PHONETIC REFLEXES OF ORTHOGRAPHIC CHARACTERISTICS IN LEXICAL REPRESENTATION

by

Jordan B. Brewer

A Dissertation Submitted to the Faculty of the

DEPARTMENT OF LINGUISTICS

In Partial Fulfillment of the Requirements
For the Degree of

DOCTOR OF PHILOSOPHY

In the Graduate College

THE UNIVERSITY OF ARIZONA

2008

THE UNIVERSITY OF ARIZONA
GRADUATE COLLEGE

As members of the Dissertation Committee, we certify that we have read the dissertation

prepared by Jordan B. Brewer

entitled Phonetic Reflexes of Orthographic Characteristics in Lexical Representation

and recommend that it be accepted as fulfilling the dissertation requirement for the

Degree of Doctor of Philosophy

____ Date: 11/16/07
Heidi Harley

____ Date: 11/16/07
Michael Hammond

____ Date: 11/16/07
Diane Ohala

Final approval and acceptance of this dissertation is contingent upon the candidate's submission of the final copies of the dissertation to the Graduate College.

I hereby certify that I have read this dissertation prepared under my direction and recommend that it be accepted as fulfilling the dissertation requirement.

____ Date: 11/16/07
Dissertation Director: Heidi Harley

____ Date: 11/16/07
Dissertation Director: Michael Hammond

STATEMENT BY AUTHOR

This dissertation has been submitted in partial fulfillment of requirements for an advanced degree at the University of Arizona and is deposited in the University Library to be made available to borrowers under rules of the Library.

Brief quotations from this dissertation are allowable without special permission, provided that accurate acknowledgment of source is made. Requests for permission for extended quotation from or reproduction of this manuscript in whole or in part may be granted by the head of the major department or the Dean of the Graduate College when in his or her judgment the proposed use of the material is in the interests of scholarship. In all other instances, however, permission must be obtained from the author.

SIGNED: Jordan B. Brewer

ACKNOWLEDGMENTS

I would like to especially thank Mike Hammond, Heidi Harley, and Diane Ohala for serving on my dissertation committee and providing valuable feedback that greatly improved the quality of this work. I couldn't have asked for a better committee or hoped to get the opportunity to sit in on so many intriguing discussions in committee meetings. I would also like to thank Adam Ussishkin, Andy Wedel, and Natasha Warner for helping me more concretely formulate my ideas in the early stages of the dissertation.

Furthermore, I would like to thank my colleagues Sumayya Racy, Polly O'Rourke, Stacy Oberly, Jaime Parchment, Peter Richtsmeier, Dave Medieros, Adam Baker, Jason Ginsburg, and Benjamin Tucker, who each provided a listening ear, helpful suggestions, and refreshing end o' semester feedings. These linguists just may comprise the kindest, smartest, and most encouraging graduate class of all time.

Lastly I want to extend my most sincere thanks to my husband Jed Brewer. Without his constant support, and sacrifice (especially in the child-rearing domain) well beyond the call of marital duty, it is very likely that this dissertation would never have been completed.

DEDICATION

I dedicate this dissertation to my children, Gavrie Elliot Brewer, and the yet to be named little one. I can't exactly say that you made the dissertation *easier* to complete, but sometimes motivation takes different forms. It's going to be a lot of fun to see you two grow up. I love you very much.

TABLE OF CONTENTS

TABLE OF CONTENTS - Continued

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

A large domain of linguistic inquiry concerns the nature of words. It is widely thought that words are stored and represented in our minds in a structure termed the lexicon, in which every word has a 'lexical representation'. Researchers conduct experiments and examine intuitions about words to determine the content and structure of the lexicon. One interesting component in lexical representation, for literate speakers, is an orthographic representation for words. It has been traditionally assumed that while this orthographic information is available and useful in such tasks as visual word recognition (i.e. reading) or in writing, orthographic information about words is not necessarily involved in non-visual linguistic tasks, like auditory word perception, or speech production.

There has been some research however, which has challenged this notion of the isolation of orthographic information to visual processes. In a seminal study Seidenberg and Tanenhaus (1979) found an influence of orthography in an auditory rhyming judgment task. Subjects were faster to judge as rhyming those pairs which shared an orthographic representation of the rhyme than those who shared only phonology (i.e. *pie-tie* vs. *rye-tie*). Additional recent research has confirmed these effects of orthography in auditory perception tasks (Taft & Hambly, 1985; Halle, Chereau, & Sequi, 2000; Ziegler & Ferrand, 1998). Even more surprisingly, some experiments have suggested effects of orthography in speech production (Tanenhaus, Finigan & Seidenberg, 1980; Lupker, 1982; Wheeldon & Monsell, 1992; Damion & Bowers, 2003). These experiments all

show facilitated naming latencies for words which share orthographic characteristics with some prime environment. As such, these results can all be explained as effects of orthography on lexical access of words rather than affecting the production process *per se*.

In contrast, the experiments and analyses described in this dissertation show an un-ambiguous effect of orthography on speech production. Orthographic characteristics of word-final sounds, and words themselves are shown to influence the durations of spoken productions of those sounds, and whole words. These effects are robust to the mode of lexical access, whether through experimentally elicited reading aloud of words, or through the spontaneous generation of words in a modified sociolinguistic interview format.

Chapter 1

INTRODUCTION

1.0 Introduction

A large domain of linguistic inquiry concerns the nature of words. It is widely thought that words are stored and represented in our minds in an abstract structure termed the lexicon, in which every word has a 'lexical representation'. What is not so clear is just what the nature of a lexical representation is: what type of information about a word we store, and how those pieces of information relate to each other and to other cognitive processes when we process and produce language. In experimental psycholinguistics, there are two ways that researchers have investigated the content and structure of the lexicon. Some studies are conducted on word perception or comprehension and yet others are concerned with people's productions of words. It seems that there is an implicit assumption that factors producing effects in subjects' performance of various tasks are reflective of factors involved in the content of lexical representations or alternatively, the organization of the lexicon itself.

1.1 What's in a word?

A wide variety of experiments in word processing and production have provided some of the candidates for information contained in lexical representation, or involved in organizing the structure of the lexicon itself. As a result of these experiments, many potential elements of lexical representation have been proposed, including various aspects of meaning (semantic content) (McRae, de Sa, & Seidenberg, 1997), the sounds which make up the word (phonological form)(McClelland & Elman, 1986), as well as the frequency with which those sounds co-occur (phonotactic probability) (Vitevitch & Luce, 1998, 1999; Vitevitch, Luce, Pisoni & Auer, 1999; Luce & Large 2001), and how the word can be used in relation to other words (syntactic content) (Miller, 1991; McRae, Spivey-Knowlton, & Tanenhaus, 1998; Rayner, Warren, Juhasz, & Liversedge, 2004). Perception tasks in particular have revealed the influence of measures of probability, such as frequency of occurrence (Savin, 1963; Vitevitch, 2002a) or likelihood of occurring near a given other word (McRae, Spivey-Knowlton, & Tanenhaus, 1998; MacDonald, 1993; Narayanan & Jurafsky, 1998). Semantic aspects of individual words have also been shown to influence word perception. Whaley (1978) showed an effect of "richness of meaning" (calculated as the number of semantic associates of a word) in a lexical decision task. When subjects are asked to quickly determine whether a given stimulus is a word or not, subjects are faster to identify as words those having a greater number of semantically related words. Using the same experimental paradigm, James (1975) found that a word's concreteness influences response times, such that concrete words are responded to faster than abstract words (BONE vs. FATE).

The grammatical category of a word has also been shown to have an influence in word perception tasks (Bradley, 1978). Taft & Forster (1976) found different effects of open and closed class words in a word-interference lexical decision paradigm. This experimental paradigm is similar to that described above with respect to the effect of semantic characteristics, but involves an additional word-interference factor which allows a particular glimpse into conflicts in perception between classes of words. An additional influence includes the phonological neighborhood of a word, which is the number of additional words that are phonologically very similar to the word in question (Luce & Pisoni, 1998). This is typically calculated using a standard edit distance of one, which means that the phonological neighborhood of a word is equal to the number of words that can be created through the addition, subtraction, or changing of one sound. Lastly, for literate individuals, the orthographic representation of the word (Muneaux & Ziegler, 2004; Ziegler & Ferrand, 1998) is also shown to affect subject's performances in perception and production tasks. Those experiments in particular show that auditory word recognition is facilitated when the rhymes of words are orthographically represented in a more frequent and consistent way. A more thorough description of the methodology and results can be seen in section 1.3 below.

The above discussion of elements of lexical representation based in part on the role that they play in people's perceptions of words contains much information while being quite brief. These characteristics are not intended to be comprehensive, but rather illustrative. In an effort to clarify the overall picture of lexical representation briefly described above, the chart in Figure 1.1 illustrates how these various pieces of information might comprise a lexical representation for two example words.

	Orthographic form	Phonological form	Phonological neighborhood	Syntactic category	Semantic concreteness	Word frequency	Syntactic content
cucumber	cucumber	/kjukʌmbɚ/	no neighbors	Open	very concrete	infrequent	noun
An	an	/æn/	many neighbors	Closed	less concrete	very frequent	determiner

Figure 1.1 A diagram of some types of information hypothesized to constitute a lexical representation for two example words.

While many various characteristics of a word have been shown to affect subjects' perception and production of words, this dissertation is concerned in particular with exploring just the effects of a subset of those. In particular, the phonological and orthographic information contained in lexical representations and how those pieces of information relate to each other is a domain ripe for inquiry. The next few sections discuss the traditional and more novel conceptions of the role of orthography in lexical representation. Similar to the discussion above concerning a variety of aspects of lexical representation, the field's understanding of the role of this particular characteristic of lexical representation has been informed through an examination of the results of studies concerned with orthographic effects in word perception and word production tasks. Consequently, discussion of the pertinent studies follows.

1.2 The role of orthography in lexical representation

The role of orthography in lexical representation is intriguing in the disparity of research considering its role in perception/comprehension versus its role in production.

The role of orthographic representation in lexical access of words in our lexicon has been well studied, particularly with respect to the visual domain. Lexical decision is a common methodology employed to explore these effects (Forster, 1976 among others). As briefly described in section 1.1 above, lexical decision is a word recognition task whereby words and non-words are presented to a subject (visually or auditorily, or sometimes even in both domains) and the subject is required to respond as quickly as possible as to whether the stimuli is indeed a word, or not.

Taft (1979) succinctly characterizes lexical access as the matching of some sensory information to lexical information. He proposes that this is done through an access code. Children learn to talk well before they learn the orthographic representations of words. Consequently, their lexicon is necessarily accessible through some kind of phonological code well before they begin to spell or read words. As a result of this, it was initially thought that visual access of words must piggyback off of that previous system of phonological access (Coltheart 1980). For this to happen, orthographic visual stimuli would have to be phonologically encoded before initiating the process of lexical access. However, starting with Forster (1976) and followed by other researchers, experiments were performed which indicate that lexical items can be accessed directly through an orthographic code without mediation through phonology. Forster's serial search model (1976) suggested that the serial orthographic structure of a visually presented word is assigned a code which directs the search of the lexicon to the appropriate storage bin (although the bins are arguably phonologically based). Other models, such as the logogen and activation models simply propose that there are orthographic units (in addition to the phonological units proposed for auditory access) which can be fired or activated and which upon reaching the appropriate absolute or relative level of activation access the lexical entry of the word provided by the sensory stimulus. The particular details of the various models and their differences are not important for discussion here. What is relevant is that despite the proposed method of access stipulated by various models, they all allow for access of a word directly from the orthographic representation of a word when presented visually (i.e. reading).

It is therefore not surprising that orthographic information contained in our lexical entries is available to us in processing tasks directly related to the visual modality, such as reading or writing. What is more surprising is that some research has attributed a role to orthography in strictly auditory tasks in language perception, when it might be assumed that activation or availability of orthographic information is not relevant to the task at hand. The next section turns to a description of those studies.

1.3 Orthographic effects in speech perception

A number of studies have considered the possibility that orthographic information may indeed be available in the processing of auditory stimuli and thus influence the access of lexical items even in that domain. These experiments exploring the role of orthography in speech perception are methodologically diverse. In a seminal study Seidenberg and Tanenhaus (1979) found an influence of orthography in an auditory rhyming judgment task. Subjects were asked to determine whether pairs of words that varied in orthographic overlap rhymed. Subjects were faster to judge as rhyming those pairs that shared an orthographic representation of the rhyme (like *pie –tie*) than those that shared only phonology (like *rye-tie*). Taft and Hambly (1985) found an effect of orthography in an auditory syllable-monitoring task. Specifically, when the syllable the subjects monitored was spelled the same way as a syllable in a word, subjects were induced to say that the syllables matched, even when the phonology differed (monitoring for *lag* in *lagoon* [ləgun]).

In a phoneme monitoring task in Dutch, Dijkstra, Roelofs, and Fieuws (1995) found an effect of frequency of orthographic representation of a phoneme. They observed faster latencies to sounds when they were represented in the word with their most frequent orthographic representation. For instance, subjects were faster to identify a /k/ presented in the word *kabouter* "goblin" than /k/ in *cabaret*, where 'k' is a much more frequent representation than 'c' of the sound /k/. Halle, Chereau, and Segui (2000) also found an effect of orthography in a phoneme monitoring task. They utilize the phenomenon in French of voicing assimilation. Subjects were presented the word /apsyrd/ *absurde* in a gated task, that is, where subjects govern their own progress through the recorded sounds one at a time. Given a choice between /p/ and /b/, subjects accurately reported hearing /p/. This phoneme /p/ is what is actually produced in the surface form of the word. In contrast with that phonetic information however, the orthographic representation of that sound is the letter 'b', which usually corresponds to the voiced stop counterpart. When the whole word was presented (non-gated), subjects were influenced to report /b/, presumably as a result of the activation of the orthographic representation of whole word, which was not present in the gated task.

Ziegler and Ferrand (1998) observe a consistency effect of orthographic representation in an auditory lexical decision task which replicates their findings of the same in a visual domain. Subjects are faster to respond to words whose rhymes can only be spelled one way (consistent) than words whose rhymes can be represented in multiple ways (inconsistent). An example of this difference can be found in /ip/, which can be spelled '*eep*' or '*eap*' vs. /ʌk/, which can only be spelled '*uck*'. Additional experimental methodologies finding results of orthographic information in auditory tasks (where stimuli are presented solely in the auditory domain) include syllable segmentation (Morais, Content, Cary, Mehler, & Segui, 1989), phoneme blending (Ventura, Kolinsky, Brito-Mendes, & Morais, 2001), and shadowing (Slowiaczek, Soltano, Wieting, & Bishop, 2003). These experiments in speech perception are able to address the question of how orthographic information might influence our access of words in our lexicon from auditory input, complementing our understanding of how orthographic information can influence perception and word access from visual input.

1.4 Orthographic influence in speech production

A influence of orthography on production seems simple and intuitive, at least to non-linguists. In a popular current fast food commercial for chipotle-spiced chicken wraps, the writers play on this intuition. Two young men show up at a young woman's doorstep. Man 1 raps a few Spanish lyrics, and the girl responds with a quizzical expression in the direction of Man 2. Man 2 answers her implicit question with a response similar to "What, you asked for a rap with some spice!" The girl replies quite succinctly, "I said a [rːæp] not a [ræp]." with a noticeable durational difference between the two /r/'s.

Despite the availability of anecdotal evidence, in contrast to the substantial research conducted on the influence of orthography on speech perception and lexical access, there are relatively few studies that address the potential influence of orthography on speech production, or more specifically on the production of words. Each of these studies has approached the question from the perspective that orthographic information might affect response latencies[1] to elicited words. This perspective is not very different from that of many studies in the influence of orthographic information in speech perception. The hypothesis throughout is that orthographic overlap between stimuli in various tasks enables the speaker to start saying a word more quickly than when there is no orthographic overlap, or when the potential area of overlap is giving conflicting orthographic information.

In a modified Stroop test (Warren, 1974) conducted by Tanenhaus, Finigan, and Seidenberg (1980) subjects are asked to name the color of a printed word (target) that is presented after a brief visual prime. Note that like the traditional Stroop test, subjects are not asked to name the target word, but rather the color of that printed target. Prime-target

[1] Response latencies in this experimental domain can be considered the amount of time it takes from presentation of stimuli to the subject's onset of auditory word production.

pairs that were orthographically similar (*good-food*), phonologically similar (*rude-food*), or both phonologically and orthographically similar (*mood-food*) all produced significant color naming interference relative to unrelated prime-target pairs (*well-food*). Though certainly not directly addressing orthographic influences on production, the results of this study nonetheless indicate that orthographic information from a visual stimulus is activated when subjects are producing words (in this case, color names) that are not visually presented and can indirectly influence that production.

Two experiments have shown an effect of orthographic information on production in picture naming tasks. Lupker (1982) provided subjects with pictures that had words superimposed over the picture. The words were of four types relative to the picture; unrelated, orthographically and phonologically related, only orthographically related, and only phonologically related. Picture naming latencies were facilitated in each of the experimental conditions relative to the unrelated control. There was facilitation when the word shared only orthography with the picture name as well as when the word shared only phonology with the picture name. Additionally, there was greater facilitation when the superimposed word shared both orthography and phonology with the picture name than when only phonology was shared.

Wheeldon and Monsell (1992) conducted a priming experiment where subjects named a picture following the presentation of a definition of the picture. Significant repetition priming was observed relative to a control for which the definition was of a word unrelated to the picture. Then they presented as primes definitions for words that were homophones of the words represented by the picture and found that homographic homophone definitions (*pipe, pipe*) but not heterographic homophone definitions (*sale, sail*) produce substantial priming in the picture naming task. These two experiments show that orthographic representations of the priming words are active and can affect lexical retrieval of target words. Importantly in this second experiment, the orthographic representations of the priming words were never visually presented to the subject.

Finally, an experiment by Damion and Bowers (2003) utilizes a form-preparation paradigm where speakers memorize pairs of words and then are instructed to produce the second member of a pair given the first member as a stimulus. Stimuli are presented in blocks of items which either share a characteristic or crucially don't share the characteristic, in this experiment that characteristic being the orthographic representation of a word initial phone. This experimental paradigm assumes a partial phonological planning of responses so that in a homogenous block, the subject is faster to say the pairs because they've already preplanned the word initial sound(s). This study finds an inhibition in pair naming in the heterogenic block of pairs when the word initial sound was spelled in a different way than all the others (/k/ represented by 'k' or 'c'). For example, in a block consisting of *kettle-king, kangaroo-key, kitchen-cabin*, inhibition is observed for the last pair relative to the first three. This study echoes the speech perception study by Dijkstra et al. (1995) where subjects are faster to monitor for /k/ when it is spelled with its most frequent representation.

These four studies raise interesting questions. It seems to be generally assumed, and sometimes even explicitly stated, that orthographic information should have no role in constraining our speech production. Current models of speech production certainly don't provide an opportunity for orthographic influence on production (for instance Levelt, Roelofs, and Meyer 1999; Dell 1986). Damien and Bowers (2003) explicitly state what many surely consider reasonable, that "*orthographic information should be irrelevant to the process of speaking*", even if interactive models of the lexicon allow for mutual feedback in any process. Taft (2006), in light of his conclusion that orthography must shape abstract phonological representations of words, goes so far as to claim that "*The suggestion that something as unnatural as orthography can play a role in shaping phonological representation is antithetical to any linguistically-based accounts of phonology…"* Taft and Damien and Bowers in these papers acknowledge the seeming incongruity between the results of their research, and those results described above, which examine the potential role of orthography in influencing phonological representations and people's productions of words, and the traditional conception that orthographic information, while relevant in some domains, is not available to speech production processes.

1.4.1 Three possible explanations

The paper by Taft (2006) cited above presents interesting results from a pseudo-homophone judgment task performed by speakers of a non-rhotic English dialect, where subjects are asked to read to themselves a nonce-word and determine whether it has a real word homophone. The results show that subjects were less likely to respond positively to r-less pseudo-homophones, when the orthographic representation of their homophone did contain an 'r'. Specifically, when subjects were not allowed to say the prompt pseudo-homophone (eg. *cawn*), aloud, they had difficulty accurately reporting that it was homophonic with a real word spelled with an 'r' (eg. *corn*). Importantly, for a non-rhotic dialect, these items are supposedly truly surface homophones.

Taft argues that this indicates that even when the phonological output of a speaker's grammar doesn't reference certain orthographic content (the surface representation of 'corn' is /kɔːn/ without an /r/), their underlying representations of the words must still contain this information. If the underlying phonological representation didn't contain this abstract orthographic information, one could not account for the observed difficulty in homophony judgments. He argues more generally that orthography must shape underlying phonological representations, creating a more 'abstract' representation than that traditionally assumed by psychological models of print-to-sound conversion as well as phonological theories.

While a fundamental modification to the content of the phonological representation of words is certainly one hypothesis that could account for the results presented above, there are perhaps at least two other possible explanations. One alternative is that whatever the nature of the underlying phonological representation, orthographic information could enter the speech-production process at some derivative step between the selection of the phonological form of a word for production, and the application of phonetic rules (like assimilation for instance).

Yet another possibility is reflected through a denial of the locus of orthographic effects in lexical representation, assuming instead that orthographic information is appearing to have an impact in speech production as a function of its impact on lexical access, or lexical organization. To explore this option in greater detail, consider this alternative explanation to the results described above. In light of the fact that the speech production effects described above relate to naming latencies (that is how quickly a word can begin to be produced), an argument can be made that the effect of orthography is affecting lexical access.

To counter this possible explanation, far more compelling evidence for the influence of orthography on speech production would be an effect that cannot be explained apart from the actual production process, specifically, after the word has been selected for production. One indication that an effect is influencing production rather than access would be if the effect appears regardless of the domain of retrieval of lexical items (reading, auditory, picture-naming, spontaneous generation). Another indication would be if the effect occurred late in the production of the word, rather than as has been shown above in the naming latency of the word. One further indication would be for the orthographic effect to have its influence in multiple areas of the production of a word, rather than simply being evident in one area.

1.5 The current research

The research presented in this dissertation provides precisely the type of evidence that was shown to be lacking in the above discussion of previous research, which generated the alternative possible explanation. The results of two initial experiments considering possible effects of orthography on sound durations in word production conducted in a lab are presented to supply some background information and suggest areas of particular interest. Secondly, a third experiment is conducted and reported which continues to narrow the scope of interest. And lastly, there are a series of regression analyses performed on a large spoken corpus. The sum of this research shows effects of orthography on spoken word productions that cannot be explained solely as impacting the selection of words, but must refer to the impact of orthography on actual speech production. Specifically, the orthographic effects presented here are evident without respect for the access code of the lexical items, and different types of orthographic information in words influence durations of various aspects of words, including durations of word final sounds. The alternative explanation allowed by the ambiguous results of previous experiments is ruled out by the current research. The two remaining alternative explanations are discussed and illustrated in greater detail in the conclusion to this dissertation.

1.5.1 Dissertation outline

Chapter Two describes two experiments designed to investigate a possible influence of orthographic information on spoken word productions. In Experiment One, subjects read a prepared random list of words out loud. The words are matched such that obstruents in the word final codas can be spelled with varying orthographic representations (/f/ in *raft, staffed, graphed, laughed*). Some of these orthographic representations are spelled with one letter, others with two, and yet others with three. An analysis conducted on the sound durations of the target word-final coda obstruents show that there is a positive effect of the number of letters on sound durations. Sounds represented by more letters are produced with longer durations even though all the sounds are represented phonologically by a single phoneme (in the case above /f/). Under a traditional account of speech production, words are produced from their phonological form, and if these items have identical phonological forms at the point of interest, the effects would seem to contradict that account of speech production.

A tempting explanation for such an unexpected result might be that the materials were visually presented. One might consider that the observed effect is no more than a shallow visual processing effect in the conversion of print to speech. Experiment Two is designed to test that alternative hypothesis. In this experiment, subjects were visually presented with a random list of nonsense words conforming to English phonotactic patterns and were asked to read these novel words aloud. The words were paired such that word final obstruent sounds could be spelled with one, two or three letters, similar to the format in Experiment One. If the results in Experiment One can be solely attributed to shallow visual processing, then the effect should also hold for non-words in Experiment Two. However, an analysis of target sound durations in Experiment Two showed no significant effect of the number of letters on the duration of a sound.

The difference in effect of orthography between Experiments One and Two suggests that the observed orthographic effect is tied to lexical representations of words because the stimuli in Experiment One were real words, while the stimuli in Experiment Two were nonce-words. This question motivates Experiment Three which is described in Chapter Three. In this experiment, subjects were asked to read aloud a random list of monosyllabic nonsense words containing word final voiceless stops which can be represented by one or two letters (in effect a replication of Experiment Two). Then the subjects are trained on meanings for these novel words through association with a picture and definition. Finally, subjects are presented a picture naming task where they are shown a series of pictures corresponding to the newly learned words. At this point, the previously nonsense words have become like real words and have lexical entries. The two sets of identical materials can be compared across the two halves of the experiment to determine the effect of orthography on non-word production and word production of the same phonological and orthographic strings.

The results of an analysis on the non-word productions are consistent with the results in Experiment Two. There was no effect of number of letters representing a word final sound on that sound's duration. For words, there was also no local effect of number of letters spelling a sound on that sound's duration, contra Experiment One. However, there was an effect of the number of letters spelling a word on that word's duration, and a global effect of number of letters spelling a word final sound on the duration of the whole word. Although the array of word-effects in Experiment Three is slightly different than those found in Experiment One, they nevertheless support the claim that orthography influences speech production.

These same questions, and considerably more, could be asked of a much larger data-set of sounds, words, and sound and word durations. In fact, it will be shown that the experimental constraints on the nature of word and non-word stimuli evidenced in Experiments One through Three prohibit a satisfactory analysis of some interesting experimental questions. However, those same questions can be asked and answered of a less experimentally governed set of materials, provided it is of a sufficient size.

This very large data-set in the form of a subset of the Buckeye corpus is examined in Chapter Four. This chapter explores in far greater detail many specific effects of orthography on durations of sounds and words in a phonetically tagged corpus. Specifically, the result found in Experiment 1, that the orthographic length of word-final sounds contributes to the duration of that sound, are confirmed. Other orthographic characteristics of word final sounds are shown to affect sound durations, such as the frequency of representation of a sound, and the variability in possible representations for a given sound. Additionally it is shown that the orthographic length of words affects the duration of the word (independent of the phonological length of the word, and the orthographic length of a word final sound has its own independent contribution to word durations. Finally, the question of the interaction of lexical status and orthographic effect is explored in an analysis of function and content words, in the influence of frequency of words on their durations and on the effects of orthography on their durations, as well as the morphemic structure of words.

These experiments and corpus analyses show effects of orthography on speech production in a task reading words aloud, a picture naming task, and in spontaneous natural speech production. These varied methodologies and results provide compelling evidence that the effect of orthography is not located in a retrieval of lexical forms for production, nor simply in access of word forms. Rather, the effect must be explained as part of our production mechanism. An access explanation is not available for the results of this research because the effects of orthography occur late in the productions of words, as well as throughout the production of target words. The effects mentioned are present regardless of access medium which again strongly indicates that the locus of effect is in production rather than access. This surprising result leads to many questions concerning the nature of language production and how orthographic information might be available to and utilized by speakers as they produce language on an everyday basis. There remain two alternative hypotheses for how orthography might impact speech productions, whether through a modification to the underlying phonological representations or through a derivative step in the speech production process. Consequently, Chapter Five concludes with a discussion of those options, the necessary modifications to current models of speech production needed to account for these results, and an exploration of the myriad research questions raised by these effects.

Chapter 2

EXPERIMENTS ONE AND TWO

2.0 Introduction

This chapter presents the results of two initial experiments. They explore what potential effects orthography might have on production with a simple methodology of reading a list aloud. Experiment One elicits native speaker productions of English words in which certain sounds are represented by different orthographic strings (*rich, ditch*). An analysis of the durations of those sounds in the speakers' recorded production of each word then follows. As described in more detail in section 1.3 above, there is a difference between an effect of orthography on the amount of time it takes to begin saying a word and an effect on the duration of sounds within words. While an effect on naming latencies might be attributable to the lexical access of the word, an effect of orthographic form on speakers' durations of sounds within words would indicate that the effect is located in peoples' production processes rather than in their retrieval or access of words.

Experiment Two addresses the nature of the observed orthographic effect in production, particularly with respect to mental representations. The previous research cited in Chapter One and Experiment One in this chapter examine influences of orthography on word production. Crucially, the items in those experiments are real words. Real words have lexical representations. The second experiment in this study examines whether a similar orthographic effect on production can be found for non-words. Though a non-word visual stimulus will necessarily create some temporary type of mental representation in order to perform the grapheme to phoneme conversion which enables a speaker to say the stimulus aloud, they presumably do not have a stored lexical mental representation. Finally, a discussion of the implications of the results of these experiments follows along with what further avenues of research and alternative methodologies would be beneficial to further our understanding of orthographic influences on production.

2.1 Experiment One: Production of English Words

Previous research investigating potential effects of orthography on production has used experimental paradigms where pictures or visually-presented words overlap in phonology and orthography, orthography alone, or phonology alone (cf. section 1.2. Tanenhaus & Seidenberg, 1980; Lupker, 1982; Wheeldon & Monsell, 1992). Words were elicited from subjects and the relative speed in naming was interpreted as an effect of the various overlapping components of representation. Specifically, it is the consistent finding that naming latencies are even more greatly facilitated when words share orthographic and phonological form than when they share only phonological form which was taken to indicate an influence of orthography on production. Recall, in the form-preparation paradigm study by Damian & Bowers (2003), a finding of speeded pair naming latencies when the word initial consonant was spelled consistently with the remaining pairs in an experimental block was taken to indicate this influence of orthographic information on production.

Though the authors of these studies claim that the results indicate orthographic influence at the level of production, a critical examination of these methodologies shows that the effects could all be attributed to lexical retrieval of the words rather than their production. This is particularly the case because subjects' naming latencies are the dependent variable. An inhibition or facilitation of naming latencies could result from a faster or slower access of the word. The locus of this effect would then be occurring prior to sending the word off to a speech production process. As mentioned in Chapter One above, more convincing evidence of an effect of orthographic information on speech production *per se* would be an effect occurring in the middle or toward the end of the word, or as part of the actual production of the word, rather than just how quickly a word can begin to be produced. This evidence could not be explained as an effect on the access of words, but would instead have to be explained as a post access effect, either the influence of orthography on underlying representations themselves, or the influence of orthography on the phonetic encoding of a word during production.

A non- or pre-literate individual is familiar with the phonological forms of words which occur in people's spoken productions. For instance, they are at some level aware that the phonological form of *happy* is [hæpi]. When the individual becomes literate, they gain experience in connecting phonological and orthographic forms of words. It is at this point that they become aware that the orthographic form corresponding to [hæpi] is *happy*. Muneaux & Ziegler (2004) claim that familiarity with a word's orthographic representation can affect one's mental representation of the word. Specifically, the results of a neighbor-generation task (where subjects are asked to provide phonological neighbors to stimuli) reveal more productions of phonological neighbors with greater orthographic overlap. The authors claim that these results show that words with more consistent orthography/phonology mappings share stronger links to words with shared orthographic form than those with inconsistent representations.

Following this claim that orthographic representations have impacted people's mental representations of words, a hypothesis for Experiment One is that this difference in representation is reflected in our actual production of words. Specifically, Experiment One explores the effects of orthography on duration of sounds in English words. As will be explained in greater detail shortly, subjects were asked to produce real words of English which contained different orthographic representations of specific sounds. This experiment tests the simple question of whether phones that are represented by a sequence of multiple letters are produced with a different duration than those same phones represented by fewer letters. I hypothesize that speakers will produce sounds with longer duration when represented by multiple letters than when represented by fewer letters. Critically, in the design of the experiment, these sounds occur word finally, allowing for critical new evidence, the results of which could argue powerfully for a post-access account of orthographic effects.

2.1.1 Methods

Due to limitations in the nature of the English orthographic system and the content of its lexicon, it is essentially impossible to find items that would allow for experimentally controlling each of the many variables in words that may influence segment duration. For example, an ideal comparison in this experiment would be between a pair of words matched in frequency, phonemic length, orthographic length, morphological complexity, etc., but that differed only with respect to the orthographic length of the final sound. These and other variables are discussed in far greater detail later in this section, but for the purposes of this illustration I briefly consider the variable of frequency, which has been shown to effect word reduction (Whaley 1991; Johnson 2005; Jurafsky 2000). This reduction includes such aspects as reduced whole-word durations. Considering two words with like phonemic length and characteristics, the more frequent of the two words is consistently shorter in duration than the less frequent. See section 2.1.2 below for a more thorough description of the above studies. It is very difficult if not impossible to find two words in the English language which are homophones or minimal pairs which have similar frequencies and for which a coda consonant varies in the number of letters used in its orthographic representation (e.g. *click* and *clique* have different frequencies). This is just one of the confounds present in the structure of English which causes difficulty in constructing perfectly matched experimental stimuli.

Because English does not have the words necessary to control for each of the potentially confounding variables across the stimuli set, statistical control for these variables was chosen. Statistical control is a good alternative to experimental control of items when that experimental control is not possible, or not desired. Statistically controlling for various effects simply means using some mathematical model which indicates what effects on the dependent measure are attributable to the various independent variables, allowing the researcher to be confident that they are accurately and without bias assessing the effect of any particular variable of interest on the dependent. The statistical control chosen here results in an experimental design that can be analyzed with a multiple regression.

A multiple regression, also known as Ordinary Least Squares (OLS) regression, is a statistical tool used commonly in non-linguistic social science disciplines, but has not yet succeeded ANOVA as the statistical tool of choice for linguists. For those unfamiliar with regression as statistical modeling tool, I briefly mention what it does, and how to understand its output. While I have mentioned explicitly the reasons for using multiple regression for the particular experiment described here, below I address more generally on what types of data a multiple regression is an appropriate statistical tool. Parts of this discussion will be referred to in passing throughout the dissertation as appropriate. The reader already familiar with this statistical tool may skip the remainder of this section.

There are three main benefits to using a regression to analyze linguistic data. An OLS regression allows you to test the effect of either continuous or discrete independent variables on a continuous dependent variable. In order to test a continuous variable on a dependent variable with ANOVA you must first bin your independent variable into discrete chunks. The other option with ANOVA is to experimentally control for continuous variables, to either have all your stimuli of the same value, or choose two values (high and low) and test the difference between those means. Choosing either option makes your model lose accuracy in prediction. Linguistic data is rife with continuous independent variables. One common question is the effect that frequency (of words, sounds, etc.) has on linguistic processing. A regression statistic is optimal for testing continuous effects.

Another benefit to using regression is its ability to treat categorical variables as scalar. A regression can test the hypothesis that as you move incrementally up a scale of discrete units, you linearly affect the dependent variable. This is useful when theory indicates that there is an increasing or decreasing relationship of effect on levels of a variable. This is more specific than simply testing whether the means for each level are different, which is the test conducted by an ANOVA. In this series of experiments, number of letters and sounds contained in words and the number of letters used to represent word final sounds are all categorical variables treated as scalar by the models.

A third exciting benefit to using a regression model to examine linguistic data is its ability to test the mutual and independent effects of each independent variable on the dependent measure. The output of an OLS regression is an F-statistic for the whole model resulting in an R-squared which indicates the percentage of variation observed in the dependent variable which can be explained through the joint effects of the independent variables. In addition to this statistic, a regression also specifies the independent contributions of each variable to the variation in the dependent measure. This information is provided through a coefficient and a t-statistic. The coefficient for each variable conveys the size of the effect that a change in one unit of the variable has on the dependent variable. The t-statistic converts to a p-value which informs as to the probability of the effect of that variable being observed by chance, rather than because it actually affects the dependent variable. As a consequence of this quality of a regression statistic, when a variable is said to significantly affect a dependent measure, that can be read as saying “independent of the effects of every other variable in the model”.

Having briefly discussed the structure of English with respect to the experimental question at hand and the constraints that this places on experimental control of stimuli and statistical tools for analysis of future results, we are prepared to delve into a description of the materials used.

2.1.2 Materials

The stimuli used in this experiment were 53 words of English, matched such that a target consonant in the coda (p, t, k, f, s, z, ʧ) could be spelled in two or more ways. In some cases a multiple contrast in orthographic representation of a given consonant sound was found, for instance: *draft, graphed, laughed, staffed*. There is a wider contrast in orthographic representations for coda consonants than, for instance, for onset consonants. For this reason, sets of words with varying coda consonant representations were created, specifically seventeen sets containing two to five items each. This location in the word provides the added benefit of enabling us to analyze the results with respect to subjects' productions rather than access of words. For the reasons mentioned above, these sets did not contain minimal pairs or homophones, except in a few cases.

In each of these sets the target consonant sound was embedded in a very similar context: the preceding vowel was identical, the location of the phone in the coda (whether final or non-final) was identical, and the following consonant was identical in those sets whose target consonant was non-final (i.e. *rapt, rapped*). It was not possible to match the items in each set for number of letters used to represent the target consonant, frequency of the word, number of syllables, number of morphemes, or the length of word.

However, under a regression model, this is not a problem. The coda location of the target sounds (in contrast to word initial or medial) maximizes the variation in orthographic representation for the various sounds. The following variables shown in Figure 2.1 were measured for each word and were analyzed through a multiple regression analysis:

a.) number of CONSONANT-LETTERS in the target consonant representation,

b.) FREQUENCY of the word (as found in the Brown corpus),

c.) number of MORPHEMES in the word,

d.) total WORD-LETTERS of the word measured by letters,

e.) number of SYLLABLES in the word,

f.) the location of the consonant within the coda (whether FINAL or the first member of a cluster), and additionally,

g.) the place and manner of articulation of the consonant sound in question.

Figure 2.1 Variables analyzed with respect to sound durations in Experiment One

I now review each of these variables.

CONSONANT-LETTERS. The variable of most interest in this paper is the number of CONSONANT-LETTERS used to represent a particular sound. As such, the number of letters used in the representation of the target sound in each word was used as a factor in the multiple regression. For instance, one letter is used for /t/ in the word *hit*, two letters are used to represent the sound /k/ in a word like *back*, and three letters are used to represent the sound /ʧ/ in words like *ditch* or the sound /t/ in *height*.[2] As a result, this variable had three values. The mean length was 1.47 letters. See Table 2.1 for summary statistics.

FREQUENCY. In addition to the number of letters used in the representation of the target sound, frequency of the word is another likely influence on segment duration. Whalen (1991) showed that infrequent words have longer durations than frequent words, though his results varied depending on many additional factors, and his study considered only durations of whole words and not parts of words or individual segments. In addition, Jurafsky (2000) found that higher probability words experienced reduction in many domains. A shorter duration (relative to some standard for that word) for the word as a whole, or a part within the word, could be considered a type of reduction. The materials in this experiment ranged in frequency from zero (not in the corpus) to 8771.

WORD-LETTERS, MORPHEMES, SYLLABLES. Additionally, the length of the word (*fuss* vs. *us*) as measured by the number of LETTERS in the word, the number of MORPHEMES it contains (*rapt* vs. *rapped*), and the number of SYLLABLES in the word (*Monarch* vs. *mark*), may influence speakers' durations of sounds. If speech is governed by a timing mechanism of sorts, then it could be that the longer a word is in terms of the above variables, the shorter each of the individual sounds will be in duration. The items in Experiment One had lengths of two to seven letters, one or two syllables, and one or two morphemes.

---

[2] Historically, it is certainly the case that *ght* sequences were consonant clusters. An informal pilot experiment in which subjects were asked to indicate "which letters make up the vowel sound and which make up the consonant sound" showed that subjects synchronically interpret *gh* as contributing to the consonant rather than the vowel.

PLACE/MANNER OF ARTICULATION. Because varying sounds have inherent durational differences (Crystal & House 1988), it is necessary to also account for the effects of the manner and place of articulation of the consonant in determining the effect of the number of letters used to represent the sound. In this experiment, duration measurements for stops end at the end of closure and don't include any of the burst or release. Duration measurements for fricatives include all of the frication, and affricate durations include the stop closure as well as the frication upon release. The limited measurement for stops relative to fricatives and affricates would lead us to expect that fricative and affricate durations would be longer than those of stops. Bivalent place of articulation variables (coded as 0 and 1) were also included: LABIAL, ALVEOLAR, and VELAR.

FINALITY. Some researchers have found differences in sound durations contingent upon their location in the word. Van Sohn & Pots (1999) found that consonants reduce acoustically in word-medial environments in terms of sound energy and duration. Crystal and House (1988) showed that consonants occurring first in a word final consonant cluster (VCs#) have shorter durations than when the consonant stands alone word finally (C#). For these reasons, the location of the consonant within the coda, (whether final or no) is taken into account. See Appendix A for a complete list of stimuli. Summary statistics for the 53 items in Experiment One with respect to each of the variables just mentioned follow in Table 2.1:

Summary statistics of Experiment One variables					
Variable	**Observations**	**Mean**	**Std. Dev.**	**Min**	**Max**
Durations	362	0.151154	0.070946	0.009412	0.385152
CONS LETTERS	355	1.470423	0.611656	1	3
FREQUENCY	355	386.7775	1345.417	0	8771
WORD LETTERS	355	4.684507	1.313408	2	7
MORPHEMES	355	1.132394	0.339397	1	2
SYLLABLES	355	1.15493	0.413333	1	3
FRICATIVE	356	0.455056	0.498677	0	1
FINALITY	356	0.808989	0.393651	0	1
ALVEOLAR	321	0.411215	0.492822	0	1
VELAR	321	0.367601	0.482905	0	1
LABIAL	321	0.221184	0.415692	0	1

Table 2.1: Characteristics of each variable measured for the English words in Experiment One. Each of these variables is used in the multiple regression for this experiment.

2.1.3 Participants

The participants in this experiment were 7 members of the Tucson community between the ages of 23 and 30, all with undergraduate degrees in various fields, and all native speakers of English. Their participation in the experiment was voluntary.

2.1.4 Procedure

The stimuli were pseudo-randomized and presented in list form on a single sheet of paper; the first and last stimuli served as fillers and were not analyzed. Members of a set were separated by at least one other item. Following Warner et al (2005), a list format was chosen in which the items are transparent to the subjects. This means that an attentive subject would notice that words with identical phonological rhymes were occurring throughout the list. A list format was nonetheless chosen because sub-phonemic durational effects reported in the incomplete neutralization literature were of such small magnitude that they could be obscured with the introduction of unnecessary noise, as might happen in sentential contexts. Prior to recording, subjects were given an opportunity to read over the list to ensure familiarity with the words. They were then instructed to read the words aloud into a hand-held (though not by the subject) microphone in a quiet room and their utterances were recorded directly to a computer at a sampling rate of 22,050 Hz.

2.1.5 Data Analysis

The speech was analyzed using Praat (Boersma & Weenink, 2005) and duration was measured for the target consonant in each word (Klatt, 1974; Hassan, 2003) using the spectrogram and waveform of each signal. Stops were in every case preceded by a vowel, and in most, but not all cases, word final. As such, stop-duration was measured from the offset of voicing as visible in the waveform to the beginning of the broadband burst (but not including the burst) indicative of release as visible in the spectrogram. Fricatives occurred most often, but not always word finally. Fricative duration was measured from the onset to the offset of frication. Because it is more difficult to accurately measure fricative duration word finally than when the fricative precedes another segment and because of potential measurement bias by the author, measurements were randomly checked and affirmed by an independent recorder. The one affricate examined in this study always occurred word finally. Affricate duration was measured from the offset of voicing to the offset of frication. In a few cases, duration measurements were unobtainable because the word final stop was not released. This only occurred for one speaker, and only for the consonant /t/.

2.1.6 Results

A summary of the distribution of duration times by CONSONANT-LETTERS is shown in Table 2.2 below. An increase in the number of letters corresponds to an increase in the mean duration of the consonant.

Mean Duration by Number of CONSONANT-LETTERS

# of Letters	Observations	Mean Duration
1	210	0.14236
2	123	0.160702
3	22	0.170529

Table 2.2: Mean durations for items in which the consonant is represented with one letter, two letters, and three letters. Mean durations increase as the number of letters used to represent the sound increase.

This table reveals a general tendency for an increase in mean durational time for the consonant as the number of letters in the orthographic representation of that sound increase. Figure 2.2 shows mean durational times broken down by consonant sounds and number of letters used to represent the sound.

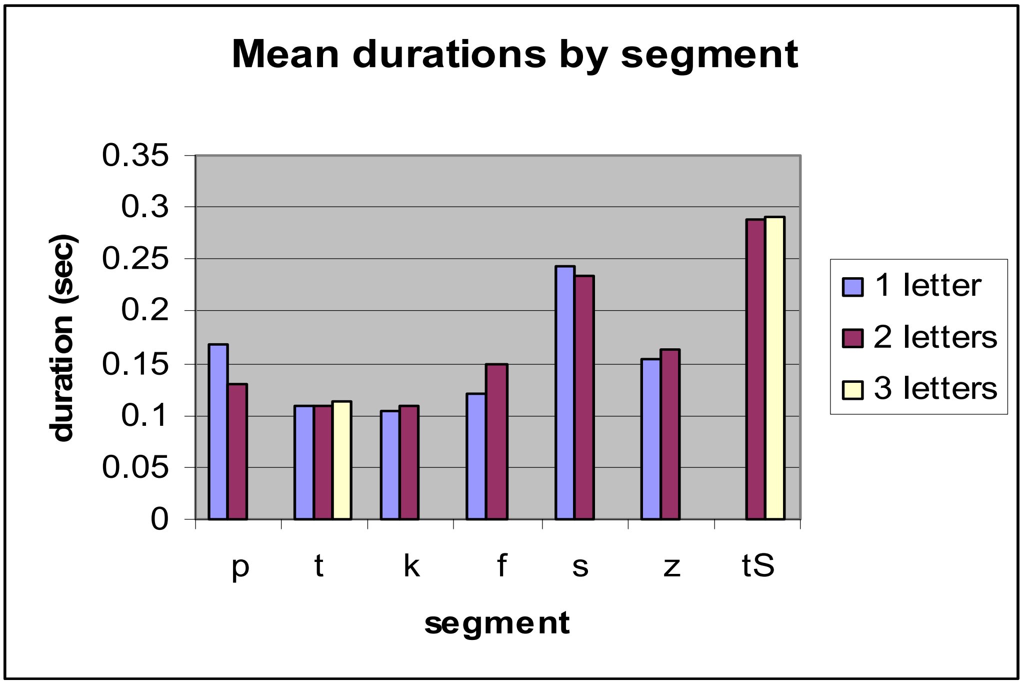

Figure 2.2: Mean durations for each segment in each letter condition. Only /t/ occurred in all three conditions. Most sounds exhibit an increase in duration corresponding to an increase in letters.

As mentioned above, the structure of the English orthographic system constrains the type of experiment that can be run. English does not have the words necessary to control for the potentially confounding variables across the stimulus set. This necessitates the use of statistical control to view the effects of those variables rather than experimental control. An Ordinary Least Squares (OLS) multiple regression with individual fixed effects tests the magnitude of the relationship between the variables CONS LETTERS (among other variables) and *duration*. Using individual fixed effects simply means that we expect individuals to behave differently because of reasons unrelated to the experiment or materials used. As long as the assumptions of OLS are satisfied[3] (Green, 2003), the coefficient β_1 will be an unbiased estimate of the relationship between letters and duration. I summarize the statistical analysis in Table 2.3 and then explain more thoroughly in the following prose.

Multiple Regression

OLS fixed-effects (within) regression
Group variable (i): individual

Variable	Coefficient	Std. Err.	t-stat	p-value
CONS LETTERS	0.017313	0.004668	3.710	0.000
WORD LETTERS	-0.00822	0.002869	-2.870	0.004
FINALITY	0.051065	0.010144	5.030	0.000
SYLLABLES	0.008017	0.007086	1.130	0.259
MORPHEMES	-0.00791	0.009274	-0.850	0.394
FREQUENCY	2.46E-07	1.71E-06	0.140	0.886
FRICATIVE	0.122807	0.005834	21.050	0.000
ALVEOLAR	0.043283	0.008308	5.210	0.000
VELAR	0.027249	0.010462	2.600	0.010
LABIAL	(dropped)			
_cons	0.040257	0.015398	2.610	0.009
R-sq:	Within	0.7152		
	Between	0.2599		
	Overall	0.6837		
F(9, 325)		90.69		
Prob>F		0.000		

Table 2.3: Results of the multiple regression model fit to the data, using the variables letters, frequency, length, morphemes, syllables, location in the coda (final or no), manner of articulation (fricative and affricate or stop), and place of articulation. Fit of the model is R-squared= .6837, and the p-value of letters is < 0.000.

Addressing first the fit of the model to the data, R-square = 0.6837 and F(9, 325) = 90.69; $p < 0.001$. An R^2 of 0.6837 indicates that the model is able to explain 68% of the variation in the variable *duration*. This model has captured much of the effects on

[3] The assumptions are linearity of effect of each variable, that each of the variables are uncorrelated with the error term, that each observation has the same variance (homoskedasticity), and the covariance between two items is equal to zero.

duration. Consider now the variable of interest—CONS LETTERS (the number of letters in the orthographic representation of the target consonant sound). The letters variable has a significant influence on the duration of the consonant sound ($p<.001$). More specifically, the model predicts that each additional letter in the representation of a particular consonant corresponds to a 17 millisecond increase (CONS LETTERS coefficient = 0.017313) in the speaker's production of that consonant. Across all word final coda consonants, the mean duration is .151 seconds (cf Table 2.2). 0.0173 of 0.151 = 11%. This means that speakers increase consonant duration by 11% for every additional letter used in the orthographic representation.

Table 2.2 also reveals an apparent non-linearity in letters. The millisecond difference between mean durations in the one and two letter conditions is greater than that between the two and three letter conditions. For this reason I also analyzed the data using an OLS estimate allowing for the CONS LETTERS variable to be interpreted non-linearly. This model is

$$D = \gamma_0 + \gamma_1 ln(\text{CONS LETTERS}) + \theta X + \eta$$

where *ln(CONS LETTERS)* is the natural logarithm of letters. This specification allows the variable CONS LETTERS to have a diminishing effect on duration as an additional letter is added. The CONS LETTERS variable is allowed to be non-linear, but the model is still linear. What this means is that the scale of the letters variable is adjusted logarithmically in order to analyze it with a linear model.

Adjusted Multiple Regression

OLS fixed-effects (within) regression Group variable (i): individual				
Variable	**Coefficient**	**Std. Err.**	**t-stat**	**p-value**
LN CONS LETTERS	0.028786	0.007532	3.820	0.000
WORD LETTERS	-0.00812	0.002825	-2.880	0.004
FREQUENCY	1.68E-07	1.71E-06	0.100	0.922
MORPHEMES	-0.01037	0.009349	-1.110	0.268
SYLLABLES	0.007765	0.007027	1.100	0.270
FINALITY	0.049358	0.010246	4.820	0.000
FRICATIVE	0.121921	0.005842	20.870	0.000
ALVEOLAR	0.044919	0.008332	5.390	0.000
VELAR	0.027963	0.010466	2.670	0.008
LABIAL	(dropped)			
_cons	0.060306	0.016595	3.630	0.000
R-sq:	Within	0.7159		
	Between	0.2454		
	Overall	0.6845		
F(9, 325)		91.01		
Prob>F		0.000		

Table 2.4: A similar regression as shown in Table 2.3, using a model that allows for non-linearity of the letters variable (that there is a greater difference in duration between 1 and 2 letters than between 2 and 3 letters). The R-square = .6845, the p-value for letters remains significant: p= .000.

At the mean duration for all consonant sounds (0.151 milliseconds), an addition of one letter corresponds to an increased duration of a statistically significant 28 milliseconds, ($p<.001$), holding other potentially confounding effects constant. A few noticeable aspects of this analysis should be pointed out. Though the statistical significance of the effect of LN CONS LETTERS on duration is similar to that shown in the linear letters model ($p < 0.001$ vs. $p < 0.001$) the meaningfulness of this significance, which is shown by the coefficient on LN CONS LETTERS, has increased in the model allowing for non-linearity of the CONS LETTERS variable ($c = 0.0173$ vs. $c = 0.0287$).

This coefficient term can be interpreted as the millisecond change in the dependent variable as you increase one unit in the independent variable.

What this indicates is that the increase from one to two letters in the representation of the sound has a greater effect on duration than the increase from two to three letters in the representation of the sound. This captures what we observe from the summary statistics of the effect of letters on duration shown in Table 2.2. Not only does the regression with a non-linear CONS LETTERS variable show an increase in the coefficient of CONS LETTERS, those variables whose effect in the first model was below the $\alpha = 0.05$ level remain statistically significant with this adjustment to letters.[4]

Only the effect of one variable (CONS LETTERS) on the duration of sounds has been addressed so far, but the above tables show effects of many other variables. A few observations are worth mentioning. Word-final sounds have longer durations than when they occur as the first member of a coda consonant cluster (FINALITY; $p < .001$). Considering the previous research by Jurafsky et al. (2000) and Whalen (1991), it is somewhat surprising to note that FREQUENCY of the word is not shown to have a significant effect on the duration of a coda consonant in the word ($p = .886$). However, neither of their studies considered specifically the duration of sounds within words. It is possible that the shorter durations of frequent words as found in Whalen (1991) is driven

[4] An alternative analysis of this apparent non-linear effect of letters on sound durations could be that the effect on sound durations is increased as the proportion of the orthographic length of the sound to the orthographic length of the word is increased. This alternative explanation should be tested, especially given results in future chapters suggesting an interaction of word and sound orthographic length on word and sound durations.

by a reduction in vowel or onset duration alone, rather than affecting coda consonant duration.[5]

The variables WORD LETTERS, MORPHEMES, and SYLLABLES are independent measures of size. An increase of one unit in total length of the word, as measured in total letters, resulted in a shorter duration for the consonant in question (WORD LETTERS; $p<.001$). However, this was the only size measurement which had a significant effect on the duration of the coda consonant.

As expected, the manner of articulation of the consonant affects the duration of that consonant. Fricatives have a longer duration than stops. The average difference in duration between fricatives or affricates and stops is 122 milliseconds (FRICATIVE; $p<.001$). Further, these data show that the place of articulation of the consonant affects the duration with which it is produced. Alveolars are produced with a longer duration than labials (ALVEOLAR; $p<.001$) and velars are also produced with a longer duration than labials (VELAR; $p<.010$). In a study of English stop consonant durations, Crystal and House (1988) find that across all environments, stop closure durations are slightly shorter for alveolars than labials and velars[6]. However, when you examine their data of word-final stops only, we see that /t/ has a longer duration than /k/, and they record no observations of /p/ word finally.

---

[5] Problems in data recovery prohibit a post-hoc analysis of the impacts of any of these variables on whole word durations.

[6] But see Zue (1976) and Luce & Charles-Luce (1985) for different results taken from more natural speech.

2.1.7 Discussion

Experiment One showed that there is a statistically significant effect of number of letters in the orthographic representation of a word on the duration of a consonant sound by a speaker. Furthermore, the coefficient on this variable has shown the size of the effect. According to the first model, one additional letter in the orthographic representation at the mean duration of 0.151 seconds causes a durational increase of 0.0173 seconds, which is a percentage increase of 11%. According to the second model, where the CONS LETTERS variable is adjusted logarithmically, an increase of one letter in orthographic representation at the mean duration causes an increase in duration of 0.0287 seconds, which is a percentage increase of 19%.

These results suggest that for literate speakers of English, production of words is a function of the orthographic representation of those words. Furthermore, differences in the number of letters used in the representation of a sound in the word result in systematic differences in speakers' pronunciations of those sounds. Previous research on orthographic influence on production tasks has showed facilitation in word elicitation and picture-naming when the orthographic form of the elicited word or picture matched the orthographic form of other words with which the subjects were presented. The current experiment has shown an effect of orthography within a single word and incidentally at the end of the word as opposed to the beginning, as shown in the previous experiments measuring naming latencies.

All these production experiments only address whether orthography influences speakers' productions of already-familiar words, for which speakers have mental representations. But what about when there is no long-term lexical representation of a given string as 'word'? Do differences in orthography influence production similarly when occurring in items without lexical entries? The discussion of Experiment 1 has so far assumed that the observed orthographic effect is localized with respect to the lexicon; orthographic information has become an aspect of our mental representations for words.

There is another potential explanation for the observed effect in Experiment One. Rather than tying this effect to the lexical entry or the lexical access of the spoken word, the entire effect could be driven by a shallower level of visual processing. What if the subjects are producing the consonant sound in question as they read the word on the page, as they might read sheet music? Just as a period following the note indicates to the musician to draw out that note, a second (or third) letter could be suggesting to the speaker to draw out that sound. In this case, the observed effect in Experiment One could be explained through some adjustments to a print-to-sound conversion model. However, if long-term mental representations are necessary to mediate or contain orthographic information which contributes to durational differences in production, then non-words, which arguably do not have mental representations, should show no effect of letter length on sound durations.

2.2 Experiment Two: Production of Non-words

It is relevant then to explore whether the observed orthographic effect occurs in non-words. Experiment Two in this chapter addresses just that question. This experiment tests the effects of orthography on non-words. It may be possible that individuals could construct a word-like mental representation for a word on their first encounter with a novel word, but this is certainly unlikely, especially given no information other than its orthographic representation. Specifically, Experiment Two is designed to determine whether sounds in non-words which are represented by a sequence of multiple letters are produced with a different duration than when those sounds are represented by fewer letters.

Unlike words, non-words have no lexical entry, and therefore the reading and production of the non-word should not access a particular lexical representation. The absence of an effect in Experiment Two would suggest that this orthographic information influencing production is associated with the lexicon and with our mental representations of words. A positive effect of orthography in this experiment, however, would not be as clear in its theoretical implications. An effect of orthography in an experiment using non-words could indicate that production differences are caused solely by the reader's visual processing system such that longer letter sequences require lengthened sounds.

This potential implication seems problematic for some previous research in speech perception, which has noted effects of orthography in word and sound perception when the testing involved no visual processing, but only auditory processing. Of course another potential explanation of a positive effect of orthography on duration of sounds in non-words could be that consistency in relation between duration and orthographic representation is the means by which orthographic information has come to be associated with people's mental representations of words.

2.2.1 Methods

Monosyllabic nonsense words were culled from Monononsense, a dictionary of monosyllabic nonsense items for English with phonotactic probabilities and neighborhood densities, which was created from Newdic[7]. Neighborhood was calculated with a standard edit distance of one, such that the total number of neighbors equals the number of real English words that can be created by adding, dropping, or replacing one phone. Phonotactic probability was calculated as the product of the probabilities of the onset occurring as an onset in real English words, and the rhyme occurring as a rhyme in real English words.

2.2.2 Materials

The non-word items for Experiment Two were paired such that a target consonant in the coda (p, t, k, f, s, ʧ) could be spelled in two ways (e.g. /geʧ/ *gech, getch)*. While Experiment One used statistical control for relevant variables due to constraints in the orthographic system of English coupled with actual lexical entries as mentioned above, this non-word experiment was free to utilize experimental control of relevant variables. The experimental control of the variables allowed using non-words allows for an ANOVA statistical analysis of the data. Non-words were chosen whose coda consonant could be represented with one or two letters (for instance /blɛk/: *blek, bleck*), one or three letters (/twɑɪt/: *twite, twight*), or two or three letters (/stæʧ/: *stach, statch*). All the non-

[7] Available from the SPAM lab at the University of Arizona.

words had the same phonotactic probability (calculated as the frequency of the onset multiplied by the frequency of the rhyme) to six decimal places (all items had probability of .000001 or .000002), and had a neighborhood density of either eight or nine. Simply put, the items in pairs are identical, except for their orthographic representation. All of the items are matched on potentially important variables like phonotactic probability, neighborhood, and syllabicity. These stimuli are also necessarily matched on frequency since, as non-words, they have a frequency of 0. A total of 16 pairs of words in three letter representation sets, for a total of 32 items, were used in this experiment.

Because there is some systematicity in the English spelling system, not every target obstruent can be represented by each number of letters. For instance, the affricate /ʧ/ can not be represented by one letter in coda position. Of course co-articulation results in some dialects' pronunciation of words like *tune* with an initial /ʧ/, thereby spelled with one letter. Furthermore, none of /p/, /k/, /f/, or /s/ can be represented by three letters. Because of this constraint, the different sets do not contain words with the same target sounds. The one vs. two letter pairs fully complete a two factorial design, in which the stops /p/, /t/, /k/, and fricatives /f/ and /s/ are represented by one and two letters each. The two vs. three letter pairs only consist of the fricative /ʧ/, represented by two and three letters. The one vs. three letter pairs only consist of the stop /t/, represented by one and three letters. Table 2.5 contains a complete list of materials.

Experiment Two Materials

Item	sound	Phonotactic probability	Neighborhood density	Spelling1	spelling2
1v2					
/strop/	p	0.000002	9	strope	stroapp
/snʌp/	p	0.000001	8	snup	snupp
/blɛk/	k	0.000001	9	Blek	bleck
/skok/	k	0.000002	8	skoak	skoack
/stuf/	f	0.000002	8	stoof	stooph
/gɛf/	f	0.000001	9	Gef	geff
/flɑʊs/	s	0.000002	8	flous	flouss
/hɑs/	s	0.000001	9	haws	hoss
/swʌt/	t	0.000001	8	swut	swutt
/krʌt/	t	0.000001	8	Krut	krutt
1v3					
/snit/	t	0.000001	9	sneet	sneeght
/drut/	t	0.000001	8	drute	drught
/twɑɪt/	t	0.000001	9	twite	twight
2v3					
/gɛʧ/	ʧ	0.000001	8	gech	getch
/nɑʊʧ/	ʧ	0.000001	8	nouch	nowtch
/stæʧ/	ʧ	0.000002	9	stach	statch

Table 2.5: The non-word materials used in Experiment Two. Three sets of sounds are shown that can be represented by 1 or 2 letters, 2 or 3 letters, or 1 or 3 letters, followed by their phonotactic probability, neighborhood density, and orthographic representations.

2.2.3 Participants

Thirty-two native English speaking undergraduates at the University of Arizona participated in this experiment for course credit. The non-word nature of the stimuli resulted in differences among subjects in accuracy of production. Data from seventeen of these subjects was excluded due to a success rate of less than 75% on production of the non-word tokens. Successful production of a non-word was defined as one which contained the string of phonemes it was intended to elicit, and was determined auditorily by the author. The remaining 15 subjects' data are analyzed below.

2.2.4 Procedure

The items for Experiment Two were pseudo-randomized and presented to the subject in list format, as in Experiment One. No member of a pair was presented directly after the other. All recordings were made in a sound-attenuated booth in the Douglass Phonetics Lab at the University of Arizona. Speakers read the lists aloud into a stand microphone, and their speech was recorded on an Alesis digital recorder at a sampling rate of 44,100 Hz. The speech was then analyzed and duration measurements were taken using Praat (Boersma & Weerink, 2005). The criteria for measurement of each sound type were the same as for Experiment One.

2.2.5 Results

The results of Experiment Two show that speakers' durations of sounds in non-words are not influenced by the number of letters used to represent the sound. Duration measurements were averaged across items for the by-subjects analyses and across speakers for the by-items analyses. Average durations appear in Table 2.6.

Table 2.6: Mean durations by Sound Type and Representation

Manner	Number of Letters	Observations	Mean
Stop	1	128	0.117111
	2	76	0.126576
	3	29	0.119874
Fricative	1	48	0.246113
Fricative/Affricate	2	90	0.280437
Affricate	3	37	0.323389

Table 2.6: Mean durations for each type of sound observed. Duration measurements were recorded for stops represented by one, two and three letters, and for fricatives represented by one, two and three letters.

In the one vs. two letter group, an ANOVA was performed with the independent variables CONS LETTERS (three levels; 1, 2, and 3 letters) and MANNER OF ARTICULATION (two levels, stops and fricatives/affricates), with subjects (F1) or items (F2) as repeated measures. Number of Letters used to represent the target sound does not significantly affect durations of that sound ($F1(1, 14) = 2.525$, $p > .1$, $F2(1, 8) = 3.189$, $p > .1$). However, the manner of articulation does significantly affect speakers' durations ($F1(1,14) = 154.784$, $p < .001$), specifically, fricatives and affricates are produced with a longer duration than stops, as is expected. There was no interaction between the two independent variables.

In the two vs. three letter group, all the items contained the affricate /ʧ/. An ANOVA was performed on this subset of the data with Number of Letters as the independent variable, and subjects (F1) or items (F2) as repeated measures. Number of Letters does not significantly affect the duration of sounds in this group either ($F < 1$ for both F1 and F2). In the one vs. three letter condition, all the items contained the stop /t/. An ANOVA was performed on this subset of the data with Number of Letters as the independent variable, and subjects (F1) or items (F2) as repeated measures. Number of Letters significantly affects duration in the by-subjects analysis ($F1(1,14) = 7.249$, $p < .05$), but not in the by-items analysis ($F2(1,2) = 2.156$, $p > .2$).

Though Table 2.6 appears to show a large increase in the fricative/affricate condition that corresponds to the number of letters used in their representations, it is not appropriate to analyze all these items together, as the one letter items contained only the fricatives /s/ and /f/, and the three letter items contained only the affricate /ʧ/, while the two letter condition contained both. An inherently longer duration for affricates would thus inappropriately sway the analysis.

2.2.6 Discussion

In sum, the analysis here does not show that the number of letters used in the representation of sounds in non-words affects speakers' productions of those sounds in non-words. In only one of the three sets of data tested did any effect of CONS LETTERS reach significance, and this one only did so in the by-subjects analysis. In this way, the results from Experiment Two contrast with those in Experiment One, which showed that for real words, the number of letters used in the representation of sounds in the word *does* affect speakers' productions of those sounds. Because the methodology and design of the two experiments are so similar, the most obvious indication from the results from these two experiments then would seem to be that the effect of orthography on speakers' productions is attached to mental representations of words. Because speakers have no lexical entries for their first encounter with these non-words, we see no effect of orthography on their productions. Logically, the next step would be to build a study that compares the durations of sounds in words and non-words in a single study.

This challenges the notion that orthographic information is separate and distinct from whatever aspect of underlying representations influences people's productions of words. One difference between words and non-words is that speakers have no long-term mental representation for non-words, whereas they do for words. There is the possibility that literate speakers could create lexical representations for new words online, even as they first encounter the item, but this is unlikely given that no information about the word was provided except for its orthographic representation. Both words and non-words in the above experiments have orthographic representations, as they were presented visually to the subject. If effects of orthography on word production occur for words, but not non-words, then one must concede that the orthographic information that is contributing to this production difference is mediated by a lexical mental representation. However, the fact that there was no effect of orthography on segment duration in the non-word production task (Experiment Two) brings up the question of how orthographic information has come to influence our mental representations of words, as shown in Experiment One.

2.3 General discussion

This chapter addressed the effects of orthographic information on the production of consonant sounds in words and non-words. The existing claims in the literature on orthographic influences on production relate to naming latencies; the productions of words with familiar orthographic representations are initiated more quickly than words with less familiar orthographic representations. The first experiment described in this chapter contributes to our understanding of these effects by showing that orthography can influence durations of sounds within words, rather than given a certain prime environment, the amount of time it takes to name a word. More specifically, these experiments show a direct rather than indirect influence of orthography on production.

The specific orthographic influence addressed in this study is the effect that the number of letters used in the orthographic representation of sounds has on speakers' durations of those sounds. Experiment One examined real words and utilized a multiple regression to test the effect of the number of letters used in the representation of a sound on the duration of that sound. There was a statistically significant increase in duration corresponding to an increase in letters. A robustness check was performed, allowing for a non-linear effect in the letters variable, which resulted in similar statistical significance, but an even greater coefficient. In addition to the effect of letters on sound durations, the regression model also tested potential effects of other variables that previous literature has suggested impact people's productions. Accounting for these variables allows our interpretation of the effects of orthography on duration to be unbiased.

Experiment Two turned to non-words to address the possibility that the orthographic effect noted in Experiment One was a result of visual processing of orthography rather than to differences in the mental representation of words in our minds. This experiment determined that similar orthographic effects are not found in non-word productions, another question which had not yet been addressed by other researchers.

An ANOVA analysis of the non-word item responses in Experiment Two showed no significant effect of the number of letters used in the representation of a sound on speakers' durations of those sounds.

One hypothesis that was put forward for the reason words and non-words behave differently with respect to their influence by orthography is that the orthography is associated with speakers' mental representations. Speakers have mental representations for words, but not for non-words. If the orthographic information resulting in the observed influence is mediated by a speaker's lexicon, or mental representations of words, then non-words' lack of lexical entries, or mental representations, would account for the lack of observed orthographic influence on production. A more thorough understanding of this possibility, and an explanation of how the orthographic information comes to be located with respect to the lexicon must wait for future studies.

Experiment Three in the next chapter tests this hypothesis that orthographic effects on production are mediated by lexical representations. The stimuli in Experiments One and Two were necessarily different, being words and non-words respectively. It is possible that the differences observed in orthographic effects on duration are not only resultant of their differences in lexical representation, but also on particular characteristics of different phonological strings themselves. In Experiment Three identical phonological strings are elicited from subjects in two conditions: at first presentation, when they are certainly non-words, and then again after lexical training when lexical representations for those novel words have arguably been created.

In Experiments One and Two, the subjects were visually presented with the words and asked to read the words aloud into a microphone. It could be the case that this presentation of the material and directions called subjects' attention to orthographic representation in a way that is not necessarily consistent with their default organization and use of language. If this is the case, then we would expect this effect to diminish or disappear when these words are elicited in an auditory domain or through a picture naming task whereby their orthographic representations are not attended to, or when they are spontaneously generated in natural speech, that is to say, when they are not elicited at all. Experiment Three, described in Chapter Three, and the corpus study found in the following Chapter Four are designed to address these concerns. In Experiment Three, the phonological strings in the word condition (after training) are elicited through a picture naming task (similar to the method used by Creel, Tanenhaus & Aslin (2006), cf. section 3.1). In the corpus study, all of the words analyzed were spontaneously produced by speakers participating in a modified sociolinguistic interview. In both these cases, the orthographic representations were not explicitly presented to the speaker. Seeing the effects of orthography in these domains would indicate that this orthographic information exerts an influence in production processes apart from visual processing and the concomitant lexical access.

Chapter 3

NOVEL WORD TRAINING EXPERIMENT

3.0 Introduction

The preliminary experiments described in the previous chapter have shown that the number of letters used in the representation of certain sounds influences the duration of the phonetic realization of that sound within a word. However, that effect did not hold for non-words, suggesting that the effect is tied to orthographic information as part of a lexical representation, rather than a purely an online visual effect triggered by the reading process. This chapter presents a third experiment which is designed to more fully examine this hypothesis.

The results reported in Chapter Two were taken from data which was elicited from subjects through a visual presentation of orthographic representations for words. It may have been the case that this visual processing drew subjects' attention to the orthographic representation of words in a way which is not directly reflective of speakers' natural production processes. The present experiment includes a training session on novel words, followed by a picture-naming task to elicit productions of these items for which a lexical representation may have been made during training. This picture naming task takes out the overt orthographic influence that the task of reading words aloud may have introduced.

The effects of orthography examined in the experiments described in Chapter Two were very specific. Those experiments tested the influence orthographic length of sounds (represented by one, two, or three letters) on the durations of those sounds. This same question can be addressed in the present experiment described below. The positive results confirming this hypothesis in Experiment One leads to the question of whether there may be additional orthographic effects on the duration of a larger unit of speech, the word. If adding a letter to the representation of a word-final consonant increases the duration of that consonant, then it might follow that adding a letter to the representation of a word (holding the phonological length constant) increases the duration of that word. In this experiment, additional measurements of the orthographic length of words and the durations of words allow for an examination of this second broader hypothesis concerning orthographic effects on speech production.

3.1 Experiment Three

In this experiment subjects are visually presented with novel phonological strings, represented with a particular orthography. Then they are trained on the meanings of those new words through picture and definition association. Finally, the novel words are elicited through a picture naming task. If the effect of orthography on durations of speakers' utterances is dependent on a lexical representation of a word, and if the experimental design has been successful in creating new lexical representations for the stimuli, then we might expect an effect to emerge after training.

In examining the role of lexical stress in representations and the use of knowledge of lexical stress in access, Creel, Tanenhaus, & Aslin (2006) test subjects on their acquisition of an artificial lexicon containing 48 words. Part B of this Experiment will follow much of the novel word learning methodology that was effectively used in that experiment. There is one critical difference between the lexical acquisition tested in that experiment and this. The novel lexicon being acquired by the subjects was not presented as novel words in their own native language, but rather as novel words in a novel language. In this way, most if not all adult lexical acquisition studies are similar. Researchers test acquisition of words in a second language, or in an entirely new language, but only in child language acquisition studies are researchers training subjects on words that belong to the subjects' native language. This is one aspect of the current experimental design which differs from other designs and the success of which remains to be evaluated.

3.1.1 Participants

A total of 65 students at the University of Arizona participated in this experiment to fulfill partial course credit. The productions of the 51 native English speaking subjects were analyzed.

3.1.2 Materials

The materials consisted of a set of non-words, taken from Mononononsense (see footnote 7 in Chapter 2), in which a word-final voiceless stop (/p/, /t/, /k/) can be spelled in two ways, and a set of pictures to be randomly paired with the novel words in Part B of the experiment. The word-final consonants can be categorized as spelled with one or two letters (for instance, a word ending in [t] will be spelled in one condition with a *t* and in another with a *tt*). These non-words are organized into four counterbalanced lists, such that any one subject receives each non-word only once, spelled either with more or with fewer letters. Half of any subject's non-word stimuli will include word-final consonants that are spelled with one letter, and half with two letters. See Table 3.1 for a list of the non-word stimuli used in the four counterbalanced lists. The phonological strings are controlled for neighborhood density (calculated with a standard edit distance of one) and phonotactic probability (calculated as the probability of the onset times the probability of the rime)[8]. Each of the phonological strings has a neighborhood density of 14 or 15. Their phonotactic probabilities are identical to 5 decimal places and range from .0000006496 to .0000062630. These two factors are controlled because they have been shown to influence performance on lexical tasks in perception (Vitevitch et al. 1999, and others), so one could imagine their potential influence in production as well.

Phonological String	One letter Orthographic Representation	Two letter Orthographic Representation
/vɪp/	vip	vipp
/θæp/	thap	thapp
/slɛt/	slet	slett
/plʌt/	plut	plutt
/ʤɛk/	jek	jeck
/flʌk/	fluk	fluck

Table 3.1: Orthographic stimuli used in Experiment Three; each phonological string can be spelled with one or two letters representing the word final consonant.

[8] Both of these variables and calculations are described in more detail in the introduction to Experiment Two, found in Chapter Two.

The pictures and definitions used in Part B of the experiment were selected by the author and were also deemed by two other linguists to be plausible referents for novel English words. Because each of the novel words on which the subjects were to be trained were monosyllabic, pictures and definitions were selected which appeared simplex in composition and function. They were also selected with the intent that they be items which the subjects would not previously have encountered and consequently either not already have lexical representations for, or not have strong ties to existing lexical representations. The pictures and corresponding definitions can be found in Appendix B.

3.1.3 Procedure

In Part A of the experiment, each subject was visually presented with three words in which the word final voiceless obstruent is spelled with one letter, and three in which it is spelled with two letters. They receive each phonological string only once for a total of six non-words. The presentation order of the non-words is random. The subject is instructed through written instructions on the screen to read the word aloud. Their utterance is recorded and analyzed for duration of the word and of the word-final consonant.

The same subjects who participated in Part A of the experiment immediately followed with participation in Part B, a training session. Each non-word was paired with a picture of an unfamiliar object. Each picture had a corresponding definition describing the function of the object pictured. Each subject only saw one of the two spellings for the novel word paired with one picture and that picture's definition. The subjects were then trained on the association of the word, picture, and definition. The goal of this training was to help the subjects create lexical representations for the novel words.

The first part of the training consisted of a presentation on a computer screen of the picture together with its definition and the orthographic representation of the novel word. The subject was then asked to read the word aloud after having learned what the word meant. The training randomly cycled through the novel words and their associated pictures and definitions. After seeing the six novel words with their associated referents and reading them aloud once, the subject was tested for accuracy of acquisition in an alternative forced choice test. The subject was presented with a picture and matching definition, and two of the non-words that they had been learning. They were asked to select the word form which matched the picture and definition. Subjects moved to the experimental trial only when they reached 100% accuracy in identifying the novel words. If a subject didn't achieve 100% accuracy even after 10 forced choice trials, they were excused from the final testing phase.

The final testing phase consisted of an elicitation of the novel words through a picture-naming task. The novel pictures were presented to each subject in random order without their definitions. The subject was asked to say aloud the word which appropriately named the picture shown. Consistent with Experiment Two, their productions were recorded and analyzed for both the duration of the whole words and the variably-spelled word-final consonants.

3.1.4 Data Analysis

In any experiment where subjects apply phonological content to a novel orthographic string, not all subjects choose the same phonological representation. For instance, some subjects read the non-words in ways not intended, invoking a non-canonical orthography-to-phone mapping. One common mistake is the exchange of lax vowels for tense counterparts (i.e. pronouncing *jek* [ʤek] rather than [ʤɛk]). Pronunciations as intended of novel words were determined auditorily by the author and only correct productions of non-word durations were analyzed for this part of the experiment. Similarly in Part B, the experiment was designed to elicit particular novel words through a picture stimulus. Sometimes speakers would incorrectly name the picture with an entirely new non-word, one which they made up on their own (often incorrectly combining onsets and rimes from the training set words). Other times speakers would substitute a different newly learned word for the intended word. Yet other times, speakers would simply respond “I don’t know”. Each of these types of responses were discarded from analysis, and only novel word responses to picture stimuli that matched the intended pronunciation were analyzed in this part of the experiment.

Measurements of sound durations were made from the offset of voicing of the pre-consonantal vowel as evidenced in waveform and spectrogram to the end of the closure of the voiceless stop, also as evidenced in waveform and spectrogram and determined by the author. As happened in Experiment One, some subjects did not release the word final sounds, therefore no duration measurements could be made for these items. This happened on at least one word throughout the whole experiment for a total of

17 subjects and predominately on /t/-final words (/slɛt/- 15, /plʌt/- 10, /vɪp/- 2, /θæp/- 1).

Following the standards established in Chapter Two for Experiment One and Two duration measurements, word durations were measured from the onset of the frication or burst of the first segment in each phonological string to the end of the closure of the word-final voiceless obstruent. In each instance of a non-released word final consonant, the word duration measurement was terminated at the end of the voicing of the pre-consonantal vowel and was determined by waveform and spectrogram by the author. As for Experiments One and Two, these measurements were spot checked for accuracy by another phonetically trained linguist.

3.2 Results

3.2.1 Orthographic effects on sound durations

The criteria above result in a smaller subset of analyzable data than may have been expected from so large a subject pool. A total of 37 non-word elicitations were dropped from analysis due to incorrect pronunciations of the orthographic stimuli. A much greater number of trained-word elicitations were dropped (124) due to subjects' inability to learn the non-words on which they had been trained. The remaining word and sound durations are divided into two sets, non-word elicitations from Part A of the experiment, and picture naming elicitations from Part B of the experiment. A summary of sound durations for both sets of data is provided in Table 3.2.

Mean sound durations and N for each phonological string

	Nonwords (Part A)				Words (Part B)			
	N	1 letter	N	2 letters	N	1 letter	N	2 letters
/jɛk/	27	0.1343	23	0.1253	15	0.121	14	0.1309
/flʌk/	18	0.1152	21	0.1134	10	0.1295	14	0.1245
/vɪp/	28	0.1321	26	0.1474	23	0.1225	17	0.1281
/θæp/	25	0.1322	24	0.1341	16	0.1168	20	0.1264
/slɛt/	24	0.1275	18	0.1472	11	0.1303	12	0.1399
/plʌt/	17	0.1304	19	0.1158	10	0.1336	8	0.1355

Table 3.2: Shows mean durations and number of instances for each of the phonological strings in each part of the experiment. Part A is reading non-words aloud, and Part B is a picture naming elicitation of novel words for which there has been training.

It is illuminating to note that while the numbers of items for non-words in Part A of the experiment is reasonably high, that N decreases dramatically in Part B of the experiment where subjects were asked to name a picture, the meaning and name of which they had been trained on. While the means shown for non-words in Table 3.2 show no overall pattern, increasing from one to two letters in some cases while decreasing in others (see also Figure 3.1), a pattern does seem to emerge for the set of trained-words. The column of means for two-letters is generally consistently higher than the corresponding one-letter column (except for /flʌk/, see Figure 3.2).

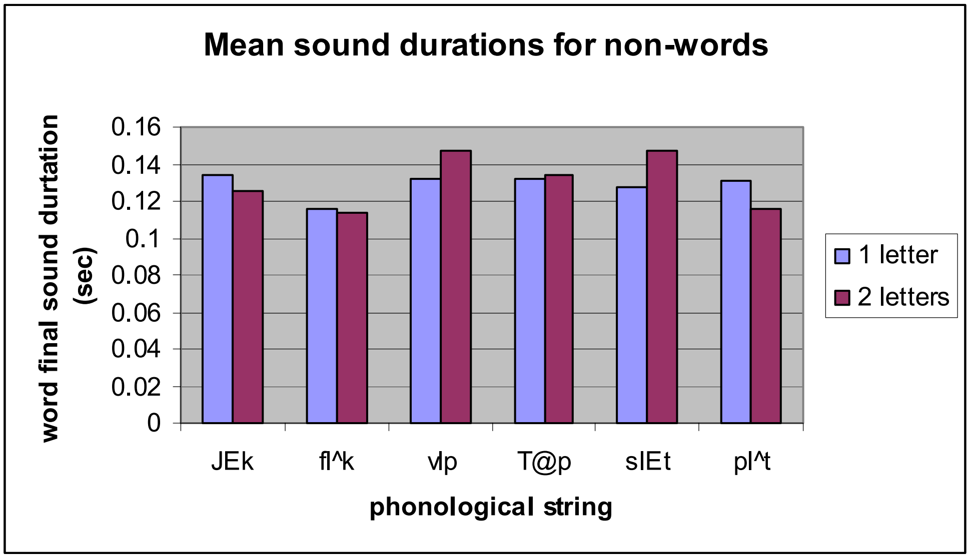

Figure 3.1: Mean word final sound durations for each of the non-word phonological strings in Part A of the experiment, sorted by number of letters used to represent the sound.

A difference in means test[9], commonly known as a t-test, shows that the differences in sound duration for the two representations of the word-final sound in each non-word is not significant ($p>.4$).

[9] The experimental control of items in this experiment indicates that this is a more appropriate statistical test than a regression for instance. See also the discussion of experimental vs. statistical control in Chapter Two.

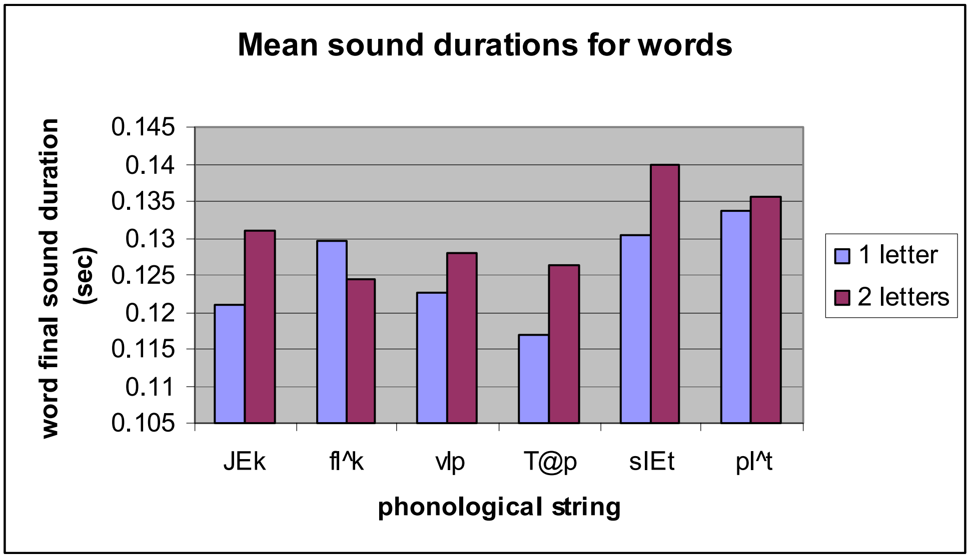

Figure 3.2: Mean word final sound durations for each of the trained phonological strings in Part B of the experiment, sorted by number of letters used to represent the sound.

In contrast with the bar chart in Figure 3.1, where no noticeable consistent overall pattern emerged, the graph in Figure 3.2 shows a consistently larger average duration in each of the two letter means for each word except /flʌk/. This fits the expectations generated by differences in Experiments One and Two described in Chapter Two. However, it is important to note that despite the suggestive differences between the two figures, a t-test performed on each word pair reveals that the difference in sound duration based on number of letters used to represent the word final sound is not significant ($p>.1$ for each analysis). These specific results do not mirror those found in Experiments One for real words and I'll return in section 3.3 to discuss why these counterintuitive null results may have been observed. But first, the discussion turns to an analysis of orthographic effects on whole word durations.

3.2.2 Orthographic effects on word durations

As mentioned in the introduction above, making measurements of word durations in addition to sound durations allows for an examination of a broader hypothesis concerning the effects of orthography on speech production. This more general hypothesis is that the number of letters (WORD LETTERS)used to spell a word contributes to the duration of that word independent of an effect of phonological length (WORD PHONES). Or more specifically, perhaps the number of letters used to spell a word final sound (CONSONANT LETTERS) can impact the overall duration of words, even if it isn't consistently shown to affect sound durations themselves.

The phonological strings in the experiment contained either three or four phonemes, and were orthographically represented with three, four, or five letters. Mean word durations for each of these combined factors along with the number of items the mean is averaged over is shown in Table 3.3.

Mean word durations for phonological and orthographic length						
	N	**3 letters**	**N**	**4 letters**	**N**	**5 letters**
3 phones	94	0.3448975	125	0.3749773	42	0.443522
4 phones	0		104	0.4194723	104	0.426098
Total	94	0.3448975	229	0.3951846	146	0.433612

Table 3.3: Mean word durations for level of letters used to spell the word by the two phonological lengths. There are no observations of three letter, four phoneme words because consistent with the structure of English, no one letter represents multiple phones.

The experimental stimuli are controlled across items in such a way as to avoid introducing a multitude of potentially duration-affecting factors. However, this same control results in the variables WORD PHONES, WORD LETTERS, and CONSONANT LETTERS being collinear. This collinearity means that in order to examine these specific variables, a regression model needs to be adopted. What is meant by this is that any two of those variables perfectly predicts the third variable. An example is the pair of /k/ final stimuli shown in Table 3.4 below.

Word	WORD PHONES	WORD LETTERS	CONS. LETTERS
jek	3	3	1
jeck	3	4	2
fluk	4	4	1
fluck	4	5	2

Table 3.4: An example of variable characteristics for two experimental items.

Within a word (holding WORD PHONES constant), the only way to increase the WORD LETTERS is to also add a letter to the representation of the word-final sound (CONSONANT LETTERS). Across words (holding WORD LETTERS constant; *jeck, fluk*) the only way to increase the WORD PHONES is to also add a letter to the representation of the word-final sound (CONSONANT LETTERS)[10].

This collinearity has ramifications for attempting to statistically model the independent effects of these variables on word durations. We are mathematically required to drop one of the three variables in our analysis. Having done that, in each possible permutation, we are left with observing effects of both remaining variables on word durations. Consider one permutation, where we find significant positive effects for both orthographic length ($p<.001$) and phonological length ($p<.05$) on word durations.

[10] This is true across all the stimuli with the exclusion of /θæp/, which has a single phonological segment onset represented with two letters.

The adjusted R-squared of such a model is .1444, indicating that these two variables alone account for more than 14 % of the variation in word durations in this experiment. Refer to Table 3.5 for a regression model showing the effects on word duration attributed to these variables.

Number of obs = 183		R-squared = 0.1538		
F(2, 180) = 16.36		Adj R-squared = 0.1444		
Prob > F = 0.0000				
variables	**Coefficient**	**Std. Error**	**t**	**P>|t|**
WORD LETTERS	.0377052	.0101919	3.70	0.000
WORD PHONES	.0301688	.0149657	2.02	0.045

Table 3.5: A regression model showing effects of WORD LETTERS and WORD PHONES on word durations.

These effects support the hypothesis generated from the results of Experiments One and Two, that orthographic characteristics may impact the duration of subjects productions of whole words. It must be noted though, that we are not observing the effect of WORD LETTERS directly. Instead that effect is subsumed through the combined effects of the other two variables. We don't know how much of the effect attributed by a model to WORD LETTERS is due to the WORD LETTERS, and how much is due to the number of letters spelling the word final sound (CONSONANT LETTERS). Similarly for the effect attributed by the model to WORD PHONES, we can not be certain how much of that effect is due to the WORD PHONES, and how much is due to the CONSONANT LETTERS. Despite the fact that the statistical model employed (and in fact any statistical analysis) cannot separate out what orthographic and phonological characteristics are contributing what amount to durational differences between words, it does convincingly support that orthographic characteristics are impacting subjects' durations of words.

3.3 Discussion

Though it is always wise to be cautious in interpreting a null-effect, there are a few aspects of this experiment which should be addressed. These unique characteristics of the experiment reassure us that despite the multi-faceted results of the experiment, the general hypothesis that orthographic characteristics impact subject's productions of words, but not non-words remains supported. These characteristics also importantly provide direction for future research including what the characteristics of a data-set likely to provide less ambiguous results might be.

First, in the analysis of orthographic effects on word durations, we see a significant positive effect of both WORD LETTERS and WORD PHONES on word durations. However, while we can be certain that an effect of orthographic characteristics on subject's productions of words is present, the experimental control of the stimuli renders a statistical analysis incapable of attributing independent effects on word duration to any of the variables WORD PHONES, WORD LETTERS, and CONSONANT LETTERS as a result of their collinearity. This is an unfortunate coincidence of this particular restrictions placed upon the stimuli set. However, a larger and less experimentally controlled data-set would enable an investigation into the independent effects of each of these variables on word durations which would nicely complement the analyses conducted so far on effects of orthography on sound durations.

The result showing no significant effect of orthographic representation on sound durations for non-words in Part A mentioned above is expected, based on the similar null-result found in Experiment Two described in Chapter Two above. However, the null-result for trained novel words in Part B of the experiment is somewhat unexpected. This lack of effect may stem from two independent characteristics of the experimental design. Recall that the N (number of items the mean is averaged across) for the sound durations of words in Part B of the experiment was quite small relative to the N for non-words in Part A as well as N for Experiments One and Two described in Chapter Two. A post hoc power analysis of this data reveals a power of less than .5 for my largest sample size word pair. Typically, a power of .8 is desired, indicating that there is only a 20% chance of failing to observe an actual effect. The power of the current experiment was such that an effect was unlikely to be observed.

Even though the qualitative analysis suggested an effect of CONSONANT LETTERS on consonant durations of words, given the number of subjects, this N for items is apparently unavoidable and is a consequence of the variation resulting from asking subjects to assign a phonological string to a novel orthographic representation as well as difficulties experienced by subjects in learning new English words. While children's language acquisition studies regularly require subjects to form new lexical representations for novel words presented in an experimental setting, lexical training in adult subjects focuses on teaching new words in second or novel languages. This experimental design is unique in that it asks native English speaking adults to create new lexical representations for English words. The subjects' difficulty in acquiring these new words despite the training contributed to the small number of items in this condition. A larger data-set, one in which the speakers are not required to learn novel words or assign phonological strings to previously unseen visual stimuli would completely avoid this issue of a small sample size, and consequently a lack of power.

A second possible reason for the null result in the analysis of orthographic effects on sound durations involves the status of the elicited words in the picture naming task in Part B of the experiment. While subjects were (with varying success) able to name pictures with the novel words on which they had been trained, a question remains as to how similar these newly created words are to previously existing words in the subjects' lexicons. As a result of their newness and the subjects' relative unfamiliarity with the words, are they somehow less word-like than other words, or do they actually have identical status with other nominal lexical entries? The subjects never used the words in sentences, they never heard anyone but themselves say the word, they never encounter the word outside the experimental setting. Perhaps it would be appropriate to say that, while lexical proto-entries may have been created for these novel words, those entries are impoverished. If this not-unlikely scenario is the case, then in fact a lack of observed result is not surprising, but rather expected. Taking all of the results shown in this experiment as a whole, it actually appears that novel words are being treated in some cases like words, and in other cases like non-words, as though they do indeed have an intermediate status.

In fact, if we refer back to the differences in means between the sound durations for those sounds spelled with one versus two letters, we see that independent of the non-statistical significance of the differences, those differences are only on the order of on average 8 milliseconds. This is quite a bit smaller than the effects of spelling on sounds found in Experiment One which tested subjects' productions of English words. That effect size was 17 milliseconds, or 29 milliseconds in the logarithmically adjusted analysis. This smaller effect size may also be suggestive of a mediation of lexical status on the effects of orthography on sound durations in speech production.

One aspect that would be interesting to consider further is the role of familiarity of a word, and experience with a word combined with the possibility of gradient lexical status as suggested above by the somewhat ambiguous results of orthography on productions of novel words. One way to test this hypothesis would be with a larger data-set of English words that vary in terms of their lexical status. Some ways in which real English words vary in terms of the content of their lexical representations, or more generally their lexical status, may be frequency of the word (perhaps as an appropriate proxy for word familiarity), morpheme structure of the word, or grammatical class of the word.

3.4 Conclusion

The results of this experiment lead to some interesting questions regarding the effects of WORD PHONES, WORD LETTERS, and CONSONANT LETTERS on word durations. A larger and less experimentally controlled data-set would enable us to tease apart these effects and determine their individual contributions to word and sound durations. Additionally, a small difference in means of sound durations of sounds spelled with one or two letters was shown to be not significant. However, the inevitable consequences of the experimental design contributed to a small number of items to be analyzed. This reduces the power of the experiment to observe effects of orthography on productions. A larger data-set of spoken words would allow for such an effect, if present, to be observed.

Furthermore, it was hypothesized that the novel words created in this experiment were impoverished with respect to their lexical representation. This impoverishment may contribute to a smaller impact of orthographic representation on spoken word production. Again, a larger data-set, containing words which vary along dimensions of lexical status would allow for testing that particular hypothesis. The obvious next step in exploring potential effects of orthography on speech production is through analyzing a large and less experimentally controlled data-set of spoken English words.

Chapter 4

BUCKEYE CORPUS ANALYSES

4.0 Introduction

The previous experiments have shown that while there appears to be an effect of the number of letters used to represent a sound or word, that effect is tied to lexicality. Experiment One shows an effect of the number of letters used to spell a word final sound on the duration of speakers' productions of that sound. In contrast, that same variable was not shown to be significant in people's productions of non-words, evidenced in Experiment Two. Experiment Three explored in greater detail this asymmetric effect of letters used in the orthographic representation of words and non-words. No effect of number of letters on speaker's sound durations was observed for the non-words, consistent with the lack of local effect in Experiment Two for non-words. For 'words' (recall that their status as words remained somewhat in question) in Experiment Three, there was a smaller difference in means between one and two letter representations of sounds than observed in Experiment One. This difference was qualitatively suggestive, but statistically not significant, and a post-hoc power analysis suggested that the experiment was lacking in statistical power to observe an effect of that size. An analysis of a potential orthographic effect on the durations of entire words showed effects of orthography on word durations, but one which required further disambiguation.

While the previous chapters have discussed effects found through lab-elicited speech in a series of planned experiments, this chapter turns to a large corpus to explore the same questions in recorded natural speech. The data were taken from the Buckeye Corpus, a corpus of conversational speech from 40 speakers in the Columbus, Ohio region (Pitt et al, 2007).

There are considerable benefits to be obtained from using such a data-set when asking the research questions described in previous chapters. It is not always the case that results obtained from an examination of lab-elicited speech correspond to the effects present in natural speech. This concern about the potential irrelevance of effects found in the lab environment is avoided by using this corpus. Any effects we see obtained in this non-elicited speech would not be subject to the same potential criticism. An additional factor that could have potentially contributed to the observed effect of the number of letters on sound duration is the visual presentation of material. This could especially be the case in Experiment One, though the absence of effect in Experiments Two and Three for non-words might diminish that concern. Nevertheless, as all the natural speech in this corpus is generated spontaneously by the participants in response to questions in an interview format, this methodology also frees us from those concerns.

Recently, corpora have been used across a number of different linguistic domains. In phonology and phonetics, researchers have examined phonetically tagged corpora to answer questions about form variation resulting from variables present in natural speech (Bell et al., 2003; Byrd, 1993), as well as to confirm findings on phonological and phonetic processes found in lab-elicited speech.

4.1 The Data set from the Buckeye Corpus

The Buckeye Corpus, containing natural speech from 40 subjects local to Columbus, Ohio, contains approximately 300,000 words and was collected, transcribed and labeled by researchers at The Ohio State University. Each recording took place during an interview at The Ohio State University and is of high quality. The recordings were orthographically transcribed by undergraduate student researchers using Soundscriber software. The sound files and transcriptions were then input into an automatic phonetic transcription program, Entropics Aligner to produce phonetic labels for each word. Research assistants selected appropriate word-forms from the program dictionary and aligned them to the words in the sound file. Phonetically trained researchers then hand-realigned the phonetic labels to words and sounds using waveform and spectrograph of the soundfiles (Kiesling, Dilley, & Raymond, 2006).

4.1.1 Materials

Experiments One through Three discussed in previous chapters all utilized materials whose word endings contained voiceless obstruents that can be orthographically represented in more than one way. To ensure continuity between those experiments and the current analyses of Buckeye corpus data, as well as to enable comparisons between results, all monosyllabic words ending in voiceless obstruents were culled from the corpus[11]. This resulted in a smaller corpus containing the spellings of words, their timestamps (where, in what recording, and by which subject the word could be found),

[11] Using a PERL script, see Appendix C.

their durations, and the word-final sound (phone) of the word as produced by the speaker. Words whose citation form ends in a voiceless obstruent but whose actual productions were produced with this word final sound deleted, or assimilated to a following sound, were removed from the data set. An example of a word of this type is *it* spoken in sequence before *can't* resulting in a production of [ɪkænt] with the /t/ elided or assimilated. After this culling, the process resulted in approximately 38,000 tokens of approximately 950 different words. Summary statistics of which word-final sounds are represented in the data-set and how many letters they are spelled with can be found in Table 4.1.

Word final sounds			
Letters	**1**	**2**	**3**
Sound			
/p/	1732	1	
/t/	10357	5	550
/k/	7349	856	
/f/	1546	717	
/s/	8408	487	
/ʧ/		1254	105

Table 4.1: Total instances in the culled data-set from Buckeye Corpus of each word-final voiceless obstruent with each number of letters representing it.

4.1.2 Variables

The amount of variation in natural speech can be massive. Some of the most notable types of variation in natural speech stem from the process of reduction as exemplified by *want to* vs. *wanna* (Johnson, 2005). These reduction processes manifest themselves as vowel reduction and vowel elision within words, shorter durations across words, and coarticulation both within and across words. Other specific recognized contributors to variation in the spoken forms of words in natural speech include word frequency (Zipf, 1929; Fidelholz, 1975; Rhodes, 1992, 1996), utterance position (Klatt, 1975; Ladd and Campbell, 1991; Crystal and House, 1990, *inter alia*Π, and temporal relation to a disfluency (Bell et al., 2003). Some of this variation will be captured through variables consistent with those used in Experiments One through Three, while other variations, like the influence of the position of a word with respect to interruptions in speech (disfluencies) on durations will be subsumed in the error term.

In light of the research described above, which suggests many additional factors that may contribute to sound and word durations, the following pieces of information were collected for each token and along with the words themselves comprise the data set. Some pieces of information pertain to the word, and others to the final obstruent sounds in the word.

WORD DURATION:

Word durations were collected from the labeled word text files provided with the corpus. The onset and offset of each word were noted by Buckeye corpus project personnel with a timestamp (the time location within each recording the onset or end occurred), and a simple subtraction calculation as part of the extraction program mentioned above provided duration measurements.

SOUND DURATION:

Sound durations were collected from the labeled phones text files provided with the corpus[12]. Duration measurements were made by Buckeye corpus project personnel according to the following standards: Measurements of released stops were made from the beginning of closure to the end of the burst, and for unreleased stops from the beginning of closure to the onset of the next segment as evidenced through wave-form and spectrograph analysis. Measurements of fricatives were made from onset to end of frication as evidenced in the spectrograph (Pitt et al., 2006).

INDIVIDUAL SOUND:

Six different voiceless obstruents are found with orthographic variation word-finally in the corpus. They are [p t k f s ʧ]. Each word is labeled with the sound that occurs in its word-final coda.

CONSONANT LETTERS:

This variable ranges from 1 to 3 letters. Not each sound is represented by the full range of potential variation, a fact of English spelling, and accurately reflected by the variation in the corpus itself, summarized above in Table 4.1.

FREQUENCY:

This information was obtained from the corpus itself[13]. It is an absolute value of the number of occurrences of a given word in the corpus.

WORD LETTERS:

The length of the word measured by the number of letters used to spell the word.

WORD PHONES:

---

[12] The sounds were collected through a modification to the previously mentioned PERL script.
[13] Appendix C.

The length of the word measured by the number of sounds or phones used in the citation (unreduced) form of the word.

FINALITY:

Some words end in consonant clusters. When a non-final member of the consonant cluster is also a voiceless obstruent (ie *risk*), all relevant measurements for the non-final member are also noted. In this case, the token is also marked for non-finality of obstruent sound. Non-final members are included in the analysis because they allow for greater variation in orthographic representation of sounds.

FUNCTION STATUS:

I mark each token with respect to its lexical status as a function or content word. An exhaustive list of all monosyllabic function words which end in a voiceless obstruent can be found in Appendix D.

4.2 Research Questions

The pieces of information above, which previous research has suggested might contribute to variation in spoken word duration, are included in the regression analyses below. As it seems necessary below, these measures will be described more completely. By using the measures above one is able to ask many of the same questions as were asked of the previous data-sets collected in Experiments One through Three. This corpus data even allows us to ask additional questions and test a broader set of hypotheses. Whereas in Experiments One through Three, we were concerned with the effect of the number of letters used to represent a word final obstruent sound on the duration of that sound within the spoken word, we can now also assess whether the number of letters used to represent a whole word has an impact on the duration of that spoken word, independent of other factors.

Experiments One through Three indicated that the effect of orthography on people's productions is mitigated somehow by lexical status. Specifically, real words, or items with lexical entries, show an orthographic effect on spoken word duration. However, that effect was not seen for non-words, items for which the speaker has no lexical entry. The data-set collected from this corpus will allow us a different angle from which to view this asymmetry tied to lexical status. The corpus contains both function and content words. Function words differ from content words on a number of dimensions. Function words belong to the theoretical class of words called closed (difficult to add new words of this class to a language), while content words belong to the open class of words (frequently and easily experiencing new additions). Additionally, function words are grammatically necessary, while content words are not. Function and content words exhibit different lexical access effects. While access of content words is speeded as the frequency of the word increases, this does not hold for function words (Bradley 1978), though some (Taft 1990) have argued that this is reflective of post-access decision processes. Studies in ERP (event related potential) methodology have revealed differences across the two classes of words also. Pulvermuller (1999) shows that lexical access for content words involves a different region of the brain associated with word meanings, than the region activated during lexical access to function words. Because of the wide variety of words in this corpus, we can compare the impact of orthography on words with different lexical status and one way to differentiate lexical status is along this dimension of grammatical class distinction.

Additionally, the wide variety of lexical items in this corpus gives rise to drastic differences in word frequency. As mentioned above in the section 3.2.1, we may be able to use this particular characteristic of this large corpus to address the issue raised by Experiments One through Three concerning the differential orthographic effects mediated by lexical status. If more frequent words can be said to have stronger lexical representations than very low frequency words, then frequency gives another measure of lexical status on which we can test the relative effects of orthography. One last dimension of lexical status present in this corpus is the morphemic composition of words. The corpus contains both mono-morphemic and bi-morphemic words, and the effects of orthography across these two classes can be compared to assess whether different effects of orthography on productions is obtained for words with different lexical statuses.

The hypothesis that the orthographic representation of words, and sounds within words, can affect people's productions of words and their sounds is relatively straightforward. However, this document, and the research of others have shown that many different variables can affect those same durations. These many factors along with the variability introduced by natural speech indicate that a regression analysis would be a prudent way to statistically test this hypothesis. Many of the research questions laid out above have been asked of the experimentally elicited data collected in Experiments One through Three and described in Chapters Two and Three. This novel data-set drawn from the Buckeye Corpus allows us to more clearly and accurately assess those questions and more, as outlined in detail below.

For natural, spontaneous speech,

1. Is there an effect of orthography on sound durations? Specifically, is it the case that the more letters used to represent a word final voiceless obstruent, the longer the spoken duration of that segment? Additionally, are there other orthographic characteristics of sounds which contribute to their durations, such as sound frequency, or amount of orthographic variation of each sound?
2. Is there an effect of orthography on word durations? Does the number of letters used to spell a word impact the duration of that word? Independent of the total number of letters spelling a word, does the number of letters in a word-final obstruent also affect the word's duration?
3. Is there an interaction of lexical status and orthographic representation? Do content and function words show differing effects of orthographic word length on word duration (see 2 above)? Does the frequency of a word interact with orthographic characteristics of the word in their impact on sound or word durations? Do morphologically complex words behave differently than simplex ones in the effects of orthography on durations?

4.2.1 Question 1: Effects of orthography on Sound durations

In this section results from 4 regression analyses are discussed. The first model is consistent with the model used for Experiment One except for the addition of some new variables, as previously discussed. Subsequent analyses are conducted and other regression models considered as a function of results found and questions raised by the first model, which will be explained in detail as they occur.

Sound Duration Analysis 1: The relevant variables which should be included in a model to test Question 1 include sound duration (the dependent variable), orthographic length as measured by the number of letters used to represent the sound, finality (whether the sound is last or a previous member of a coda cluster), and information about the sound itself. Orthographic length is the variable of interest, however prior research has shown the effects of the other variables as well, so it is necessary to isolate their effects in the model too. Crystal & House (1988) show that final sounds are longer in duration than non-final sounds and that members of a consonant cluster are shorter than their singleton counterparts. Those same researchers and many others have also shown inherent durational differences between sounds. It is of course not surprising to consider that fricatives or affricates may have greater inherent length than their stop counterparts, but there are also differences in duration within stops for the various places of articulation. For stop closures, Zue (1976) found /p/ to be shorter than /t/ and /k/, and Luce & Charles-Luce (1985) found durations to decrease as one progresses from bilabial to velar to alveolar place. Byrd (1993) also found an effect of place of articulation on stop closure durations such that bilabial>velar>alveolar place.

Table 4.2 shows the model used to test this question. The variables mentioned above; WORD LETTERS, FINALITY, and INDIVIDUAL SOUND, account for 13 percent of the variation in sound durations in the data-set. However, each of these variables has a significant independent effect. In line with previous research (Crystal & House 1988) we see a significant effect of coda final placement of the obstruent. Specifically, word final obstruents are 14 milliseconds longer than non-final coda consonant members.

Furthermore, and also consistent with the research conducted on the topic (Zue, 1976; Luce & Charles-Luce, 1985; Byrd, 1993), we see inherent durational differences for each of the consonant phones. The coefficients associated with the phones themselves should be read in relation to the duration of the sound /p/, conspicuously absent from the model. This variable with the fewest observations is automatically dropped from the model in order to provide a point of comparison from which to interpret the coefficients. These significant values show that inherent durational differences between sounds are evident also in this corpus. For coda consonants, /t/ is shorter than /p/ by 20 milliseconds, /k/ is sorter than /p/ by 9.5 milliseconds, /f/ is longer than /p/ by 4.5 milliseconds, /s/ is longer than /p/ by 17.5 milliseconds, and /ʧ/ is longer than /p/ by 32.5 milliseconds. An analysis of all the sound durations in the corpus in all word positions reveals the same relative lengths as described in this model for these 6 word final voiceless obstruents.

Number of obs = 36356			R-squared = 0.1318	
F(7, 36348) = 788.39			Adj R-squared = 0.1316	
Prob > F = 0.0000				
variable	**Coef.**	**Std. Err.**	**t**	**P>|t**
CONS LETTERS	.0095452	.0006881	13.87	0.000
FINALITY	.0138444	.0008852	15.64	0.000
t	-.0204961	.0011363	-18.04	0.000
k	-.0094076	.0011571	-8.13	0.000
f	.0045494	.0014244	3.19	0.001
s	.0175063	.0011439	15.30	0.000
ʧ	.0326553	.0017616	18.54	0.000

Table 4.2: A model of the effects of orthographic length, finality, and sound type on sound duration.

This model's confirmation of previous results is expected and confirms the validity of the methodology for identifying relevant factors. This makes it particularly noteworthy that there is also a significant effect of the number of letters used in the representation of the coda consonant (CONSONANT LETTERS). Specifically, an increase of one letter in the orthographic representation corresponds to a 9.5 millisecond increase in consonant duration.

Sound Duration Analysis 2: The effect of orthographic length of a word-final sound on that sound's duration, which was observed in Experiment One has now been confirmed. Independent of the inherent durational differences between sounds, adding a letter to the representation of a word-final sound increases the duration of that sound by 9.5 milliseconds. There are different durations associated with the various sounds themselves, information which is currently encoded as dummy variables for each sound.

However, an examination of the summary statistics for total word-final sound instances found in Table 4.1 reveals that the frequency of occurrence of the various sounds corresponds well to their relative durations. For instance, /t/ is the most common sound and is also the shortest in duration. On the other end of the continuum, /ʧ/ is the least common sound, and also the longest in duration. The only sound which doesn't conform to this pattern is /s/. A goal in regression analyses is to minimize the number of unique variables needed to predict the dependent variable. Consequently, one might consider replacing the 6 dummy variables with one sound frequency variable and examining whether the model increases or decreases in predictive capacity.

A regression analysis performed on sound durations with CONSONANT LETTERS, FINALITY, and SOUND FREQUENCY as independent variables results in an adjusted R-squared of .021, indicating that the variables in the model account for 2 percent of the variation in sound durations. SOUND FREQUENCY has a significant effect such that more frequent sounds correspond to shorter durations, and each of CONS LETTERS and FINALITY remain significant with effects in the same direction as above, but the lower R-squared indicates that this model loses considerable predictive capability relative to the model with dummy variables for each of the sounds. This model can be seen in Table 4.3.

Number of Observations	= 36360		R-squared	= 0.0210
F(3,36356)	= 259.81		Adj R-squared	= 0.0209
Prob > F	= 0.0000			
variable	**Coefficient**	**Std. Error**	**t**	**P>|t|**
CONS LETTERS	0.0146949	0.0006656	22.080	0.000
FINALITY	0.0080573	0.0009307	8.660	0.000
SOUND FREQUENCY	-0.0000006	0.0000001	-7.350	0.000

Table 4.3: A regression model replacing sound dummies with one continuous sound frequency variable.

Sound Duration Analysis 3: Although this frequency analysis shows no improvement over the previous model with dummy variables for sounds, it raises the question of whether the frequencies of representations might impact sound durations. Sounds are orthographically realized in different ways in English. For instance, the sound /k/ is realized as 'k', 'ck', 'ch', 'c', and 'que' word finally in various monosyllabic English words (i.e. *yak, stack, tech, tic, clique*). However /k/ is not realized equally often in each of the representations; some are more frequently used than others. The number of times each orthographic representation is used to represent each of the sounds examined was entered into the data-set and coded as a continuous variable called REALIZATION. I coded this variable as the raw number of ways that sound was spelled a particular way in the corpus. There are 20 different ways to spell the 6 sounds. The values for this variable range from 1 (when /f/ is represented by 'ph') to 11647 (when /s/ is represented by 's').

Adding this variable to the first model predicting sound durations (cf. *Sound Duration Analysis* 1), we observe few changes. As shown in Table 4.4, the adjusted R-squared for the model is .1332, indicating that the variables listed account for over 13% of the variation in sound durations in the data-set. Each variable previously noted to be significant remains so, with effects in the same directions. Interestingly, this new variable, REALIZATION, which addresses the frequency with which each representation for each sound is instanced, also has a significant effect on sound durations. Specifically, as the frequency of the orthographic realization of a sound increases, the duration of that sound decreases. This effect is independent of the effect of orthographic length of the sound, or number of letters used to represent the sound. This model can be seen in Table 4.4.

Number of observations = 36327 R-squared = 0.1333
F(8, 36318) = 698.47 Adj R-squared = 0.1332
Prob > F = 0.0000

variables	Coefficient	Std. Error	t	P>|t|
CONS LETTERS	.0038697	.0009173	4.22	0.000
t	-.0117194	.0014554	-8.05	0.000
k	-.0033132	.0013177	-2.51	0.012
f	.00582	.0014293	4.07	0.000
s	.0268817	.0015002	17.92	0.000
ʧ	.0374	.0018611	20.10	0.000
FINALITY	.0128639	.0008899	14.45	0.000
REALIZATION	-1.10e-06	1.14e-07	-9.65	0.000

Table 4.4: A regression model for sound durations with frequency of orthographic realizations (REALIZATION) of sounds added.

The coefficient for REALIZATION here looks tiny, but it's not really all that small. The variable is coded in total number of instances, and that variable spans over 11,000 in count. Consequently, that small coefficient is the change in sound duration for an increase of 1 in the variable REALIZATION. Increasing the frequency of the realization of a sound by 1000 instances decreases the duration of the sound by 1.1 milliseconds.

Sound Duration Analysis 4: The very suggestion that some sounds are represented in many different ways leads to the question whether the possible variation in orthographic representation might affect sound durations. As mentioned above, some sounds like /k/ can be represented a large variety of ways. Other sounds, conversely, have only a few options for orthographic representation in word-final occurrences. Consider the two sounds /k/ and /p/. As described above, the sound /k/ can be spelled in word final codas in monosyllabic English words in five distinct ways ('k', 'ck', 'ch', 'c', and 'que'), while /p/ can be spelled in only two ways ('p', and 'pp'). Each of the other sounds in this analysis falls within this range. A regression including a variable for variability in the representation of a given sound (VARIATION*)* allows for an analysis of the potential independent contribution of variability of orthographic representation on sound durations.

The relevant changes and advantages to this model, shown in Table 4.5, are worth mentioning. The adjusted R-squared is .1332 indicating that the variables still account for over 13% of the variation in sound durations in the data-set. The fit of this model, then, is effectively the same as the previous. However, an advantage of this model is that it takes some information out of the error term, or subsumed by previously mentioned variables, and can show its independent contribution to sound durations. Each of the variables previously assessed as independently contributing to sound durations remains so, with effects in the same direction. Additionally, this new variable addressing the orthographic variability of each sound independently contributes to sound durations. Specifically, an increase in the number of ways to represent a sound corresponds to a decrease in sound duration of that sound.

Number of obs = 36327		R-squared = 0.1333		
F(8, 36318) = 698.47		Adj R-squared = 0.1332		
Prob > F = 0.0000				
Variables	**Coefficient**	**Std. Error**	**t**	**P>|t|**
CONS LETTERS	.0038697	.0009173	4.22	0.000
VARIATION	-.0011044	.0004392	-2.51	0.012
FINALITY	.0128639	.0008899	14.45	0.000
t	-.0095107	.0008048	-11.82	0.000
f	.0080288	.001102	7.29	0.000
s	.0279861	.0011393	24.56	0.000
ʧ	.0374	.0018611	20.10	0.000
REALIZATION	-1.10e-06	1.14e-07	-9.65	0.000

Table 4.5: A regression model for sound durations including the variable *variation* capturing the number of unique orthographic representations for each sound in a word-final position. Note that 'k' is dropped from the model, mathematically necessary to uniquely identify the added variable VARIATION.

So we see there are at least three unique effects of various orthographic characteristics of word-final sounds on the durations of those word-final sounds. First, there is an effect of CONSONANT LETTERS, that is, the number of letters used to spell the word-final sound: the more letters are used in the representation of the sound, the longer the duration of that sound. Secondly, there is an effect of the frequency of REALIZATION of a sound, such that more frequent orthographic representations for sounds correspond to shorter durations of that sound. Lastly, we see an effect of the potential VARIATION in orthographic representation of a sound. Sounds which can be spelled in a greater variety of ways are produced with shorter durations than those sounds which have a more limited variety of orthographic representations.

All of these effects independently contribute to sound durations. This naturally leads to a somewhat broader question, whether the orthographic representation of a sound is limited in its scope of effect, or whether the orthographic representation of a sound can contribute to durational differences at the word level. Given the results of the previous analyses, an intuitive hypothesis would be that the number of letters used to spell a word might affect word durations, independently of the phonological length of words. However, these previous results also suggest that there may be a global effect of the number of letters used to spell a word-final sound on the duration of the whole word, independent of the orthographic length of the whole word. Recall that results from Experiment Three indicated an effect of WORD LETTERS, CONSONANT LETTERS, and WORD PHONES on word durations. In that case, the individual effects of these variables could not be teased apart, leaving the question of whether all contributed or some and not others. With the current data set from the Buckeye Corpus we can now address the individual contributions (or not) of these variables.

4.2.2 Question 2: Effects of Orthography on Word durations

This section turns first to the broad and intuitive hypothesis that the orthographic length of a word (WORD LETTERS) might influence word durations. Considering the effects shown in the previous section, it seems reasonable that adding a sound to a word would increase a speaker's production of that word, accounting for the addition of phonological length as a variable theoretically expected to impact word duration. This section considers the results of two regression analyses on word durations.

Word Duration Analysis 1: As mentioned above, previous researchers have shown that more frequent words experience reduction, one manifestation of which is a decrease in word duration (Whalen, 1992; Jurafsky et al. 2001). Hence, word FREQUENCY is a reasonable variable to add to the predictive model. Based on the discussion above, some relevant variables which might reasonably be included in a model to test Question 2 include word duration (the dependent variable), FREQUENCY, WORD PHONES, and WORD LETTERS. WORD LETTERS is the variable of interest and specifically mentioned in our hypothesis. Consider the output of this simple Ordinary Least Squares regression model on the data-set, shown in Table 4.6.

Number of obs = 33441			R-squared = 0.2285	
F(3, 33437) = 3300.61			Adj R-squared = 0.2284	
Prob > F = 0.0000				
variable	**Coef.**	**Std. Err.**	**t**	**P>|t|**
WORD LETTERS	.0144495	.0008536	16.93	0.000
WORD PHONES	.0363913	.0013322	27.32	0.000
FREQUENCY	-.0000157	3.25e-07	-48.28	0.000

Table 4.6: A model of the effects of WORD LETTERS, WORD PHONES, and word FREQUENCY on word durations.

The three variables discussed above; WORD LETTERS, WORD PHONES, and word FREQUENCY, account for nearly 23 percent of the variation in word durations in the data-set as evidenced by the adjusted R-squared of .2284. Notice also the significance of each of these variables shows that each independently affects word duration. For WORD

PHONES, adding one segment to a word increases the duration of that word by 36 milliseconds. For word FREQUENCY, the negative sign indicates that the effect is in a downward direction: an increase of one unit of frequency results in a decrease of duration by .01 milliseconds. Both of these effects are significant at the $p<.001$ level. While the effects of phonological length and word frequency are to be expected, given the results of the previous research, the model also reveals a significant positive effect of orthography on word durations. The coefficient for orthographic length shows that as we increase the length of the word by 1 letter, the duration of that word is increased by 14 milliseconds. This is independent of phonological length, which means that words that have the same number of phones still differ in duration depending on the number of letters used to spell the words. This effect is also significant at the $p<.001$ level.

Word Duration Analysis 2: While the broad hypothesis that the orthographic length of a word contributes to the word's duration is confirmed with the above analysis, there remains a question of whether the orthographic characteristics of a particular word-final sound can independently contribute to word durations. We can test this question with the addition of a variable for orthographic length of a word-final sound (letters) to our regression model above.

Adding this variable to the regression shows some interesting results, see Table 4.7. Notably, each of the variables previously shown to be significantly affecting word durations continue to do so with effects in the same direction. The adjusted R-squared for the model is .2404 indicating that the variables in the model account for over 24% of the variation in word durations. Additionally, we can note that there is a significant independent effect of the orthographic length of a word-final sound on word durations such that adding a letter to the orthographic representation of a word-final sound increases the duration of a word. The coefficient on this variable is .036 which shows that the word-duration increase per word-final letter increase is 36 milliseconds.

Number of obs = 33441   R-squared = 0.2405
F(4, 33436) = 2646.25   Adj R-squared = 0.2404
Prob > F = 0.0000

variables	**Coefficient**	**Std. Error**	**t**	**P>|t|**
WORD LETTERS	.0052359	.0009372	5.59	0.000
WORD PHONES	.0500901	.0014503	34.54	0.000
FREQUENCY	-.0000131	3.42e-07	-38.48	0.000
CONS LETTERS	.0362284	.0015777	22.96	0.000

Table 4.7: A regression on word durations in Buckeye with the variable CONSONANT LETTERS added.

One significant difference between this model and the model not including the variable for orthographic length of a word-final (CONS LETTERS) sound is the considerable decrease in the coefficient for WORD LETTERS, a variable encoding the number of letters used to spell the word (.0144 vs. .0052). One possible reason for this would be if not all the words in the data-set behave the same with respect to the influence of orthographic length on word durations when also accounting for the number of letters used to spell a word final sound.

The previous chapters have shown suggestive evidence that subjects behave differently with respect to the influence of orthography on word productions when comparing elicitations of words and non-words. While the Buckeye corpus is a spoken word corpus of only actual words, there are nonetheless distinctions present in the lexical status of the items present in the corpus.

4.2.3 Question 3: Interaction of lexical status and orthographic effects on durations

One final research question that this data-set addresses here has to do with the role of lexical status on the impact of orthographic representation on production. Experiments One and Two reported in Chapter Two have shown that the effect of orthography is different for words and non-words. Although this corpus of natural spontaneous speech consists almost exclusively of real words, there are many ways in which the words vary. Some of this variability will be exploited, to address the question of the apparent interaction of lexical status with orthographic effects on speakers' productions. This section will examine the results of four regression analyses concerned with that issue.

Lexical Status Analysis 1: One of the ways that words in this corpus vary is with respect to their grammatical status. Specifically, there are two classes of words represented. The corpus contains both function words and content words. Their differences lie not only in varying average frequencies of occurrence, but also in their status in the lexicon with respect to the rest of the grammar. It can be argued that function words have different kinds of lexical entries than content words do. Instead of encoding contingent, conceptual, real-world information, they encode grammatical relationships and structure (Harley, pc). This lexical difference allows us to examine whether durations of words with different lexical statuses are differentially influenced by orthographic representation.

In order to address this question, two identical models are used on two halves of the data set. One half (16260 tokens) contains the function words only, and the other, the content words (17181 tokens). We can then compare the impact of orthographic word length, on word durations across the two datasets. The variables in this model are word duration, WORD LETTERS, WORD PHONES, and word FREQUENCY. In short this model, shown in Tables 4.8 and 4.9 is identical to the model shown in Table 4.6, but is run on two parts of the data-set in order to compare the effects of orthography on content vs. function words.

The interpretation of the model below on the two parts of the data-set is not exactly simple. Turning first to Table 4.8, the adjusted R-squared informs us that the variables in the model account for 19.5 percent of the variation in content word durations. WORD LETTERS, WORD PHONES, and word FREQUENCY all significantly affect content word durations at a level of $p<.001$. The signs of the coefficients conform to the expected effects, and match those found in Table 4.6, section 4.2.2.

Table 4.9, which shows the results of the same model applied to the set of function word tokens, reveals some differences. The variables in the model account for 12 percent of the variability in function word durations, indicating that these variables do not predict function word durations as well as content word durations. WORD PHONES and FREQUENCY have similar significant effects on function word duration as they were seen to have in Table 4.8 for content word durations.

The effect size of FREQUENCY is smaller for function words than content words, and the effect size of WORD PHONES is greater for function words than content words.

The variable of WORD LETTERS is not so clearly significant. Applying the arbitrary standard of p>.05 for significance would show this variable to be not significant. For those tempted to consider a p-value of .051 "close enough", consider the sign of the coefficient. If orthographic length has a significant effect on function word durations, then that effect is negative, such that an increase of one letter results in a *decrease* of approximately 3 milliseconds in word duration, completely the opposite direction as the effect previously discussed.

Number of obs = 17181			R-squared = 0.1950	
F(3, 17177) = 1386.81			Adj R-squared = 0.1948	
Prob > F = 0.0000				
variable	**Coefficient**	**Std. Err.**	**t**	**P>|t|**
WORD LETTERS	.0208295	.00128	16.27	0.000
WORD PHONES	.0145015	.0016135	8.99	0.000
FREQUENCY	-.0000233	5.14e-07	-45.32	0.000

Table 4.8: A model showing the impact of orthographic length on content word durations.

Number of obs = 16260			R-squared = 0.1200	
F(3, 16256) = 739.06			Adj R-squared = 0.1199	
Prob > F = 0.0000				
variable	**Coefficient**	**Std. Err.**	**t**	**P>|t|**
WORD LETTERS	-.0027311	.0013971	-1.95	0.051
WORD PHONES	.0555398	.0026324	21.10	0.000
FREQUENCY	-6.51e-06	4.14e-07	-15.74	0.000

Table 4.9: A model showing the impact of orthographic length on function word durations.

Because such an interesting result hinges on the difference between the coefficients for orthographic length across the two models, it is necessary to ensure that the difference is statistically significant. A Wald test can test any linear hypothesis (Green, 2003). The one we're interested in is whether the two WORD LETTERS coefficients above are equal. The test statistic has a Chi-squared distribution with df_1. For the coefficients in the two models above, the Wald statistic = 121.61 resulting in a $p<.0001$. The effects of orthographic length on word durations of content and function words are indeed different. This difference in effect is reflective of the theoretical differences between function and content words as discussed above.

Lexical Status Analysis 2: This analysis confirms the possibility which was raised in the previous section that the effects of orthography are different across one distinction in lexical status in the data-set. However, we can see this difference even more clearly when we additionally test the effects of orthographic length of the word-final sound (CONSONANT LETTERS) in addition to orthographic length of the word (WORD LETTERS), phonological length of the word (WORD PHONES), and word FREQUENCY across the same distinction. Adding to the model above a variable (CONSONANT LETTERS), representing the number of letters used to spell the word final sound, reveals some interesting results.

The variables shown to be significant in the previous analysis continue to be so here with effects in the same direction. The R-squared for the analysis run on function words = .1449, while the R-squared for content words = .1994. This indicates that the identical variables included in the two models better account for the variation in word durations of content words (19.94% of the variability) than function words (14.49% of the variability). For function words, the new variable CONSONANT LETTERS also has a significant effect on word durations ($p<.001$) such that adding a letter to the representation of a word-final sound increases the duration of the word by 62 milliseconds. For content words that effect is still significant (p<.001) with a positive effect of letters of 19 milliseconds on word durations. Additionally, we see similar size effects of orthographic length on word durations across the two halves of the data. However, these effects are in opposite directions. The number of letters used to spell a function word now is revealed to have a significant *negative* effect on function-word durations of approximately 19 milliseconds. In contrast, the number of letters used to spell a content word has a significant positive effect on content word durations of approximately 15 milliseconds. These results can be seen in Tables 4.10 and 4.11.

Number of obs = 16260			R-squared = 0.1451	
F(4, 16255) = 689.51			Adj R-squared = 0.1449	
Prob > F = 0.0000				
variables	**Coefficient**	**Std. Error**	**t**	**P>|t|**
WORD LETTERS	-.0189543	.001565	-12.11	0.000
WORD PHONES	.084911	.0029232	29.05	0.000
CONS LETTERS	.0622333	.0028522	21.82	0.000
FREQUENCY	-3.17e-06	4.36e-07	-7.27	0.000

Table 4.10: A regression on word durations of function words with CONSONANT LETTERS added.

Number of obs = 17181			R-squared = 0.1996	
F(4, 17176) = 1070.53			Adj R-squared = 0.1994	
Prob > F = 0.0000				
variables	**Coefficient**	**Std. Error**	**t**	**P>|t|**
WORD LETTERS	.0151689	.0013984	10.85	0.000
WORD PHONES	.0217299	.0017666	12.30	0.000
CONS LETTERS	.0190383	.0019216	9.91	0.000
FREQUENCY	-.0000219	5.33e-07	-41.03	0.000

Table 4.11: A regression on word durations of content words with CONSONANT LETTERS added.

The results of this distinction in lexical status, taken with the lack of orthographic effects found in experimentally elicited non-words reported in Chapters Two and Three strongly indicate that this effect of orthography is tied to the lexical representations of words. This grammatical class distinction of function vs. content words is not the only dimension along which the words in this Buckeye corpus vary in terms of lexical status. One may also claim that more frequent words have a stronger lexical status than those with lower frequencies. In fact, non-words can be considered to have a frequency of zero.

Lexical Status Analysis 3: The regression models used above can test whether the effect of orthography revealed in the previous analyses has a different effect for words of different frequencies. Given the results of the previous analyses and experiments, one would expect a smaller effect of orthography for words with lower frequencies on the assumption that word-frequency is positively indicative of lexical status or strength. A differential effect on the dependent variable (sound durations) of one independent variable (in this case, CONSONANT LETTERS) at different levels of a second independent variable (in this case, word FREQUENCY) is called an interaction. The question of whether this interaction is indeed occurring can be analyzed through the addition of an interaction term of orthography and frequency to the model (INT FREQ*LET), in addition to the individual variables. This term can be added to the previous model, which revealed an effect of orthography on durations. Then, provided there is a significant effect of the interaction on the dependent variable, it is a matter of interpretation of the coefficient to determine just what that interaction is.

A simple test is to add this interaction term to the model that shows an effect of the number of letters used in the representation of a word final sound (CONSONANT LETTERS) on the duration of that sound. We can modify the model shown in Table 4.2, section 4.2.1. This has the potential to be particularly revealing because it is the same effect tested for in Experiments One, Two, and Three described in Chapters Two and Three. Adding the interaction term, and the individual variable FREQUENCY which is mathematically necessary for an accurate estimate of the interaction term, we see similar overall results, see Table 4.12 below. The adjusted R-squared of the model is .1325 indicating that the variables in the model account for more than 13% of the variation in sound durations in the data-set. The variables which were previously shown to have significant effects on sound durations continue to do so with effects in the same directions. Interestingly, this model also shows a significant negative effect of frequency on sound durations ($p<.001$) and a significant interaction between letters and frequency ($p<.001$).

Number of obs = 36356		R-squared = 0.1327		
F(9, 36346) = 618.15		Adj R-squared = 0.1325		
Prob > F = 0.0000				
variable	**Coefficient**	**Std. Error**	**t**	**P>|t|**
CONS LETTERS	.0062457	.0009739	6.41	0.000
FREQUENCY	-8.84e-06	2.23e-06	-3.96	0.000
FINALITY	.014399	.000902	15.96	0.000
t	-.0190831	.0011725	-16.28	0.000
k	-.0079985	.0011877	-6.73	0.000
f	.004899	.0014262	3.44	0.001
s	.0182273	.001151	15.84	0.000
ʧ	.03376	.0017696	19.08	0.000
INT FREQ*LET	8.03e-06	2.22e-06	3.61	0.000

Table 4.12: A regression on sound durations with an interaction term of frequency and letters, INT FREQ*LET, added.

This interaction term has a positive coefficient. One way to understand the direction of this interaction is to consider the simpler mathematic model predicting sound durations, see Figure 4.1. The coefficients for the three variables of interest have been included, and the other known contributors to sound duration have been omitted. We see that sound duration is determined by (excluding for now the other significant factors in the regression above) the difference between the coefficients of letters and frequency plus the coefficient of the interaction term.

Sound duration = .0062457(*CONS LETTERS)* - .00000884(*FREQUENCY*) + .00000803(*INT FREQ*LET*)

Figure 4.1: A simple mathematical equation showing the effects of two variable and their interaction on sound duration.

One can simply plug into this equation the various values of the independent variables CONS LETTERS and FREQUENCY and get outcomes for the sound duration. Holding FREQUENCY constant at for instance 1, we see that the difference between sound durations for a change from 1 to 2 in the value of CONS LETTERS is approximately 6.2 milliseconds. However, if we consider the same change from 1 to 2 in the value of CONS LETTERS while holding FREQUENCY constant at 1000, we find a difference between sound durations of approximately 14.3 milliseconds. Consequently we see that this interaction term can be interpreted as an increasing effect of letters on sound durations for words with increasing frequencies. This is precisely what we predict based on the results of the previous experiments and Buckeye analyses. It is notable that while this example model was performed on the entirety of monosyllabic Buckeye data, this effect is robust across just the content words, and also within the function words.

Lexical Status Analysis 4: One final lexical distinction among the monosyllabic words in the Buckeye corpus is the morphemic content of the words. While the majority of words are mono-morphemic (34696 tokens) there are some which are bi-morphemic (4163 tokens). Derived words which are made from the combination of multiple morphemes are lexically distinct from underived words. We can test whether orthographic content differentially affects word durations of mono-morphemic versus bi-morphemic words in the Buckeye data-set through the inclusion of an interaction term between morphemes and orthographic length. This model reveals that there is a significant effect of MORPHEMES on word durations such that bi-morphemic words are longer in duration than mono-morphemic words (p>.001). Interestingly, there is also a significant interaction of MORPHEMES and WORD LETTERS, (INT MORPH*LET). Using the equation above, it is evident that the direction of this interaction is such that there is a greater effect of orthographic length of a word on word durations for mono-morphemic words than there is for bi-morphemic words (p<.001).

Number of obs = 33441			R-squared = 0.2434	
F(6, 33434) = 1792.58			Adj R-squared = 0.2433	
Prob > F = 0.0000				
variables	**Coefficient**	**Std. Error**	**t**	**P>|t|**
WORD LETTERS	.0025712	.0009642	2.67	0.008
WORD PHONES	.0519081	.0014817	35.03	0.000
CONS LETTERS	.0373903	.0015895	23.52	0.000
FREQUENCY	-.0000132	3.48e-07	-38.00	0.000
MORPHEMES	-.1106423	.0098125	-11.28	0.000
INT MORPH*LET	.0217841	.0019178	11.36	0.000

Table 4.13: A regression on word durations with an interaction term for WORD LETTERS and MORPHEMES included.

To summarize, four different analyses of interactions between orthographic content and lexical status within the Buckeye corpus, and asymmetric effects of orthographic information on words and non-words in the Experiments conducted in Chapters Two and Three all together support the hypothesis that effects of orthography on speech production are tied to lexical entries. It is the nature of the lexical entry which determines in what ways orthographic information can come to impact a person's production of that word. Words with strong, salient, and prototypical lexical entries have robust effects of orthography on productions, but words that have no or less stable and weaker lexical entries, for whatever reason, have mitigated effects of orthography in productions.

4.3 Discussion

The results of this examination of natural speech contained in the Buckeye corpus further illuminate the results of the previous experiments. The various regression models are able to tease apart a number of variables which contribute to word durations and sound durations in spontaneous speech. Most broadly, the above analyses show that orthographic information impacts speakers' productions of words. Specifically, the number of letters used to spell a word positively affects the duration of the word, and the number of letters used in the representation of a word final voiceless obstruent positively affects the duration of that particular sound as well as the duration of the entire word in which it is found. Other orthographic characteristics of word-final sounds also contribute to the durations of those sounds. More frequent orthographic representations of sounds correspond to shorter durations, all else constant. Additionally, sounds that have a greater variability in potential orthographic representation correspond to shorter sound durations, all else constant.

The interaction of lexical status and orthographic representation seen across the previous experiments (non-words versus words) was also further discussed in this chapter, albeit with a variety of different variables (including grammatical class, word frequency, and number of morphemes). Specifically, there was a significant positive effect of orthographic length on the word durations of content words, but a negative effect of orthographic length on word durations of function words. Additionally, words with lower frequencies had a weaker effect of orthographic length of a word-final sound on the duration of that sound, and mono-morphemic words had greater effect of orthographic length on word durations than bi-morphemic words in the data-set. This differential behavior of orthography between the two classes of words reflects the differences in effects of orthography for words and non-words in previous experiment.

Chapter 5

CONCLUSIONS

5.0 Summary

This dissertation reported the results of three experiments and a series of analyses of a large spoken word corpus all examining potential effects of orthographic information on speech production. Chapter Two presented the results from two exploratory experiments. In Experiment One, a significant effect of orthography on sound durations was found, such that coda consonants spelled with more letters corresponded to longer sound durations. In Experiment Two, that same effect was not observed for non-words, or novel orthographic strings. This asymmetry was taken to be suggestive of an interaction between lexical status and orthographic influence, whereby the orthographic representations of sounds in non-words are unable to affect sound durations, perhaps because non-words have no lexical entry.

To further explore this hypothesis, in Experiment Three described in Chapter Three, productions of novel words were elicited, where the novel words varied in their orthographic representation; six monosyllabic non-words with voiceless obstruent codas were spelled with one or two letters representing the final phone. Subjects were trained on meanings for these words and then performed a picture-naming task to elicit the newly learned words. No significant effect of orthography was found on sound durations, for non-words, nor for newly learned words. The lack of effect for non-words was expected, and the lack of effect for newly trained items could have arisen from a variety of factors, one of which was the acknowledged difficulty in getting adults to learn new words. Thus, the "words" in Experiment Three may not have been very word-like. In fact, the smaller difference in means between sound durations of sounds spelled with one versus two letters compared to the differences found in Experiment One led to a discussion of the possible impoverished lexical representations of the newly learned words. Additionally, Experiment Three found an effect of the number of letters spelling a word (WORD LETTERS), the number of phonemes in the word (WORD PHONES), and number of letters used in the representation of a word-final consonant (CONSONANT LETTERS) on word durations. This suggested an avenue for further research, undertaken in Chapter Four, to tease apart the nature of this effect. One question asked was whether the variables WORD LETTERS, WORD PHONES, and CONSONANT LETTERS independently contribute to duration or whether the effect is a result of the combination of these variables.

The analyses conducted in Chapter Four on spoken word productions from the Buckeye corpus follow up many of the avenues of interest generated by the results of the first three experiments. The Buckeye corpus is a spoken corpus of 40 speakers whose speech was spontaneously generated in a modified sociolinguistic interview format. As a result, the data-set is very large, and relatively experimentally uncontrolled. This necessitates statistical control for potential duration-affecting variables, but allows for an investigation of particular hypotheses not examined in the previous experiments.

These analyses conducted on the Buckeye corpus confirm and clarify the effects seen in the results from the previous chapters. Specifically, a significant effect of orthography was found on sound and word durations for monosyllabic, voiceless obstruent-final words. As previously shown, the more letters are used to spell a sound or a word, the longer the duration of that sound or word. This effect is independent of other known or anticipated effects such as frequency of a word, number of sounds in the word, or inherent durational differences of sounds. Further analyses show that a word's lexical status interacts with this orthographic effect on sound durations. Content and function words have different effects of orthography on sound durations, and less frequent words have a smaller effect of orthography on sound durations than more frequent words. Effects of frequency in domains other than overall word frequency were also observed. The more frequent ways to orthographically represent each of the sounds correspond to faster sound durations, independent of number of letters. Furthermore, sounds which have more possible representations have faster durations than sounds with fewer possible representations.

There were a few concerns raised about the visual presentation of stimuli in the majority of the experimental methodology used in Experiments One through Three discussed above. This potential criticism is the same as can be made about all the previous experiments examining potential orthographic effects on speech production (cf. section 1.4). If the effects were only found when people are looking at words, then we could not be confident that the effects are part of people's productions of words. They could easily be explained as effects on access of words when mediated by orthography rather than effects occurring at some post-access point in the production process. However, the confirmation of the results of orthographic effects on production apart from visual presentation of any stimuli as shown in Chapter Four generates even more confidence that the effect is driven by the lexical representations of words themselves rather than generated by lexical access. This chapter continues the discussion of these orthographic effects on speakers' productions of words, and examines in particular how these effects could be incorporated into current models of speech production.

5.1 General Discussion

The conclusion of this research is clear. Primarily, it is apparent that the orthographic representations of lexical items impact our pronunciations of those words. This orthographic effect reveals itself in the durations of word-final sounds, and the durations of whole words. This effect is not simply a more to more correspondence whereby more orthographic content (a greater number of letters) corresponds to a longer duration. Rather, there are other orthographic characteristics that also impact sound durations. Specifically, the frequencies of particular orthographic representations of sounds also have effects on sound durations, such that the most frequent orthographic representations of sounds produce the shortest durations.

A second conclusion of the research described here distinguishes this work from other experiments claiming to find effects of orthography in word production. Because the methodologies of previous experiments have necessitated an analysis of naming latencies across a variety of tasks, the orthographic effects found there may be argued to stem from relative speed of lexical access of an item rather than production per se. The series of analyses conducted here show orthographic effects across actual word durations and within the durations of word-final coda consonants. That is, the effects are found in the words themselves in contrast to reaction times (naming latencies). Furthermore, the analyses from the Buckeye Corpus show orthographic effects in spontaneous speech, where subjects are not directly reading or responding to a prompt to find a particular lexical item. Instead that lexical search is initiated by the cognitive processes of communicative intent. These two characteristics of the data collected and analyzed in this dissertation allow for the unique claim that the observed orthographic effects are realized in word production, rather than being simply explained through a process of lexical access.

The absence of effect for non-words and the various Buckeye Corpus analyses on varying effects for function and content words, for high vs. low frequency words, and for multi-morphemic words all indicate in various ways that the orthographic effect is tied to lexical status. The notion that lexical representations contain orthographic information is not in itself all that exciting. However, the idea that this information should be affecting the speaker as they say words is more so. It is hard to imagine why it would be necessary or important for the linguistic system to behave this way. Either there is some important reason for the use of this information by our speaking processes, or the information is coincidentally reflected in production due to the structure of the various aspects of the linguistic system. These two possibilities are discussed in greater detail in section 5.4 on future research directions.

5.2 Two Theories

There are, it seems, two possible hypotheses that could account for the effects shown above. One hypothesis is that underlying phonological representations of words are influenced by orthographic information. This option is actually reminiscent of early generative phonology work exemplified in SPE (Chomsky & Halle, 1968), where the underlying phonological representation of [ʤəræf] was postulated to be /giræffe/! There is also more recent support for this notion. Recall from the introduction, section 1.3, results from a series of pseudo-homophone judgment experiments by Taft (2006) also suggested that underlying phonological representations are more abstract than traditionally assumed.

This hypothesis about a more detailed underlying phonological representation of words certainly can account for the interesting effects seen throughout this dissertation. Specifically, for the duration differences seen here, one might consider that the phonological representation of 'hit' is something like /h/ /ɪ/ /t/, and the representation of 'mitt' something like /m/ /ɪ/ /t/ /t/, or /hɪt/(2 morae) and /mɪt/(3 morae).

Recall, however, there is an alternative explanation of the effects shown above. Rather than directly influencing underlying phonological representations, orthography might affect a step in the derivation of a surface phonetic form. In theoretical phonology, alternations in the surface structure of words are not usually considered part of the underlying representation. For example, in English there is a process of vowel reduction pre-consonantally which sometimes results in elision in higher frequency words

(Fidelholtz, 1975 and Hopper, 1976). The difference is exemplified between *memory* [mɛmri] and *mammory* [mæmori]. Nonetheless, the underlying phonological representation is still considered to contain that elided vowel. Another example of phonetic content of an output form which is usually considered to be represented differently at the level of underlying phonological representation is obstruent voicing alternations. In German, among many other languages, there is a phenomenon of obstruent final devoicing, where any word-final obstruent surfaces as voiceless. This results in surface homophones but supposedly different underlying phonological representations ([ʀat] /ʀad/ 'wheel' vs. [ʀat] /ʀat/ 'advice'); the difference emerges phonetically in non-word-final contexts.

Following this type of analysis of surface variations not being represented in the underlying phonological representation of a word, but rather somewhere else in the representation of the word, one might consider that the duration differences reported here which correspond to orthographic characteristics of sounds and words are similarly just surface alternations. To claim this, one would need to postulate that the phonological process is affected by the orthographic characteristics of the word during the derivation. This type of alternation however is critically quite different from the types of surface variation mentioned above. Derivational and Optimality Theoretic accounts of phonological rules or constraints on the surface forms of words operate with two sometimes conflicting motivations. One motivation is to make the word easy to say, and the other is to make the word easy to understand/hear. Some examples of the former might be a prohibition against word-final voiced obstruents, an assimilatory process where adjacent consonants share place features. An example of the latter is the aspiration of stressed syllable onsets in English. These types of rules or constraints typically don't refer back to information contained in the lexical entry of the word. In order for the orthographic characteristics of a word to have an impact on its surface structure through a phonological rule or constraint, that rule or constraint would have to access the information contained in the lexical entry. This type of reference doesn't have a huge precedent in phonological theory.

The idea that orthographic information impacts our productions is so unexpected for formal linguistics, that this potential is not currently reflected in models of speech production. This is quite reasonable. Why would we build into a model of speech production a capacity for access of orthographic information which we presuppose should have no influence on productions? However, given that such an effect is apparently present in natural as well as lab-elicited speech, it becomes necessary to modify our current model to take this new information into account. The next section shows some of the implications that this research has for current speech production models and differentiates between the two theoretical accounts of how orthographic information impacts our productions.

5.3 Production model implications

The current dominant model of word production is WEAVER++ by Levelt, Roelofs, and Meyer (1999). This model describes the process of speaking from intention to articulation. It is modular, with various levels of processing. The progression starts with the cognitive system as the speaker desires to say something. The semantic characteristics of the items of meaning in what is to be said trigger the activation or selection of a Lemma. Lemma is the technical term for the semantic and syntactic lexical entry. This in turn causes the Lexeme for the selected word to be activated. Lexeme is the technical term for the phonological lexical entry. At this point the underlying phonology of a word is accessible, and this underlying representation is forwarded to the articulating output process. At this point surface alternations driven by phonological rules, syntactic position, and other phonetic features not stemming from underlying representations, are applied. The last step is the actual articulation of the word. See figure 5.1 below for a schematic of that model.

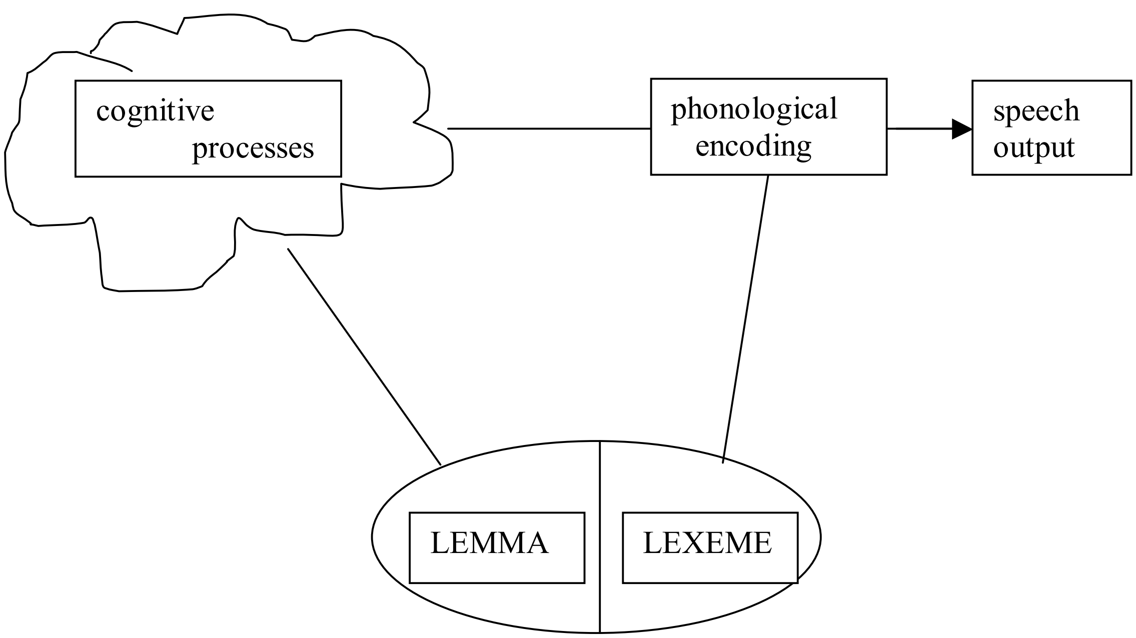

Figure 5.1: A simple schematic of a speech production model based on Levelt (1992).

What is interesting about a model of production like this is that there is no built in component to allow for orthography (the spelling of a word) to impact the production of that word. This is not surprising however, because until now there has been little overt evidence for the influence of orthography on production. After all, writing systems have been established to allow us to write, not to speak. The discussion above can inform our understanding of where such a component might be added to such a model to account for the orthographic effects found here. Two options are available. The first option, described above, is that the orthographic information impacts the phonological representation at a point before the UR is sent to the articulating processes in a process called phonological encoding. This would mean that the underlying phonological representation of words is different, more detailed, than previously considered. The second option follows also from the previous discussion. The phonological constraints or rules which affect the UR and cause the various surface variations which are currently in the model part of the articulating process must be sensitive to orthographic information located in the lexeme and modify the output according to the influence of particular orthographic characteristics. Consider the graphic in figure 5.2 detailing where these two options are located with respect to the model detailed in the previous figure.

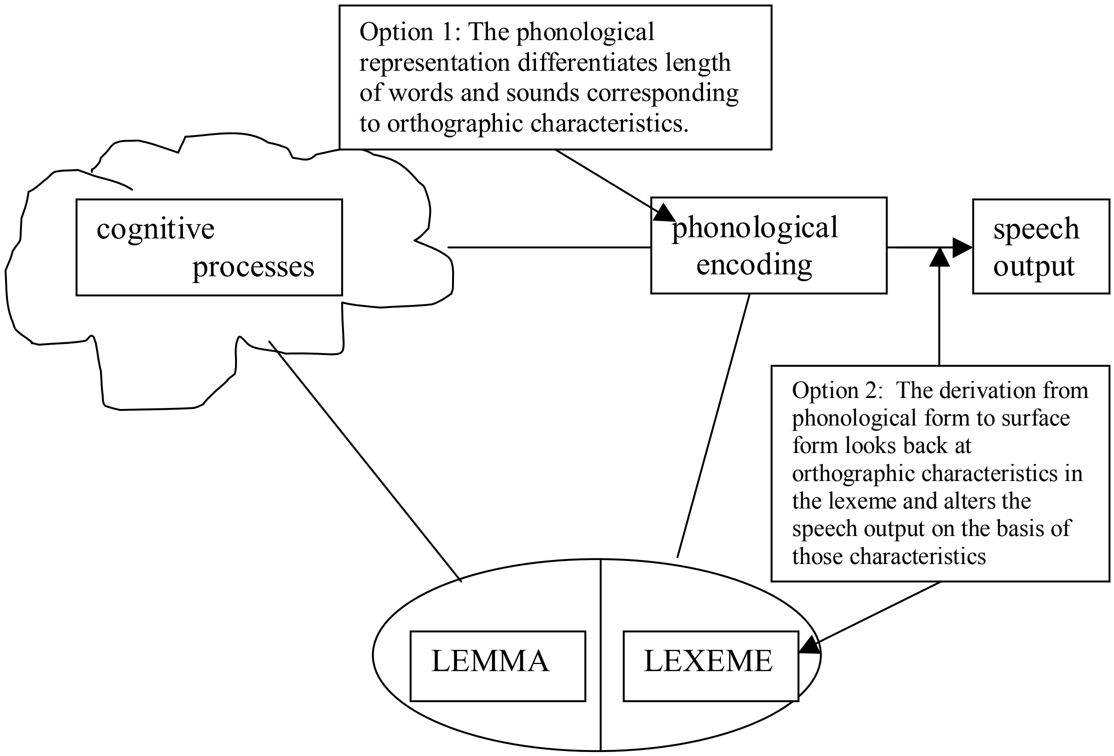

Figure 5.2: A speech production model graphic showing where the two theoretical possibilities for the effect of orthography on speech output reside.

The two potential modifications to the production model are according to the data presented here empirically equivalent. However, we have already seen independent motivation for Option 1, that is, an influence of orthography on the nature of underlying phonological representations of a word, in work by Taft (2006), in a non-production task and by Chomsky and Halle (cf. section 5.2 above). Although there appears to be independent motivation for Option 1, the data presented in this dissertation do not advocate one option over the other. A better understanding of these options, and a clearer picture of where this orthographic effect is located with respect to the rest of the production processes must await further research, to some directions of which we now turn.

5.4 Research Directions

The question of the locus of orthographic effects in speech production is not the only interesting issue raised by the results presented here and which future research may help disambiguate. One very interesting aspect of the current research and deserving of more attention is that it studies orthographic effects on productions of words for literate speakers of a literate society. It would certainly be illuminating to examine the productions of words by non-literate speakers to see if similar types of durational differences corresponding to orthographic characteristics are present in their speech as well[14]. The intuitive hypothesis, if these durational differences are indeed caused by and/or perpetuated by the orthographic information contained in lexical entries, is that then they should not be present in the speech of illiterate speakers simply because that orthographic information is presumably not present in their lexical entries.

However, interestingly enough, a study by Ohala and Zamuner (1999) has suggested that orthographic effects can be found in preliterate children in syllable segmentation tasks. Specifically, in a pause-break task where children are asked to pause between the two syllables of a bi-syllabic word, children divide syllables differently depending on the different orthographic characteristics of a word-medial consonant. Items were tested where the medial consonant was spelled with one or two identical

[14] On a similar note, it may be of use to examine these effects as a function of literate English speakers reading ability.

letters. A pair differing on this criterion for instance is *cabin* vs. *cabbage*. Words with double letters were treated differently than words with just one letter. Specifically, the preliterate children paused between two instances of the sound represented by the doubled letter, but did not repeat the letter's sound in the second syllable in words with a single letter (e.g. /kæ/ pause /bɪn/ vs. /kæb/ pause /bɪʤ/). If the children are indeed preliterate and have no explicit orthographic reference in their lexical entries for these words, the question remains how this orthographic information has come to impact their language production.

One explanation is that the orthographic effects found in literate people's productions of words (a few of which have been described in this dissertation) are prevalent and regular enough that children are sensitive to them as they acquire their native language. This would mean that the effects are of sufficient size and prominence that they are perceptible by the child's linguistic cognitive system. As a consequence of noticing phonetic reflexes of orthography in literate adult speech, the children incorporate this phonetic variation into their lexical representations of words. Critically, they do this prior to acquiring the orthographic representations for those words. This would mean that the effects attributed thus far to orthographic characteristics of words would be present in the production of children whose synchronic knowledge of English doesn't explicitly contain orthographic information. This gives rise to an interesting related question, one which is perhaps the most interesting question posed by the results described in this dissertation.

Specifically, are these durational effects noted in speech production truly the result of the orthographic characteristics of words and sounds within words? Or, are these two pieces of information merely correlated with one another? This would mean that there is some other driving force causing the durational differences between sounds and words, and the orthographic characteristics simply happen to line up with that other variable. The hypothesis stated at the beginning of this research was that orthographic characteristics, such as the number of letters used to spell a sound, determine durational characteristics of that sound. The results shown so far support that hypothesis, but do not distinguish that analysis from the alternative where some other factor correlated with the orthographic characteristics of English words is actually causing the durational differences noted.

So, are the effects shown here actually effects of orthography, effects which continue to exist in English because its speakers are literate? Or, is the effect simply an artifact of some other perhaps historical reason which happens to be reflected in orthography? For instance, English orthography currently contains letters which used to correspond to sounds which were fully articulated in the speech of English speakers (*knight* - [knɪχt]). Perhaps these sounds dropped out of consciously perceptible speech, but left artifacts of their past presence in various phonetic features, like exaggerated length, for instance, of the consciously perceptible sounds around them. We may even consider other characteristics of the English orthography system, like the relative paucity of tense vowels before consonants represented by double letters, or the prevalence of double ‘l’ word finally over single ‘l’.

If the cause of these effects is the first alternative alone, then if a subset of speakers stopped reading, raised their children without reading, and only interacted with other non-literate people, then we would expect the effect to eventually disappear. If the second, or some combination of the two, then we might expect that this effect continue to linger. This is not however a particularly easy (or ethical) experiment to conduct. Fortunately, there are perhaps some other potential ways to test these two causal alternatives.

One potential experiment could test production differences between words whose diachronic orthographic changes diverged. Specifically, one might examine pairs of individual lexical items which both lost an historic English sound, but for whatever reason aren't currently spelled the same way. One word may retain the letters once used to represent that sound, while the other word lost those letters in addition to the sound. If there are durational differences between the sounds of those two words, and the durations of sounds or words represented synchronically with more letters are longer than those missing the letters, then this might be support for the hypothesis that the orthographic characteristics of sounds or words is indeed contributing independent of any historical correlations to phonetic variation.

Alternatively, one could examine two different languages whose orthographies span the deep-to-shallow continuum. Italian is an example of a language with a relatively shallow orthography, a language whose orthographic representations of words are nearly one to one with the sounds contained in the words. French, conversely, has a deep orthography whereby many letters in the orthographic representations of words are not considered to have any synchronic phonological correspondents, but which represent sounds that did exist in older varieties of French. Words that in the two languages have undergone sound changes would thus have different orthographic realizations. While a language with a shallow orthography would also have lost the orthographic symbols which represent those lost sounds, a language with a deeper orthography would retain those letters. Differences in production of lexical pairs in both languages may also shed light on the question of coincidental or causal orthographic characteristics effects on productions.

In addition to the question of causality vs. correlation in the results found here and attributed to orthography, there are other veins of research that can be explored. These experiments and corpus analyses show an effect on production of a facet of lexical representation which according to standard views of production shouldn't matter. Other domains of lexical representation which do not intuitively have phonetic reflexes in production are the syntactic or semantic characteristics of words. In a study of plural marking in O'odham, Zepeda & Hill (1992) found that the phonological realization of a plural reduplicative morpheme varied with respect to the length of the syllable's vowel. Interestingly, an explanation of this variation could not be found in the phonological structure of the base, or even the morphological characteristics. Rather, Zepeda & Hill argue that the length of the vowel in the plural reduplicative morpheme is determined by the culturally informed categorization of the noun as protruding or bumpy. Protruding or bumpy nouns (within a given culture) get plural reduplicative morphemes with a long vowel and the less prominent or flat nouns (within that culture) get a plural reduplicative morpheme with a short vowel. This type of analysis for surface variations is quite rare. Following this precedent however, it would be interesting to examine what phonetic effects certain aspects of syntactic or semantic representation might have in speech production.

In conclusion, this dissertation finds effects of orthographic information on speakers' productions of words, a finding that is not currently accounted for in theories of speech production. This finding, though suggested by previous research, is confirmed in a far more compelling way here. These effects raise many interesting questions, particularly regarding the impact of literacy, and determining the real 'cause' behind the effects (if not exclusively orthography). It will be very enlightening to our understanding of the complexity of structure and interactions of language in our brains to continue exploring the unexpected effects of aspects of lexical representation in speech production.

Appendix A: Experiment One materials

k	t	P	F	s	z	ʧ
arc	date	rapt	draft	bass	has	rich
mark	wait	rapped	graphed	vase	jazz	switch
monarch	state		laughed	lace		which
ache	weight		staffed	concise		
fake	mitt		raft	lice		
make	It		miffed	fuss		
attack	grit		rift	us		
cognac	height		stuff	Gus		
Kodiak	spite		tough	lost		
Iraq				frost		
ransack				tossed		
click						
clique						
tic						
hike						
like						
mic						

Appendix B: Picture and definition stimuli for Experiment 3

a silicon bowl scraper

a vegetable used primarily in Thai recipes

an endangered type of lemur

a hard cheese breaker

a type of spice

a musical instrument from the historical US

Appendix C: Perl Program for Extraction of the Buckeye Data-set

```
use strict;

#vowels
my $vowels = "(iy|ih|ay|ah|ae|aa|ao|ey|eh|er|uh|uw|U|ow)";

my @filelist;

my @dirs = `dir`;

#get all the file names
foreach my $dir (@dirs) {
	if ($dir =~ /Buckeye S/) {
		$dir =~ s/^.*(Buckeye.*)$/$1/;
		chomp $dir;
		my @subdirs = `dir \"$dir\"`;
		foreach my $subdir (@subdirs) {
			if ($subdir =~ /s\d\d\d\d/) {
				chomp $subdir;
				$subdir =~ s/^.*(s\d\d\d\d.)$/$1/;
				push @filelist, "$dir\\$subdir\\$subdir\.words";
			}
		}
	}
}

#use this loop to iterate over all the filenames;
foreach my $filename (@filelist) {


	#open the file
	open F, "$filename" or die "Can't open $filename\n";

	#save the subject number
	my $prefix = $filename;
	$prefix =~ s/^.*(.....)\.words/$1/;

	#flag to eliminate metadata
	my $hatchflag = 0;

	#is the word polysyllabic
	my $poly;

	#keep track of durations
```

```
        my $lastnumber = 0;

        while (my $line = <F>) {
                #we're past the metadata
                if ($hatchflag) {
                        chomp $line;
                        my @flds = split / +/, $line;
                        #for checking monosyllabicity
                        my @parts = split /;/, $line;
                        #calculate the duration of the current word
                        my $dur = $flds[1] - $lastnumber;
                        $flds[3] =~ tr/;//d;
                        #print if a word ending in an obstruent
                        if ($flds[3] =~ /^[a-zA-Z]/ && $flds[$#flds] =~ /^([ptksf]|ch|sh)$/) {
                                if ($parts[2] =~ /$vowels.*$vowels/) {
                                        $poly = 'yes';
                                } else {
                                        $poly = 'no';
                                }
                                print "$prefix\t$flds[1]\t$flds[3]\t$flds[$#flds]\t$dur\t$poly\n";
                        }
                        $lastnumber = $flds[1];
                }
                #marks the end of the metadata
                if ($line =~ /#/) {
                        $hatchflag = 1;
                }
        }

        close F;

}
```

Appendix D

An exhaustive list of all monosyllabic function words ending in a voiceless obstruent.

AREN'T
AT
BACK
ELSE
EACH
BUT
CAN'T
DON'T
EIGHT
FIRST
GET
GETS
GOT
HENCE
IF
IT
ITS
IT'S
LAST
LESS
MIGHT
MOST
MUCH
MUST
NEXT
NOT
OFF
ONCE
OUGHT
OUT
QUITE
SHAN'T
SINCE
SIX
SUCH
THAT
THAT
THAT'S
THENCE
THIS
THRICE

THUS
TWICE
UP
US
WEREN'T
WHAT
WHENCE
WHICH
WON'T
YES
YET

http://www.marlodge.supanet.com/museum/funcword.html

References

Bell, A., Jurafsky, D., Fosler-Lussier, E., Girand, C., Gregory, M., & Gildea, D. (2003). Effects of disfluencies, predictability, and utterance position on word form variation in English conversation. *Journal of the Acoustic Society of America 113* (2).

Boersma, P. & Weenink, D. (2005). Praat: doing phonetics by computer (Version 4.3.19) [Computer program]. Retrieved April 20, 2005, from http://www.praat.org/

Byrd, D., (1993) 54,000 American Stops. *UCLA Working Papers in Linguistics. 83* 97-117.

Creel, S., Tannenhaus, M., & Aslin, R. (2006). Consequences of Lexical Stress on Learning and Artificial Lexicon. *Journal of Experimental Psychology. 32,* 15-32.

Crystal, T.H.; House, A.S. (1988a) Segmental durations in connected-speech signals: Current results. *JASA 83***,** 1553-1573.

Crystal, T.H.; House, A.S. (1988c) The duration of American-English stop consonants: an overview. *Journal of Phonetics 16(3),* 285-294.

Crystal, T. H., and House, A. S. (1990). ‘‘Articulation rate and the duration of syllables and stress groups in connected speech,’’ *J. Acoust. Soc. Am. 88*, 101–112.

Damian, M., & Bowers, J. (2003). Effects of orthography on speech production in a form preparation paradigm. *Journal of Memory and Language, 49,* 119-132.

Dell, G. S. (1988). The retrieval of phonological forms in production: Tests of predictions from a connectionist model. *Journal of Memory and Language, 27*, 124-142.

Dell, G. S. (1990). Effects of frequency and vocabulary type on phonological speech errors. *Language and Cognitive Processes, 5*, 313-349.

Dijkstra, T., Roelofs, A., & Fieuws, S. (1995). Orthographic effects on phoneme monitoring. *Canadian Journal of Experimental Psychology, 49*(2), 264-271.

Dinnsen, D., & Chrles-Luce, J. (1984). Phonological neutralization, phonetic implementation and individual differences. *Journal of Phonetics, 12,* 49-60.

Fidelholz, J. (1975). ‘‘Word frequency and vowel reduction in English,’’ in *CLS-75* (University of Chicago, Chicago), pp. 200–213.

Fourakis, M., & Iverson, G. (1984). On the incomplete neutralization of German final obstruents. *Phonetica, 41*, 140-149

Greene, William H. (2003). Econometric Analysis. Pearson Education International. 5th Edition. pp 9-43.

Halle, P., Chereau, C., & Segui, J. (2000). Where is the /b/ in "absurde" [apsyrd]? It is in French listener's minds. *Journal of Memory and Language, 43,* 618-639.

Hassan, Z. M. (2003). Temporal compensation between vowel and consonant in Swedish & Arabic in sequences of CV:C & CVC: and the word overall duration. *Phonum 9* (2003), 45-48.

Hooper, J. B. (1976). Word frequency in lexical diffusion and the source of morphophonological change. In W. Christie (Ed.), *Current progress in historical linguistics* (pp. 96-105). Amsterdam: North Holland.

Jakimik, J., Cole, R., & Rudnicky, A. (1985). Sound and spelling in word recognition. *Journal of Memory and Language, 24,* 165-178.

Johnson, K. (2006). Massive Reduction in conversational American English. (ms)

Jurafsky, D. (1996). A probabilistic model of lexical and syntactic access and disambiguation. *Cognitive Science, 20*, 137-194.

Jurafsky, D., Bell, A., Gregory, M., & Raymond, W. (2000). Probabilistic Relations between Words: Evidence from Reduction in Lexical Production. In Bybee, Joan and Paul Hopper (eds.). *Frequency and the emergence of linguistic structure.* Amsterdam: John Benjamins.

Klatt, D. H. (1975). ''Vowel lengthening is syntactically determined in a connected discourse,'' *Journal of Phonetics 3*, 129–140.

Klatt, D.H. (1976). Linguistic uses of segmental duration in English: Acoustic and perceptual evidence. *Journal of the Acoustical Society of America* 59, 1208-1221.

Ladd, D. R., & Campbell, N. (1991). ''Theories of prosodic structure: Evidence from syllable duration,'' in *Proceedings of the 12th International Congress of Phonetic Sciences*, Aix-en-Provence, France, pp. 290–293.

Levelt, W. (1992). Accessing words in speech production: Stages, processes and representations. *Cognition 42*, 1-22.

Levelt, W., Roelofs, A., & Meyer, A. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences, 22*, 1-75.

Luce, P.A.; Charles-Luce, J. (1985) Contextual effects on vowel duration, closure duration, and the consonant/vowel ratio in speech production. *JASA* , *78*, 1949-1957.

Lupker, S. (1982). The role of phonetic and orthographic similarity in picture-word interference. *Canadian Journal of Psychology, 36,* 349-376.

McClelland, J, & Elman, J, (1986). The TRACE model of speech perception. *Cognitive Psychology, 18*, 1-86.

McRae, K, de Sa, V, & Seidenberg, M (1997). On the nature and scope of featural representation of word meaning. *Journal of Experimental Psychology, 126*, 99-130.

McRae, K., Spivey-Knowlton, M. J., & Tanenhaus, M. K. (1998). Modeling the influence of thematic fit (and other constraints) in on-line sentence comprehension. *Journal of Memory and Language, 38*, 283-312.

Miller, G (1991). *The science of words.* New York: Scientific American.

Morais, J., Content, A., Cary, L., Mehler, J., & Segui, J. (1989). Syllabic segmentation and literacy. *Language and Cognitive Processes, 4*, 57-67.

Muneaux, M., & Ziegler, J. C. (2004). Locus of orthographic effects in spoken word recognition: Novel insights from the neighbour generation task. *Language and Cognitive Processes, 19*(5), 641-660.

Oldfield, R., C., & Wingfield, A. (1965). Response latencies in naming objects. *Quarterly Journal of Experimental Psychology, 17*, 273-281.

Perfetti, C., Wloltko, E., & Hart, L. (2005). Word Learning and Individual Differences in Word Learning Reflected in Event-Related Potentials. *Journal of Experimental Psychology. 31,* 1281-1292.

Pitt, M.A., Dilley, L., Johnson, K., Kiesling, S., Raymond, W., Hume, E. and Fosler-Lussier, E. (2007) Buckeye Corpus of Conversational Speech (2nd release) [www.buckeyecorpus.osu.edu] Columbus, OH: Department of Psychology, Ohio State University (Distributor).

Port, R., & O'Dell, M. (1985). Neutralizintion of syllable-final voicing in German. *Journal of Phonetics, 13,* 257-282.

Rayner, K., Warren, T., Juhasz, B., & Liversedge, S. (2004). The effect of Plausibility on Eye Movements in Reading. *Journal of Experimental Psychology. 30,* 1290-1301.

Rhodes, R. A. (1992). ‘‘Flapping in American English,’’ in *Proceedings of the 7th International Phonology Meeting*, edited by W. U. Dressler, M. Prinzhorn, and J. Rennison (Rosenberg and Sellier, Turin), pp. 217–232.

Rhodes, R. A. (1996). ‘‘English reduced vowels and the nature of natural processes,’’ in *Natural Phonology: The State of the Art*, edited by B. Hurch and R. A. Rhodes (Mouton de Gruyter, The Hague), pp. 239–259.

Savin, H. B. (1963). Word-frequency effects and errors in the perception of speech.. *Journal of Acoustical Society of America, 35,* 200-206.

Seidenberg, M.S., & Tannenhaus, M.K. (1979). Orthographic effects on rhyme monitoring. *Journal of Experimental Psychology: Human Learning and Memory*, 5, 546-554.

Shibahara, N., Zorzi, M., Hill, M., Wydell, T., & Butterworth, B. (2003). *Journal of Experimental Psychology. 56A,* 263-286.

Slowiaczek, L., & Dinnsen, D. (1985). On the neutralizing status of Polish word-final devoicing. *Journal of Phonetics, 13*, 325-341.

Slowiaczek, L. M., Soltona, E. G., Wieting, S. J., & Bishop, K. L. (2003). An investigation of phonology and orthography in spoken-word recognition. *Quarterly Journal of Experimental Psychology, 56A*(2), 233-262.

Strain, E., Patterson, K., & Seidenberg, M. (1995). Semantic Effects in Single-Word Naming. *Journal of Experimental Psychology. 21,* 1140-1154.

Taft, M., & Hambly, G. (1985). The influence of orthography on phonological representations in the lexicon. *Journal of Memory and Language, 24*, 320-335.

Taft, M. (2006). Orthographically influenced abstract phonological representation: Evidence from non-rhotic speakers. *Journal of Psycholinguistic Research, 35*, 67-78.

Van Sohn, R. & Pots, L. (1999). An acoustic description of consonant reduction. *Speech Communication, 28*(9), 125-140.

Ventura, P., Kolinsky, R., Brito-Mendes, C., & Morais, J. (2001). Mental representations of the syllable internal structure are influenced by orthography. *Language and Cognitive Processes, 16*(4), 393-418.

Vitevitch, M. S. (2002a). Naturalistic and experimental analyses of word frequency and neighborhood density effects in slips of the ear. *Language & Speech, 45,* 407-434.

Vitevitch, M. S., & Luce, P. A. (1998). When words compete: Levels of processing in spoken word perception. *Psychological Science, 9,* 325-329.

Vitevitch, M. S., & Luce, P. A. (1999). Probabilistic phonotactics and spoken word recognition. *Journal of Memory and Language, 40,* 374-408.

Vitevitch, M. S., Luce, P. A., Pisoni, D. B., & Auer, E. T. (1999). Phonotactics, neighborhood activation and lexical access for spoken words. *Brain and Language, 68,* 306-311.

Warner, N., Jongman, A., Sereno, J., & Kemps, R. (2004). Incomplete neutralization and other sub-phonemic durational differences in production and perception: evidence from Dutch. *Journal of Phonetics,* 32, 251-276.

Whalen, D. H. (1991). Infrequent words are longer in duration than frequent words. *Journal of the Acoustical Society of America, 90*(4), 2311.

Wheeldon, L., & Monsell, S. (1992). The locus of repetition priming of spoken word production. *Quarterly Journal of Experimental Psychology A, 44*, 723-761.

Zamuner, T.S., & Ohala, D. (1999). Preliterate children's syllabification of intervocalic consonants. In A. Greenhill, H. Litterfield, & C. Tano (Eds.), *Proceedings of the 23rd Boston University Conference on Language Development Vol. 2.* (pp. 753-763). Somerville, MA: Cascadilla Press.

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

Zipf, G. K. (1929). ''Relative frequency as a determinant of phonetic change,'' *Harvard Studies in Classical Philology. 15,* 1–95.