out a series of basic demographic questions (*Table 2*): sex, age, years of education, native language(s), second language(s), years of learning second language(s), and accent in English. There were also two yes-or-no questions, which asked if the participant had any reading, writing, or learning disabilities or if they had a neurological condition. The only screening criteria for this experiment were that all participants are above the age of 18 and were proficient in English. There was also an auditory test before the actual experiment began, which asked the participants “What is the last letter in the English alphabet?” If participants responded with any answer other than Z, they were automatically closed out of the survey and thanked for their participation. This was intended to make sure 1) the participants could hear the audio files inserted into the survey, and 2) to disable response-bots from being able to take the survey and collect payment. I designed this particular stopping measure because bots can do basic transcription to auto-fill blank spaces meant for typing a response, but are generally unable to interpret the auditory questions.

Participants were paid $7.50 upon completion of their surveys for their participation in the research.

Table 2. Experiment 1 Demographics

	Education		Age	
	M	*s*	*M*	*s*
Female (N = 48)	15.4	*2.1*	44.2	*12.7*
Male (N = 48)	15.4	*1.7*	34.9	*9.6*
	M	*s*		
Second Language	16.33%	-		
Language Years	8.65	*10.16*		

Note: Age and Education in years (high school = 13). All participants responded that English was their first language or was one of their first languages but was the primary language in school. Second language refers to the percentage of participants who spoke a second language. Language Years indicates the average number of years participants who spoke a second language had known the language. No participants responded that they had a reading, writing, or learning disability or some kind of neurological deficit.

VI.3 Procedures

Participants read and signed a consent form before proceeding with the experiment. After this, they were requested to disable any spellcheck applications they were running, as these had the potential to bypass the anti-spellcheck script. Participants then filled out the demographic questions and answered the audio test. If participants completed this, they were then given written instructions to go through the word list at their own pace, listen to each recording as many times as necessary, and attempt to spell the heard words as accurately as possible. After this, participants were met with a page informing them that they were going to attempt to spell either real or pseudowords words. The experiment was block designed so that 50 participants encountered pseudowords first, and 50 encountered real words first. This counterbalancing was done because real words may bias participants toward generating certain pseudoword spellings. For pseudowords, participants were asked to provide their best guess as to how the word was spelled, and for real words they were asked to attempt to spell it accurately (e.g., Folk, 2014). Participants were also informed that they could complete the list at their own pace and listen to each recording as many times as necessary.

Participants listened to an audio file in which either the real word or pseudoword was pronounced twice, and then typed their spelling attempt before proceeding on to the next word. After working through the list of 30 real/pseudowords, participants were then informed they were switching over to the word type they had not spelled for yet. After completing the second block, a final page informed them that the survey was complete and thanked them for their participation.

VI.4 Data Processing

Typing and auditory errors must be ruled out as they cause unwanted variance in responses due to being nonrepresentative of a participant's best attempt at spelling a heard word. I constructed two criteria to exclude likely typographical and auditory errors:

1) The spelling includes a trigram (sequence of three letters) that does not exist in English, verified with MCWord database (Medler & Binder, 2005).

2) The spelling includes extraneous graphemes (example: /eɪft/ spelled AIFTER) or has fewer graphemes than phonemes (example: /eɪft/ spelled AIF).

In specific cases, some substitutes were accepted as highly probable "mishearings" which can be scored as alternative targets, in particular /f/ versus /θ/, as in GWARFS/GWARTHS and FROAGS/THROAGS (see *Appendix B*).

VI.5 Real Word Measures

Irregular and Regular Segment Scores: Each segment of each real word was first categorized as either regular or irregular based on Hanna et al. (1966). The threshold was set at 16% probability. Each participant's mean accuracy was computed separately for the 90 regular and 51 irregular segments across the 30 words. These values were used as the dependent variable assessing the ability to predict individual differences in spelling accuracy.

Spelling Errors: Incorrect spelling were categorized as:

1) **Phonologically plausible errors (PPE):** misspellings that could still be pronounced as the target word (e.g., CAYENNE as KAYEN) or correct spelling of incorrect target (e.g., CAYENNE as CANON).
2) **Phonologically nonplausible errors (NPE):** possible typographical or auditory errors (e.g., CAYENNE as CANANYE OR CAYN).

Real word spelling accuracy was calculated based only on PPE spellings. NPEs, which consisted of ~7% of all trials, were excluded from any further analysis.

Additional measures:

1) **Regularity Effect:** this is measured by the difference in accuracy on Regular vs. Irregular word segments (e.g., CZ in CZAR is irregular, AR is regular).

2) **Regularization:** the extent to which a PPE is made more or less "normal," based on PG probability (e.g., ARRAIGNED misspelled ARRAIND = 51.73% probability vs. ARREIGNED = 16.2 % probability).

VI.6 Pseudoword Measures

Pseudoword Scores: Each pseudoword spelling was scored on several dimensions:

1) **P-G Mappings:** position-specific PG probabilities. Each target phoneme was mapped onto a grapheme produced by the participant, and the probability of each mapping within a spelling response was used to generate an average probability for the entire word. In cases of ambiguity, where more than one graphemic parsing was possible, the one resulting in the higher score was chosen.
2) **Neighbor Mappings:** based on phonological neighbors; for example, /raɪ nə/ spelled RHINA is consistent with the phonological neighbor /raɪ noʊ/, RHINO. The previously aforementioned phoneme-graphing mappings were used, but with mapping probabilities provided only by phonological neighbors.
3) **Oncleus Mappings:** for example, /bruk/ has the oncleus /bru/, most often spelled BROO-. The linguistic database CLEARPOND (Marian et al., 2019) was used to find all real words that shared an oncleus with the given pseudoword. A probability was then assigned to each oncleus based on how frequently they appeared among this list of words, which was then used to score the probability of participant spelling responses.
4) **Rime Mappings:** for example, /bruk/ has the rime /uk/, most often spelled -UKE. The same method as Oncleus Mappings was used, except CLEARPOND (Marian et al., 2019) was used to find all real words that shared a rime with the given pseudoword.

Segments that do not occur in English were assigned a score of 0 (e.g., /bruk/ spelled BROOCK receives a score of 0 for the rime, as -OOCK does not occur in English). Each score was then normalized within items to account for differences in which items participants responded to and the fact that phonemes, oncleuses, and rimes vary in their maximum possible probability (e.g., a word-initial /b/ is spelled B 100% of the time, but the single most likely spelling of word-initial /ɑ/, O, still only occurs 41% of the time). Specifically, each individual response was compared to the maximum value observed across participants for that item. For example, the highest PG probability observed for /eɪft/ was AFT 72.76%, so the spelling AIFTH 28.56% received a score of 28.56%/72.76% = 0.39. These measures were tested as potential additional predictors of Irregular/Regular segment accuracy on real words.

Five pseudowords were dropped from further analysis for either lacking a phonological neighbor or extant oncleus/rime. These words were /gwɑrfs/ (and its accepted alternate /gwɑrθs/), /hɚkʃ/, /haʊks/, /ʃraʊnθ/, and /sprʌmf/.

VII. Analyses and Results

VII.1 Analysis 1

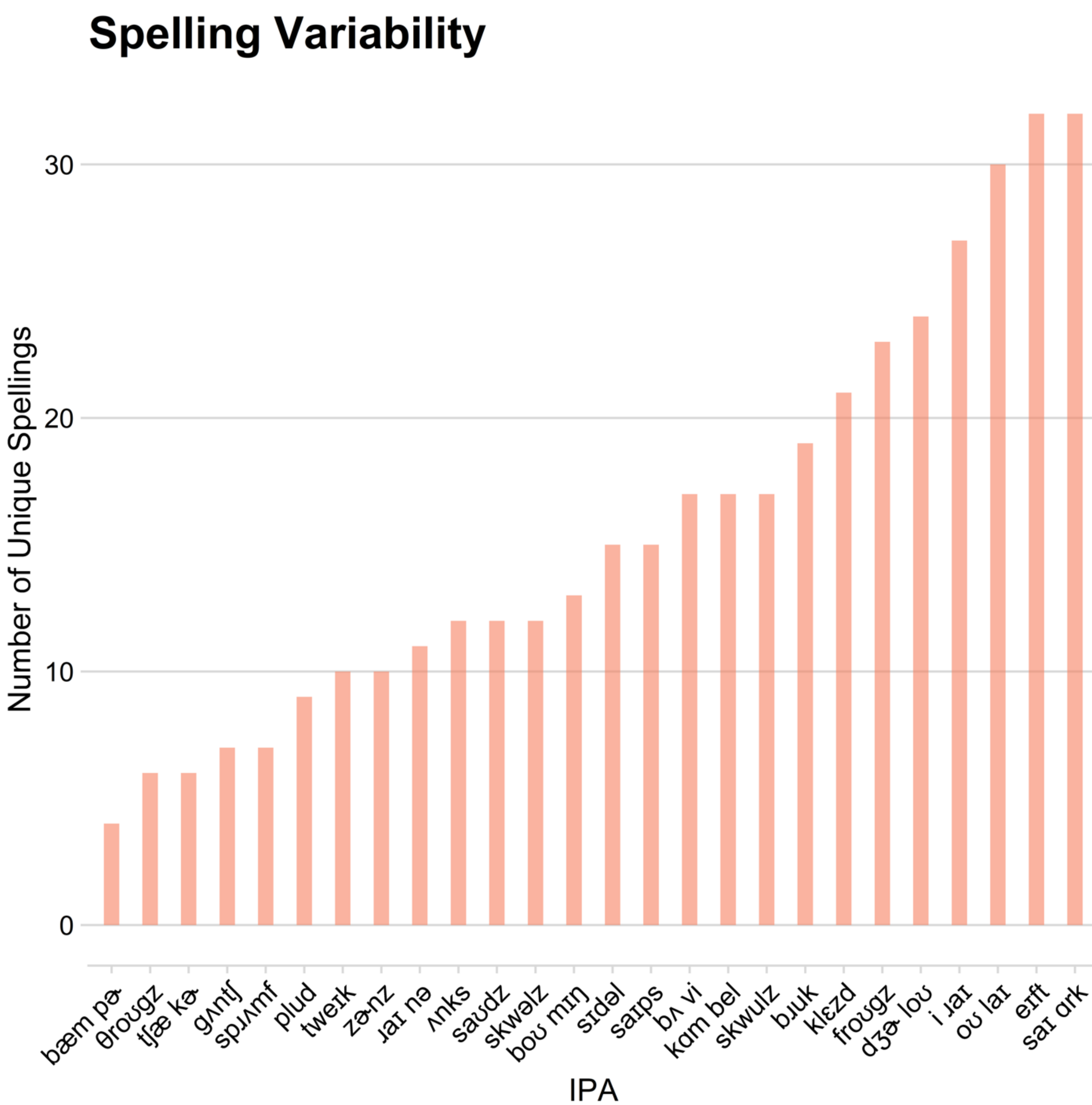
Poisson Regression: Poisson regression is a type of generalized linear model developed to model counts or frequencies using a Poisson distribution, which is a discrete distribution that is defined for positive integers (Coxe et al., 2009). Poisson regression was used to model the outcome of the pseudoword spelling task, i.e., the counts of each observed spelling were regressed on the four pseudoword probability measures (see section *VI.6 Pseudoword Measures*). First, I constructed generalized linear mixed models (GLMM) using the *glmmTMB* package in R to determine if the mechanisms accounted for a significant amount of unique variance (Brooks et al., 2017). Specifically, the model formula was count ~ PG + Neighbor +

Oncleus_Coda + Onset_Rime + offset + (1 | target). Here, offset indicates the number of observations kept after removing NPEs.

We used the *shiny* app *RcountD* (Coxe, n.d.) to interpret the effect sizes for the predictors in the Poisson regression. An effect size of 1 indicates no effect, while values greater than 1 indicate that an increase in the predictor is associated with an *increase* in the outcome, and values less than 1 indicate that an increase in the predictor is associated with a *decrease* in the outcome.

VII.2 Results

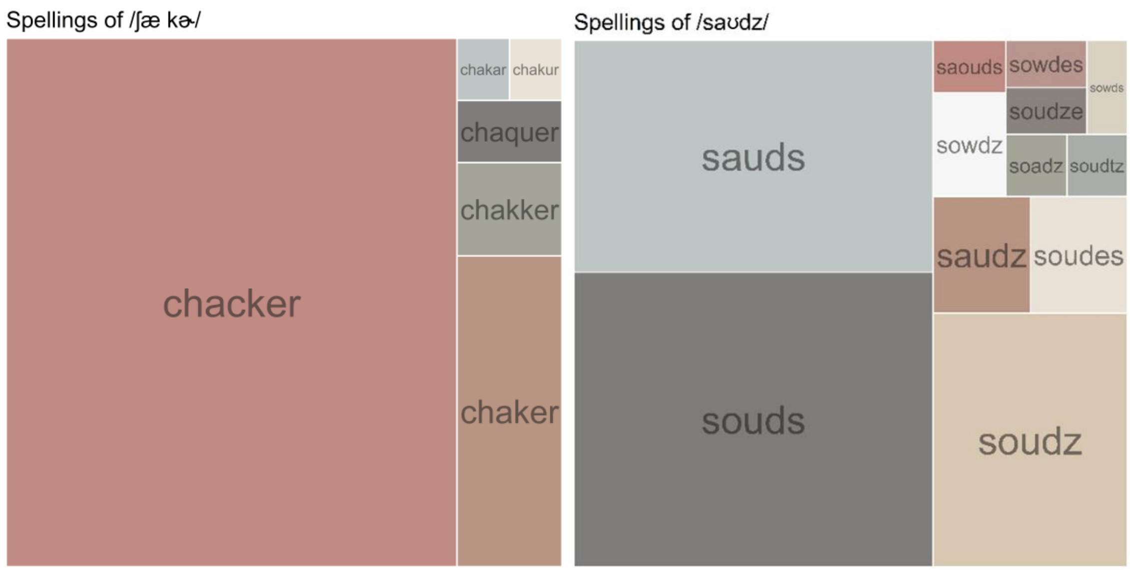
Figure 5. Number of Unique Spellings per Pseudoword Stimulus



Note: X-axis presents target items in IPA.

The number of unique spelling responses ranged widely across pseudowords, with the lowest being 4 and the highest being 32 (see *Figure 5*). This indicates that, like Coltheart and Ulicheva's (2018) findings regarding pseudoword reading, pseudoword spellings present with a wide range of variation. Additionally, there was a noticeable degree of within-word spelling variation, with some words having an even distribution of spelling and others showing that certain spellings of a pseudoword were indeed more popular than others *(Figure 6*).

Figure 6. Spelling Distribution Within Two Pseudowords



Note: Note that responses for /tʃæ kɚ/ converge on a single spelling, while /saʊdz/ demonstrates participants preferred to use one of three spellings.

Table 3. Model Summary of GLMM

	Estimate	Std. Error	*Z*-value	*Pr* (>|z|)
Intercept	-3.16	.12	-26.34	< 2e-16 ***
P-G	.01	.09	.11	.91
Neighbor	.56	.09	6.53	< .001 ***
Oncleus-Coda	.30	.09	3.52	< .001 ***
Onset-Rime	.42	.08	5.33	< .001 ***
Signif. Codes:	< .001'***'	< .01'**'	< .05'*'	

Note: Summary of GLMM to determine whether each of the four sublexical mechanisms significantly account for unique variance. Marginal R2 = .25

Table 4. Results of Poisson Regression

Predictor	Effect Size	CI (95%)
P-G	1.01	[.84, 1.22]
Neighbor	1.76	[1.48, 2.07]
Oncleus-Coda	1.35	[1.14, 1.6]
Onset-Rime	1.31	[1.31, 1.76]

Note: Results of Poisson regression on the four sublexical mechanisms.

As seen in *Table 3*, Neighbor, Oncleus-Coda, and Onset-Rime mechanisms were all able to account for a unique amount of pseudoword spelling variation ($p < .01$). Additionally, the exponential effects of these three mechanisms were positive (*Table 4*). The Poisson regression in *Table 4* indicates that probabilities generated by the Neighbor mechanism best account for the previously described variation in pseudoword spelling ($p < .01$). To a lesser extent, probabilities from both Oncleus-Coda and Onset-Rime account for this spelling variation as well ($p < .01$). P-G’s exponential effect is non-significant, meaning it does not account for any unique variance ($p \approx .91$).

VII.3 Analysis 2

Generalized Linear Mixed Models: This analysis sought to examine if better lexical knowledge is associated with better sublexical knowledge. A GLMM was conducted using the R package *lme4* (Bates et al., 2015). Binomial family was used as the distribution type. First, rate of NPEs and correct spellings of non-target words were removed from analysis. Then, a model was conducted with word accuracy (correct/incorrect) as the dependent variable (DV) and two kinds of independent variables (IV). The first type of IVs was based on stimulus properties: word length (number of letters), frequency, and P-G probability of the target real word. The second type of IVs was participant properties: age, education, average P-G, Neighbor, Oncleus-Coda, and Onset-Rime scores on pseudowords, and error rates. Error rates were included to account for potential bias from excluding trials that contained NPEs and correct non-target errors. A second GLMM was conducted that repeated the procedure but took as the DV accuracy only on *irregular* word segments, with accuracy on regular segments included as a covariate. Note that the DV in this case was weighted to indicate how many irregular segments each word had.

VII.4 Results

Table 5. Results of Real Word GLME

	Estimate	Std. Error	*Z*-value	*Pr*(>|z|)	
(Intercept)	1.90	.30	6.40	.001	***
Length	-.13	.29	-.46	.64	
Frequency	-.14	.29	-.46	.64	
Target P-G	1.52	.25	6.01	.001	***
Age	.13	.11	1.16	.25	
Education	.32	.11	2.91	.001	**
Error rate: NPE	-.39	.11	-3.62	.001	***
Error rate: Correct Non-Target	-.49	.12	-4.00	< .001	***
P-G	-.25	.19	-1.34	.18	
Neighbor	.50	.15	3.28	.001	**
Oncleus-Coda	-.12	.15	-.83	.41	
Onset-Rime	.01	.17	.08	.94	
Signif. Codes:	< .001'***'	< .01'**'	< .05'*'		

Note: Results of GLME with word accuracy as the DV. Marginal $R^2 = .36$.

Table 5 indicates that real word accuracy was significantly correlated with multiple real word predictors, including target P-G and both error rates. Additionally, accuracy was correlated with years of education. Interestingly, of the sublexical mechanisms, Neighbors are the only ones correlated with better real word spelling.

Table 6. Results of Irregular Segment GLME

	Estimate	Std. Error	*Z*-value	*Pr*(>|z|)	
(Intercept)	2.40	.26	9.36	< 2e-16	***
Regular Acc. Percentage	.61	.06	10.26	< 2e-16	***
Length	.12	.25	.47	.64	
Frequency	-.13	.25	-.50	.62	
Target P-G	.88	.21	4.11	< .001	***
Age	.20	.09	2.26	.02	*
Education	.27	.09	3.11	.001	**
Error rate: NPE	-.22	.09	-2.57	.01	*
Error rate: Correct Non-Target	-.39	.09	-4.18	< .001	***
P-G	-.38	.15	-2.57	.01	*
Neighbor	.29	.12	2.43	.02	*
Oncleus-Coda	-.03	.12	-.29	.77	
Onset-Rime	.15	.14	1.07	.29	
Signif. Codes:	< .001 ***	<.01 **	<.05 *		

n = 98

Note: Results of GLME with irregular segment accuracy as the DV. Marginal R^2 = .35.

Table 6 demonstrates that accuracy on irregular segments was significantly correlated with the same real word predictors from the previous analysis, along with age and education. Again, Neighbors are significantly correlated to irregular segment accuracy ($p \approx .015$). Additionally, the P-G pseudoword measure is also significantly correlated to irregular segment accuracy ($p \approx .01$). However, whereas Neighbors demonstrate a positive correlation with accuracy, P-G demonstrates a negative one. This indicates that participants who adhere to Neighbor probabilities when spelling pseudowords also tend to be better at spelling real words, while participants who use P-G probabilities to spell pseudowords tend to be worse at spelling irregular real words.

VIII. Summary

Experiment 1 examined whether it is possible to a) use the four proposed sublexical mechanisms to account for variance in pseudowords spelling and b) detect any relationships between real and pseudoword measures. Much like Coltheart and Ulicheva (2018) detected a large degree of variability in pseudoword pronunciation, a high degree of variability both within and across pseudoword spellings was found. The Neighbor, Oncleus-Coda, and Onset-Rime mechanisms all contributed to accounting for the detected variation in pseudoword spelling, with Neighbor probabilities being the best at accounting for a unique amount of variance. The P-G mechanism, however, did not account for a significant level of unique variance.

In addition, I discovered that there is indeed some relationship between real and pseudoword spelling. Specifically, individuals whose pseudoword spellings most reflected Neighbor probabilities tended to be better real word spellers. However, those who used spellings most aligned with P-G probabilities tended to be worse at real word spelling.

In Experiment 2, I continue to explore the relationship between pseudoword and real word spelling by categorizing participants based on their relative lexical and sublexical strength. I then make predictions regarding how each cluster will perform during a new-word learning task in terms of both first-guess and final recall accuracy.

CHAPTER III: EXPERIMENT 2

With Experiment 2, I wanted to further investigate whether the four mechanisms can help explain differences in learning (across both items and participants). This was achieved by using a learning paradigm, in which participants are able to give their “first guess” at spelling a pseudoword and are then exposed to training trials to teach them a target spelling. I examined whether the four mechanisms can explain who learns new pseudoword spellings better, and which target spellings were better learned. This was done by having participants complete a final recall of the trained pseudoword spellings, followed by a real word spelling task similar to the one used in Experiment 1. This allowed for further investigation regarding the relationship between real and pseudoword spelling behaviors, and how they influence new word learning.

For Experiment 2, I elected to only utilize monosyllabic pseudowords to simplify the pseudoword measures. While the limitations of doing so have often been noted (e.g., Mousikou et al., 2017; Lee, 2001), the general framework for multisyllabic words is less well-developed and raises additional issues that are outside the scope of this work.

One major sticking point I discovered in Experiment 1 was the great deal of auditory errors participants tended to make, especially on pseudowords. Obvious auditory errors interfere with the data and must be removed, as they do not provide any telling information outside of the fact that errors generated can be chalked up to a mishearing. Due to this issue, for Experiment 2 I decided to put the real and pseudowords into sentences. Words are typically heard within sentences, and the context of surrounding words is known to aid in their ability to understand what is being said (Roverud et al., 2020; Nittrouer S. & Boothroyd A., 1990). Additionally, sentences give an indication as to what part-of-speech the word is and, in the case of real words, its inflection. The real word “fuels” may be perceived as “fuel,” but within the context of the sentence “He fuels his car” it becomes obvious that an *s* should be present. Therefore, I elected to generate sentences for all to-be-learned words in the hopes of mitigating these types of auditory errors. In Experiment 1, 91 unique real word spellings were identified as likely typographical errors and removed; this number was reduced to 53 unique real word spellings removed in Experiment 2, which I attribute to the benefit of embedding the target real words into sentences.

Another limitation of Experiment 1 is that not every pseudoword had both a neighbor and an oncleus/rime (see section *VI.6 Pseudoword Measures*). In generating this round of pseudowords, every pseudoword had to have at least one neighbor, oncleus, and rime to ensure that no pseudowords would need to be dropped from analysis.

I hypothesize that 1) I will continue to find evidence for the four proposed mechanisms based on participants spelling responses and 2) explain who learned new words the best using both real word and pseudoword spellings.

IX. Methods

IX.1 Materials

A list of 101 orthotactically legal, monosyllabic pseudowords and sentences were generated. This was done taking each potential pseudoword and using CLEARPOND (Marian et al., 2019) to ensure that each word had at least one neighbor and oncleus/rime. From this list I selected 46 pseudowords to use in the survey. For each of the 46 target phonological word forms, I selected two spellings such that, based on previous experience with pseudoword spelling tasks, I could expect at least 15% of participants would generate each one as their first guess. I also refrained from including phonological words forms that I expected would lead to nearly all participants generating the same spelling (e.g., as in Experiment 1 with /gʌntʃ/ spelled GUNCH by ~90% of participants). This was to ensure that at least some, but not all, participants would guess the correct spelling the first time, and that the chosen spoken words didn't generate extreme amounts of response variability. Each of the two possible pseudoword spellings obtained from the pilot were counterbalanced, with each spelling appearing on two of the four lists.

I chose two target spellings per pseudoword to avoid confounding the benefit of a correct first guess with inherently "better" spellings. For the pseudoword /braɪ/, people guessing BRY might be more accurate because it is easier to memorize the spelling BRY than those guessing BRIGH. I can demonstrate that this is not the case by showing that guessing BRY only helps if it actually matches the target, whereas guessing BRIGH does increase accuracy if it is the target. This further enables analyses to detect *item differences*, i.e., which spellings are easier to learn. For example, the spelling BRY is more probable than the spelling BRIGH according to PG measures (77% versus 67%), but vice versa according to Oncleus measures (0% versus 32%). Contextual sentences were also generated for the real words used in Experiment 1 (Appendix C).

IX.2 Participants

One hundred and one participants (*Table 7*) were recruited through Mturk to complete one of four surveys on Qualtrics. Requirements for participation and payment remained the same as Experiment 1. Two participants' data were removed due to failure to follow instructions. All participants responded that English was their first language or was one of their first languages but was the primary language in school.

Table 7. Experiment 2 Demographics

	Education		Age	
	M	*s*	*M*	*s*
Female (N = 49)	15.2	*2.3*	45.7	*12.6*
Male (N = 52)	14.7	*2.1*	48.3	*10.1*

	M	*s*
Second Language	17.65%	-
Language Years	6.58	*6.91*
Disability	0.98%	-
Neurological	1.96%	-

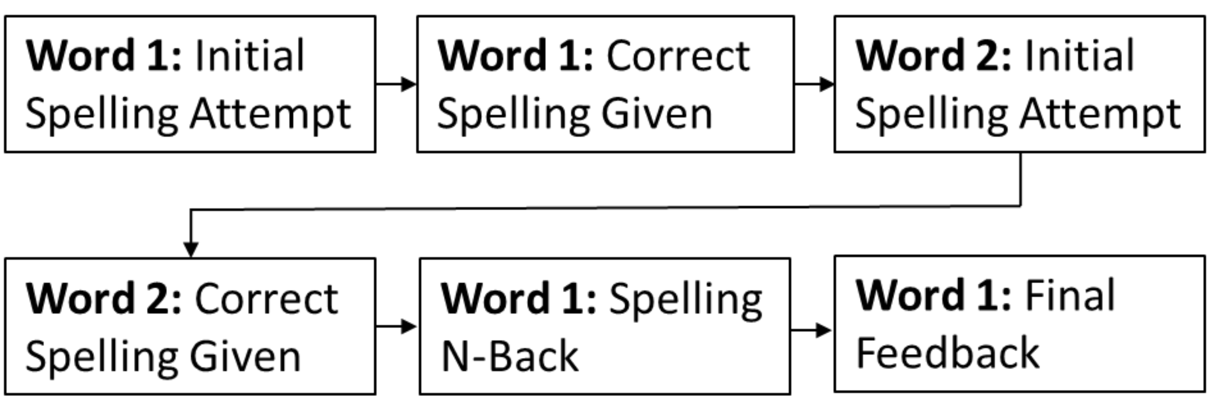
Note: Age and Education in years (high school = 13). Second language refers to the percent of participants who spoke a second language. Language Years indicates the average number of years participants who spoke a second language had known the language. Disability and Neurological indicate the percentage of participants who responded that they had a reading, writing, or learning disability, or some kind of neurological deficit.

IX.3 Procedure

Demographic questions were the same as Experiment 1. A question about what device was being used to complete the survey was also added, as the anti-spellcheck measures I used in Qualtrics could theoretically be bypassed on iPads. No participant responded that they utilized an iPad while completing the survey.

After completion of demographic questions, participants were given an example of the training trials to complete before moving on to the pseudoword training section of the survey. First, participants listened to an audio recording which stated the pseudoword, followed by the sentence containing the pseudoword, and finally the pseudoword again one final time. Participants could play this audio as many times as they wished. Each segment of the survey that contained an audio file also displayed the written sentence, with a blank space indicating which word they were meant to be spelling (e.g., "I went down to the docks for an _____."). Participants then had to type how they thought the pseudoword should be spelled twice, to confirm their response (reducing the likelihood of typos). Participants were then presented with the "correct" spelling of the word and asked to copy it by typing it into a response box before proceeding to the next trial. They then moved on to the next pseudoword and repeated that process. They were then given an N-back trial where they heard the audio of the previously encountered pseudoword and were asked to recall the correct spelling. They were then provided the correct spelling one final time after their N-back (*Figure 7*), regardless of whether their response was correct.

Figure 7. Example of Qualtrics Slide Order



At the end of the pseudoword section, there was a final recall test with the pseudowords presented auditorily in the same order as they were originally listed. Participants were asked to attempt to correctly spell the words one final time, again typing their responses twice to reduce typos.

After completion of the pseudoword portion of the survey, participants were presented auditorily with the 30 real words in the same manner as the pseudowords (word, sentence, word). Again, participants were permitted to listen to the audio as many times as they needed. They were then asked to provide the correct spelling of the word, typing their responses twice to confirm their spelling attempt.

X. Analyses and Results

X.1 Analysis 3

K-means Clustering Analysis: This analysis explored whether individuals demonstrated systematic patterns in spelling ability based on their lexical and sublexical knowledge. K-means clustering aggregates data into groups based on underlying patterns in the data. Eleven predictors from both real and pseudoword data were used to generate clusters, seen in *Figure 9*. Clusters were generated in R using the kmeans function from the *NbClust* Package (Charrad et al., 2014). I then used the fviz_cluster function from the *factoextra* package to visualize the clusters (Kassambra & Mundt, 2020). See *Figure 8* for a cluster plot.

The number of desired clusters was chosen *a priori* to be four so that participants could be put into 1 of 4 possible categories: Bad Lexical-Bad Sublexical (BL-BS), Bad Lexical-Good Sublexical (BL-GS), Good Lexical-Bad Sublexical (GL-BS), and Good Lexical-Good Sublexical (GL-GS). For example, BL-BS indicates that the participant has both a weak lexical knowledge base and is also a weak sublexical speller, making them the worst spellers overall. By that logic, GL-BS indicates the participant is weak at sublexical spelling but has a strong lexical knowledge base, and therefore should be better at spelling real words, and so on. These categories were determined based on participant real word spellings and first guess pseudoword spelling attempts. Specifically, this analysis had 9 variables, 6 sublexical and 3 lexical (See *VI.5. Real Word Measures* and *1-VI. Pseudoword Measures*).

Generalized Linear Mixed-Effects Model: I used the *glmer* function from the *lme4* (Bates et al., 2015) for a GLME model to analyze whether the previously generated clusters make accurate predictions about participants' learning of pseudowords, specifically exploring their performance on both first-guess and final-recall accuracy. The family for the model was set to binomial and the "bobyqa" algorithm was used for the optimizer. The model formula for the fixed effects was accuracy ~ cluster * outcome + (1 + outcome | ID) + (1 | target), where outcome was the number of responses that remained for each pseudoword after errors had been removed. This allowed for assessment of differences between the clusters in terms of first-guess and final recall accuracy. The model also included random effects by-participant and by-items. I used "simple coding" with the BL-GS cluster as the reference group to compare to each other cluster. In the simple coding scheme, a contrast matrix is generated that compares each level to the reference group by assigning a value of ¾ to the level being compared and –¼ to all the other levels. The intercept of the regression is the grand mean of the contrast matrix. The model tested the following predictions:

1) **BL-GS:** Should be second best on first-guess, but third on final-recall.
2) **GL-BS:** Should be third best on first-guess, but second best on final-recall.

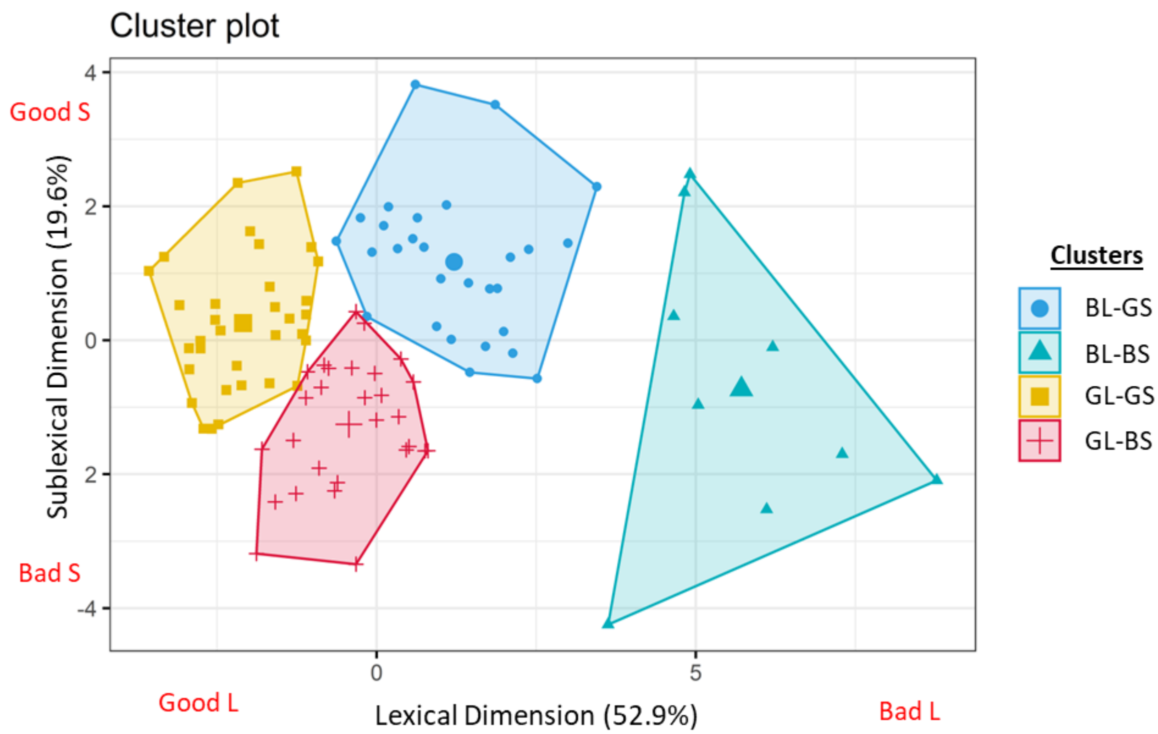
Prediction 1 assumes that the BL-GS group, who are better sublexical spellers, should have more accurate first-guesses than the GL-BS. However, Prediction 2 assumes that the GL-

BS group, who are stronger lexical spellers, will be better at final recall accuracy than BL-GS despite the latter groups' initial advantage. This prediction was made because I anticipate that strong lexical spellers will be better at learning and recalling new words than those who are lexically weak spellers.

We also conducted a GLME to explore the contribution of correct first-guess and N-back spellings. The model formula for these fixed effects was final accuracy ~ first-guess accuracy + N-back accuracy + (1 + first-guess accuracy + N-back accuracy | ID) + (1 | target). The model is constructed in this manner to allow for the fact that a participant may be correct for one spelling attempt but not the other. If first-guess accuracy does not contribute to final accuracy, then the differences in between clusters regarding their first-guess accuracy would be essentially meaningless.

X.2 Results

Figure 8. Cluster Plot



Note: The four clusters generated based on the spelling profiles. The clusters are plotted here along the dimensions of the first two principal components, explaining 52.9% and 19.6% of the variance, respectively.

Figure 8 demonstrates that the clustering analysis were able to successfully categorize participants into four separate profiles with fairly minimal overlap. This indicates that, overall, the attempt to detect systematic patterns of spelling behavior was successful.

Figure 9. Measure Means

Pseudoword Measures	**BL-BS**	**BL-GS**	**GL-BS**	**GL-GS**
PG	0.859	0.899	0.881	0.907
Neighbor	0.85	0.91	0.889	0.921
Oncleus-Coda	0.659	0.73	0.69	0.74
Onset-Rime	0.7	0.752	0.749	0.781
Real Word Measures				
Regularization	0.241	0.298	0.218	0.206
Irregular %	0.492	0.699	0.883	0.927
Regular %	0.909	0.957	0.985	0.993
Irreg - minus - Reg	-0.416	-0.258	-0.102	-0.066
Overall % Correct	0.352	0.593	0.819	0.884
Plausible Error	0.567	0.383	0.164	0.11
Nonplausible Error	0.081	0.024	0.017	0.007
n =	*9*	*28*	*28*	*35*

Note: The means for each cluster of each variable that was used to generate them. Gradients indicate the relative performance of each cluster on that particular measure, with green being strong and red being poor.

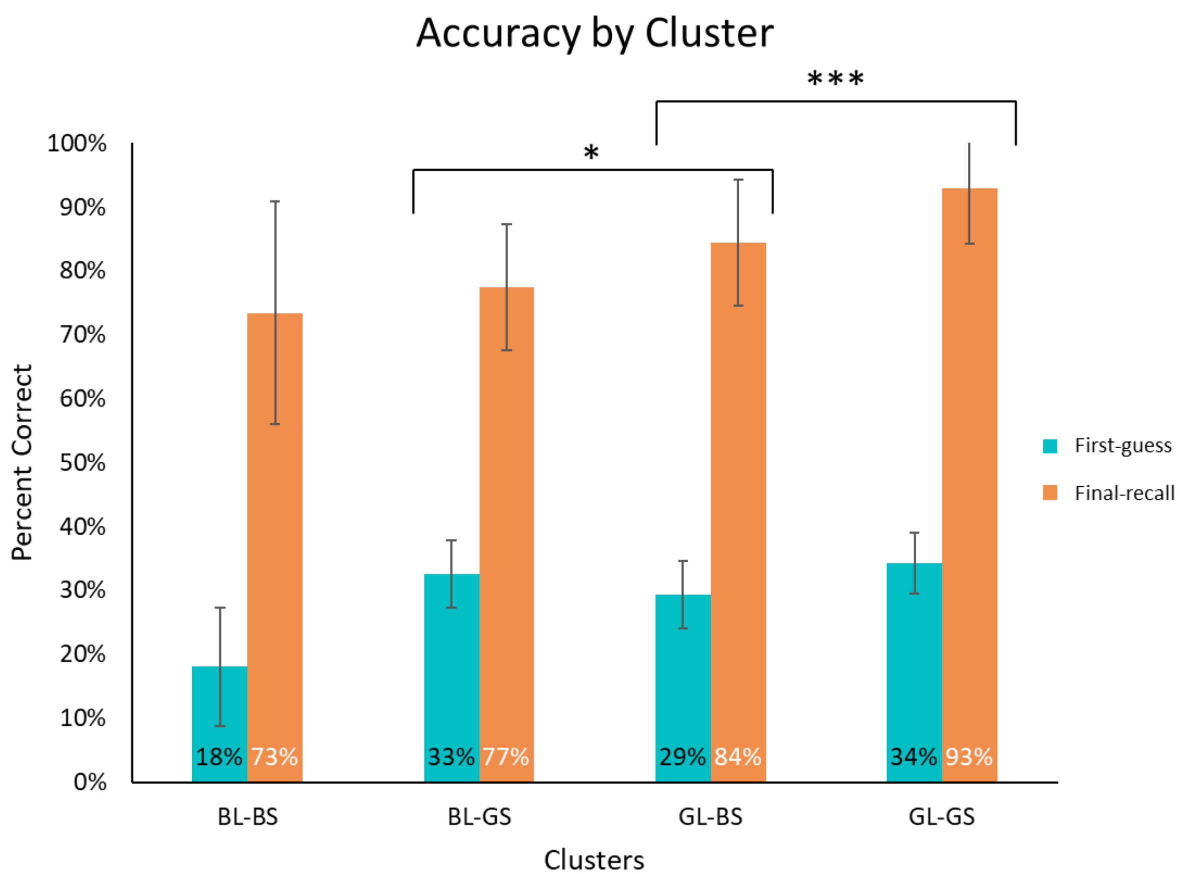
Notice in *Figure 9* that the GL-GS group tended to have the highest means for the predictor variables, while the BL-BS group the lowest. More notable, however, is that the GL-BS group has higher means for real word measures than the BL-GS, but the reverse is true for sublexical measures. This provides initial evidence that Predictions 1 and 2 are correct.

Table 8. Summary of First-and-Final Accuracy GLMM

	Estimate	Std. Error	*Z*-value	*Pr*(>|z|)	
Intercept	0.34	0.12	2.83	0.001	**
BL-GS v GL-BS	0.15	0.16	0.96	0.34	
BL-GS v BL-BS	-0.50	0.22	-2.25	0.02	*
BL-GS v GL-GS	.71	.15	4.74	< .001	***
Outcome	1.29	0.06	22.16	< 2e-16	***
O: BL-GS v GL-BS	.30	.13	2.33	.02	*
O: BL-GS v BL-BS	.28	.19	1.51	.13	
O: BL-GS v GL-GS	.64	.13	5.03	< .001	***
Signif. Codes:		< .001‘***’	< .01‘**’	< .05‘*’	

Note: Summary table of GLMM comparing the relative change in accuracy from first to final spelling between clusters. Marginal R2 = .29

Figure 10. Accuracy by Cluster



Note: First guess and final recall spelling accuracy by cluster. The '*' indicates a significant interaction of first-guess to final-recall accuracy between two clusters. Signif. Codes: '***' $p < .001$, '**' $p < .01$, '*' $p < .05$. Error bars are the 95% CI.

Data from *Table 8* regarding the change between first-guess and final accuracy within clusters is visualized in *Figure 10*. Notice that the predictions of accuracy by group holds true; the BL-GS group is more accurate at first-guesses (33%) than GL-BS (29%). However, the BL-GS (77%) and GL-BS (84%) groups have switched places in regards to which group produces the most accurate spellings in final recall. This interaction was significant, $p \approx 0.02$, indicating that the GL-BS group made more improvements from first-guesses to final recall than did the

BL-GS, despite having less success in first-guesses. There was no such significant interaction when comparing BL-GS and BL-BS clusters ($p \approx 0.13$), presumably because they were both weak in terms of lexical knowledge and therefore did not outperform one another in a learning task. However, the interaction was significant when comparing the GL-BS and GL-GS clusters ($p < 0.001$), suggesting that GL-GS's superior performance first-guesses boosted their final recall performance.

Table 9. Summary of First-guess and N-back GLMM

	Estimate	Std. Error	*Z*-value	*Pr*(>|z|)	
Intercept	.10	.17	.56	.58	
First guess	.76	.12	6.11	9.7e-10	***
N-back	1.60	.15	10.97	<2e-16	***
Signif. Codes:	< .001 '***'	< .01'**'	< .05'*'		

Note: Summary table of GLMM exploring contribution of correct first-guess and N-back on final accuracy.

Table 9 indicates that both correct first-guess and correct N-back recalls made positive contributions to accuracy on final-recall ($p < .01$), demonstrating that correct first-guesses are indeed associated with better final accuracy. If this was the only association that mattered during learning, then the BL-GS group (which had more correct first guesses) would have more correct spellings on final recall. However, despite the initial advantage the BL-GS group has, the GL-BS was able to surpass them in final accuracy, indicating that group's lexical strength more than compensated for incorrect first guesses.

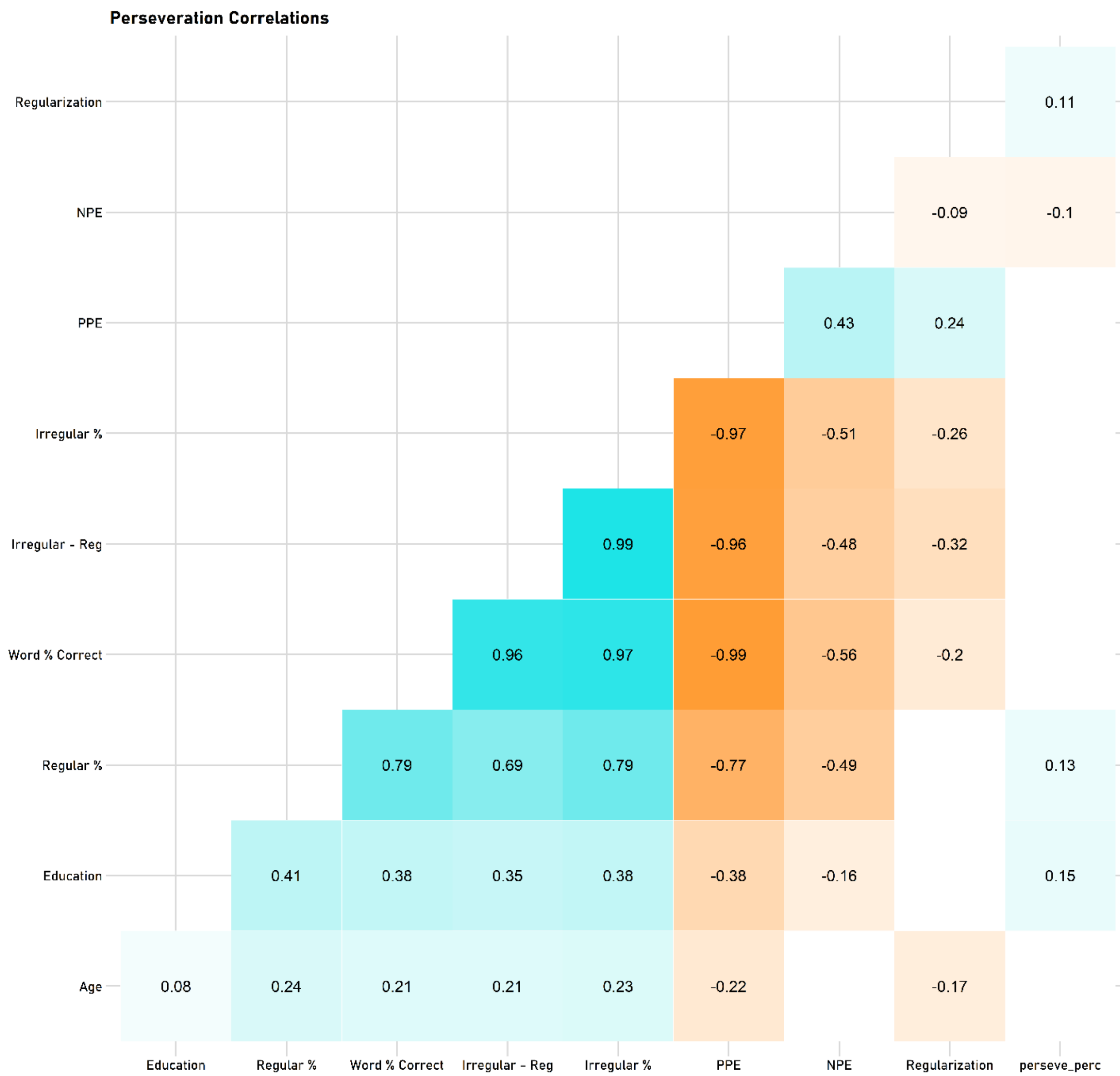
X.4 Analysis 4

This exploratory analysis was conducted due to questions that arose during data collection for Experiment 2. If a participant's N-back was incorrect, was it a perseverated first-guess spelling? Additionally, if a participant's final recall was incorrect, was it a perseverated first-guess or N-back spelling? In essence, I wanted to examine whether or not there was a relationship between real word spelling and perseveration. Additionally, I decided to examine whether certain clusters were more likely than others to revert back to an incorrect spelling. I predicted that the BL-GS group will be the most likely to revert, as they are likely to rely on their strong sublexical system to compensate for their lacking lexical system.

Throughout this analysis, I refer to the percentage of perseverations. Percentage of perseverations refers to the number of times someone perseverated out of their total number of errors. For example, two participants could have 10 perseverations, but one could have 10 reversions out of 10 errors (100%), while the other only had 10 perseverations out of 20 (50%).

X.5 Results

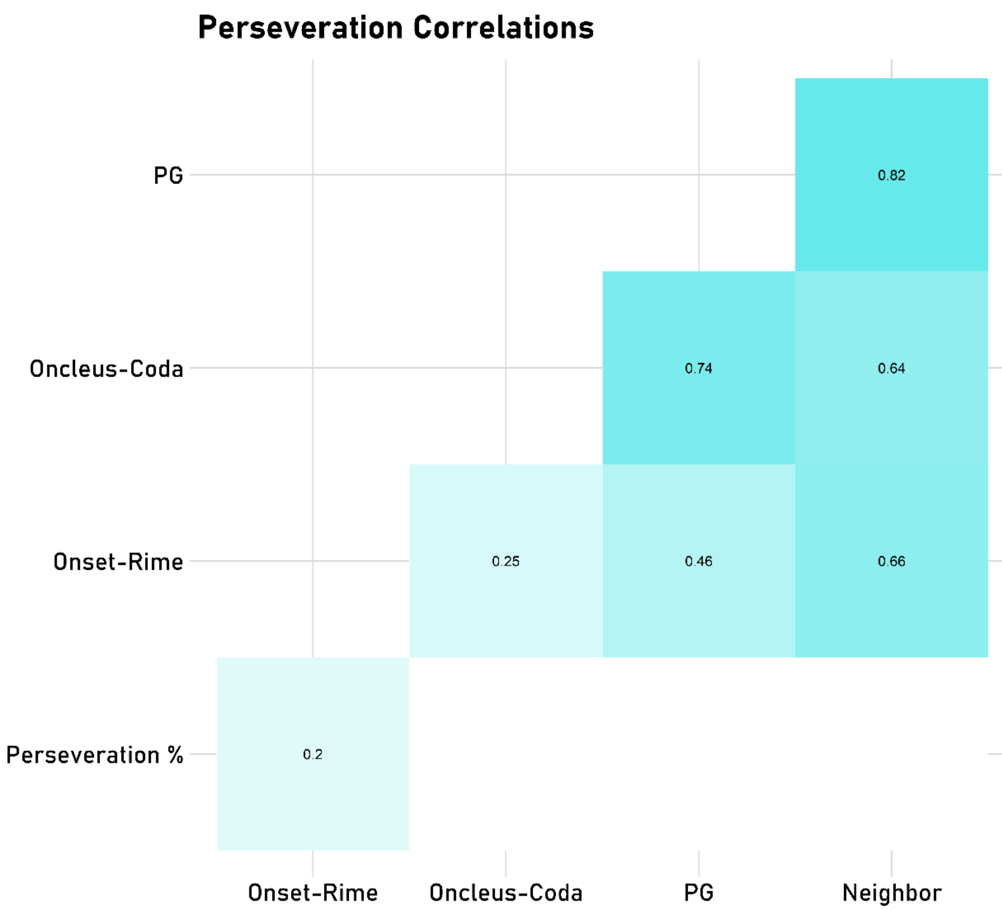
Figure 11. Perseverations and Lexical



Note: Significant Pearson correlations ($p < 0.05$) between percentage of perseverations and multiple real word spelling measures, as well as two demographic measures. Note that a perseveration can be an N-back repeating the incorrect first-guess, a final-recall repeating the incorrect first-guess, or a final-recall repeating an incorrect N-back. Correlation values below |0.198| were not significant and are not depicted here.

Figure 11 demonstrates that there were no significant correlations between percentage of perseverations and real word measures (alpha = .01). Specifically, percent of accuracy of regular segments, regularization, and years of education all positively correlated positively with higher percentage of perseverations. NPE's were negatively correlated with percentage of perseverations. This is logical, as it indicates that when "worse" spelling errors are made, they are less likely to be perseverated during later spelling attempts.

Figure 12. Perseverations and Sublexical



Note: Significant Pearson correlations ($p < 0.05$) between percentage of perseverations and the four sublexical mechanisms. Note that a perseveration can be an N-back repeating the incorrect first-guess, a final-recall repeating the incorrect first-guess, or a final-recall repeating an incorrect N-back. Correlation values below |0.198| were not significant and are not depicted here.

Figure 12 demonstrates participants who tended to adhere to Onset-Rime probabilities tended to have a higher rate of perseverations ($p < .05$). No other sublexical mechanism was significantly correlated with perseveration percentage.

Table 10. Perseverations and Clusters

Clusters	% of Perseverations	# of Perseverations
BL-BS	.469	13.5
BL-GS	.530	11.71
GL-BS	.525	13.64
GL-GS	.561	7.34

Note: The first column indicates the average rate for percentage of perseverations within the four clusters. The second column indicates the average number of perseverations within each cluster.

As seen in *Table 10*, the BL-BS group, perseverations occur at a rate of 46.9%, or slightly less than half of the time. This indicates that this group does indeed experience perseverations, but at a lower rate than all other clusters. Over half the time, this group is not repeating what they were previously told was wrong, but are instead generating new wrong spellings.

CHAPTER IV: DISCUSSION AND CONCLUSIONS

XI. Discussion

Despite its prominence in our everyday lives, the ability to spell is a little understood phenomenon. A popular account of how spelling is achieved are dual route models, which assert that there are two routes of spelling: lexical and sublexical. It is generally accepted that the lexical route involves accessing an individual's known lexicon in order to produce spellings of words they have previously encountered. However, the sublexical route, responsible for our ability to generate spellings for words we are unfamiliar with, has been the subject of intense debate among psycholinguists for decades. Prevailing theories state that the sublexical route operates as a probabilistic system. There is abundant discourse regarding how such probabilities are generated, resulting in multiple proposals as to the nature of the mechanisms employed by sublexical route throughout the years. The P-G account asserts that individual phoneme-to-grapheme probabilities, influenced by an individual's entire knowledge of the English lexicon, are utilized when attempting to spell a previously unknown word. The Neighbor account is generally the same, but asserts probabilities are generated from phonological neighbors rather than the sum of an individual's lexical knowledge. Other proposed mechanisms contend that higher level, sub-syllabic probabilities derived from the Oncleus-Coda or Onset-Rime are critical for new word spelling. The goal of this research was to examine the P-G Mapping, Phonological Neighbor, Oncleus-Coda, and Onset-Rime accounts by observing how well they were able to account for pseudoword spelling behaviors.

We suspected that sublexical spelling is not universally the same across participants. This expectation stems from the lexical quality hypothesis proposed by Perfetti and Hart (2002), which posits that different people likely use different spelling mechanisms depending on their relative lexical and sublexical strength. Evidence for individual differences in sublexical processing was provided by Coltheart and Ulicheva (2018), who demonstrated that people pronounced pseudowords variably both within and across items. I sought to replicate these results with pseudoword spelling and test the four mechanisms to see if they were able to explain any unique levels of spelling variation.

An additional goal of this research was to explore new-word learning, a subject which is again mired in controversy. Share (1994) hypothesized that the sublexical route is responsible for building up the lexical knowledgebase over time, such that dependency upon lexical knowledge steadily increases until the sublexical system may no longer be necessary past a certain threshold. Burt and Blackwell (2008) adopted an extreme position on this hypothesis by asserting that only the lexical system is at work in literate adults, with no sublexical contributions to new word learning at all. To further our understanding of the sublexical system and its relationship to the lexical, I explored whether it is possible to systematically relate real word spelling to pseudoword spelling behaviors. Additionally, I examined whether such spelling behaviors could contribute to explaining variability in new word learning.

In sum, these experiments were conducted with the goal of testing the four sublexical mechanisms based on how well they a) accounted for variability in pseudoword spelling, b) related to real word spelling behaviors, and c) explained new word learning behaviors.

XI.1 Experiment 1

Experiment 1 explored individual differences in pseudoword spelling, in much the same manner as Coltheart and Ulicheva (2018) did for pseudoword pronunciation. Specifically, I examined whether variation was present in pseudoword spelling and, if so, which of the proposed mechanisms best accounted for it. Indeed, in Analysis 1 (*Section VII*) I found that there was a wide variation of pseudoword spellings, ranging from only 2 unique responses to 31 unique responses. It was evident there was within-item variation too, in that some spellings of a given pseudoword were used more frequently than others. Using Poisson regression, I examined whether any of the four proposed sublexical mechanisms contributed to explaining both within- and across-word variation. I found that Neighbors, Oncleus-Coda and Onset-Rime were all able to account for significant unique variance. Of these, probabilities generated by the Neighbors mechanism accounted for the most unique variance. The regression also indicated that probabilities generated by the P-G mechanism did not explain any *unique* variance. Given that P-G's are a component part of the Oncleus-Coda and Onset-Rime mechanisms, there is likely overlap in terms of what variance they account for. The participants' predominant use of Neighbor probabilities aligns with the assertions of Share's self-teaching hypothesis in that people draw upon their lexical knowledge to spell even when encountering new words (Share, 1994).

Experiment 1 also investigated whether I could demonstrate a systematic relationship between the lexical and sublexical systems. Analysis 2 (*Section 1-IX*) provided evidence that there is indeed an association between pseudoword and real word spelling behaviors. Specifically, people who tend to use Neighbor probabilities when spelling pseudowords also tend to be better real word spellers, both in overall accuracy and irregular segment accuracy. A possible interpretation of these results is that these participants have an extensive number of words stored in their O-LTM and are therefore inclined to use this sizeable database of knowledge when spelling new words. In essence, those with more years of education likely have a larger vocabulary and are therefore more likely to have encoded irregular mappings that appear extremely infrequently in English.

Use of P-G probabilities, however, was associated with worse real word spelling. This makes sense, as P-G's are a lower order representation that draw from the broader known lexicon rather than a specific grouping of words. These results suggest that those who use individual P-G probabilities likely have a limited lexical database to draw from, resulting in a restricted vocabulary that provides little useful information regarding irregular segment spelling.

XI.2 Experiment 2

The purpose of Experiment 2 was to examine individual differences in new word learning. I successfully generated clusters in Analysis 3 (*Section 2-III*) based on a combination of real and pseudoword predictors. In doing so, I was able to categorize individuals as one of four types of spellers reflecting relative lexical and sublexical strength. I wanted to see if these clusters would support predictions about learning the spellings of new words. The main prediction was that the BL-GS group would initially provide more accurate guesses as to the correct spelling of the pseudowords than the GL-BS group. By the final-recall, however, I believed that the GL-BS group would overtake the BL-GS group in spelling accuracy despite the initial advantage the stronger sublexical spellers have. This is because the GL-BS group has a stronger lexical knowledge base and should therefore be more successful at memorizing words. Indeed, while BL-GS started ahead of GL-BS on first-guess accuracy, they fell behind GL-BS significantly on final accuracy ($p \approx 0.02$).

While there was no significant difference in the change in accuracy from first-guess to final-recall between BL-BS and BL-GS ($p \approx 0.13$), the BL-BS group did perform worse in final-recall accuracy than the BL-GS group, despite both being lexically weak. This indicates that the sublexical route does still play a role in new word learning even in literate adults, counter to beliefs put forth by Burt and Blackwell (2008). Finally, I found that GL-GS improved significantly more from first-guess to final-recall than BL-GS (p < .001); despite both groups being good at sublexical spelling and producing more accurate first-guesses than the other two groups, GL-GS still learned significantly more words than BL-GS. This provides further confirmation that the GL-GS group has better lexical knowledge and may leverage their sublexical processing to better support integrating new words into O-LTM.

These results indicate that both lexical and sublexical strength independently benefit the learning of new words. The sublexical route allows for processing and integration of new information into the O-TLM; logically, a stronger sublexical system will more efficiently contribute to lexicalization (Share, 1995). Alternatively, it is possible that strong sublexical spellers may be able to consistently produce the same spelling via "guessing" during each spelling attempt, so that even in cases of memory failure they are still able to generate the correct spelling. The relative strength of the lexical system speaks to the degree of past success in learning new words, with stronger systems indicating increased rates of success. In essence, those who have strong lexical systems obtained their broad pool of vocabulary because they are good at memorizing and recalling spellings.

As seen in Experiment 1 (*Section 1-X. Results)*, those who tended to use Neighbor probabilities also tended to be better real word spellers. This suggests that those who possess a stronger lexical system more often draw from O-LTM to inform new word spellings. The more neighbors the sublexical system has to compare a new word to, the more accurate the probabilities generated are likely to be.

I also examined the rates of perseveration errors among the four groups, which I interpret as failures to recall the correct spellings despite the fact that feedback was provided throughout the learning trials. As previously established, the GL-BS group is inclined to perform better on lexical tasks, indicating they are likely better at remembering previously seen spellings than the BL-GS group. However, as indicated by Analysis 4 (*Section 2-V*), when they are unable to recall the correct spelling, the GL-BS group defaults back to their original guess over half of the time (52.5%). Despite the fact that they are not necessarily strong at sublexical spelling, it is probable that members tend to remember the initial guess they made and default back to it instead of attempting to generate an entirely new spelling. For the BL-GS group, perseverations occur over half the time an error is made (53%). This rate of perseverations may be occurring because this group tends to make good initial guesses, and therefore tends to default back to their original guess when they are unable to recall the correct spelling of a word. It is also possible that this group generates a new guess at the spelling, but due to their strong sublexical abilities, they tend to generate the same incorrect guess as before instead of producing new spellings at random. Unsurprisingly, the GL-GS group tends to make less errors overall compared to the other clusters. When they do make an error, they tend to default back to their previous incorrect spelling attempt at a rate of 56.1%. This high rate of perseverations may be due to a combination of aforementioned tendencies: GL-GS participants tend to make good first guesses and are more likely to remember those guesses when recall of the correct spelling fails.

We also found that adherence to Onset-Rime probabilities was correlated with a higher rate of perseverations. One possible explanation for this result is that the GL-GS group tends to have higher-level representations of spelling, in this case adhering to sub-syllabic structures.

Finally, the BL-BS group had the lowest rate of perseverations (46.9%), despite making the most errors overall. I interpret this somewhat surprising finding as evidence that, given to their weak lexical knowledge, the BL-BS group struggles to encode or retrieve spellings from memory and are therefore less likely to remember the either correct spelling or their own previous response. In such cases of forgetting, the BL-BS group turn to their sublexical system to produce a new response. However, since their sublexical system is also weak (and therefore inconsistent), they are likely to produce a spelling that differs from their previous response.

XI.3 Implications and Limitations

Our research provides evidence that there are indeed individual differences in pseudoword spelling behavior. Moreover, probabilities generated from Neighbors, Oncleus-Coda, and Onset-Rime are all able to account for spelling variability, with Neighbors being the best. Generally, psycholinguists research and support only a singular sublexical mechanism, yet the results indicate that it is likely multiple mechanisms are at play when people spell new words. Such data suggests that researchers should not be so restrictive in their research regarding the sublexical system, and instead embrace the possibility that different people do indeed operate under different mechanisms. The next logical step is to investigate which individuals use what mechanism or combination of mechanisms, and how those varying mechanisms may interact with one another.

We also discovered that there is some sort of relationship between lexical and sublexical spelling. Specifically, those who use Neighbor probabilities tended to be better real word spellers, while those who used P-G probabilities were worse. This indicates that those who have a larger base of lexical knowledge are more likely to recruit said knowledge to aid in the processing of new words. Those with a restricted vocabulary may rely more on their sublexical spelling system, limiting their ability to successfully spell the plethora of irregular words that appear in the English language. Future research should examine exactly why those who use Neighbors probabilities rather than Oncleus-Coda or Onset-Rime (which are both higher-order representation that draw heavily from lexical knowledge) are better real word spellers. In terms of real word application, the possibility of explicitly instructing people to use Neighbors when spelling new words to mitigate their sublexical weakness should be explored. In sum, I have established that there is a relationship between the lexical and sublexical systems, but the precise nature of said relationship and what that means from an educational standpoint should be the subject of future studies.

Finally, I found that it is indeed possible to categorize people based on their spelling abilities and make accurate predictions about their learning outcomes. Such ability to classify people by their relative strengths and weaknesses in lexical and sublexical processing is extremely promising in terms of real-world applications. For example, clustering could be conducted on students in a classroom environment, allowing instructors identify where each group of students may need targeted instruction. Teachers or special tutors could formulate explicit plans of study for each of these groups that addresses where weaknesses in their respective spelling systems lie. Given the promising results regarding potential predictive clusters (See section *X.1. Analysis 3*), it is clear that further research into how clustering could be made into a feasible diagnostic tool is necessary. Identifying what blend of tests work best to accurately identify types of spellers, which teaching methods allow for the greatest level of improvement for each of these groups, and how to unobtrusively implement clustering in a real-world environment is also critical.

Our research also provides insight into where current weaknesses may lie in our diagnostic tools for spelling disorders. Pseudowords are often used to diagnose disorders such as dyslexia and dysgraphia to explore patients' sublexical processing (e.g., Tilanus et al., 2013). However, some researchers contend that using pseudowords provides no unique information beyond that of what real words can provide (e.g., Burt and Blackwell, 2008). Irregular word tasks are generally considered to explore lexical processing by using pronunciations and/or spellings that are extremely rare and require prior knowledge for accuracy. Patients are often classified by accuracy on different aspects of real words, e.g., irregular vs. regular or high vs. low frequency (Rapp, 2015). My research indicates that lexical knowledge and sublexical knowledge can be relatively strong or weak independently of one another (See section *X1. Analysis 3*).

Analysis 2 found that there is indeed a relationship between sublexical and lexical processes; those who tend to use Neighbor probabilities (which are by definition derived from known lexical items) are better real word spellers, while those who use P-G probabilities are worse. However, we are only just beginning to explore the nature of this relationship, and the practical implications thereof. It is extremely likely that many disorders span both the lexical and sublexical routes. Patcon et al. (2014) demonstrated that when participants made recall errors on pseudowords with doubled letters, they typically moved the doubling to a letter that is more commonly a doublet. This demonstrates interference from the lexical system during a sublexical conversion task, supporting the argument that both P-G conversion and O-LTM play a role in new-word learning. The tools we commonly use for diagnostics are useful insofar as to explore possible disorders of the sublexical route, but generally ignore possible disorders of the lexical knowledgebase. Further research is most certainly needed to more thoroughly extrapolate how pseudoword and real word spelling behaviors interact, so that we can determine which combination of tools constitutes the best battery of tests.

A limitation of this research is that certain contributions to variation, such as dialectical differences, were outside the scope of these particular experiments. Dialectical variation can be thought of in terms of how varying accents tend to pronounce and perceive phonemes differently. For example, British English speakers pronounce the words pin and pen as /pɪn/ and /pɛn/, while certain groups of American English speakers have experienced a language "merger" in which the distinction is lost (both are pronounced as /pɛn/). Further research is needed to determine to what extent, if any, these linguistic variations contribute to individual differences in spelling. Another notable limitation is that I excluded proper nouns when generating Neighbor, Onset, and Rime probabilities. This is because proper nouns are highly regional, and words (specifically names) that are highly frequent and salient in one area may be completely unheard of in another. However, it must be recognized that it is entirely possible participants included proper nouns in their consideration of how to spell a pseudoword (e.g., /jɑnd/ spelled JOHND after JOHN, as opposed to JAWND or JAUND). Other limitations include excluding the influence orthographic feedback has on spelling, and using exclusively monosyllabic words after Experiment 1 (Mousikou et al., 2017; Lee, 2001). Future research could explore the possible impact of orthographic feedback and determine whether the current findings remain true for multisyllabic words.

One might ask why researchers bother to study pseudoword spelling in fully literate adults. Wouldn't researchers time be better spent examining the spelling behaviors of new language learners, or exploring real words knowledge? In reality, even literate adults do not have complete knowledge of the English language and may encounter new words that they need to be able to accurately process and integrate into their knowledge base. Neologisms, or newly coined words, are extremely common given the ever-changing nature of a living language; most dictionaries announce a "word of the year" that is often a neologism. Each generation tends to develop their own slang that is nigh uninterpretable to older generations, and new technologies or commercial products also receive new, unique names as an identifier. Word learning is a lifelong endeavor, and pseudoword studies allow for further exploration how this occurs so that we might one day fully understand how we achieve such learning.

XII. Conclusions

Spelling is not as simple as hearing a word and writing it down. A highly complex, little understood sublexical system is at play that allows people to systematically generate and learn new spellings. In better understanding this system, researchers can discover why spelling can vary so greatly from person to person and what makes someone a better or worse speller and. My hope is that this line of research will eventually be able to inform better, scientifically founded tools for teaching spelling within the general population, alongside more targeted and effective treatments for those with dsylexia and dsygraphia.

REFERENCES

Barton, K. (2009) Mu-MIn: Multi-model inference. R Package Version 0.12.2/r18. http://R-Forge.R-project.org/projects/mumin/

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

Berndt, R. S., Reggia, J. A., & Mitchum, C. C. (1987). Empirically derived probabilities for grapheme-to-phoneme correspondences in English. *Behavior Research Methods, Instruments, & Computers*, 19(1), 1-9.

Bolger, D. J., Hornickel, J., Cone, N. E., Burman, D. D., & Booth, J. R. (2008). Neural correlates of orthographic and phonological consistency effects in children. *Human Brain Mapping*, *29*(12), 1416–1429.

Brooks ME, Kristensen K, van Benthem KJ, Magnusson A, Berg CW, Nielsen A, Skaug HJ, Maechler M, Bolker BM (2017). glmmTMB balances speed and flexibility among packages for zero-inflated generalized linear mixed modeling. *The R Journal*, 9(2), 378–400. https://journal.r-project.org/archive/2017/RJ-2017-066/index.html.

Burt, J. S., & Blackwell, P. (2008). Sound–spelling consistency in adults' orthographic learning. *Journal of Research in Reading*, *31*(1), 77-96.

Charrad M, Ghazzali N, Boiteau V, Niknafs A (2014). NbClust: An R package for determining the relevant number of clusters in a data set.” *Journal of Statistical Software*, 61(6), 1–36. https://www.jstatsoft.org/v61/i06/.

Chee, Q. W., Chow, K. J., Yap, M. J., & Goh, W. D. (2020). Consistency norms for 37,677 english words. *Behavior Research Methods*, *52*(6), 2535-2555.

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.

Coltheart, M., & Ulicheva, A. (2018). Why is nonword reading so variable in adult skilled readers? *Peerj*, *6*, 4879.

Coxe, S., West, S. G., & Aiken, L. S. (2009). The analysis of count data: A gentle introduction to Poisson regression and its alternatives. Journal of personality assessment, 91(2), 121-136.

Coxe, S. (n.d.). Effect size for Poisson regression. Retrieved September 17, 2022, from https://stefany.shinyapps.io/RcountD/

CMU Dictionary, Carnegie-Melon University, (1995)

Folk, J. (2014). Lexical influences on the sublexical system during nonword spelling. *Soj Psychology*, *1*(3), 01–08. https://doi.org/10.15226/2374-6874/1/3/00112

Fry, E. (2004). Phonics: A large phoneme-grapheme frequency count revised. *Journal of Literacy Research, 36*(1), 85-98.

Gilhooly, K. J., & Watson, F. L. (1981). Word age-of-acquisition effects: A review. *Current Psychological Reviews*, 1(3), 269-286.

Grömping U (2006). "Relative Importance for Linear Regression in R: The Package relaimpo." *Journal of Statistical Software*, **17**(1), 1–27.

Hanna, P. R., Hanna, J. S., Richard E. Hodges, & Edwin H. Rudorf, Jr. (1966). Phoneme-grapheme correspondences as cues to spelling improvement.

Jared, D., McRae, K., & Seidenberg, M. S. (1990). The basis of consistency effects in word naming. *Journal of Memory and Language*, 29(6), 687–715.

Karimi, H., & Diaz, M. (2020). When phonological neighborhood density both facilitates and impedes: age of acquisition and name agreement interact with phonological neighborhood during word production. *Memory & Cognition*, *48*(6), 1061–1072.

Kassambara, A. and Mundt, F. (2020) Factoextra: Extract and Visualize the Results of Multivariate Data Analyses. *R Package Version 1.0.7.*

Kuznetsova A, Brockhoff PB, Christensen RHB (2017). "lmerTest package: Tests in linear mixed effects models." *Journal of Statistical Software,* **82**(13), 1–26.

Lee, C. H. (2001). Absence of syllable effects: monosyllabic words are easier than multisyllabic words. *Perceptual and Motor Skills*, *93*(1), 73–77.

Lenth, Russel. (2019). Emmeans: estimated marginal means, aka least-squares means. *R package version 1.4.1*.

Liberman, A. M. (1992). The relation of speech to reading and writing. *Advances in Psychology*, 94, 167–178. http://dx.doi.org/10.1016/S0166-4115(08)62794-6

Lumley, Thomas, based on Fortran code by Alan Miller. (2020). Leaps: Regression subset selection. *R package version 3.1*. https://CRAN.R-project.org/package=leaps

Marian, V., Bartolotti, J., Chabal, S., & Shook, A. (2012). CLEARPOND: Cross-linguistic easy-access resource for phonological and orthographic neighborhood densities. *PLoS ONE*, 7(8), e43230. https://doi.org/10.1371/journal.pone.0043230

Marjou, X. (2019). OTEANN: Estimating the transparency of orthographies with an artificial neural network. *arXiv preprint arXiv:1912.13321*.

Meng, H. M., Seneff, S., Zue, V. W., US Dept of the Navy Funding, & MASSACHUSETTS INST OF TECH CAMBRIDGE LAB FOR COMPUTER SCIENCE Performing. (1994). In *Phonological parsing for bi-directional letter-to-sound/sound-to-letter generation*. essay, Defense Technical Information Center : Fort Belvoir, VA. https://doi.org/10.21236/ADA458591

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

Nittrouer S., Boothroyd A. (1990). Context effects in phoneme and word recognition by young children and older adults. *J. Acoust. Soc. Am.* 87, 2705–2715. 10.1121/1.399061

Oldfield, R. C., & Wingfield, A. (1965). Response latencies in naming objects. *Quarterly Journal of Experimental Psychology*, 17(4), 273281.

Pacton, S., Borchardt, G., Treiman, R., Lété, B., & Fayol, M. (2014). Learning to spell from reading: General knowledge about spelling patterns influences memory for specific words. *Quarterly Journal of Experimental Psychology*, 67(5), 1019–1036.

Pang Dan-ling & Yang Hui (1997) The phonological processing of Chinese phonograms. *Asia Pacific Journal of Speech, Language and Hearing*, 2:3, 177-194, DOI: 10.1179/136132897805577323

Patterson, T. J., & Folk, J. R. (2014). Lexical influences on the sublexical system during nonword spelling. *SOJ Psychology*. https://doi.org/10.15226/2374-6874/1/1/00112

Pecher, D., Boot, I., van, D. S., Madden, C. J., Huber, D. E., & Zeelenberg, R. (2011). The sound of enemies and friends in the neighborhood. *Experimental Psychology*, *58*(6), 454–63. https://doi.org/10.1027/1618-3169/a000113

Perfetti, C. A., & Hart, L. (2002). The lexical quality hypothesis. In L. Verhoeven, C. Elbro, & P. Reitsma (Eds.), *Studies in Written Language and Literacy* (Vol. 11, pp. 189–213). John Benjamins Publishing Company. https://doi.org/10.1075/swll.11.14per

Pexman, P. M., Cristi, C., & Lupker, S. J. (1999). Facilitation and interference from formally similar word primes in a naming task. *Journal of Memory and Language*, *40*(2), 195–229. https://doi.org/10.1006/jmla.1998.2617

Pritchard S, Coltheart M, Palethorpe S, & Castles A. (2012). Nonword reading: comparing dual-route cascaded and connectionist dual-process models with human data. *Journal of Experimental Psychology*. 2012;38(5):1268–1288. doi: 10.1037/a0026703.

Rapp, B. (2015). Handbook of cognitive neuropsychology: What deficits reveal about the human mind. *Psychology Press.*

Rapp, B., Epstein, C., & Tainturier, M. J. (2002). The integration of information across lexical and sublexical processes in spelling. *Cognitive neuropsychology*, *19*(1), 1-29.

Rapp, B., Folk, J. R., & Tainturier, M.-J. (2001). Word reading. In *The Handbook of Cognitive Neuropsychology: What deficits reveal about the human mind* (pp. 233–262).

Rapp, B., & Lipka, K. (2011). The literate brain: The relationship between spelling and reading. *Journal of cognitive neuroscience*, *23*(5), 1180-1197.

Rastle, K., Harrington, J., & Coltheart, M. (2002). 358,534 nonwords: The ARC nonword database. *Quarterly Journal of Experimental Psychology,* 55A, 1339-1362.

Roverud, E., Bradlow, A. R., & Kidd, G. (2020). Examining the sentence superiority effect for sentences presented and reported in forwards or backwards order. *Applied Psycholinguistics*, *41*(2), 381–400. https://doi.org/10.1017/S014271642000003X

Sadat, J., Martin, C. D., Costa, A., & Alario, F. X. (2014). Reconciling phonological neighborhood effects in speech production through single trial analysis. Cognitive psychology, 68, 33-58.

Schüürmann, G., Ebert, R.-U., Chen, J., Wang, B., & Kühne, R. (2008). External Validation and Prediction Employing the Predictive Squared Correlation Coefficient—Test Set Activity Mean vs Training Set Activity Mean. *Journal of Chemical Information and Modeling*, 48(11), 2140–2145. https://doi.org/10.1021/ci800253u

Seidenberg, M. S., & Tanenhaus, M. K. (1979). Orthographic effects on rhyme monitoring. *Journal of Experimental Psychology: Human Learning and Memory*, 5, 546 –554. http://dx.doi.org/10.1037/0278-7393.5.6.546

Share, D. L. (1995). Phonological recoding and self-teaching: sine qua non of reading acquisition. *Cognition*, *55*(2), 151–218.

Shea, J., Wiley, R. W., Moss, N., & Rapp, B. (2020). Pseudoword spelling ability predicts response to word spelling treatment in acquired dysgraphia. *Neuropsychological Rehabilitation*, 38.

Siew, C. S. Q., & Vitevitch, M. S. (2019). The phonographic language network: Using network science to investigate the phonological and orthographic similarity structure of language. *Journal of Experimental Psychology*, *148*(3), 475–475. https://doi.org/10.1037/xge0000575

Stemberger, J. P. (2004). Neighbourhood effects on error rates in speech production. *Brain and Language*, *90*(1), 413–422. https://doi.org/10.1016/S0093-934X(03)00452-8

Tainturier Marie-Josèphe, Bosse, M.-L., Roberts, D. J., Valdois, S., & Rapp, B. (2013). Lexical neighborhood effects in pseudoword spelling. *Frontiers in Psychology*, *4*. https://doi.org/10.3389/fpsyg.2013.00862

Tilanus, E. A., Segers, E., & Verhoeven, L. (2013). Diagnostic profiles of children with developmental dyslexia in a transparent orthography. *Research in Developmental Disabilities*, *34*(11), 4194-4202.

Ulicheva, A., Coltheart, M., Grosseck, O., & Rastle, K. (2021). Are people consistent when reading nonwords aloud on different occasions? *Psychonomic Bulletin & Review*, *28*(5), 1679–1687. https://doi.org/10.3758/s13423-021-01925-w

Wickham H (2016). ggplot2: Elegant graphics for data analysis. *Springer-Verlag New York.* ISBN 978-3-319-24277-4, https://ggplot2.tidyverse.org.

Yates, M., & Slattery, T. J. (2019). Individual differences in spelling ability influence phonological processing during visual word recognition. *Cognition*, *187*, 139-149.

Zevin, J. D., & Seidenberg, M. S. (2006). Simulating consistency effects and individual differences in nonword naming: a comparison of current models. *Journal of Memory and Language*, *54*(2), 145–160.

APPENDIX A: IRREGULAR REAL WORD STIMULI USED IN EXPERIMENT 1

Irregular words
Arraigned
Awning
Barrack
Bidet
Cayenne
Chasm
Comely
Czar (or Tsar)
Eunuch
Eyesore
Frauds
Fuels
Gauche
Girder

Knelt
Liquored
Loathing
Macabre
Masseur
Onslaught
Plagues
Plume
Psalms
Rogues
Schlep (or Shlep, Schlepp, Shlepp)
Sorrel
Sovereign
Spheres
Trousseau
Whirling

APPENDIX B: PSEUDOWORDS USED IN EXPERIMENT 1

Note that some pseudowords have accepted alternates in parenthesis.

Pseudowords (phonemes)	***most frequent spelling***
bæm pɚ	BAMPER
boʊ mɪŋ	BOMING
bɹuk	BRUKE
bʌ vi	BUVVY
dʒɚ loʊ	GERLOW
eɪft	AIFT
froʊgz (θroʊgz)	FROGUES
gʌntʃ	GUNCH
gwɑrfs (gwɑrθs)	GWARFS
haʊks	HOWKS
hɚkʃ	HURKS
i ɹaɪ	ERYE
kɑm bel	COMBLE
klɛzd	KLEZZED

oʊ laɪ	OLIE
plud	PLUDE
pɹoʊ bjut	PROBUTE
pʌm bəl	PUMBLE
ɹaɪ nə	RHINA
saɪ ɑrk (saɪ jɑrk)	SYARK
saɪps	SIPES
saʊdz	SOUDS
sɪdəl (sɪtəl)	SIDDLE
skwulz (skwəlz)	SQUOOLS
spɹʌmf	SPRUMPH
ʃraʊnθ	SHROUNTH
tʃæ kɚ	CHACKER
tweɪk	TWAKE
ʌnks	UNKS
zɚnz	ZERNS

APPENDIX C: SENTENCES USED FOR REAL WORDS IN EXPERIMENT 2

Irregular word sentences
He was formally **arraigned** in court.
The **awning** provided shelter from the rain.
It was crowded in the **barrack.**
The bathroom **bidet** is perfectly clean.
Cayenne is tasty, but very spicy.
The **czar** wore a glittering crown.
Those salesmen were a bunch of **frauds.**
She **fuels** her car.
The man **knelt** over the grave.
The rum cake was heavily **liquored**.
He suffered a verbal **onslaught.**
Plagues can be highly lethal.
The book contained many old **psalms.**
They **schlep** up the staircase.
He reins over a **sovereign** nation.
Spheres are my favorite shape.
The plane is **whirling** around in the air.

A massive **chasm** lies between the mountains.
She is such a **comely** woman.
The Roman emperor's favorite servant was his **eunuch**.
That ugly neon sign is an **eyesore**.
It was **gauche** of you to mention her accident.
The **girder** groaned under the weight.
Her **loathing** of him goes beyond hatred.
She has a **macabre** sense of humor.
The **masseur** rubbed the sore muscle.

APPENDIX D: PSEUDOWORDS, SENTENCES, AND TWO ALTERNATE SPELLINGS

USED IN EXPERIMENT 2

Note that some pseudowords have accepted alternates (words in parenthesis in IPA column).

IPA	***Sentences***	***1st List***	***2nd List***
wɝɹp	I fell but was able to /wɝɹp/ back up.	wurp	werp
zɚnz	They /zɚnz/ for way too long.	zurns	zerns
pjut	I own a fairly small /pjut/.	pute	pewt
gaɪt	The computer came with a /gaɪt/.	geit	gite
stɪv	Whether a rectangle, square, or /stɪv/, they are all shapes.	stiv	stive
braɪ	The show came to a /braɪ/ end.	bry	brigh
poʊθ (poʊf)	He made /poʊθ/ on the winter night.	poath	poth
stɝp	He took the /stɝp/ through the woods.	stirp	sturp
troʊ	The best /troʊ/ come from that region.	trow	trough
θil (fil)	The /θil/ spun around violently.	theel	thiel
dæs	Don't /dæs/ your eggs before they hatch.	das	dass
bɪmf	Please ask where she put the /bɪmf/.	bimf	bimph
hɚk	I cannot believe you can /hɚk/ on an airplane.	herc	hurk
snaɪt	Did you find a /snaɪt/ in the basement?	snite	snight
fɹunt	The man will /fɹunt/ everyone at the party.	fruint	froont
snoɪ	I would /snoɪ/ more often if I had time.	snoi	snoy
dʒaɪn	He picked only the /dʒaɪn/ roses.	gine	jine

dʒɚ	She may be ugly, but she's no /dʒɚ/.	jer	ger
dʒɜrst	The /dʒɜrst/ followed the rabbit.	jurst	jerst
bɛlf	Find the /bɛlf/ and correct it immediately.	belf	belph
bju	The /bju/ burned quickly.	bew	bue
driθ (drif)	The team was set to destroy the /driθ/.	dreath	dreeth
skɪmf	That little /skɪmf/ took him by surprise.	skimf	skimph
daʊð (daʊv)	I saw that /daʊð/ behind the building.	douth	dowth
mɪp	Try not to /mɪp/ the entire night.	mip	mipp
kɪbz	She likes to /kɪbs/ when it's raining.	kibbs	kibs
rɑlt (rɔlt)	The wall collapsed, revealing the /rɑlt/.	ralt	rolt
bʌv	Bees are completley unable to /bʌv/.	bove	buv
soʊf	He had to /soʊf/ it twice before it worked.	soph	soaf
blɜrm	The door made a /blɜrm/.	blerm	blurm
ʃɹaʊn	He /ʃɹaʊn/ away from his mother.	shroun	shrown
ditʃ	That is the /ditʃ/ this evening.	deetch	deach
tɪv	The man's /tɪv/ bike made every head turn.	tive	tiv
dʒɑnd (dʒɔnd)	I just really like my /dʒɑnd/.	jond	jaund
zul	You /zul/ the box.	zool	zule
eɪft	She was going down to the docks for an /eɪft/.	aift	afte
klæf	There’s a heard of /klæf/ on the mountainside.	claff	klaff
tʃuk	The /tʃuk/ leaves were raining from the sky.	chook	chuke
blɑd (blɔd)	Your /blɑd/ interior looks great with that color.	blod	blad

kɛlf	The /kɛlf/ took only seconds to pass.	kelph	kelf
bIltʃ	Don't /bIltʃ/ your chances to get in.	bilch	biltch
floʊk	There was a huge /floʊk/ yesterday.	floke	floak
fin (θin)	I think I saw a /fin/ in the garage.	feen	fien
poʊg	The /poʊg/ was wrapped neatly in paper	poge	pogue
skweɪt	Turtles can /skweɪt/, but only slowly.	squate	squait
tɹitʃ	Your father made /tɹitʃ/ remarks after the debate.	treach	treech

APPENDIX E: EXCEPTIONS TO THE MOP WHEN CONDUCTING P-G MAPPING

The grapheme X – Under the maximum onset principle, X can be a problematic grapheme. For example, MOP demands that the word EXPLAIN would be broken apart as /ɪk `splaɪn/. As we can see, the X crosses the boundary between the syllables. This is corrected using the same approach as Chee et al. (2020), who determined that in these specific cases the X should be put into the coda: EX → /ɪks/ + PLAIN →/plaɪn/.

Final E – Some parsings would force the silent E away from its vowel. For example, the MOP says that the word BASEBALL would be parsed as BA // SEBALL. However, the E in BASEBALL is mapped with A as A_E, so this causes it to cross the boundary between syllables. Instead, we map it as BASE // BALL.

-When following a consonant, final E maps onto preceding vowel grapheme in the syllable if it carries no phonetic value of its own (e.g., /u/ → U_E in DUKE, *not* /k/ → KE). This is also known as the 'silent E'. When following a vowel, silent E is included with vowel graphemes (e.g., /u/ → UE in GLUE, *not* U_E)

Morphemes – Our system does do not have a code specific to morphemes, as they are instead handled implicitly.

- When a silent E precedes a D, this is always mapped as /d/ → ED as in BAGGED or /t/ → ED as in WALKED. This remains true even in cases where ED follows a vowel (e.g. /d/ → ED as in ALLOWED, not /aʊ/ → OWE) or the word stem had a final E (e.g. PLANED is still /d/ → ED).
- Plural morphemes, which include the phonemes /s/, /z/, or /əz/ (or /ɪz/ depending on the accent), are also handled implicitly. When ES is pronounced /əz/ (where you can

	still ‘hear’ the vowel), such as in BUSSES, ES is mapped as two graphemes, /ə/ →E and /z/ → S. When the vowel in ES is silent, such as in GOES or CAKES, it is mapped as /s/ → ES. Note that again, ES is always used even in cases where the original stem had a final E.
Other silent letter - Other letters that seem to be ‘silent’ are mapped to phonological vowels whenever possible (e.g., CAUGHT is C-AUGH-T (not C-AU-GHT)… IRON is I-RO-N not I-R-ON).	