ORTHOGRAPHIC QUALITY IN ENGLISH AS A SECOND LANGUAGE

by

Susan Dunlap

Bachelor of Arts in Psychology, University of Denver, 2002

Master of Science in Cognitive Psychology, University of Pittsburgh, 2006

Submitted to the Graduate Faculty of

the Kenneth P. Dietrich School of Arts and Sciences

in partial fulfillment

of the requirements for the degree of Doctor of Philosophy

University of Pittsburgh

2012

UNIVERSITY OF PITTSBURGH

KENNETH P. DIETRICH SCHOOL OF ARTS AND SCIENCES

This dissertation was presented

by

Susan Dunlap

It was defended on

August 2, 2012

and approved by

Julie Fiez, Professor, Department of Psychology

Natasha Tokowicz, Associate Professor, Department of Psychology

Alan Juffs, Associate Professor, Department of Linguistics

Dissertation Director: Charles A. Perfetti, University Professor, Department of Psychology

ORTHOGRAPHIC QUALITY IN ENGLISH AS A SECOND LANGUAGE

Susan Dunlap, PhD

University of Pittsburgh, 2012

Learning new vocabulary words in a second language is a challenge for the adult learner, especially when the second language writing system differs from the first language writing system. According to the lexical quality hypothesis (Perfetti & Hart, 2001), there are three constituents to word-level knowledge: orthographic, phonological, and semantic. A set of studies investigated the nature of orthographic knowledge in advanced learners of English as a second language. In a data mining study, students' spelling errors were analyzed. Results showed that first language background and second language proficiency have an effect on the rates and types of spelling errors made. In two training interventions, students showed learning gains from two different types of spelling instruction: a form focus condition and a form-meaning integration condition (Norris & Ortega, 2000). In a separate audio dictation task, non-native English speakers were shown to be sensitive to word frequency and age of acquisition but not regularity. In a cross-modal matching task, the same students were most susceptible to transposition foils that preserved target letters but in an incorrect order, and least susceptible to phonological foils that preserved phonological but not orthographic form of the target word. In a spell checking task, students had more difficulty rejecting misspelled words that maintained the phonological form of the target word than misspelled words that did not preserve phonology of the target. Overall, findings suggest that intermediate to advanced learners of English as a second language still show difficulty with the language's deep orthography, but that they can benefit from minimal amounts of instruction. Furthermore, these students appear to be acquiring orthographic

knowledge via exemplar-based rather than rule-based strategies. This research expands upon the lexical quality hypothesis and finds support for the arbitrary mapping hypothesis.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

PREFACE

This work was supported by the Pittsburgh Science of Learning Center, and has been funded by the National Science Foundation, awards #SBE-0354420 and #SBE-0836012. Data were collected with approval from Institutional Review Boards both at Carnegie Mellon University and at the University of Pittsburgh. I gratefully acknowledge the guidance of my dissertation chair, Charles A. Perfetti, and committee members, Julie Fiez, Natasha Tokowicz, and Alan Juffs. Yinz taught me more than this document can reflect.

I also wish to thank the students and teachers of the English Language Institute for their participation. For help with task preparation, data collection, and analyses, I thank Ben Friedline, Carol Harmatz, Rob Mucklo, Krista Williams, Kelly Kirk, Lindsay Harris, Jeanine Sun, and especially Hannah Bartle. For other technical support, I thank Lisa Fisher and Claire Bradin Siskin.

Many people in Pittsburgh provided moral support along the way. For unforgettable kindnesses, I thank Pauline Kraly, Shelly Tavis, Jo Bodnar, Queenie Kravitz, Barb Kucinski, Keiko Koda, David Klahr, Martin Greenberg, Mim Schwartz, Annie Robertson, Judy Callan, Mike Greenwald, Mary Grace Fitzgerald, Ginger Boneysteele, Helen Boyce, Baoguo Chen, Juan Zhang, wǒ de xǔe mèi Li-yun Chang, wǒ de mèi mèi Zhao Yun, and my most excellent colleague Derek Ho Leung Chan. Special thanks to Craig Ferguson, the best non-academic reason for visiting Greensburg, and to Sidney Crosby, the best non-academic reason for temporarily residing in Pittsburgh. And of course, for many reasons, God bless Andrew Carnegie.

I am most thankful for the support of my friends and family back home: Deb, Michael, Nick, Lora, and Ellie; Kimball, Rich and Tom; Karen, George, and Allison, Rachel and Justin, and of course Byron, Amanda, and Angela.

1.0 OVERVIEW

Learning to read and write in a second language is a challenge for the adult learner, especially when the second language (L2) writing system differs from the first language (L1) writing system (Red, 1999). The student must learn new decoding rules and form new mappings among orthographic, phonological, and semantic representations. English, in particular, has a deep orthography (Frost, Katz, & Bentin, 1987), and presents particular challenges for reading and spelling.

Reading involves two key steps: word recognition and word integration (Fender, 2001). Word recognition requires identifying a printed word form and retrieving meaning and syntax. Word integration requires using semantic and syntactic information of individual words to build a meaningful discourse. According to the lexical quality hypothesis (Perfetti & Hart, 2001), the quality of lexical representations at the word recognition step will affect reading comprehension at the word integration step. We know about some lexical factors (e.g., frequency, age of acquisition, concreteness) that affect word identification for native English speakers. However, less is known about what factors affect the quality of orthographic, phonological, and semantic representations for non-native English speakers.

As is the case for poor decoders of English as a native language (Landi, Perfetti, Bolger, Dunlap, & Foorman, 2006), learners of English as a second language (ESL) can get by with underspecified orthographic representations when reading. They can do so by relying on context, background knowledge, word shape, or other cues for successful decoding. But producing correct word spellings is a more demanding task that can reflect a student's overall word knowledge better than most oral tasks or receptive tasks (Wade-Woolley & Siegel, 1997). Spelling requires more robust knowledge of sound-spelling rules and, in a deep orthography, word-specific exceptions. The aim of this set of studies was to explore how adult learners of English as a second language (ESL) develop and utilize orthographic representations in English.

I used a three-pronged approach to investigating ESL spelling: a data mining study, two in-class interventions, and three behavioral experiments. The goals of the data mining study were to confirm teacher observations about L1 background effects on student spelling in L2 English, and to identify potential target words for the subsequent spelling interventions. The goals of the spelling interventions were to determine which type of instruction—form focus or form-meaning integration—would lead to better learning gains, and to investigate effects of L1 background and L2 proficiency on gains from spelling interventions. The goals of the behavioral experiments were to investigate effects of lexical characteristics on orthographic representations in ESL, and to test a proposed model of the cognitive mechanisms of spelling.

1.1 FRAMEWORK

In language instruction, much emphasis is placed on learning new vocabulary words (e.g., Nation, 2001; VanPatten, Farmer, & Clardy, 2009). But what, exactly, does it mean to *know* a word? According to the lexical quality hypothesis (Perfetti & Hart, 2001), there are three constituents to word-level knowledge: orthographic, phonological, and semantic (see Figure 1). The orthographic constituent refers to the written form of a word. In English, as in other alphabetic languages, this means knowledge of word spellings. The phonological constituent refers to the spoken form of the word, that is, how a word is pronounced. The semantic constituent refers to the meaning of the word and how it is used in context.

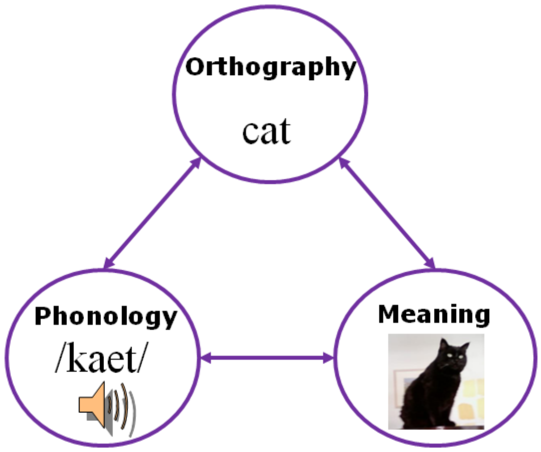

Figure 1. Constituents of word knowledge according to the lexical quality hypothesis (Perfetti & Hart, 2001).

For each individual, the strength of each constituent can vary from word to word. Furthermore, the strength of the connections among constituents may vary across words. As an example, a person might know the meaning and pronunciation of the word “definitely” but be unsure how to spell it. Or, a reader may encounter the word “awry” and be able to glean the meaning from context but mistakenly pronounce it as “AW-ree.”

The relative strength or weakness of each of the three lexical constituents, as well as the connections among them, has an impact on reading comprehension (Perfetti & Hart, 2001). The lexical quality hypothesis was first developed to account for reading performance of higher-skilled and less-skilled native English readers. It made predictions about how quickly readers of different abilities resolve lexical level ambiguity, homophony, and homography during reading.

If underdeveloped lexical representations negatively affect reading comprehension, it should follow they would have a negative effect on writing and spelling as well. Language production tasks are almost always more difficult than language receptive tasks (e.g., Cocking & McHale, 1981; Lee & Muncie, 2006; Wade-Woolley & Siegel, 1997). Therefore, producing correct word spellings would serve as a more reliable indicator of strong lexical knowledge than word recognition (e.g., lexical decision, decoding).

The constituent model can easily be extended to non-native speakers of English as well. The same predictions would hold: namely, that weaker constituent knowledge and weaker links among constituents lead to poorer reading comprehension. Again, there are likely many factors affecting the strength of lexical representations for non-native English speakers. In addition to factors affecting native speakers (e.g., frequency, length, concreteness, imageability, age of word acquisition), factors specifically affecting non-native speakers include: at what age the learner first began learning the second language, translation equivalence between languages, a word's cognate status between English and the learner's native language. For word reading (decoding) and spelling (encoding), the regularity of grapheme-phoneme mappings in the target language is also a factor affecting strength of lexical quality.

1.2 BACKGROUND

1.2.1 Spelling Acquisition in English

There are several proposed models for how native English speakers learn to decode words (Chall, 1979; Ehri & Wilce, 1985; Frith, 1985; Gough & Juel, 1991) and to encode words (Gentry, 1982; Henderson, 1981; Henderson & Beers, 1980; Henderson & Templeton, 1986; Treisman & Bourassa, 2000) in their native language. The stages of spelling development typically begin with a pre-communicative or pre-alphabetic stage, where children have yet to discover that graphemes represent phonemes. This stage is followed by the emergence or understanding of the alphabetic principle. Later stages are usually a refinement of the language-specific phoneme-grapheme mappings.

Although this may be relevant for native learners of English, these stage models exclude adult L2 learners who already have well-formed grapheme-phoneme relationships based on their native language writing system. The similarity or difference between first and second languages can have an effect on how successfully the learner masters the second language (MacWhinney, 2005a, 2005b; Tokowicz & MacWhinney, 2005). Much less research has been devoted to determining how later learners of English as a second language form mappings among graphemes and phonemes in a deep orthography. Some even question whether it is worthwhile to teach spelling during second language instruction (Polak & Krashen, 1988). And if it is worthwhile, what instructional methods would lead to the best learning gains?

The current set of studies aimed to investigate how ESL students represent the links among orthographic, phonological, and semantic constituents of English L2 words (Perfetti & Hart, 2001).

1.2.2 Effects of Word Frequency and Regularity

Much research has been done investigating the effects of word frequency and sound-spelling regularity for native English processing tasks such as reading comprehension, word naming, and lexical decision (e.g., Macizo & Van Petten, 2006; Morrison & Ellis, 1995; Rice & Robinson, 1975), as well as the neural activation during such tasks (e.g., Al-Hamouri et al., 2005; Fiez, Balota, Raichle, & Petersen, 1999). Many of these findings have been replicated in other native languages, such as Spanish (Conrad, Carreirras, Tamm, & Jacobs, 2009) and French (Lété, Peereman, & Fayol, 2008). However, less is known about how these factors affect the way that non-native speakers of English develop, utilize, and produce orthographic representations (Nassaji, 2005). To what extent are ESL students sensitive to frequency and regularity in terms of developing and using orthographic representations?

High frequency of written and spoken exposure is predicted to improve quality of individual lexical entries. Furthermore, just as is the case for native speakers, frequency of written input ought to help learners determine sublexical patterns and distributional properties (Krashen, 1989). Assuming students are paying attention to the input (Schmidt, 1990) at the sublexical level and detecting patterns or rules, then orthographic representations would show effects of word frequency and regularity (e.g., Martinet, Valdois, & Fayol, 2004). Then it also follows that more proficient non-native speakers might show greater frequency effects. Being exposed to a greater range of types and tokens of words could allow a more advanced learner to form analytic decoding—and eventually encoding—strategies. Therefore, word frequency is hypothesized to have an effect on spelling in English as a second language.

Word regularity is an important lexical factor in deep, or opaque, orthographic writing systems. English has a deep orthography (Frost, Katz, & Bentin, 1987), meaning sound-spelling mappings are not consistent. Arabic, Korean, and Spanish writing systems are shallower orthographies compared to English. Orthographic depth has been shown to have effects on literacy acquisition (Seymour, Aro, & Erskine, 2003; Spencer, 2007), development of phonological awareness (Goswami, Ziegler, & Richardson, 2005), and spelling acquisition (Caravolas, 2004). Because of the unpredictable mappings from phoneme to grapheme in encoding, and from grapheme to phoneme in decoding, English words can range on a continuum from highly regular (e.g., “cat”) to highly irregular (e.g., “yacht”). Dual-route models of word decoding (e.g., Bates, Castles, Luciano, Wright, Coltheart, & Martin, 2007) posit separate cognitive paths for reading regular versus irregular words. Regular words can be read via an assembled route; irregular words must use an addressed route.

Frequency and regularity have been shown to interact in naming and lexical decision tasks (e.g., Balota & Ferraro, 1993; Hino & Lupker, 2000). In English, more common words are more often irregular (e.g., the, one) but can be accessed easily due to their higher frequency. Hence, irregularity has less of an effect on processing of high frequency words. Again, less is known about how this interaction might play out in non-native English speakers. Will they show the same decreased effect of regularity for higher frequency words? Also, does this interaction extend to development and usage of orthographic representations in production tasks (e.g., writing, spelling)? The current project aimed to explore these questions further.

1.2.3 Effects of Age of Acquisition

Here, age of acquisition (AoA) refers to how early a word is learned, rather than at what age a person begins to learn a language (as is used in the critical period literature, for example). Earlier learned items are hypothesized to have an advantage over later learned items. Age of acquisition has been found to predict ease of processing in lexical decision, word naming, picture naming, and semantic categorization tasks (Chen, Zhou, Dunlap, & Perfetti, 2007; Morrison & Ellis, 1995).

According to the arbitrary mapping hypothesis (Ellis & Lambon Ralph, 2000; Monaghan & Ellis, 2002a; 2002b; Zevin & Seidenberg, 2002, 2004), age of acquisition effects are susceptible to the nature of the mappings among orthography, phonology, and semantics. If the mappings are inconsistent or unpredictable, as they are in English phoneme-grapheme mappings, earlier learned items give less of a boost to learning subsequent items. This means that learning earlier-acquired words such as "have" and "one" do not help a person subsequently learn "pave" and "cone." When mappings are inconsistent, age of acquisition effects will be increased, with earlier learned items being advantaged. When mappings are highly consistent, age of acquisition effects will be reduced (Chen et al., 2007). In English, the mappings between orthography and phonology are highly inconsistent. Thus, learners of English as a second language are predicted to show sensitivity to age of acquisition, with earlier learned items being strongly advantaged over later learned items.

1.2.4 Cognitive Mechanisms of Spelling

What cognitive mechanisms are involved in retrieval of orthographic representations? Several cognitive models of spelling have been proposed to account for the underlying mechanisms of retrieving orthographic forms during verbal or written spelling tasks (Bates et al., 2007; Caramazza, Miceli, Villa, & Romani, 1987; Margolin, 1984; Miceli and Capasso, 2006; Perfetti, Rieben, & Fayol, 1997; Rapcsak, Henry, Teague, Carnahan, & Beeson, 2007). The majority of these models were based on English as the speaker's primary or only language, which is a problem for making generalizations or universal theories of lexical representations (Share, 2008; Zhou, Ye, Cheung, & Chen, 2009). First I review several of these models, and then I describe my own cognitive model, a hybrid of the relevant components of these various spelling models.

Rapcsak and colleagues (2007) developed a dual-route model to predict the spelling and reading performance of brain damaged patients with either alexia, the inability to understand written matter, or agraphia, the inability to produce written language (see Figure 2). This model allows for both written and spoken input and output. The hypothesized orthographic and phonological lexicons are separate stores, which are both separate from semantic representations. Because the model assumes English as the native language, there are two routes to accessing word constituents: one for regular and one for irregular words. As is the case in English, regularity can vary in the grapheme-to-phoneme direction as well as the phoneme-to-grapheme direction. The model accounts for the orthographic depth of English, but does not include other languages, and also does not include any mention of bilinguals or second language learners.

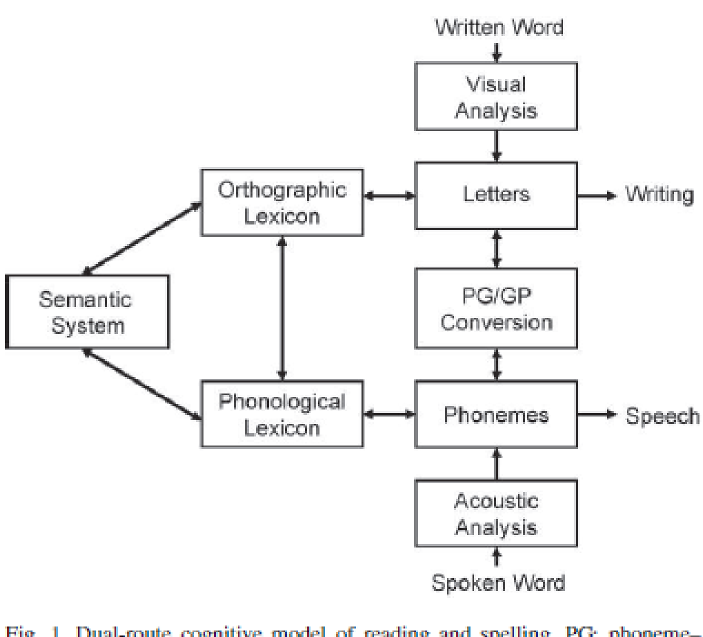

Figure 2. Dual-route cognitive model of reading and spelling, from Rapcsak et al. (2007).

Bates and colleagues (2007) also proposed a dual-route model (see Figure 3), but theirs was intended to determine the genetic and environmental predictors of spelling and reading in English. The authors conducted a test of the genetic effects on reading aloud and spelling of regular and irregular words in English. They proposed separate cognitive routes for spelling of words that follow regular phoneme-grapheme spelling rules than for irregularly spelled words. The regular word path, represented by the right side of the model, goes from phonological input directly to a rule-based phoneme-to-grapheme encoding process. The irregular word path, represented by the left side of the model, also begins with phonological input, but then goes through two steps: a phonological lexicon look-up, then an orthographic lexicon look-up. Both routes end up with the written or typed production of grapheme units. Like the Rapcsak et al. (2007) model, the Bates model has dual routes that account for both regular and irregular spellings English. Their model does not consider languages other than English, and also does not include any mention of bilinguals or second language learners.

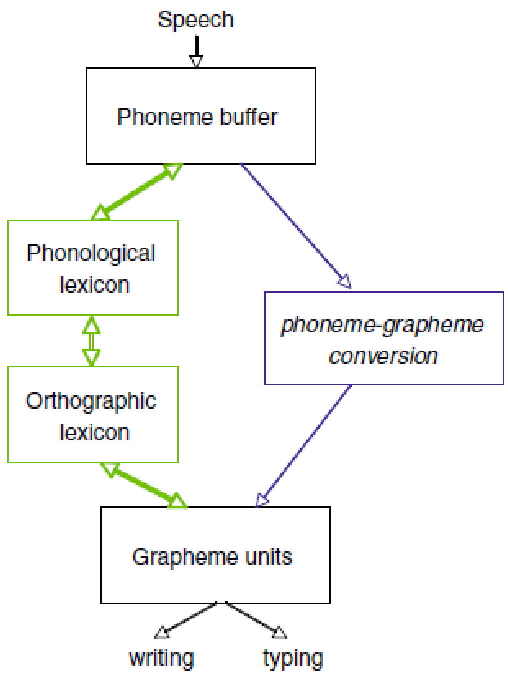

Figure 3. Dual-route cognitive model of spelling, from Bates et al. (2007).

Caramazza and colleagues (1987) developed a model derived from a neuropsychological case-study of an individual with acquired dysgraphia, a pronounced difficulty in producing written language (see Figure 4). The model assumes Italian, a shallow orthography, as the native language and does not mention other languages or bilingualism. The distinguishing piece of this model is the phonological buffer, which was hypothesized to be impaired in the case-study patient. In this model, there are separate routes for familiar versus novel words.

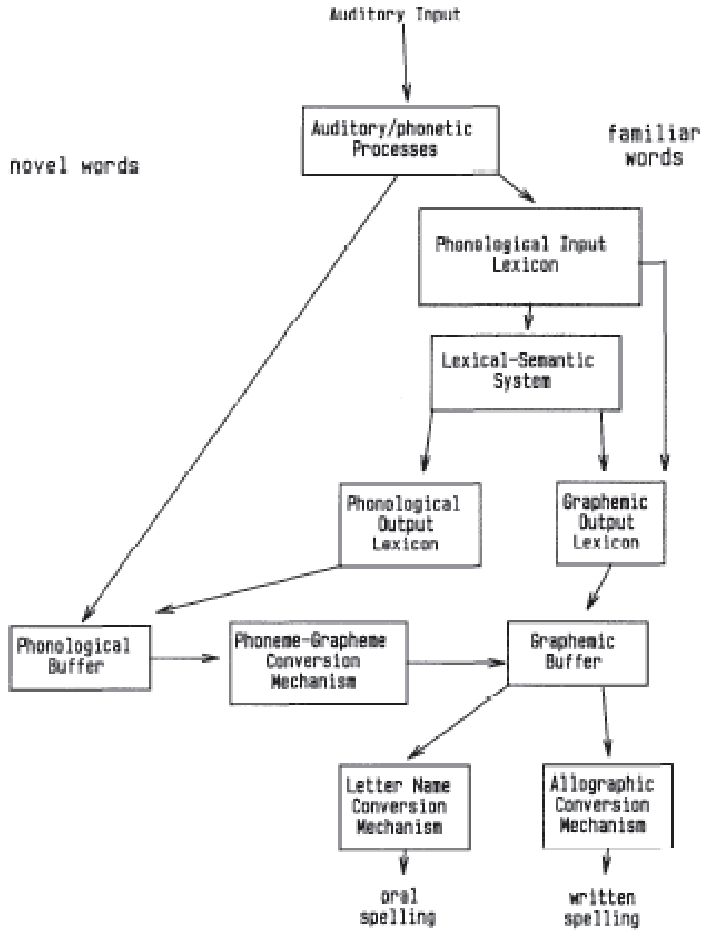

Figure 4. Schematic representation of a model of the spelling process, from Caramazza et al. (1987).

Miceli and Capasso (2006) also used case studies to develop their model (see Figure 5). One advantage of this model is that it is not language-specific. They used English and Italian, which are both alphabetic writing systems, but Italian is orthographically much shallower than English. However, it still does not consider the bilingual's spelling processes. Like Caramazza and colleagues (1987), they include a graphemic buffer.

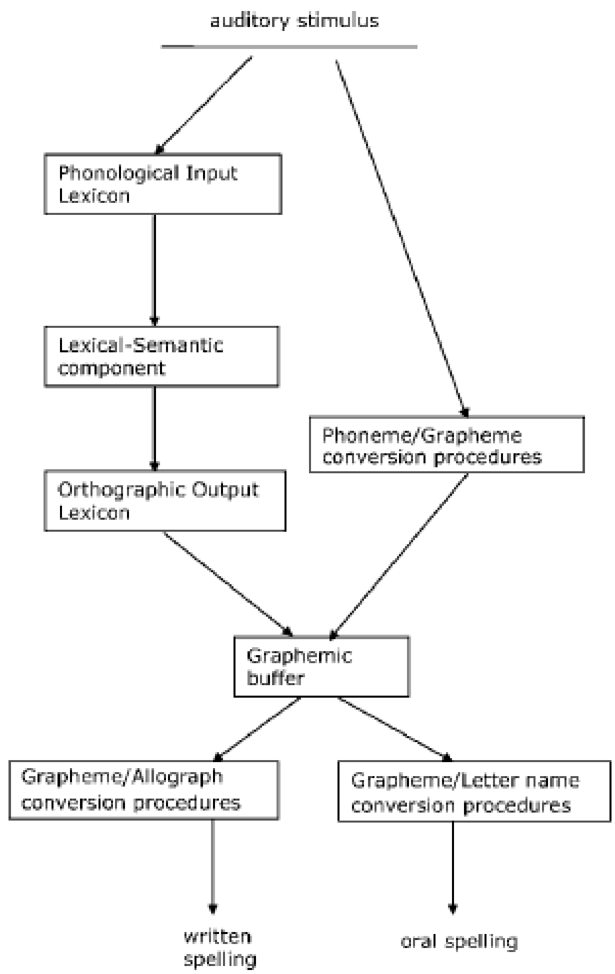

Figure 5. Schematic representation of the spelling system, from Miceli and Capasso (2006).

Margolin (1984) devised an intricate model of spelling processes, including semantic, phonological, motor, and perceptual steps (see Figure 6). This model's emphasis is on handwriting, and how apraxia disrupts the manual production of spelling, but the model does include oral spelling output as well. Margolin assumes English as the native language. It does not consider other languages or bilingualism.

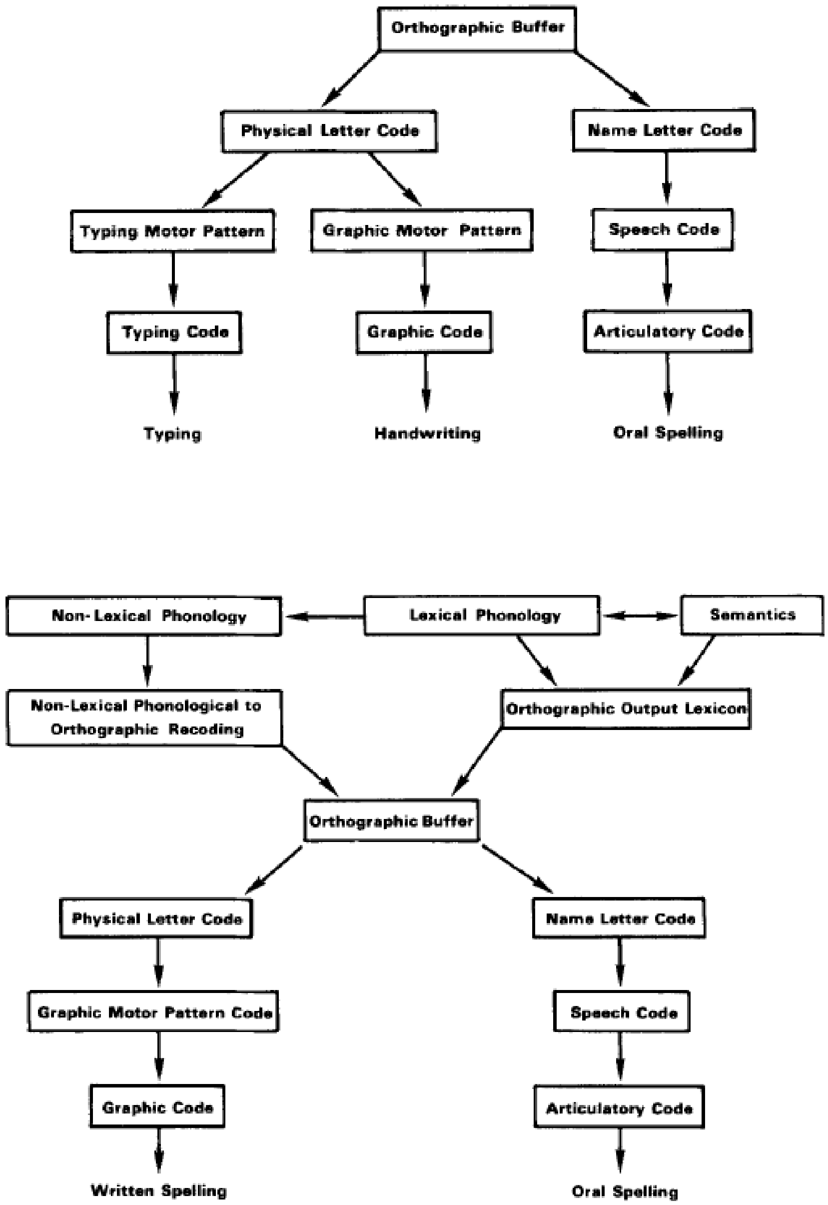

Figure 6. A model of the peripheral aspects of cognitive processes in spelling (top) and handwriting control (bottom), from Margolin (1984).

To summarize, the existing cognitive models go a long way toward explaining the mechanisms of spelling in English, but they also have their shortcomings. For one thing, they are predominantly Anglocentric (Share, 2008). This could be due to the fact that English has inconsistent sound-spelling mappings which necessitate special encoding and decoding processes. These models could be applied to languages with shallow orthographies, but might not be such parsimonious explanations of their processing. Furthermore, with one exception (Miceli & Capasso, 2006), the models reviewed do not account for the speller having knowledge of more than one language. Nor do they consider task-specific processes, as in the bilingual interactive activation model (Dijkstra & Van Heuven, 2002). In the most recent version of this model, the BIA+, it is assumed that a bilingual's two languages are integrated not only at the at the semantic level, but also at the phonological and orthographic levels (see Figure 7). The model accounts for word recognition in a variety of experimental tasks, as well as for task-specific decision processes. However, the model does not allow for investigation of language production tasks such as word spelling.

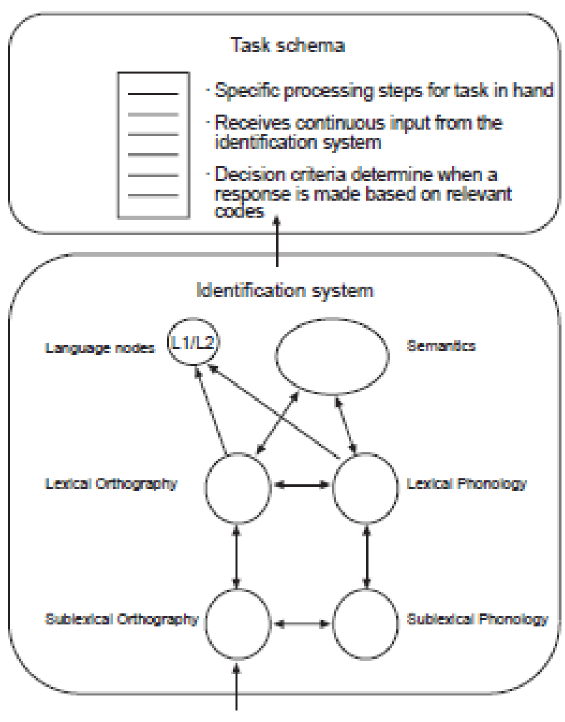

Figure 7. The BIA+ model of bilingual word recognition from Dijkstra and van Heuven (2002).

1.2.5 The Proposed Model of Orthographic Retrieval

The model I propose is meant to combine the useful pieces of the previous models while specifically explaining the cognitive mechanisms involved in orthographic retrieval by ESL learners. First, I take into account the fact that the speller is bilingual. Having knowledge of two language means having additional sets of knowledge for each lexical constituent. Like Dijkstra and van Heuven (2002), I assume some—though not total—overlap between orthographic, phonological, and semantic stores (see Figure 8). Depending on the writing systems of the two languages, there may be more or less overlap of the L1 and L2 orthographic stores. The same holds for phonological stores. Importantly, the influence of L1 is can negatively or positively transfer to L2, depending on the similarity or difference between languages. Additionally, proficiency might play a role such that the less native-like a learner is, the stronger influence from L1 is.

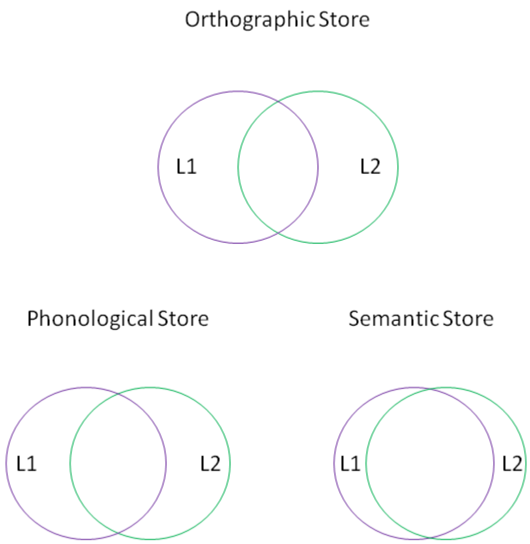

Figure 8. A bilingual's representations of orthographic, phonological, and semantic information in each language (L1 and L2).

Because task demands vary, the process of retrieving information from the orthographic store will depend on the nature of the task as well as the student's goals. In a receptive spelling task, such as spell checking or lexical decision, the orthographic form is provided and the goal is to verify its correctness. This may be done using an orthographic verification within the orthographic store. Or it may be done by going from orthography to phonology via grapheme-phoneme mappings, then searching the semantic store for a match. Or it may be done by going from orthography directly to the semantic store, bypassing phonology (see Figure 9). In an audio dictation task—a spelling production task—the phonological representation is provided and the goal is to produce the orthographic representation. This may be done directly from phonology to orthography using phoneme-grapheme mappings, or via semantics using a word lookup strategy (see Figure 10).

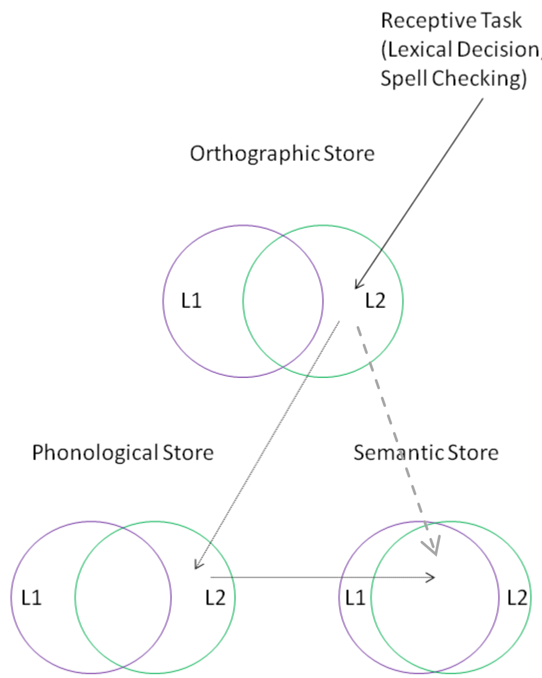

Figure 9. Possible paths to orthographic verification in a spelling recognition task.

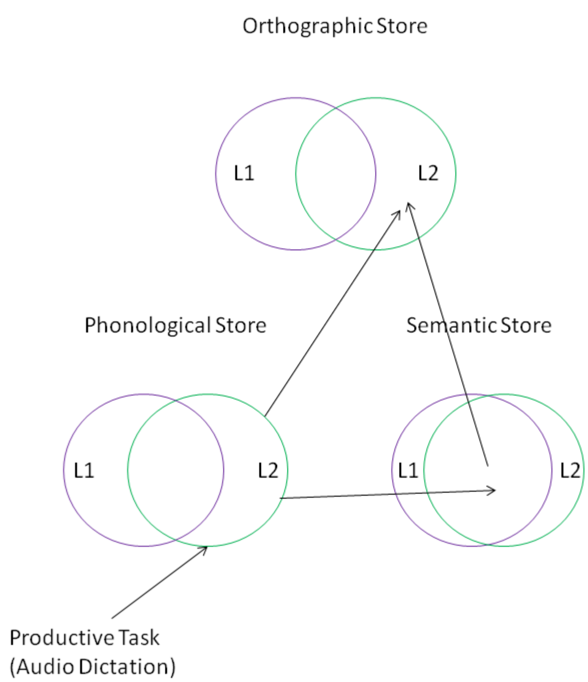

Figure 10. Possible paths from phonology to orthography in a spelling production task.

In my model of orthographic production (see Figure 11), the nature of the input is acoustic, phonetic, or subvocal. The source of the input is either internal or external. Internal input comprises self-driven spelling. Basically, a communicative message to be encoded graphemically originates with the writer, and there are no external orthographic or phonological inputs. Internal input is found in tasks such as writing an essay for class, sending an email message, jotting down a shopping list, or telling another person the correct spelling of your name. External input is used when the content to be written originates from an outside source, such as another speaker, another writer, or a digital recording to transcribe. External input occurs in tasks such as in-class dictation tests, transcriptions, spelling bees, and copying down an other person's writing.

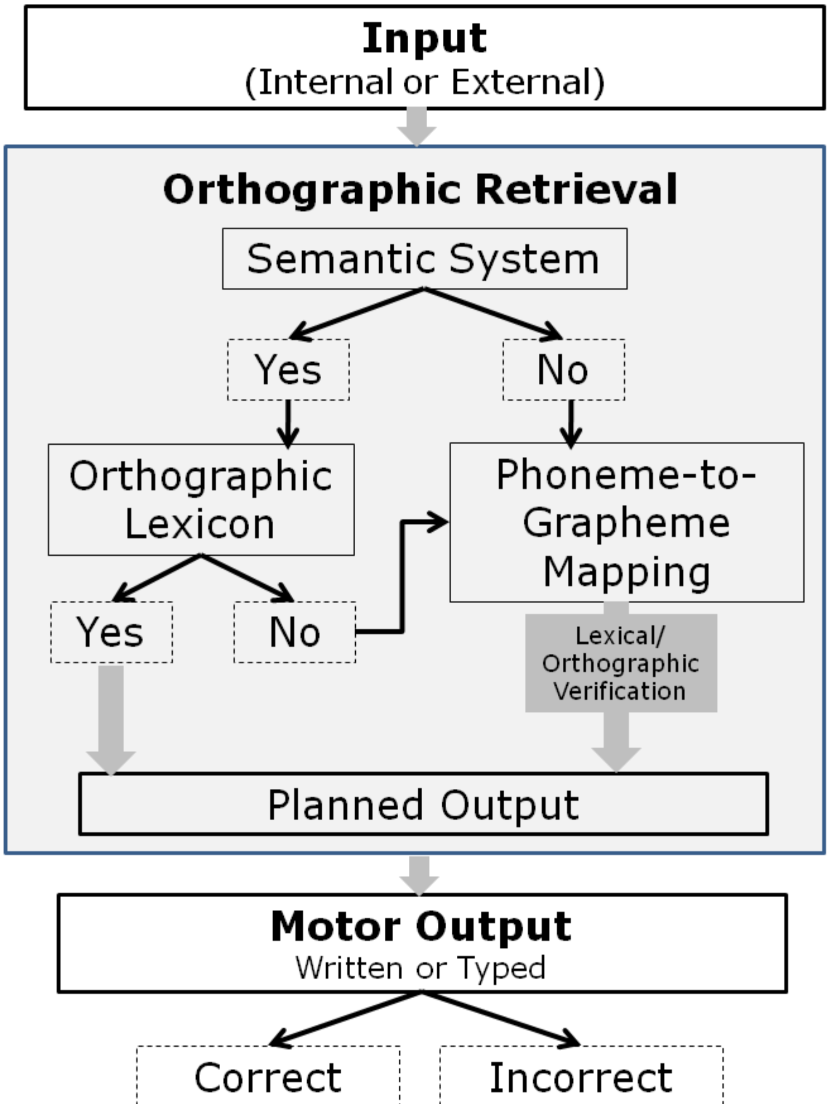

Figure 11. Proposed cognitive process model of orthographic retrieval and production.

At the orthographic retrieval stage, the student draws upon background knowledge of word constituents. With bilinguals, this knowledge contains information about both known languages. When input is external, the student must first search the semantic system for a target entry. If successful, the next step is retrieving the written form of the recognized word. If the student cannot retrieve a semantic or orthographic entry, the alternative route is attempting a phoneme-grapheme lookup. This can be followed up on with a verification that what the phoneme-grapheme mapping produced is a legal or familiar form. When input is internal, the semantic system entry has already been generated by the speller. The remaining steps are the same as for external input. The result of the orthographic retrieval stage is the planned output.

Motor output is the process of producing the planned output in oral or manual (e.g., handwritten, typed, signed). Written output is done manually with a pen, pencil, stylus, or other such writing implement. The motor response entails forming the correct shape of each grapheme and in the correct sequence. Typed output is done manually onto a computer keyboard, phone pad, or similar device. The motor response involves pressing the correct sequence of buttons. Spoken output is done orally. The motor response involves oral production of the correct sequence of letters, such as in a spelling bee or when telling somebody the correct spelling of an unfamiliar word. When people are writing, they might also vocally or subvocally utter the letter names or phonemes of what they are encoding. For example, when typing the word "Wednesday," I subvocally pronounce the units "Wed," "nes," and "day." This phenomenon is apart from spoken output; rather it can be considered a cognitive offloading mechanism for particular word-specific difficulties.

As suggested by the BIA+ model (Dijkstra & van Heuven, 2002), each of the components in the orthographic retrieval stage would be impacted by a multilingual's two or more languages. The semantic system would contain concepts shared by both languages as well as concepts unique to each language, and the orthographic lexicon would allow for word forms in either language. The phoneme-grapheme mappings would be derived from each language's rules about word spellings. The competition model (MacWhinney. 2005a, 2005b) would then make predictions about the positive or negative transfer that would ensue, depending on cross-linguistic similarities or differences in writing systems, orthographic depth, and so on.

The proposed cognitive model of orthographic production in English as a second language can account for performance on several language tasks, including audio dictation, lexical decision, and cross-modal word matching. In an audio dictation task, the participant hears a word (external, auditory input), conducts orthographic retrieval, and then plans and executes motor output (typed). In a lexical decision task, the participant sees a letter string (external, visual input), does a lookup and verification, compares the written form with their mental orthographic representation, and makes a response contingent upon the comparison results. In a cross-modal word matching task, the participant hears a word (external, auditory input), possibly converts it to an orthographic representation, and holds the phonological and/or orthographic form in working memory. When the visual word is presented, the participant compares its form with the mental representation formed, and makes a response contingent upon the comparison results. These three tasks were used in the behavioral experiments (discussed in Section 5) to investigate how non-native English speakers retrieve orthographic representations in their L2.

1.2.6 Effects of First Language Background

When students with different first language backgrounds show differential performance on English second language tasks, it could be attributed to important cross-linguistic contrasts or to differences in educational and cultural attitudes (Bialystok & Hakuta, 1994). First language effects have been found for learning to read in a second language (e.g., Bruder & Henderson, 1985; Chikamatsu, 1996; 2006; Hayes-Harb, 2006; Perfetti & Dunlap, 2008; Red, 1999; Thompson-Panos & Thomas-Ruzic, 1983; Wade-Woolley, 1999; Wade-Woolley & Siegel,

1997) and to spell in English as second language (Figueredo, 2006). Regarding spelling acquisition, one important cross-linguistic factor is the first language writing system. Arabic, Korean, and Spanish each use an alphabetic writing system. However, Spanish is the only one of these three to use a subset of the 26 letters used in English. Hence, written Arabic and Korean differ more from English than written Spanish does. Chinese uses a non-alphabetic writing system. Hence, written Chinese differs even more greatly from English than Arabic, Korean, and Spanish do.

The system accommodation hypothesis (Liu, Dunlap, Fiez, & Perfetti, 2003; Perfetti, Liu, Fiez, Nelson, Bolger, & Tan, 2007) posits that native Chinese speakers learning English as a second language will assimilate the L2 alphabetic writing system into their non-alphabetic L1 way of reading and word learning. The converse does not occur, however. A native English speaker learning Chinese as a foreign language cannot use an alphabetic principle to read and write Chinese characters. In this case, the learner must accommodate to the new writing system, rather than assimilate into current learning mechanisms. Brain imaging research supports the system accommodation hypothesis. While there are some brain regions that are activated during reading of any language, there are specific brain regions involved only in reading Chinese or only in reading English (Perfetti et al., 2007).

In most Arabic-speaking countries, there is typically an emphasis on oral culture, and this preference seems to transfer to English as a second language (e.g., Fender, 2001; 2003; 2008, Figueredo, 2006). In Chinese and Korean schools, students spend more time practicing English reading and writing skills than speaking and listening. Because Spanish-speaking students might come from European, Central American, or South American countries, their cultural and educational factors are more diverse. Effects of native culture and educational attitudes, as well as individual differences within each L1 background group, could also lead to different kinds of spelling strategies or difficulties in English as a second language. Therefore, first language background is one subject variable that was included in the present study.

1.2.7 Effects of Second Language Proficiency

Greater proficiency in English means that a learner has become familiar with a larger range of words. In doing so, the student has the opportunity to fine-tune orthographic representations. Similar to a native English speaker progressing from sight word reading to phonics-based decoding, a non-native English speaker can benefit from having to distinguish among a larger pool of words (Chall, 1979; Ehri, 1991: Ehri & Wilce, 1985; Frith, 1985; Gentry, 1982; Gough & Juel, 1991; Henderson, 1981; Henderson & Beers, 1980; Henderson & Templeton, 1986; Treisman & Bourassa, 2000). Once this occurs, the student can rely more on analytical decoding and encoding rather than holistic processing. Also, with more variety of input comes the opportunity to notice spelling consistencies, where they exist. Therefore, second language proficiency is another subject variable that was included in the present study.

2.0 DATA MINING STUDY

Because English has highly variable phoneme-grapheme mappings, spelling is difficult for native and non-native speakers alike. Recognition of correct written forms is an easier task than production of correct written forms. When an ESL student does spell a word correctly, we can be reasonably sure hat his or her lexical representation of that word is well-developed (Perfetti & Hart, 2001). However, when a student spells a word incorrectly, it can be for any number of reasons. We cannot be sure if the orthographic and/or phonological constituents of the word are underspecified, or if the error was due to carelessness or inattention.

First language background can have an effect on the rate of spelling mistakes students typically make. Anecdotal reports from teachers of English as a second language suggest that students from Arabic-speaking countries tend to have much poorer writing and spelling skills than students from other language backgrounds (e.g., Chinese, Spanish). This deficit occurs despite comparable speaking and listening skills across groups (A. Juffs, personal communication, 2009). In terms of the lexical quality hypothesis (Perfetti & Hart, 2001), this means that their phonological and semantic constituents are strong, but there are weaknesses in the system at some combination of the orthographic constituent, the link from phonology to orthography, and the link from semantics to orthography. First language background can also have an effect on the type and severity of spelling mistakes students typically make. Again, Arabic L1 students are reported to have particular difficulty with English vowels in both encoding and decoding (e.g., Martin, 2011). Also, Arabic L1 students have been reported to make spelling errors that are further off the mark. For example, one Arabic student misspelled “audience” as “oneiouns” in an in-class sentence dictation task (Dunlap, unpublished data).

According to a writing system accommodation account (e.g., Liu et al. 2003), students whose first language writing system is also alphabetic (e.g., Arabic, Korean, Spanish) should have better spelling in English than students whose first language is non-alphabetic (e.g., Chinese, Japanese). However, not all alphabets are identical. English uses 26 letters in a Roman alphabet. Spanish uses the same 26 letters, and sometimes adds diacritical markings (e.g., ñ, é). Korean and Arabic are alphabetic writing systems, but each has a different set of graphemes visually different than those used in English and Spanish. According to a graphemic familiarity account, students whose alphabet is closer in appearance to English (e.g., Spanish) should have better spelling than those whose alphabet differs from English (e.g., Arabic, Korean).

Proficiency in the target language is also hypothesized to affect spelling rates. The more experience a student has in reading and writing in English, the better their spelling should be (e.g., Treisman & Bourassa, 2000).

In the data mining study, student transcriptions of recorded spoken utterances were analyzed for spelling error rates and types. The goals of the data mining study were:

(1) to find empirical confirmation for teacher observations that spelling difficulty varies by L1 background and L2 proficiency; and

(2) to identify specific areas of improvement for intervention studies.

2.1 DATA MINING STUDY – METHOD

2.1.1 Participants

Participants were 88 non-native English speakers enrolled in intermediate to advanced English-as-a-second-language courses through the English Language Institute (ELI) at the University of Pittsburgh in Spring 2008. Of the 88 students, 23 were intermediate (ELI Level 3, equivalent to scoring 45-59 on the Michigan Test of English Language Proficiency [MTELP]); 34 were intermediate-advanced (ELI Level 4, equivalent to scoring 60-79 on the MTELP); and 31 were advanced (ELI Level 5, equivalent to scoring 80-100 on the MTELP).

Each student's first language was one of the following: Arabic, Mandarin Chinese, Korean, or Spanish. The Spanish students sometimes reported learning additional languages besides English. The Arabic, Chinese, and Korean students, however, did not. In general, these students began studying English as a foreign language in their home countries in elementary school. Thus, they are not considered balanced bilinguals. Nonetheless, they were fairly proficient in English as a second language, and were motivated to continue studying in an intensive English program.

2.1.2 Materials

As part of their curriculum, ELI students were occasionally required to do recorded speaking activities (RSAs). While in the computer lab, the student would log in to a program that provided an open-ended prompt, such as "What did you do this weekend?" or "How do people in your home country celebrate the New Year?" The student would have one minute to prepare an answer, and then would record an oral response. After completing the recording, each student was asked to transcribe the content of the recording so that the teacher could follow along with the response. However, the transcriptions were not graded for accuracy, so students did not need to devote a lot of time being careful to reproduce speeches verbatim or to spell every word correctly.

These RSA transcriptions produced 68,882 tokens, which were the source of data for coding and analyses.

2.1.3 Coding

Each transcribed word form was coded as either correct or incorrect. Foreign words, proper nouns, slang words, and recently coined words were counted as correct if they had an accepted spelling. Examples of these items include food items (e.g., falafel) and personal names that can vary in their spelling (e.g., Mohammed). British spellings of words (e.g., neighbour, theatre, realise) were also coded as correct.

Incorrect spellings were further categorized by type of error: form errors, transpositions, encoding errors, and morphological errors. *Form errors* contained the correct letters in the correct order, but lacked proper capitalization, spacing, or punctuation. *Transpositions* contained all the correct letters but in an incorrect order. Both form errors and transposition errors likely represent adequate orthographic knowledge but entail some carelessness in typing. *Encoding errors* were items in which one or more letters were added, deleted, or substituted. These errors likely represent a failed attempt at retrieving orthographic representations directly, or a weak link from the phonological to orthographic constituents. *Morphological errors* were not real words in English, yet they exhibited an attempt at expressing a combination of morphemes. Overregularizations of verb tenses, such as “heared” for “heard,” as well as misapplication of plural morphemes, such as “ourselfs” for “ourselves” or “musics” for “music,” were coded as morphological errors. These errors are different from encoding errors in that they typically were faithful to phoneme-grapheme mappings. For examples of each coding category, see Table 1.

2.1.4 Analyses

Error rates and types were categorical, so chi-square tests were conducted to test whether the observed data differed from a pattern found merely by chance.

Table 1. Error Coding Schema for Data Mining Study

Error Type	Description	Attempt	Correct Form
Form Error	Capitalization	forbes	Forbes
	Spacing	infact	in fact
	Punctuation	couldnt	couldn't
Transposition	Transposition	afetr freind	after friend
Encoding Error	Consonant – addition	affraid	afraid
	Consonant – deletion	acount	account
	Consonant – substitution	concider	consider
	Vowel – addition	fiew controle	few control
	Vowel – deletion	devlop favorit	develop favorite
	Vowel – substitution	becose docters	because doctors
	Multiple consonant and/or vowel additions, deletions, or substitutions	knowlige neberghood	knowledge neighborhood
Morphological Error	Morphological	heared musics truthable	heard music trustworthy

2.2 DATA MINING STUDY – RESULTS

2.2.1 Error Rates

The proportions of correctly and incorrectly spelled words by L1 background (Arabic, Chinese, Korean, Spanish) and L2 proficiency (Level 3, Level 4, Level 5) are shown in Figure 12. Proportions of only incorrectly spelled words by L1 and L2 are shown Figure 13. A two-way chi-square test was done on the number of incorrectly spelled items, with first language as one factor and class level as the other factor. Results showed that the error rate pattern did not occur by chance, $\chi^2(6) = 188.35$, $p < .05$.

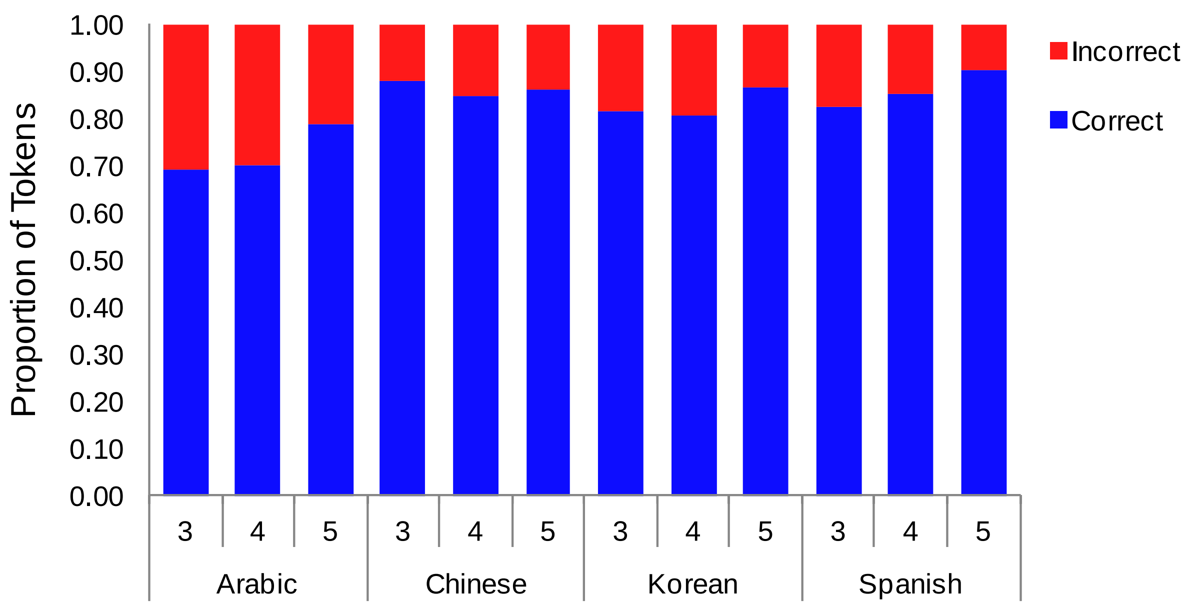

Figure 12. Proportion of correctly spelled and incorrectly spelled words in the data mining study.

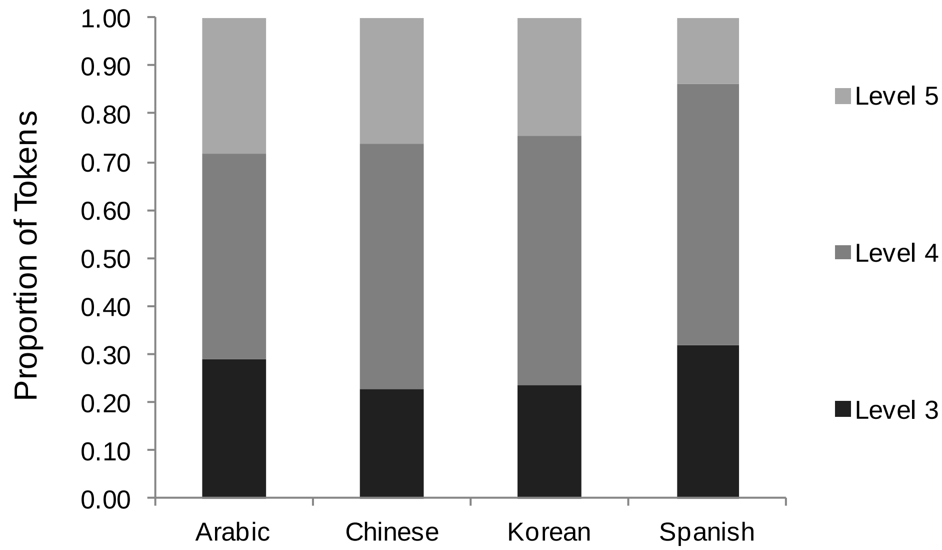

Figure 13. Proportion of incorrectly spelled words by L1 proficiency and L2 background.

2.2.2 Error Types

For each L1 group and each L2 proficiency level, the proportion of items in each error category was calculated (see Figure 14). A separate chi-square test was performed on a subset of error items, dubbed "true encoding errors." This subset excluded form errors (capitalization, punctuation, and spacing errors), transposition errors, and morphological errors (e.g., musics, truthable). Thus, true encoding errors included vowel and/or consonant deletions, additions, or substitutions. A two-way chi-square was done on the number of items coded as true encoding errors, with L1 background as one factor and L2 proficiency as the other factor. Proportion of true encoding errors are shown in Figure 15. Results showed that the error rate pattern did not occur by chance, $\chi^2(6) = 70.94$, $p < .05$.

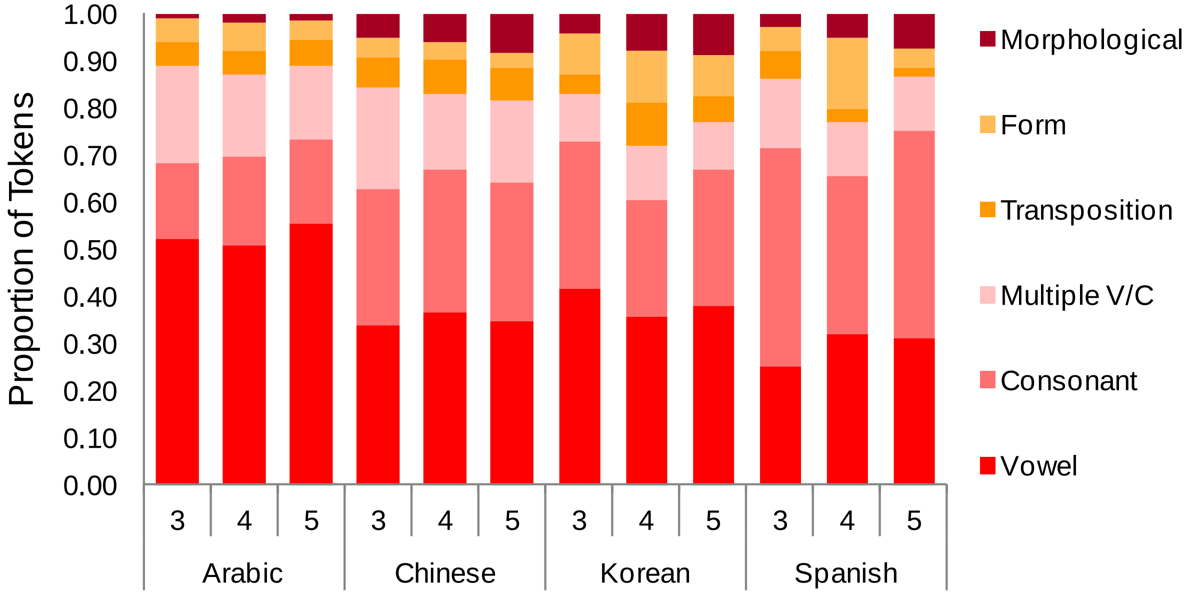

Figure 14. Proportion of error types by L1 proficiency and L2 background.

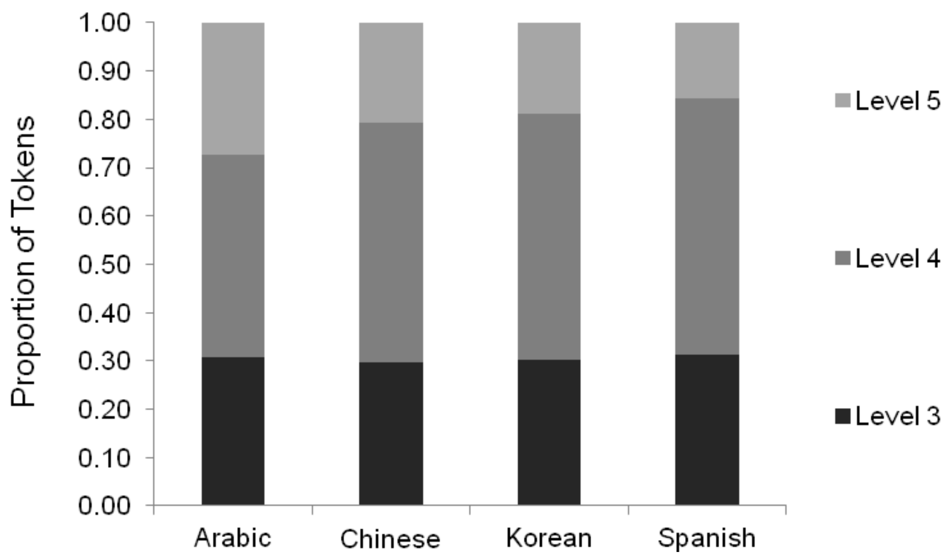

Figure 15. Proportion of true encoding errors by L1 proficiency and L2 background.

2.2.3 Correctly Spelled Words

Of the correctly spelled words, items were identified as being in sub-groups of the Academic Word List (Coxhead, 2000). This list comprises 570 word families (each word family is a lemma plus its morphological variations, for example: analyze, analytical, analyses, etc.), with a total of 3111 words, which accounted for approximately 10% of the word tokens in an academic text corpus. Academic Word List words are less frequent than the most common 2000 words in written English in West's (1952) General Service List. AWL K1 represents the most common words on the list (sublists 1 and 2); K5 the least common words (sublists 9 and 10). Proportions of words in each frequency band are shown in Figure 16 . Results of a two-factor chi-square test, with each of the 12 participant groups as one factor and each of the 5 frequency bands as the other factor, showed that the distribution pattern did not occur by chance, $\chi^2(44) = 988.52$, $p < .05$.

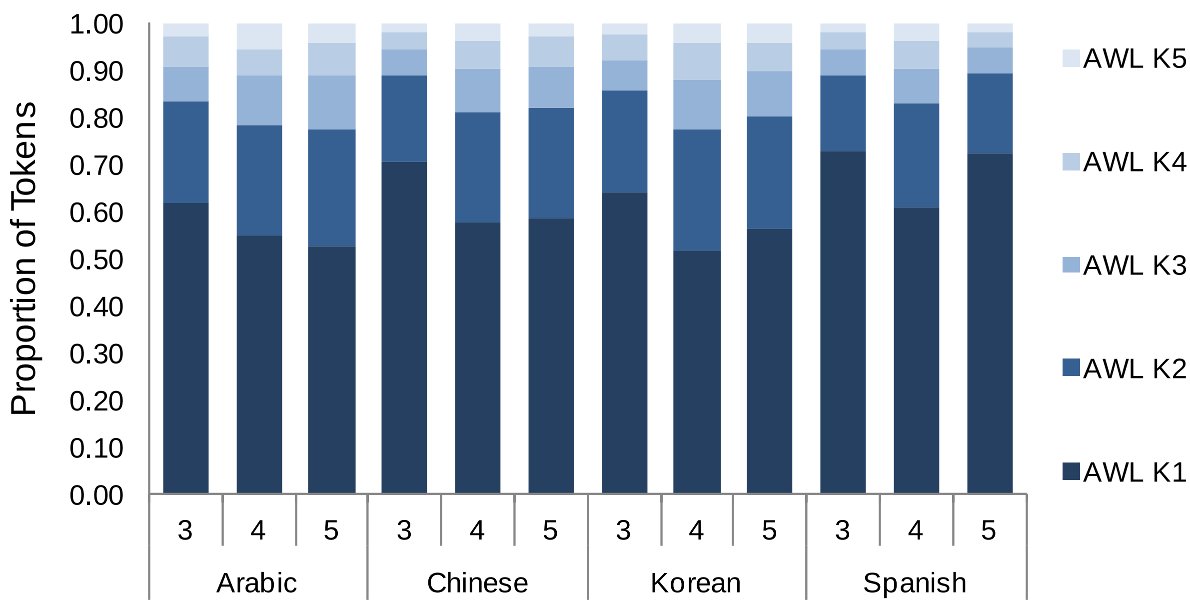

Figure 16. Proportion of correctly spelled words in each Academic Word List (AWL) frequency band (K1 = higher frequency; K5 = lower frequency).

2.3 DATA MINING STUDY – DISCUSSION

Student transcriptions of recorded spoken utterances provided 68,882 word tokens which were coded as correctly spelled or incorrectly spelled. Analyses by first language background and second language proficiency showed the following results.

Each group of students (Arabic, Chinese, Korean, Spanish) at each level of proficiency (Levels 3-5) had some difficulty with spelling. Between 2% and 28% of attempted words were misspelled. Overall, error rates were highest for Level 4 (intermediate-advanced) students. Teacher observations were empirically confirmed: Arabic L1 students had the highest rate of spelling errors. This difficulty persisted across proficiency levels such that at Level 5, Arabic students still had higher error rates than Level 3 students for other first language backgrounds. Consonant errors were most prevalent for the Spanish L1 students, especially at Level 3. Vowel errors were most prevalent for the Arabic L1 group (consistent with Abu-Rabia, 1997; Abu-Rabia & Siegel, 2002). Because short vowel sounds are not usually marked in written Arabic, this "learned neglect" of written vowels might transfer to spelling in L2 English (Martin, 2011; Ryan & Meara, 1991).

One possible reason Arabic students made more spelling errors is that they were attempting more difficult vocabulary during the recorded speaking activities. Analysis of correctly spelled words by Academic Word List subgroups supports this to some extent. Arabic L1 students used lower rates of more common (AWL K1) words, compared to other L1 background students. Of course, this is an indirect measure of the frequency of incorrectly spelled words. Nonetheless, it suggests there is a mismatch between spoken and written production (Hofman & Habib-Allah, 1982). Arabic L1 students seem to be attempting more advanced words in their spoken utterances, then have more difficulty spelling them in their transcriptions.

The transcription data show a mismatch between students' spoken production and their written production. Students were free to use semantic items and syntactic structures of their choosing. Even so, they encountered difficulty correctly encoding a portion of these items. In terms of the lexical quality hypothesis (Perfetti & Hart, 2001), the link from meaning to phonology is stronger than the link from meaning to orthography and the link from phonology to orthography. In terms of Krashen's (1989) input hypothesis and Schmidt's (1990) noticing hypothesis, ESL students do not appear to acquire or strengthen orthographic representations merely through exposure during reading or classroom vocabulary instruction. Instead, Swain's (1985, 1993) notion that comprehensible output is key in developing communicative competence is supported.

The system accommodation hypothesis predicts different reading (decoding) strategies for alphabetic versus non-alphabetic writing systems. However, there did not appear to be marked deficits for Chinese L1 students in English spelling (encoding), relative to other L1 background students. A graphemic familiarity account would predict that Spanish L1 students would have the least difficulty in English decoding and encoding, relative to Chinese, Arabic, and Korean students. In the data mining analyses, this was not borne out. Because these groups of students are so advanced already, it is possible that any effects of writing system familiarity have been equalized. Yet the Arabic L1 students still lag behind their classmates in quality of orthographic representations, especially vowel spellings. A tenable explanation for this is that their first language writing system omits many vowel markings. Therefore, Arabic L1 students bring a sort of neglect for vowel markings to their English L2 spelling. It has already been shown that vowel neglect transfers to English reading (Martin, 2011; Ryan & Meara, 1991).

Since there was clearly room for improvement for all L1 groups and all proficiency levels, a series of spelling interventions was justified. Target items for the subsequent spelling interventions were words on which a majority of students across L1 backgrounds and ELI levels made true encoding errors, not merely carelessness or form errors.

3.0 INTERVENTION STUDY #1

Analysis of students' spelling rates and types showed that intermediate to advanced learners of English as a second language have room for improvement in production of orthographic representations. This was shown to be the case even when the students' semantic and phonological representations were adequate for formulating spoken utterances in a class assignment. In terms of the lexical quality hypothesis (Perfetti & Hart, 2001), the link from meaning to phonology is stronger than the link from meaning to orthography and the link from phonology to orthography (see Figure 17).

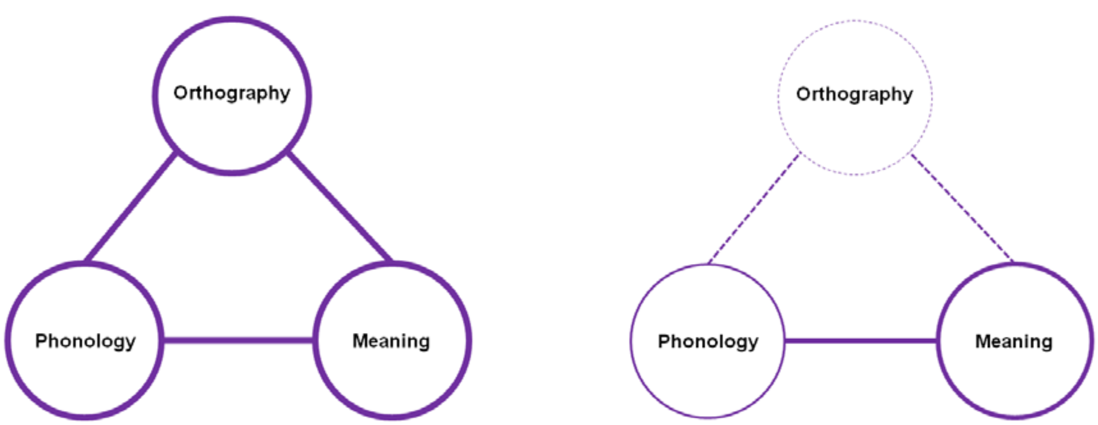

Figure 17. Graphic depiction of a high quality lexical representation (left) and an ESL student's underspecified lexical representation (right).

The primary aim of the two spelling intervention studies was to compare the benefits of two different types of spelling training—a form-focus condition and a form-meaning integration condition (Norris & Ortega, 2000). A focus on forms alone involves targeting a structure to be learned but in the absence of meaning. Form-meaning integration involves some feature focus, particularly to troublesome formal properties, but with the intent of learning what will be useful in future communication. Results of Norris and Ortega's (2000) meta-analysis strongly suggest that instruction that utilizes an explicit focus on form integrated with meaning is most effective. While the topic of their analyses was grammatical structures, I assert that the same concepts can be applied to other aspects of language instruction, such as acquisition of phoneme-grapheme mappings.

Previous studies of vocabulary instruction with native English speakers (Balass, 2004; Balass, 2011; Balass, Nelson, & Perfetti, 2009; Nelson, 2010) showed that skill differences have an effect on learning when training provides incomplete information. In their studies, participants learned two of the three lexical constituents (orthography and meaning; phonology and meaning; or orthography and phonology) of rare English words. The orthography plus meaning condition was considered to be the deepest form of encoding; the orthography plus phonology condition was considered the shallowest. Behavioral and electrophysiological evidence showed both training effects and skill differences in how well words were learned.

For the spelling interventions, the two types of training conditions were operationalized at the lexical rather than grammatical level. In the form-focus training condition, the assumption was that the student has an adequate representation of the semantic constituent of a word, but a weak orthographic representation. The goal of this kind of training was to focus on strengthening the weak pieces: the orthographic knowledge as well as the link between the phonological and orthographic forms of a word (see Figure 18). During training trials, focus was drawn toward the spelling of the word as the individual letters were shown visually and pronounced orally. When asked to produce the written form of the target word, they were provided with the phonology and had to retrieve the relevant orthography.

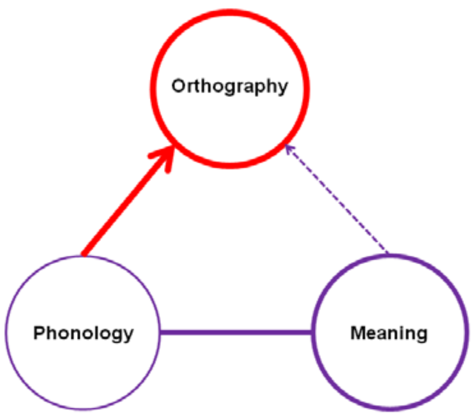

Figure 18. Target of focused instruction in the form-focus training condition.

In the form-meaning integration condition, the assumption was that students benefit most from developing well-integrated lexical representations. The goal of this kind of training was to take advantage of and build upon existing strengths (see Figure 19). During training, students were provided with a meaningful sentence context. Importantly, when asked to type the written form of the target word, students were prompted with the same sentence context. Thus, they had to retrieve the orthographic constituent with no external phonological cues.

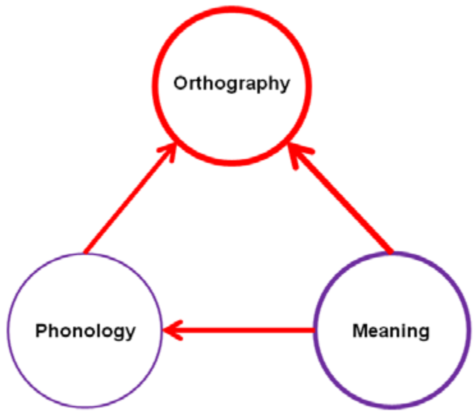

Figure 19. Target of focused instruction in the form-meaning integration training condition.

In the present study, the two training interventions were designed to answer the following research questions:

1. Does the benefit of form-meaning integration, found for learning of second language grammatical structures (Norris & Ortega, 2000), extend to the development of orthographic representations in English as a second language?
2. Are there differential effects of form-focus and form-meaning integration training as a function of first language background or second language proficiency?

It is not yet known which of the two types of training might be better for spelling instruction in English as a second language. According to a depth of processing account (Craik & Lockhart, 1972), the form-meaning integration would be superior. Deeper processing of a word's semantic, phonological, and orthographic constituents would be hypothesized to lead to better long term memory for that word. Focusing on phonological and orthographic forms in the absence of meaning is akin to shallow processing. However, a divided attention account would predict that form-focus instruction would be superior. When a student already has adequate semantic representations, including word meanings in the training trials distracts from the key feature to be learned, namely spelling.

3.1 INTERVENTION STUDY #1 – METHOD

3.1.1 Participants

Participants were 56 adult learners of English as a second language studying at the University of Pittsburgh's English Language Institute in Spring 2009. Only students who completed at least two of the three test sessions were included in the analyses. Of the 56 students, 17 were intermediate (Level 3), 26 were intermediate-advanced (Level 4), and 13 were advanced (Level 5). Their first languages included: Arabic, Chinese, German, Japanese, Korean, Thai, and Turkish (see Table 2).

As with the data mining participants, most of these students did not report learning other languages besides their native language and English as a second language. The exceptions are those listed as other/multiple in Table 2. In general, these students began studying English as a foreign language in their home countries in elementary school (approximately age 10). By the time they were accepted into the intensive English program, they have had several years of English instruction.

Table 2. L1 background and L2 proficiency for participants in Intervention Study #1

	Level 3	Level 4	Level 5	Total
Arabic	5	10	2	17
Chinese	0	5	4	9
German	1	1	0	2
Japanese	0	2	4	6
Korean	4	5	2	11
Thai	2	0	1	3
Turkish	2	1	0	3
Other/Multiple	3	2	0	5
Total	17	26	13	56

3.1.2 Materials

Training items were 60 real English words selected from the set of misspelled words in the data mining study. These were items that were consistently difficult for spellers across first language backgrounds and across levels of proficiency. The words were divided into two lists, which were matched for length in letters, number of syllables, number of phonemes, frequency (Kučera and Francis, Thorndike-Lorge written, and Brown verbal, and Zeno), familiarity, concreteness, and imageability.

Audio dictation pre-test and post-test items were 40 real English words. Of the 40 words, 15 were selected from training Set 1, 15 were selected from training Set 2, and 10 were untrained control items of equivalent length and frequency.

Lexical decision pre-test and post-test items were 40 letter strings. Of the 40 items, 20 were yes responses (correctly spelled real English words, 10 from training Set 1 and 10 from training Set 2) and 20 were nonwords that were actual misspellings of real English words made by students in the data mining study. These misspellings had all been categorized as encoding errors, with either addition, deletion, or substitution of either a vowel or a consonant.

Materials were programmed and presented in RunTime Revolution 3.0. Training and testing items, as well as their lexical characteristics, are listed in Appendix A.

3.1.3 Design

The independent variable was training condition, with the two levels being form focus training and form-meaning integration training. Training was manipulated within-subjects, such that each student learned half of the target vocabulary words in one training condition and half in the other. Dependent variables were accuracy on the audio dictation and lexical decision tasks. Gains in learning were measured as the difference between post-test and pre-test scores. Subject variables were first language background and second language proficiency.

3.1.4 Procedure

Students participated in the testing and training tasks as part of in-class activities over the course of one semester of ELI Writing classes, as follows:

- Week 1: Two pre-test tasks were administered—the audio dictation task and the lexical decision task.

- Week 2: Study session #1, 15 items from training Set 1 were presented in the form focus training condition to half the students (Group A) and in the form-meaning training condition to the other half of the students (Group B).
- Week 3: Study session #2, the remaining 15 items from training Set 1 were presented in the form focus training condition to half the students (Group A) and in the form-meaning training condition to the other half of the students (Group B).
- Week 4: The audio dictation and lexical decision tasks were administered again.
- Week 5: Study session #3, 15 items from training Set 2 were presented in the form-meaning integration training condition to half the students (Group A) and in the form focus training condition to the other half of the students (Group B).
- Week 6: Study session #4, the remaining 15 items from training Set 2 were presented in the form-meaning integration training condition to half the students (Group A) and in the form focus training condition to the other half of the students (Group B).
- Week 7: The audio dictation and lexical decision tasks were administered again as post-tests.

Audio dictation task. The student sat at a computer monitor while wearing headphones and was instructed to click on an audio icon to hear the target word pronounced by a native English speaker, then to type the target word into a text box shown on the screen (see Figure 20). The student was allowed to click the audio icon as many times as needed, and to correct and retry spelling by using the backspace key on the computer. However, after proceeding to the next word, the student was not allowed to go back to any previous word.

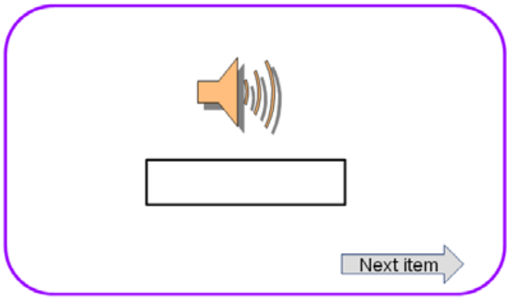

Figure 20. Screen shot of an audio dictation trial.

Lexical decision task. On the computer monitor, the student was presented with a letter string, and was instructed to click "CORRECT" if the word was a correctly spelled, real English word or "INCORRECT" if the word was not a correctly spelled, real English word (see Figure 21).

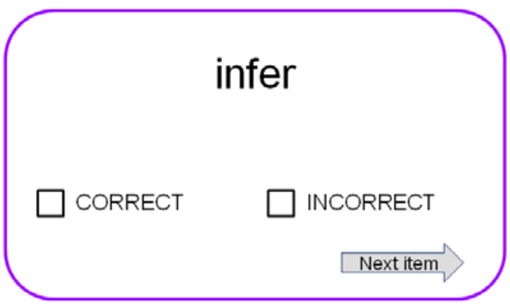

Figure 21. Screen shot of a lexical decision task trial.

Study sessions. Students studied 15 vocabulary words each session. In both training conditions, students were presented with the written form, the spoken form (pronounced by a native English speaker), and a brief definition of the word on an initial slide (see Figure 22).

In the form focus training condition, this was followed by a slide highlighting the individual letters (see Figure 23). When the student clicked on the audio icon, a native English speaker spelled the word aloud then repeated the word. On the next slide, the student was instructed to type the vocabulary word. An audio icon was provided, so that the student could hear the word repeated if necessary (see Figure 24). No feedback was provided on their responses.

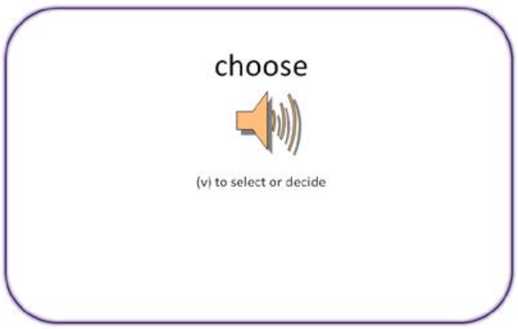

Figure 22. Screen shot of the first slide for each study trial.

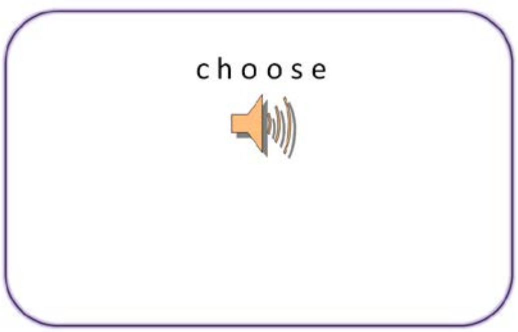

Figure 23. Screen shot of the second slide for each form focus study trial.

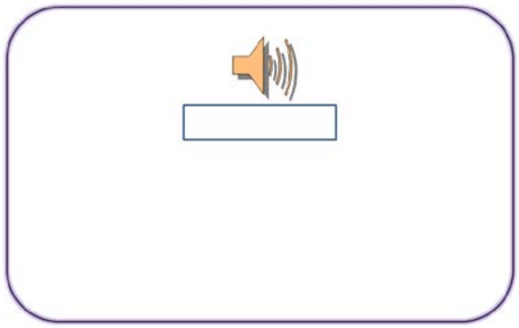

Figure 24. Screen shot of the third slide for each form focus study trial.

In the form-meaning integration training condition, the initial slide was followed by the vocabulary word used in a sentence context (see Figure 25). When the student clicked on the audio icon, a native English speaker read the sentence aloud. On the next slide, the student was provided the sentence context with the target word removed, and was instructed to type the vocabulary word into the text box (see Figure 26). A brief definition was provided as a hint, but no audio icon was provided. No feedback was provided on their responses.

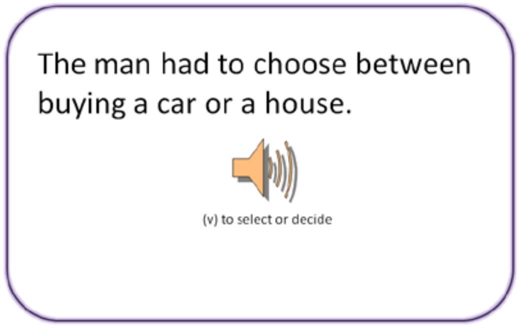

Figure 25. Screen shot of the second slide for each form-meaning integration study trial.

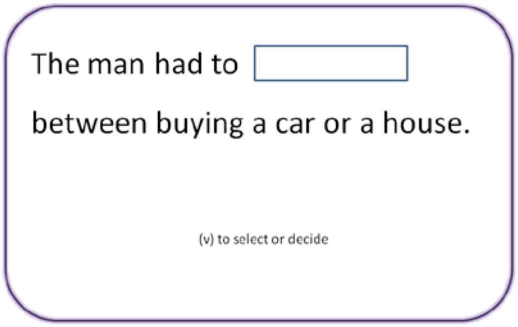

Figure 26. Screen shot of the third slide for each form-meaning integration study trial.

All sessions were held in a computer classroom so that each student could open and proceed through a version of the executable file programmed in RunTime Revolution 3.0. Completed data files were collected by the experimenter at the end of each class session.

3.2 INTERVENTION STUDY #1 – RESULTS

3.2.1 Results

On the lexical decision pre-test, students scored an average of 80.79% accuracy (*SD* = 11.73). Mean gains in accuracy from pre-test to post-test are shown in Figure 27. For words in the form focus training condition, students gained an average of 6.16 percentage points (*SD* = 6.84). For words in the form-meaning integration training condition, students gained an average of 4.97 percentage points (*SD* = 9.88). For untrained words in the control condition, students gained an average of 5.72 percentage points (*SD* = 9.29) According to results of planned paired *t*-tests, there were no statistically significant differences in gains between the two training conditions, or between trained and untrained words, all *p*-values > 0.05.

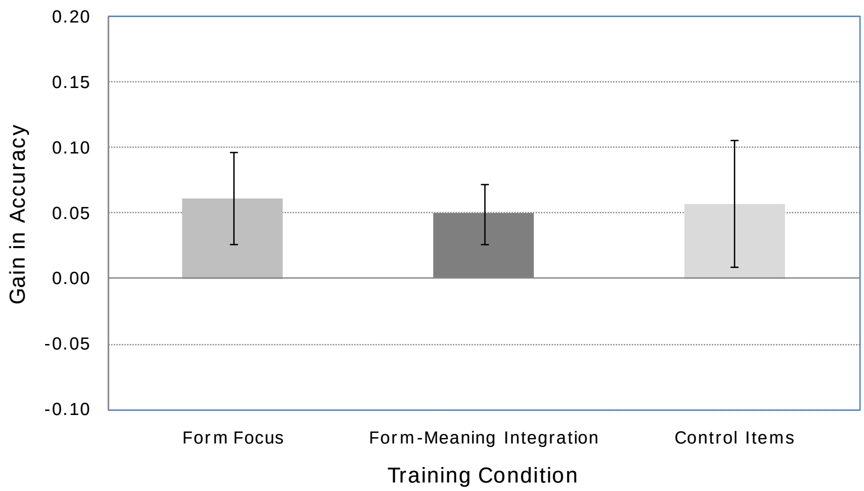

Figure 27. Mean accuracy gains on the lexical decision task.

On the audio dictation pre-test, students scored an average of 58% accuracy (*SD* = 16.11). Mean gains in accuracy from pre-test to post-test are shown in Figure 28. For words in the form focus training condition, students gained an average of 14.91 percentage points (*SD* = 5.54). For words in the form-meaning integration training condition, students gained an average of 13.71 percentage points (*SD* = 22.11). For untrained words in the control condition, student scores actually decreased by an average of 2.15 percentage points (*SD* = 14.78). However, this decrease was not significantly different from zero, $p > 0.10$. There was no statistically significant difference in gains between the two training conditions, $p > 0.10$, but there was a significant difference between trained and untrained words, $t(158) = 4.43$, $p < .001$.

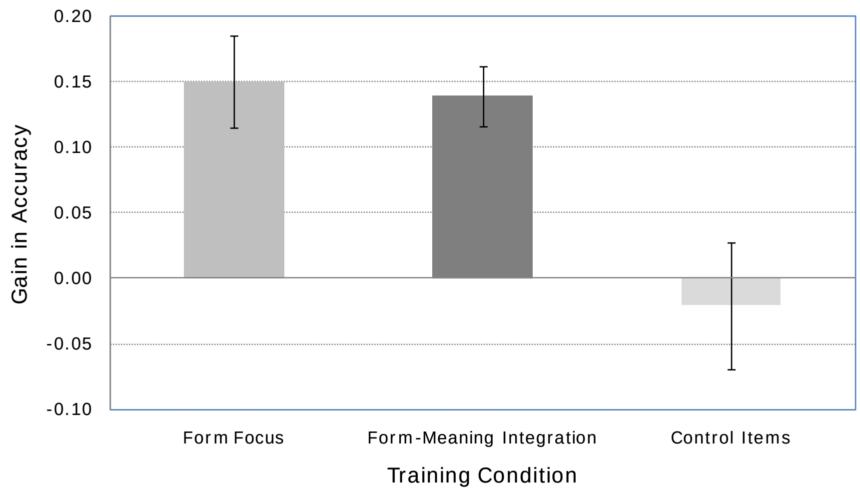

Figure 28. Mean accuracy gains on the audio dictation task.

Five of the first language backgrounds with adequate sample sizes (Arabic, Chinese, German, Japanese, and Korean) and all three ELI writing class levels (3, 4, and 5) were used in comparing effects of L1 background and L2 proficiency. Mean gains in accuracy on each task by training condition, first language background, and second language proficiency are shown in Figures 29-32.

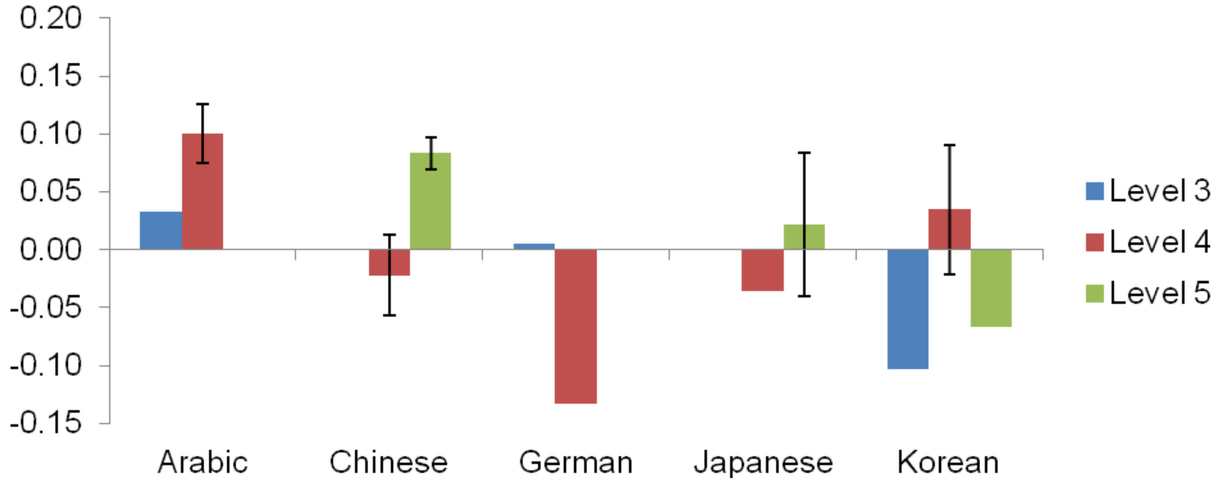

Figure 29. Mean accuracy gains on form-meaning integration training items in the lexical decision task, by L1 and ELI level.

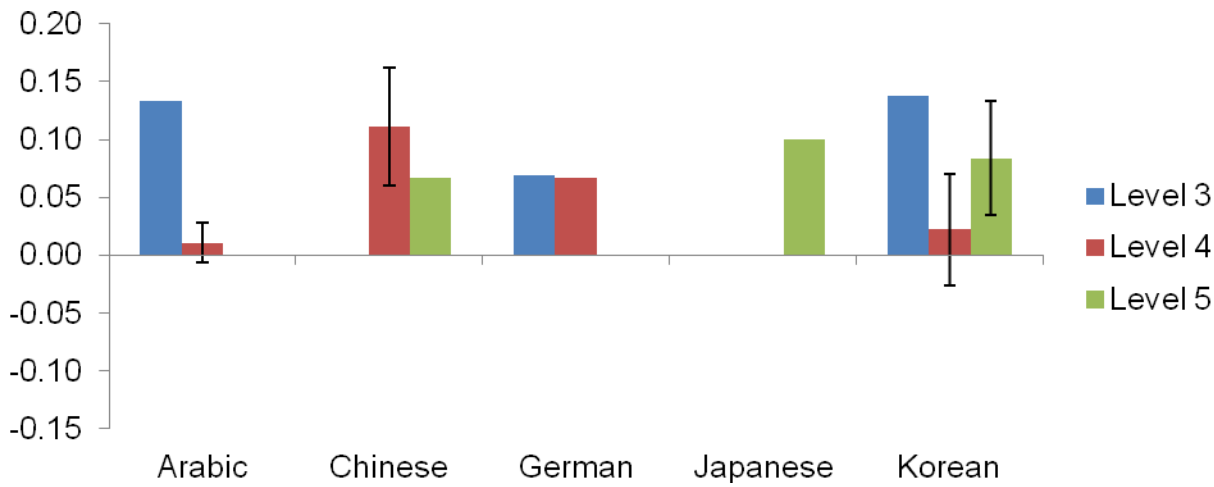

Figure 30. Mean accuracy gains on form focus training items in the lexical decision task, by L1 and ELI level.

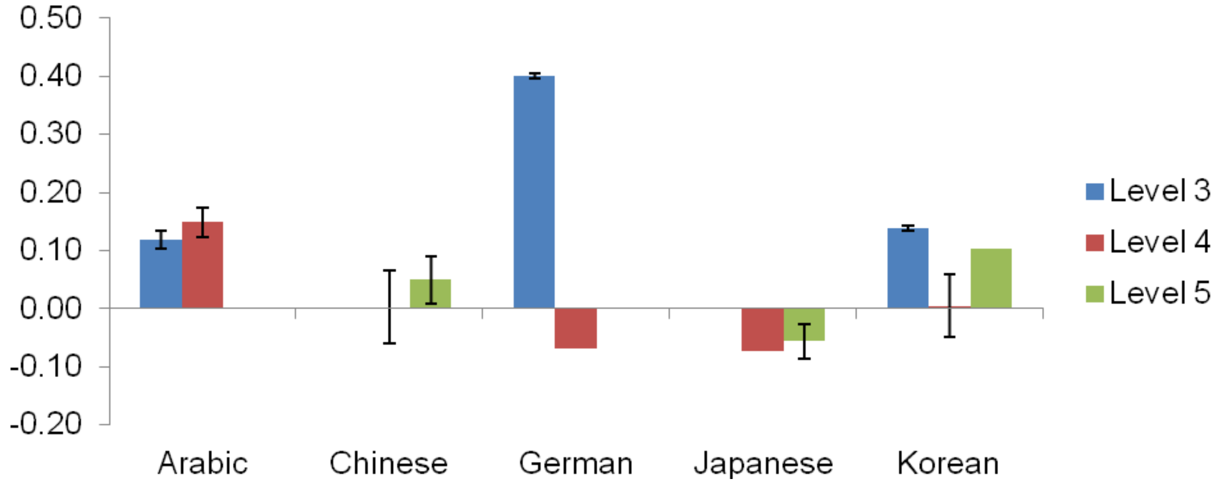

Figure 31. Mean accuracy gains on form-meaning integration training items in the audio dictation task, by L1 and ELI level.

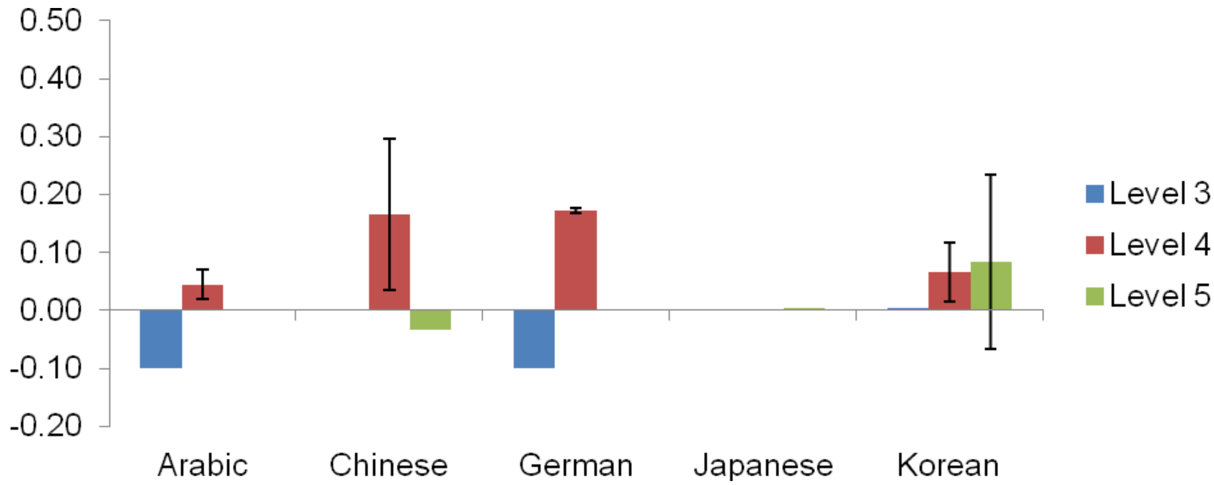

Figure 32. Mean accuracy gains on form focus training items in the audio dictation task, by L1 and ELI level.

3.3 INTERVENTION STUDY #1 – DISCUSSION

3.3.1 Discussion

After approximately two hours of training in either a form focus condition or a form-meaning integration condition, students showed small gains in a lexical decision task and moderate gains in an audio dictation task. There was no main effect of training condition; gains were equivalent for the form-focus training and the form-meaning integration training on both outcome measures.

In the lexical decision task, students had the same gains—approximately five percentage points—for trained as well as untrained items. This suggests that maturation effects, rather than experimental treatments, account for improvement on this task. Throughout the course of the semester, students encounter myriad words. Exposure to printed English words during this time seems to be enough for students to make small gains in the ability to recognize orthographic patterns, an arguably easier task than audio dictation.

There were no clear-cut main effects of L1 background or L2 proficiency (ELI level) on the lexical decision task. However, more students had decreases rather than gains after training in the form-meaning integration condition (see German and Japanese Level 4, Korean Levels 3 and 5 in Figure 29), whereas no groups had negative gains in the form focus training condition (see Figure 30). This finding suggests that, for many students, the form focus training is actually better than form-meaning integration when the outcome measure is a receptive rather than a productive task.

In the audio dictation task, students showed greater gains—approximately 15 percentage points—for trained items, and no gains for untrained items. These gains are especially remarkable because the audio dictation task—a production task—is more difficult than the lexical decision task—a receptive task.

Again, there were no clear-cut main effects of L1 background or L2 proficiency (ELI level) on the audio dictation task. Arabic, German, and Korean Level 3 students, Arabic and Chinese Level 4 students, and Korean Level 5 students all gained more from the form-meaning integration training than from the form focus training. Only one student (German Level 4) benefited more from form focus training.

Because no main effect of training condition was found, it could be the case that that the instructional difference found for grammatical forms (Norris & Ortega, 2000) might not extend to orthographic forms. However, students in the first intervention study often wrote down the list of vocabulary words during the study sessions so that they could continue to study the word meanings, pronunciations, or translations outside of class time. This could potentially wash out any differences between training conditions. In the second intervention study, students were more strongly encouraged to follow the training materials strictly as presented. Also, in order to satisfy the students' and teachers' preference for feedback, the second spelling intervention incorporated several feedback and review options during training.

4.0 INTERVENTION STUDY #2

To develop robust lexical representations, learners must be able to recognize words not just at a holistic or shallow level, but at a fine-grained level. For instance, a learner should know that the words "beer" and "bear" are two distinct lexical entries. The more words a person knows, the less she can rely on shallow features, such as word shape, in word identification (Treisman & Bourassa, 2000). During training in the second intervention study, target vocabulary words were presented in groups of three. The relationship among the words in each triplet was either similar in form (shared orthographic and phonological but not semantic constituents, e.g., labor, label, lapel) or random (little to no overlap of orthography, phonology, and semantics, e.g., function, survey, transfer). Grouping words together this way was hypothesized to draw attention to the importance of making fine-grained distinctions between lexical entries. Accordingly, students would have better learning for these items. On the other hand, words that share more form overlap might by their nature cause more difficulty and confusion for a learner. Accordingly, students would have worse learning for these similar items.

In the second intervention study, a sentence completion task (rather than a lexical decision task) was used as the recognition task. The sentence completion task was a two-choice cloze task intended to assess how well a student knows the appropriate lexical entry within a meaningful context. Foils were selected such that they shared orthographic and phonological but not semantic features with the target word. So, students would have to have fine-grained distinctions between potentially confusable vocabulary words (e.g., adapt and adopt).

The second training intervention was designed to replicate the first intervention study, while addressing some of its limitations, and to answer a third research question:

3. Can either form focus training or form-meaning integration training help learners of English as a second language develop fine-grained distinctions among orthographically similar words?

4.1 INTERVENTION STUDY #2 – METHOD

4.1.1 Participants

Participants were 48 adult learners of English as a second language studying at the University of Pittsburgh's English Language Institute in Spring 2010. Of the 48 students, 9 were intermediate (Level 3), 24 were intermediate-advanced (Level 4), and 15 were advanced (Level 5). Their first languages included: Arabic, Chinese, Japanese, Korean, Portuguese, Spanish, Thai, and Turkish (see Table 3).

Most of these students did reported their native language and English as the only two languages they knew. The exceptions are those listed as other/multiple in Table 3. In general, these students began studying English as a foreign language in their home countries in elementary school (approximately age 10). By the time they were accepted into the intensive English program, they have had several years of English instruction

Table 3. L1 background and L2 proficiency for participants in Intervention Study #2

	Level 3	Level 4	Level 5	Total
Arabic	4	9	3	16
Chinese	3	4	1	8
Japanese	0	1	1	2
Korean	0	4	4	8
Portuguese/Spanish	1	4	2	7
Thai	0	0	2	2
Turkish	0	1	1	2
Other/Multiple	1	1	1	3
Total	9	24	15	48

4.1.2 Materials

Audio dictation pre-test and post-test items were 15 target words from the training sets plus 9 control items. All items were 2-syllable words from 4 to 8 letters long (*M* = 6.67). Sentence completion pre-test and post-test items were 24 two-choice cloze sentence contexts. Distractor items were orthographically and phonologically, but not semantically, similar to the correct items.

Training items were 30 real English words. All items were 2-syllable words from 4 to 8 letters long (*M* = 6.40). The words were divided into two lists, which were matched for length in letters, number of syllables, number of phonemes, frequency (Kučera and Francis, Thorndike-Lorge written, and Brown verbal, and Zeno), familiarity, concreteness, and imageability. Half of the training items were word triplets (e.g., labor, label, lapel) which share orthographic and phonological but not semantic overlap.

Materials were programmed and presented in RunTime Revolution 3.0. Training and testing items, as well as their lexical characteristics, are listed in Appendix B.

4.1.3 Design

The second spelling intervention used a 2 (training condition: form focus or form-meaning integration) x 2 (set grouping: form similarity or random) within-subjects design. Dependent variables were accuracy on the audio dictation and sentence completion tasks. Gains in learning were measured as the difference between pre-test and post-test scores. Subject variables were first language background and second language proficiency.

4.1.4 Procedure

Students participated in the testing and training tasks over the course of one semester of ELI Writing classes, as follows:

- Time 1: Two pre-test tasks were administered—the audio dictation task and the sentence completion task.
- Time 2: Study session #1, 15 items were presented in the form focus training condition to half the students (Group A) and in the form-meaning training condition to the other half of the students (Group B).

- Time 3: Study session #2, a second set of 15 items were presented in the form-meaning integration training condition to half the students (Group A) and in the form focus training condition to the other half of the students (Group B).
- Time 4: The audio dictation and sentence completion tasks were administered again as post-tests.

Audio dictation task. The student sat at a computer monitor while wearing headphones and was instructed to click on an audio icon to hear the target word pronounced by a native English speaker, and then to type the target word into a text box shown on the screen. The student was allowed to click the audio icon as many times as needed, and to correct and retry spelling by using the backspace key on the computer. However, after proceeding to the next word, the student was not allowed to go back to any previous word.

Sentence completion task. Two-choice cloze sentences were presented on a computer, one at a time. The student was instructed to click on the word that correctly completed the sentence (see Figure 33).

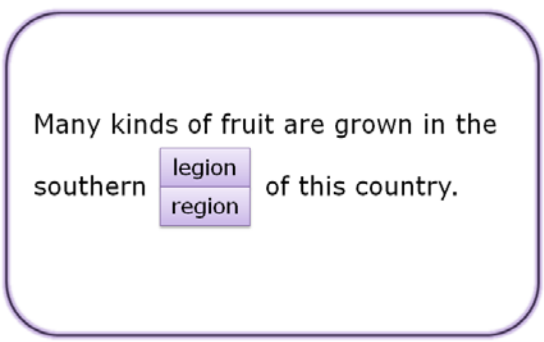

Figure 33. Screen shot of the sentence completion task.

Training sessions. Training sessions were essentially the same as in the first intervention study, with three additional features: Words were grouped into sets of three, with a brief review preceding the production tasks slides; students were allowed to check and retry their spellings on the production task slides; and at the end of each week's session, students were offered the chance to review the complete set of 15 training items.

Pre-test and post-test sessions were held in a computer lab, and students were tested individually. All training sessions were held in a computer classroom so that each student could open and proceed through a version of the executable file programmed in RunTime Revolution 3.0. Completed data files were collected by the experimenter at the end of each class session.

4.2 INTERVENTION STUDY #2 – RESULTS

On the sentence completion pre-test, students scored an average of 78.55% accuracy (*SD* = 15.78). Mean gains in accuracy from pre-test to post-test are shown in Figure 34. For words in the form focus training condition, students gained an average of 5.86 percentage points (*SD* = 2.37). For words in the form-meaning integration training condition, students gained an average of 3.20 percentage points (*SD* = 2.35). For core words learned in class but not in the intervention, students gained an average of 4.86 percentage points (*SD* = 3.31). For untrained words in the control condition, students gained an average of 5.67 percentage points (*SD* = 4.99). According to the results of planned paired *t*-tests, there were no statistically significant differences in gains between the two training conditions, or between trained and untrained words, all p-values > 0.05.

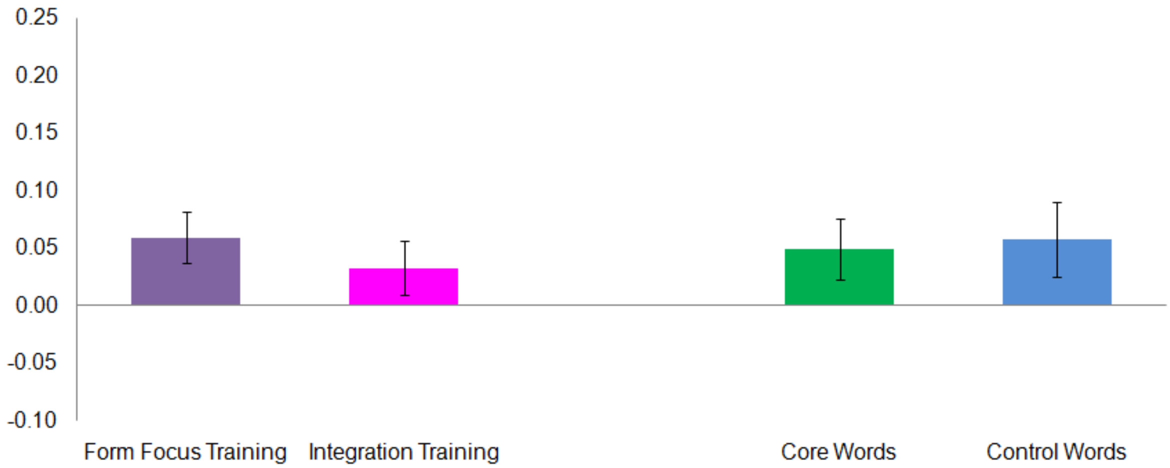

Figure 34. Mean accuracy gains on the sentence completion task.

On the audio dictation pre-test, students scored an average of 61.63% accuracy (*SD* = 25.04). Mean gains in accuracy from pre-test to post-test are shown in Figure 35. For words in the form focus training condition, students gained an average of 13.13 percentage points (*SD* = 3.25). For words in the form-meaning integration training condition, students gained an average of 12.99 percentage points (*SD* = 3.29). For core words (vocabulary words learned in class but not in the intervention), students gained an average of 12.15 percentage points (*SD* = 4.82). For untrained words in the control condition, students lost an average of .01 percentage points (*SD* = 6.38). There was a statistically significant difference between trained words and control words, $t(138) = 3.72$, $p < .001$.

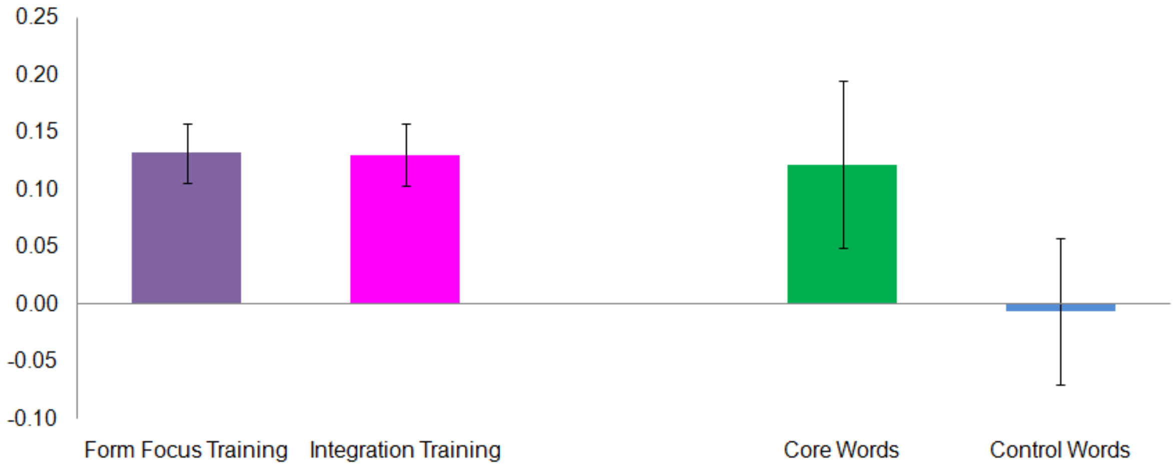

Figure 35. Mean accuracy gains on the audio dictation task.

Mean gains in accuracy on each task by training condition, first language background, and second language proficiency are shown in Figures 36-39.

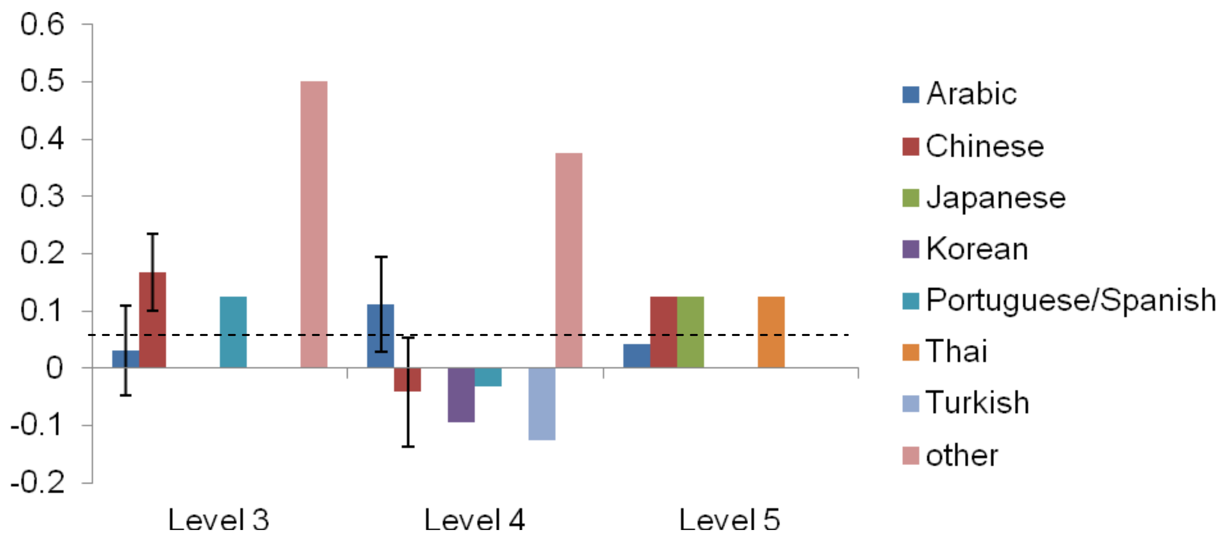

Figure 36. Mean accuracy gains on form-meaning integration training items in the sentence completion task, by L1 and ELI level. The dashed line represents the overall mean gain.

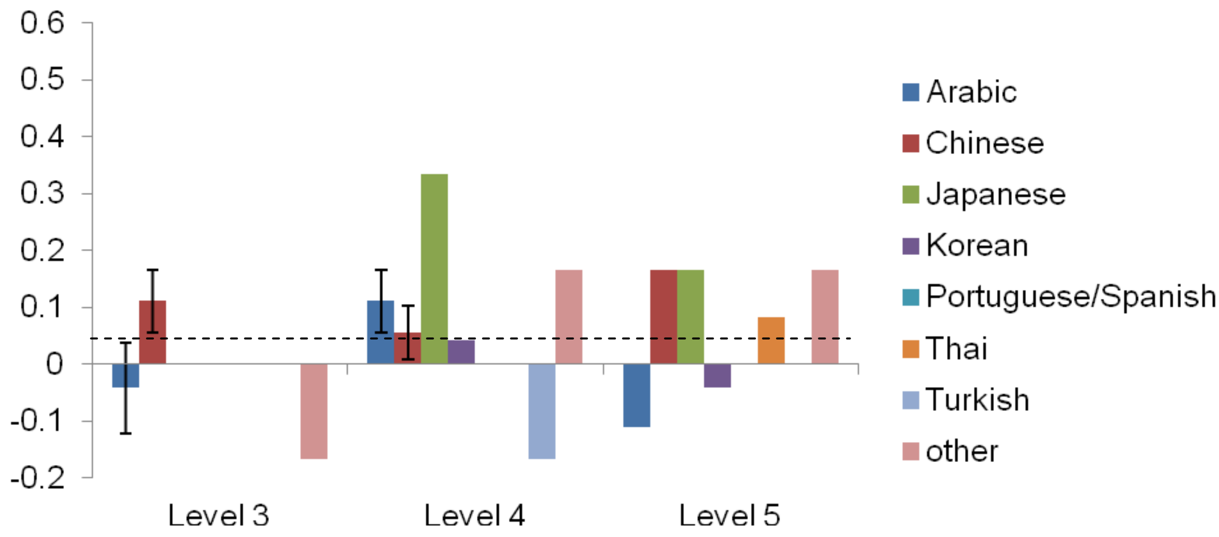

Figure 37. Mean accuracy gains on form focus training items in the sentence completion task, by L1 and ELI level. The dashed line represents the overall mean gain.

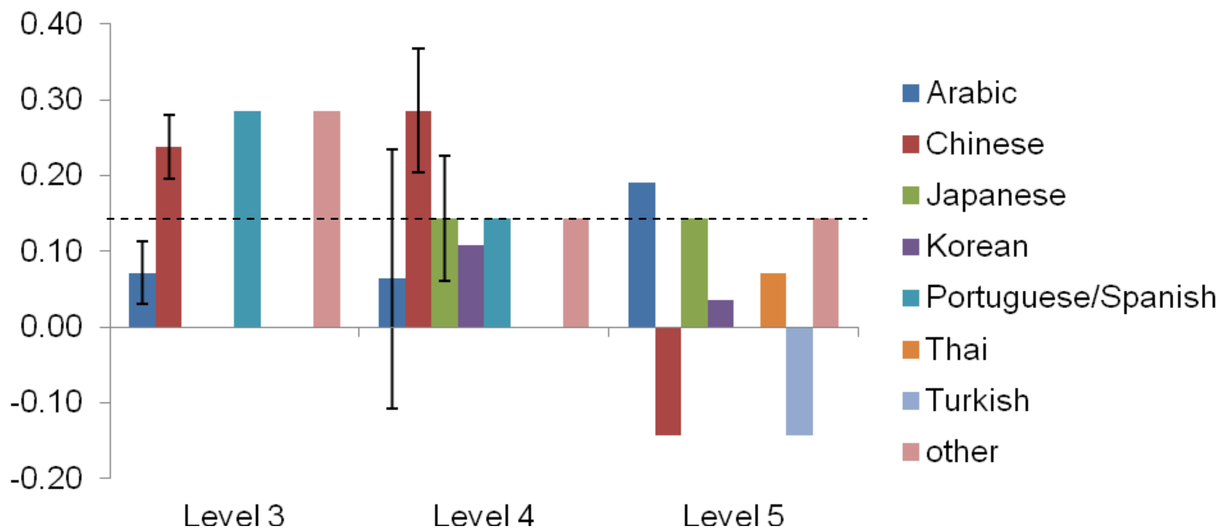

Figure 38. Mean accuracy gains on form-meaning integration training items in the audio dictation task, by L1 and ELI level. The dashed line represents the overall mean gain.

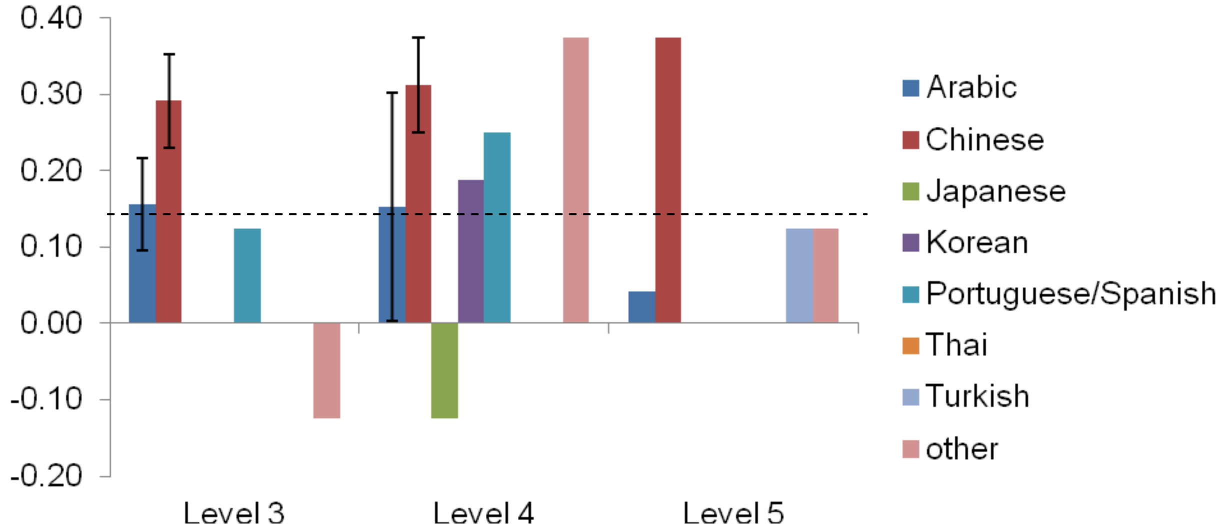

Figure 39. Mean accuracy gains on form focus training items in the audio dictation task, by L1 and ELI level. The dashed line represents the overall mean gain.

4.3 INTERVENTION STUDY #2 – DISCUSSION

After approximately one hour of training in either a form focus condition or a form-meaning integration condition, students showed small gains in a sentence completion task and moderate gains in an audio dictation task. Again, there was no main effect of training condition; gains were equivalent for the form-focus training and the form-meaning integration training on both outcome measures. As in the first training intervention, students showed greater gains in a production task (spelling dictation) than a recognition task (sentence completion).

There were no clear-cut main effects of L1 background or L2 proficiency (ELI level) on either the sentence completion task or the audio dictation task. On the sentence completion task, gains were greater in the form-meaning integration condition for a small number of students: Portuguese and Spanish students in Levels 3 and 5. Gains were greater in the form focus condition for Level 4 Japanese and Korean students. On the audio dictation task, gains were greater in the form-meaning integration condition for Japanese students in Levels 4 and 5, as well as for Arabic, Korean, and Thai Level 5 students. Gains were greater in the form focus condition for Arabic students in Level 3 and Chinese and Turkish students in Level 5.

Students in the English Language Institute are already intermediate to advanced in their English proficiency, so it is possible that they are able to take maximal advantage of any kind of training. The two training conditions, which were intended to differ in direction of focus and depth of encoding, were in essence equivalent when students were highly motivated to improve their English language skills. While this population of learners might already be too proficient for investigating developmental processes (cf Verhoeven, 2000), they are well-positioned for investigating lexical factors effecting retrieval of orthographic representations in English as a second language.

5.0 COGNITIVE EXPERIMENTS

Three cognitive tasks—audio dictation, cross-modal matching, and spell checking—were used to test the effects of lexical characteristics on orthographic recognition and production in English as a second language, according to the framework of the proposed cognitive model (see Figure 11). Word frequency, regularity, and age of acquisition were manipulated in the audio dictation task in order to test knowledge of word-specific forms as well as phoneme-grapheme mappings. In the cross-modal matching task, type of mismatch between auditory and visual presentations was manipulated in order to test the influence of orthographic and phonological similarity on word knowledge. In the spell checking task, type of spelling error was manipulated in order to test the influence of phonological information on orthographic knowledge. First language background and second language proficiency were also predicted to have effects on task performance.

5.1 COGNITIVE EXPERIMENTS – METHOD

5.1.1 Participants

Participants were 27 adult learners of English as a second language, recruited from the University of Pittsburgh's English Language Institute in Spring and Summer 2012. Of the 27 students, 14 were female and 13 were male; 25 were right-handed and 2 were left-handed. Their ages ranged from 18 to 37 years, *M* = 26.47. First languages represented were: Arabic, Chinese, Korean, Italian, Japanese, Montenegrin, Portuguese, and Thai (see Table 4). Only one of these students reported knowing multiple languages. In general, they first started studying English as a foreign language in elementary school (approximately age 10) in their home countries.

Table 4. L1 background and L2 proficiency for participants in the cognitive experiments

	Level 3	Level 4	Level 5	Total
Arabic	6	2	3	11
Chinese	1	2	5	8
Korean	0	0	3	3
Italian	1	0	0	1
Montenegrin	0	1	0	1
Japanese	0	1	0	1
Portuguese	0	1	0	1
Thai	0	1	0	1
Total	8	8	11	27

5.1.2 Materials

Materials for the *audio dictation task* were 40 real English words—5 apiece in each of the cells in the 2 (frequency: higher, lower) x 2 (regularity: higher, lower) x 2 (age of acquisition: earlier, later) design. Items ranged in length from 4 to 8 letters, number of syllables from 1 to 3, and number of phonemes from 2 to 8. None were homophones (e.g., blew, blue) or homographs (e.g., lead), and none were words that have different spellings in British English and American English (e.g., colour/color, realise/realize). Testing materials were programmed in RunTime Revolution 3.0. For a complete list of stimuli and their lexical characteristics, see Appendix C.

Materials for the *cross-modal matching task* were 160 word pairs: an auditory stimulus paired with a visually presented identity match (“yes” trials), or an orthographic foil, a phonological foil, an orthographic and phonological foil, or a transposition foil (“no” trials). Auditory stimuli were all real English words; none were homophones or proper nouns. Identity matches were the correctly spelled written form of the auditory stimulus (e.g., wager-wager). Each orthographic foil was an orthographic but not phonological neighbor of its paired auditory stimulus (e.g., wager-lager). Likewise, each phonological foil was a phonological but not orthographic neighbor of its paired auditory stimulus (e.g., wager-major). The orthographic and phonological foil was both an orthographic and a phonological neighbor with its paired auditory stimulus (e.g., wager-pager). Transposition foils were real words that shared all the same letters as the auditory stimulus, but in a different order (e.g., trial-trail). None of the stimuli had been used as target words in the audio dictation task. Testing materials were programmed in E-Prime 2.0 (Schneider, Eschman, & Zuccolotto, 2000). For a complete list of word pairs in the cross-modal matching task, see Appendix D.

Materials for the spell checking task were a subset of those used by Harris (2012). From the 837 words in her task, I selected 240 real English words, paired with their incorrect spellings. The 240 items were divided into two lists. In version A, half of the words (120) were spelled correctly. Of the incorrectly spelled items, half (60) were errors that preserved the phonology of the word, and half (60) were errors that altered the phonology of the word. In version B, words that had been spelled correctly in version A now comprised the incorrect items, again half with phonology preserving errors; words that had been incorrectly spelled in version A now comprised the correct items. Word lists were counterbalanced such that half of the participants received version A; the other half received version B. None of the spell-checking items appeared in either the audio dictation task or the cross-modal matching task. Words ranged in length from 5 to 10 letters, *M* = 7.10. Testing materials were programmed in E-Prime 2.0 (Schneider, Eschman, & Zuccolotto, 2000). For a complete list of items used in the spell checking task, see Appendix E.

5.1.3 Design

Subject variables

For all the experimental tasks, subject variables were the participants' first language (L1) background (e.g., Arabic, Chinese, Korean) and English second language (L2) proficiency (Level 3, 4, or 5 in ELI writing classes).

Audio dictation task

On the audio dictation task, independent variables were word frequency (higher, lower), word sound-spelling regularity (higher, lower), and word age of acquisition (earlier, later). *Word frequency* was defined according to values from the Subtlex-US database (Brysbaert & New, 2009). This database contains ratings on 74,286 English words, with frequency values ranging from .02 to 29,449.18, and with a mean of 13.12. For the audio dictation tasks, words with a frequency rating (from Brysbaert & New's 2009 SubtlexUS database) greater than 13.12 were categorized as higher frequency; words with a frequency rating less than 13.12 were categorized as lower frequency. *Sound-spelling regularity* was defined according to a modified version of Lange's (1997/2001, 2002) grapheme-phoneme association strength. Words with higher scores were categorized as higher regularity. *Age of acquisition* was defined using values from the MRC database (http://www.psy.uwa.edu.au/mrcdatabase/uwa_mrc.htm). Words with age of acquisition ratings less than 300 were categorized as earlier acquired; words with age of acquisition ratings greater than 300 were categorized as later acquired. This division is equivalent to words acquired by native speakers either before or after 3 years of age.

Dependent variables on the audio dictation task were: the number of times a participant listened to the target word, and the participants' accuracy on the completed spelling of the target word, scored as 0 for incorrect and 1 for correct.

Cross-modal matching task

On the cross-modal matching task, the independent variable was type of mismatch between the auditory stimulus and the visual stimulus. Visually presented words were either an identity match to the auditory stimulus (match), or one of four kinds of mismatches: orthographic foil, phonological foil, orthographic and phonological foil, or transposition foil.

Dependent variables on the cross-modal matching task were response times and accuracy scores.

Spell checking task

On the spell checking task, the independent variable was error type (none, phonology preserving, or phonology altering).

Dependent variables on the spell checking task were the participants' response times and accuracy.

5.1.4 Procedure

Participants were tested individually. They completed the three experimental tasks and then a brief demographic questionnaire. All data were collected in one session, which took approximately one hour per subject.

Spelling dictation task

Stimuli were 40 English words presented orally, one at a time, via headphones. Participants were instructed to click on the audio icon as many times as needed to hear the target word, and to spell the word they heard by typing into a text box on the computer screen. Responses were tracked using RunTime Revolution 3.0, which allowed for recording of the number of listening attempts made on each word, each keystroke (including backspaces), and the student's final answer. Each participant completed three practice trials prior to the experimental items.

Cross-modal matching task

Stimulus pairs comprised a real English word presented orally via headphones, followed by a real English word presented visually on a computer screen. Participants were instructed to press a "yes" button if the visual word was the same as the spoken word, or a "no" button if it was not (see Figure 40).

Spell checking task

Stimuli were 240 letter strings, presented visually one at a time on a computer screen. Participants were instructed to press a "yes" button if the item was a correctly spelled word in English, or a "no" button if the item was not a correctly spelled English word. Each stimulus remained on the screen until the participant made a yes or no response; skipping items was not possible. After making each response, the participant saw a hash mark symbol (###) and had to press the space bar to initiate the subsequent trial. This allowed for a self-paced progression through the task with breaks whenever needed.

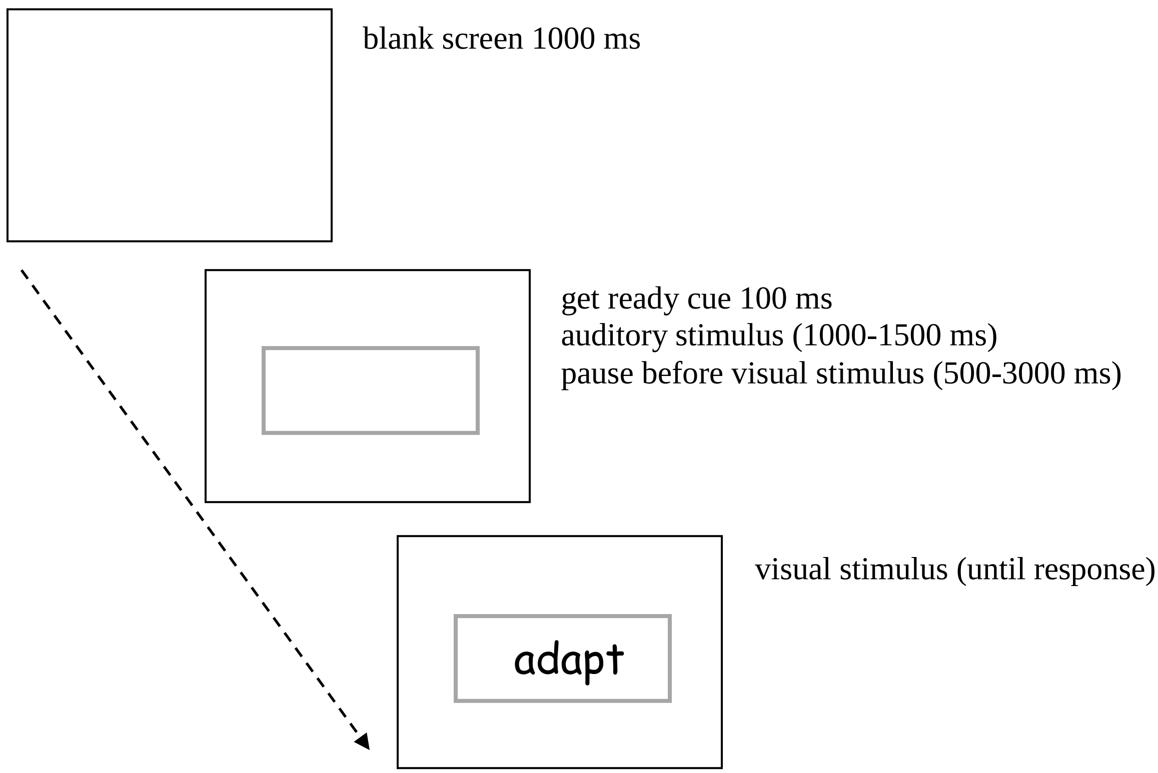

Figure 40. Trial procedure for the cross-modal matching task.

5.2 COGNITIVE EXPERIMENTS – RESULTS

5.2.1 Audio dictation task

The average number of listening attempts per word on the audio dictation task was 2.18. Mean listening rates by age of acquisition, frequency, and regularity are shown in Table 5. Analyses of variance showed a significant main effect of age of acquisition, $F(1,39) = 23.29$, $p < .001$, and a main effect of frequency, $F(1,39) = 36.88$, $p < .001$, but no effect of regularity. There was also a significant interaction between age of acquisition and frequency, $F(1,39) = 9.36$, $p < .01$. This interaction was synergistic; the effect of frequency was greater for later acquired words than earlier acquired words.

Table 5. Mean (and SD) number of listening attempts on the audio dictation task

	Higher Frequency	Lower Frequency
Earlier Age of Acquisition	Higher Regularity	Higher Regularity
	1.63 (.37)	1.57 (.16)
	Lower Regularity	Lower Regularity
	1.92 (.49)	2.18 (.66)
Later Age of Acquisition	Higher Regularity	Higher Regularity
	1.99 (.51)	1.73 (.31)
	Lower Regularity	Lower Regularity
	3.19 (.71)	3.26 (.29)

Overall accuracy on the audio dictation task was 60.65%. Accuracy rates by age of acquisition, frequency, and regularity are shown in Table 6. Analyses of variance showed a significant main effect of age of acquisition, $F(1,39) = 7.62$, $p < .01$, and a main effect of frequency, $F(1,39) = 50.87$, $p < .001$, but no effect of regularity. There was also a significant interaction between age of acquisition and frequency, $F(1,39) = 4.93$, $p < .05$. Again, this interaction was synergistic; the effect of frequency was greater for later acquired words than earlier acquired words.

Table 6. Mean accuracy (and standard deviations) on the audio dictation task

	Higher Frequency	Lower Frequency
Earlier Age of Acquisition	Higher Regularity .83 (.13)	Higher Regularity .63 (.24)
	Lower Regularity .82 (.13)	Lower Regularity .46 (.29)
Later Age of Acquisition	Higher Regularity .76 (.13)	Higher Regularity .37 (.23)
	Lower Regularity .83 (.11)	Lower Regularity .15 (.12)

Listening rates and accuracy rates were weakly correlated, -0.23. To some extent, students clicked more often to hear words that were more difficult to spell.

5.2.2 Cross-modal matching task

Response times greater than 3 standard deviations (per subject) were excluded from analyses. Additionally, any trial for which a subject reported not hearing the auditory stimulus was excluded from analyses. Together, this resulted in loss of 3% of trials. For remaining trials, the mean response time on the cross-modal matching task was 1670 ms. Response times by mismatch condition are shown in Figure 41. There was a main effect of foil type such that transposition foils were responded to more slowly than all other items, and phonological foils were responded to fastest, $F(4) = 11.70$, $p < .01$.

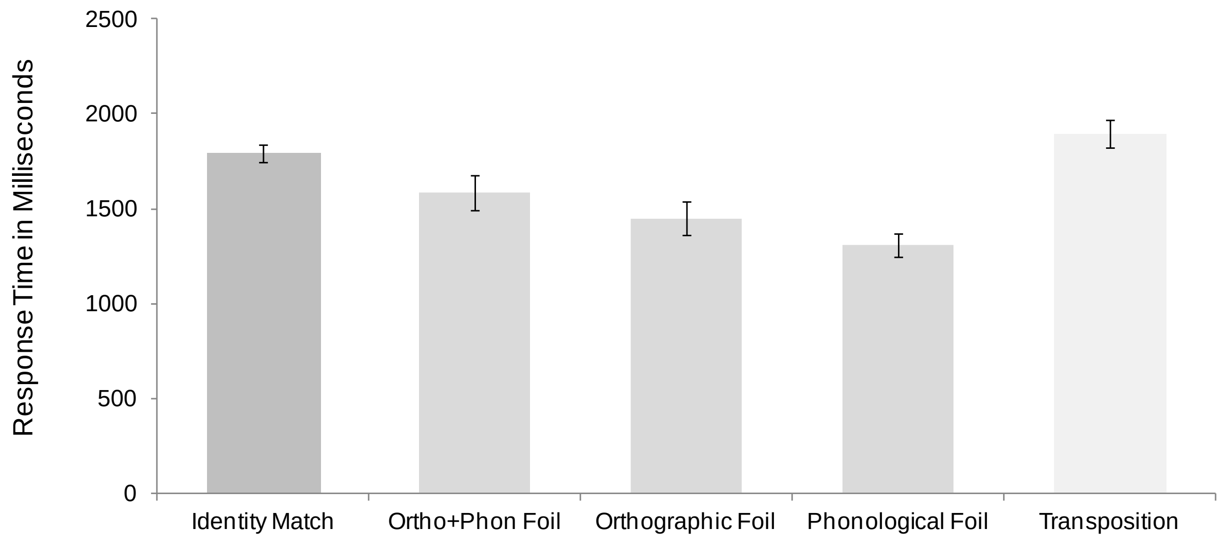

Figure 41. Mean response times on the cross-modal matching task.

Overall accuracy on the cross-modal matching task was 87.65%. Accuracy rates by mismatch condition are shown in Figure 42. Again, there was a main effect of foil type such that students made more errors on transposition foils than all other items, and they had the highest accuracy for phonological foils. $F(4) = 10.76$, $p < .01$.

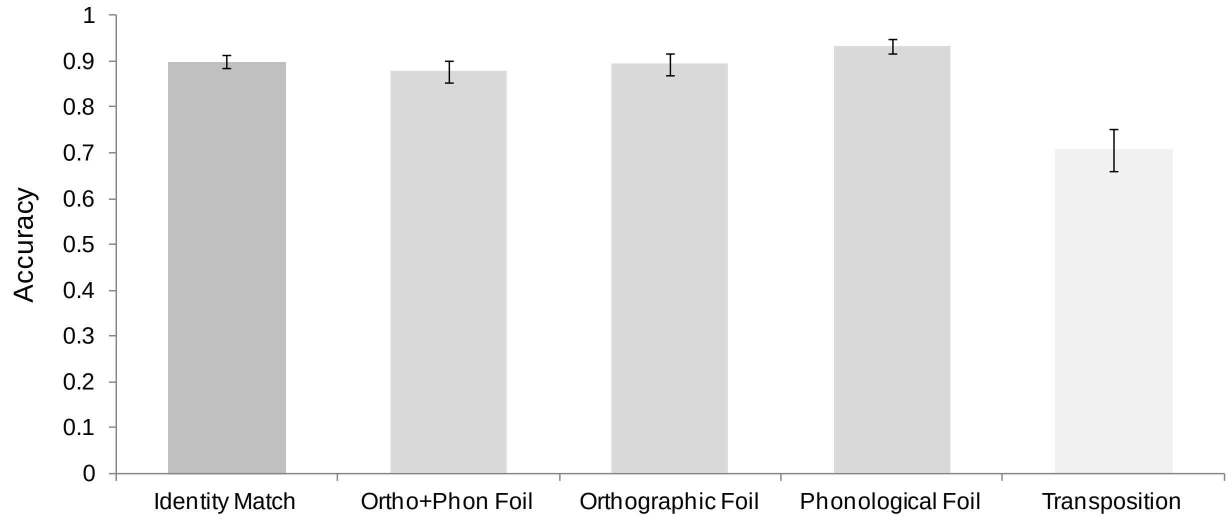

Figure 42. Mean accuracy on the cross-modal matching task.

The differences between accuracy on the identity matches and each mismatch condition were calculated. Mismatch effects are shown for first language background groups (Arabic versus non-Arabic) and second language proficiency (Level 3, 4, or 5 in ELI writing classes) in Figure 43.

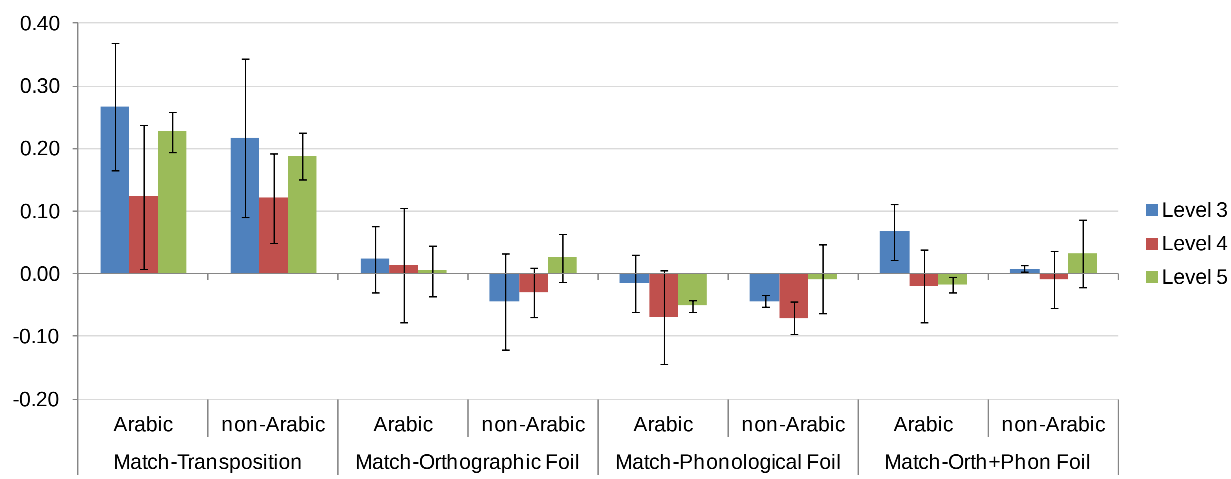

Figure 43. Mismatch effects by L1 group and L2 proficiency.

5.2.3 Spell checking task

For each participant, response times greater than 3 standard deviations were excluded from analyses. This resulted in loss of 2.34% of trials. For remaining trials, the mean response time on the spell checking task was 2818 ms. Response times by condition are shown in Figure 44. A paired sample *t*-test showed a main effect of word type such that correctly spelled items were responded to faster than incorrectly spelled items, $t(119) = 3.35$, $p < .001$. There was no difference in response times between error conditions, $t(59) = 1.20$, $p = 0.12$.

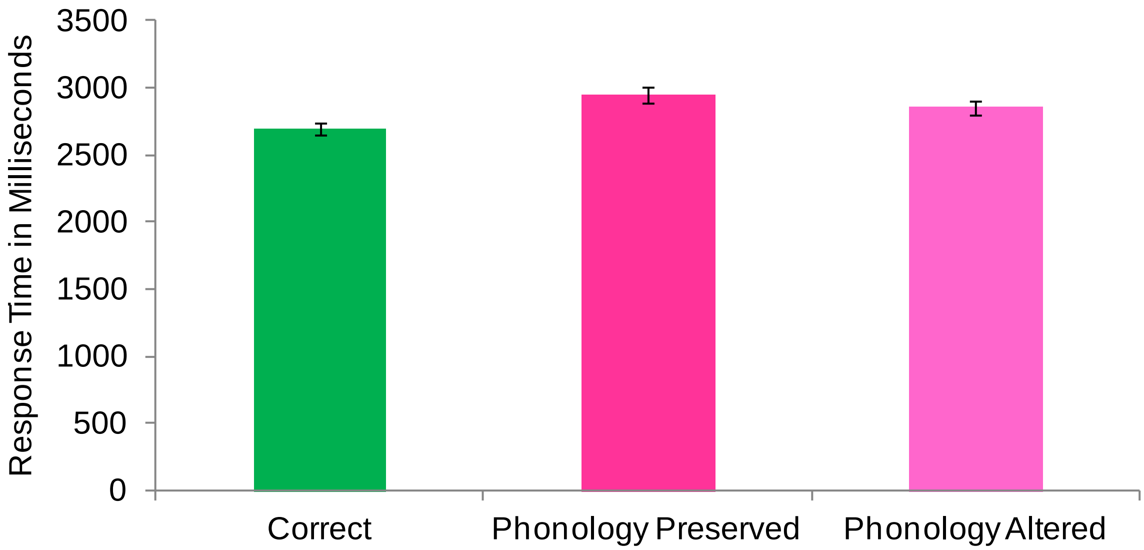

Figure 44. Mean response times on the spell checking task.

Overall accuracy on the spell checking task was 80.19%. Accuracy rates by condition are shown in Figure 45. A paired sample *t*-test showed no difference between hits and correct rejections, $t(119) = 0.95$, $p = 0.17$. However, students had significantly higher accuracy for incorrectly spelled items when the phonology was altered items than when phonology was preserved, $t(59) = 8.05$, $p < .001$.

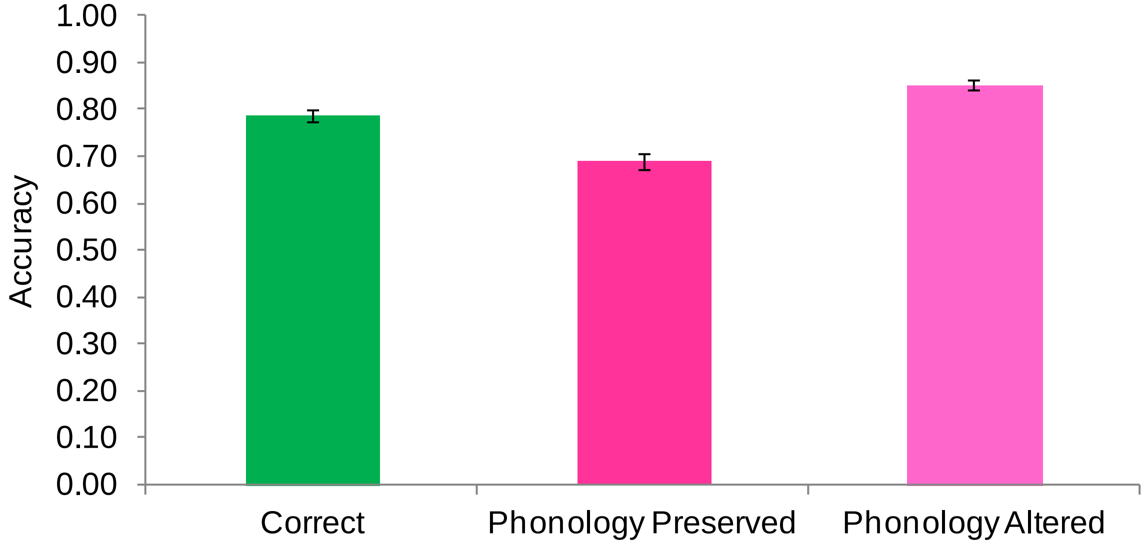

Figure 45. Mean accuracy on the spell checking task.

Mean accuracy scores for first language background groups (Arabic versus non-Arabic) and second language proficiency (Level 3, 4, or 5 in ELI writing classes) are shown in Figure 46.

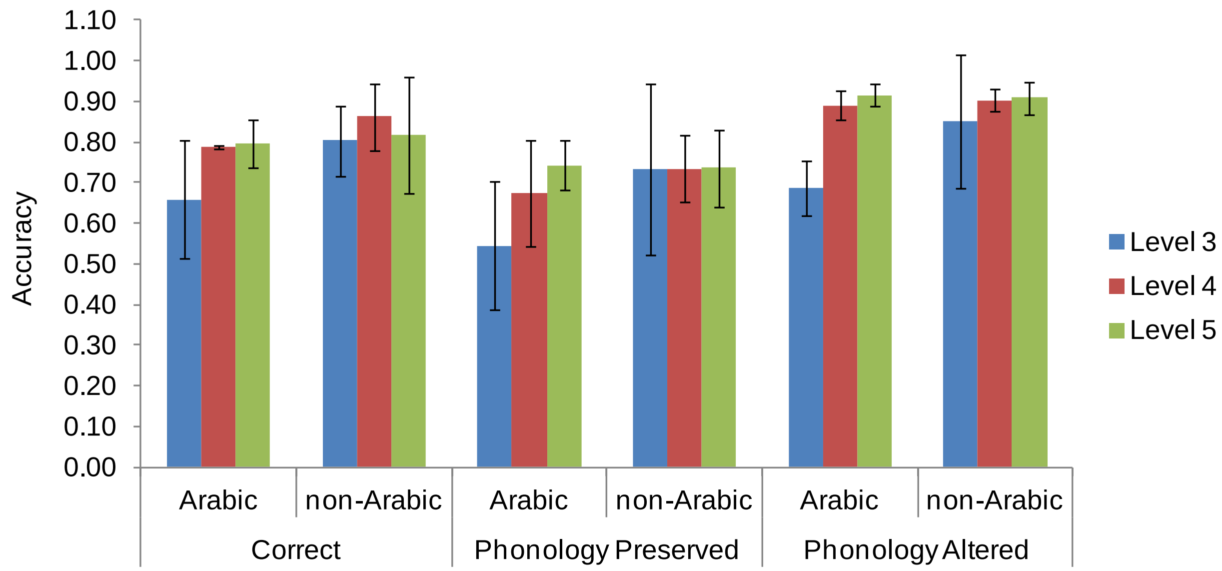

Figure 46. Mean accuracy on the spell checking task by L1 group and L2 proficiency.

5.3 COGNITIVE EXPERIMENTS – DISCUSSION

Non-native English speakers were tested for effects of lexical characteristics on audio dictation, cross-modal matching, and spell checking tasks.

On the *audio dictation task,* participants listened most often to words that were later acquired and lower frequency, regardless of regularity. Repeated listening is taken as an indication that the student was unfamiliar with the word. Participants had lowest spelling accuracy on words of lower frequency and lower regularity. Interestingly, spelling errors could be characterized as relying either on top-down or on bottom-up processes. If a student does not have a semantic or orthographic representation for the target word "siege," for instance, it could be misspelled via phoneme-grapheme mappings as "seej" or "seege." This would be evidence of a bottom-up strategy. On the other hand, the student might search for the closest lexical entry, and produce “see” as the incorrect answer. This would be evidence of a top-down strategy. Both kinds of errors were made by students in the present study, but the majority of misspellings were attempts at encoding the word according to some phoneme-grapheme mapping rules. The finding of an interaction between word age of acquisition and frequency supports the arbitrary mapping hypothesis (Ellis & Lambon Ralph, 2000; Monaghan & Ellis, 2002a; 2002b; Zevin & Seidenberg, 2002, 2004).

On the *cross-modal matching task*, participants had increased response times and decreased accuracy in the transposition mismatch condition, relative to all other mismatch conditions. This finding suggests that learners have knowledge of the individual letters comprising a word, but the order of the letters is fragile and susceptible to confusion. The fact that orthographic and phonological neighbors did not disrupt accuracy suggests that the task was relatively easy given unlimited time to respond. After students heard the auditory stimulus, there was a brief interval (500-3000 ms) in which they could formulate a tentative encoding of that stimulus. Then, when the visual stimulus appeared, the student would need to compare their generated form to the presented form. A mismatch would necessitate either a “no” response or further analysis of the two presented forms. Contrarily, a student could merely hold the auditory stimulus in memory, as in Baddeley’s phonological loop, until the visual stimulus appears. At that point, their task is to decode the written form and verify its match to the spoken form. It appears that students used the former strategy and hence were least susceptible to foils that maintained phonology but not orthography, and most susceptible to foils that maintained all the target letters but not in their proper order.

On the *spell checking task*, participants had slower response times and lowest accuracy for incorrectly spelled words that preserved phonology. This suggests that phonology is indeed activated during decoding, and subsequently disrupts retrieval of the already weak orthographic knowledge. In terms of the lexical quality hypothesis, the link from phonology to semantics is stronger than the link from orthography to semantics. In accord with Andrews and Lo (2012), even advanced speakers of English as a second language would be categorized as “less skilled” in the sense that the quality of their L2 lexical representations is not solid.

6.0 GENERAL DISCUSSION

6.1 SUMMARY OF RESULTS

A set of studies investigated the nature of orthographic knowledge in advanced learners of English as a second language (ESL). In a data mining study, students’ spelling errors were analyzed. Results showed that first language (L1) background and second language (L2) proficiency affect the rates and types of spelling errors made. In particular, Level 4 students (intermediate-advanced) made more errors than Level 3 or Level 5 students. Arabic L1 students made the most errors overall and had the most difficulty with vowel spellings in English; Spanish L1 students had the most difficulty with consonant spellings in English. These findings confirmed teacher observations of first language differences in spelling ability, and also motivated two spelling interventions designed to test the benefits of two different types of instruction on improving orthographic knowledge in ESL learners.

In the two training interventions, students showed equivalent learning gains from two different types of spelling instruction: a form focus condition and a form-meaning integration condition. Students’ gains were greater for a productive task than a receptive task, suggesting that mere exposure throughout the course of a semester accounts for improved recognition of orthographic forms, but that some instruction dedicated to production of orthographic forms is needed to improve production of orthographic forms.

In the three cognitive experiments, non-native English speakers were shown to be sensitive to several lexical characteristics in both receptive and productive spelling tasks. In an audio dictation task, non-native English speakers were shown to be sensitive to word frequency and age of acquisition but not regularity. This finding is consistent with the arbitrary mapping hypothesis, which states that earlier learned words will be more advantaged than later learned words when the mapping from one lexical constituent to another is inconsistent (Ellis & Lambon Ralph, 2000; Monaghan & Ellis, 2002a; 2002b; Zevin & Seidenberg, 2002, 2004). In the present set of studies, it is the mapping from phonology to orthography which is very irregular in English's deep orthography. The absence of regularity effects suggests that students develop orthographic representations via exemplar-based rather than rule-based strategies. Again, because in English spelling there are as many exceptions as there are rules, a rule-based approach is not any more efficient as it might be in shallower writing systems.

In a cross-modal matching task, students were most susceptible to transposition foils that preserved target letters but in an incorrect order, and least susceptible to phonological foils that preserved phonological but not orthographic form of the target word. This suggests that learners' orthographic knowledge consists of sublexical units (i.e., letters) but that their knowledge about order of these units is unstable. Students easily rejected phonological foils, which shared phonology but not orthography with the target word. This suggests that participants did encode the auditory stimulus prior to appearance of the visual stimulus.

In a spell checking task, students had more difficulty rejecting misspelled words that maintained the phonological form of the target word than misspelled words that did not preserve phonology of the target. This finding suggests that, provided with an orthographic form, ESL readers do activate phonology even in a task that does not require them to do so. This is consistent with Harris's (2012) findings with native English speakers on a similar version of this task. Furthermore, this phonological activation is disruptive to the retrieval of orthographic representations, which are relatively weak. For these students, the link from phonology to semantics appears to be stronger than the link from orthography to semantics.

6.2 THEORETICAL IMPLICATIONS

This research expands upon the lexical quality hypothesis (Perfetti & Hart, 2001) by applying it to learners of English as a second language, and by focusing on quality of the orthographic constituent of lexical knowledge. Future research could be dedicated to the next logical step: the effects of quality of orthographic knowledge on reading comprehension in a second language. For example, one could make predictions about the importance of having developed fine-grained distinctions so that a reader is not prone to decoding errors which would inhibit comprehension. Evidence from eye-tracking studies in native English readers (e.g., Nelson, 2010) suggests that non-native English readers would also show effects of lexical characteristics as well as proficiency on reading ability. Previous electrophysiological studies (e.g., Balass, 2011; Harris, 2012) suggest that the ERPs of less proficient non-native English readers would show evidence of decreased awareness of textual errors, including syntactic violations, semantic violations, and spelling errors.

This research did not find that the benefit of form-meaning integration extends to learning of orthographic forms in English as second language. The meta-analysis done by Norris and Ortega (2000) showed that explicit instruction of form-meaning relationships was ideal for learning grammatical structures. In English, as in most languages, there are many rules governing syntactic structure. When rules are easy to deduce, implicit instruction is sufficient for acquisition (Reber, 1989). However, when rules are complicated, explicit instruction is better suited. Sound-spelling mappings in English are highly inconsistent. When there are rules, they apply selectively to sets of words which still might have exceptions (e.g., lint, mint, pint) or they depend on other factors such as the word's etymology (e.g., /f/ being spelled "ph" in Greek-derived words). Beyond basic phonics instruction, it becomes quite convoluted to encompass the many nuances of English spelling. The instructional methods in the spelling training interventions were implicit, in that they did not explain rules or explicitly point out phoneme-grapheme mappings. When this was the case, there was no difference found between a form-focused training condition and a form-meaning integration training condition.

6.3 PEDAGOGICAL IMPLICATIONS

There were equivalent gains in each of the two training interventions, but students reported liking the form-meaning integration training much more than the form focus training. One of the pedagogical implications of the present research would be to take this preference into consideration. Given limited classroom time, if both conditions are equally beneficial, then I would recommend the learning task that would keep students engaged and motivated to learn.

Another pedagogical implication relates to the training task format. Both training conditions in both training interventions required students to produce the spelling of the target vocabulary words. It is possible that this partly accounted for students' learning gains, particularly in the audio dictation outcome measure (e.g., Bosman, 1994, van Hell, Bosman, & Bartelings, 2003). In terms of the proposed model of orthographic retrieval and production

(Figure 11), reaching the planned output step completes the process of retrieval. Any path taken to reach that step ought to strengthen lexical representations.

6.4 FUTURE DIRECTIONS

One of the challenges of doing cross-linguistic research is assuring equivalence in linguistic factors across languages. In the current set of studies, English language learners from a dozen different first language backgrounds were represented. It was beyond the scope of this project to measure or control for factors such as instructional conditions in the students' home countries, or reading and spelling ability in the students' first languages, as interesting as it might be to know more about the effects of such factors on spelling in English.

In the set of cognitive experiments, word frequency, regularity, and age of acquisition were manipulated in an audio dictation task. There are many available databases for determining the frequency of occurrence of word type and tokens in written or spoken English. These databases are derived from a multitude of sources meant to represent the typical exposure to language a native speaker might experience. However, the databases might not be representative of a non-native speaker's typical experience with written or spoken English, even if living in an English speaking country. Nonetheless, the Subtlex-US database (Brysbaert & New, 2009) used in the present project appeared to serve as a good enough proxy for frequency of exposure for the sample of intermediate to advance learners of English as a second language.

Determining the age of acquisition for individual English words was also a challenge. The ratings for native English speakers was again used as a proxy for non-native speakers because no good databases exist yet for the latter. In future research, I would strongly recommend first creating a more accurate and representative set of ratings to reflect the order or word learning in second language instruction. Admittedly this would be an immense undertaking. But if done, it would allow research on age of acquisition effects to make stronger claims about the arbitrary mapping hypothesis, effects of frequency trajectories, and so on.

Defining regularity of English spellings was also a challenge. Previous work with has been limited to early learned words or one-syllable words (e.g., Lange, 1997/2001, 2002) and to decoding not encoding. Most measures failed to capture the variety of potential spelling options for each phoneme (e.g., the grapheme-to-phoneme ratio measure of word complexity, Saz, Lin, Eskenazi, 2012). Much of the irregularity of English spellings arises when morphemes—and usually syllables—are added to root words, altering grapheme-phoneme correspondences. For instance, more people might spell "definitely" correctly if they know that it is related to the word "finite." The current project made due with a modified version of Lange's measure of grapheme-phoneme association strength. But lack of systematic descriptions of irregularity in English is problematic for research both with native and with non-native speakers. In future research, I would strongly suggest developing a metric based not just on word bodies (e.g., cat, hat, bat) or word families with exceptions (e.g., gave, cave, have), but rather on the pool of graphemes matched to the pool of phonemes. More concretely, the vowel sound /i/ can be spelled with the letters *e*, *ea*, *ee*, *ei*, *eo*, *ey*, *i*, *ie*, *is*, *oe*, or *y*, as in the words: *be*, *tea*, *see*, *receive*, *people*, *key*, *ski*, *believe*, *debris*, *subpoena*, and *any*, respectively. Adding a silent *e* to a word increases spelling options for just this one example phoneme. A comprehensive measure of English sound-spelling irregularity would consider how frequently each sound is spelled with each potential letter or set of letters (including silent letters). Then, I predict, clearer irregularity effects could be seen in processing of orthographic and phonological neighbors.

In the data mining study, analyses were done only on tokens produced in the recorded speaking activities. If further analyses were done by errors on types as well as tokens, a more refined picture might emerge. For instance, it would be possible to see which words students consistently struggle with, or which misspelled words are usually spelled correctly by a student (implying carelessness rather than underspecified knowledge). Type/token ratios can be used to measure linguistic complexity and diversity, so it would also be possible to determine how much each student is challenging himself to use more difficult vocabulary.

6.5 CONCLUSION

Despite sporadic attempts to reform English spelling (e.g., Andrew Carnegie's American Simplified Spelling Board), it remains an orthographically deep writing system. Any learner or English, whether native or non-native, will face the challenge of encoding and decoding words with inconsistent grapheme-phoneme correspondences. In the present set of studies, the findings suggest that even intermediate to advanced learners of English as a second language still show difficulty with the language's deep orthography, but that they can benefit from minimal amounts of instruction, provided there the training task requires some orthographic production on the student's part. Furthermore, these students appear to be acquiring orthographic knowledge via exemplar-based rather than rule-based strategies. They show sensitivity to word frequency and age of acquisition, but surprisingly not regularity. This research expands upon the lexical quality hypothesis (Perfetti & Hart, 2001) by applying it to learners of English as a second language, with a particular focus on the orthographic constituent of lexical knowledge.

APPENDIX A

STIMULI FROM INTERVENTION STUDY #1

Number of syllables and length in phonemes were taken from the MRC Database (http://www.psy.uwa.edu.au/mrcdatabase/uwa_mrc.htm). The Zeno frequency values listed are the Standard Frequency Index values from *The Educator's Word Frequency Guide* (Zeno, Ivens, Millard, & Duvvuri, 1995). This value is a logarithmic transformation of a dispersion-weighted frequency of type per million tokens. BNC frequency values represent the combined written and spoken frequency taken from the British National Corpus (http://www.natcorp.ox.ac.uk/).

Table 7. Intervention #1 training items, set 1

ITEM	Letters	Syllables	Phonemes	Zeno Frequency	BNC Frequency
accustom	8	3	7	37.2	34
administer	10	4	9	45.4	538
celebrate	9	3	8	49.7	1388
choose	6	1	4	59.6	6707
column	6	2	5	55.7	2776
counsel	7	2	6	45.9	1285
definite	8	3	7	53.8	1555
dimension	9	3	8	47	1601
earlier	7	3	4	59.3	16370
endure	6	2	5	48.2	516
exotic	6	3	7	44.7	1119
family	6	3	6	66	33761
forbidden	9	3	6	50.2	891
freight	7	1	4	48.8	963
grammar	7	2	5	50	2414
historical	10	4	10	52.7	5513
immigrant	9	3	8	47	352
increase	8	2	6	60.2	16796
innovation	10	4	8	43.1	1693
intermediate	12	5	10	47.8	1353
irresponsible	13	5	11	40.8	418
nutrition	9	3	9	49.2	503
recreation	10	4	9	50.6	898
refinery	8	4	8	39.9	171
regrettable	11	4	9	32.3	234
remember	8	3	7	63.5	18448
sacred	6	2	6	51.2	1254
tuition	7	3	6	42.2	461
vegetable	9	4	8	51.8	955
widespread	10	2	8	52.2	3221

Table 8. Intervention #1 training items, set 2

ITEM	Letters	Syllables	Phonemes	Zeno Frequency	BNC Frequency
acquire	7	3	5	50.7	1986
actually	8	4	7	60.8	25430
assess	6	2	4	46.4	2662
attribute	9	3	8	45	731
cease	5	1	3	45.9	1002
cite	4	1	3	40.4	290
confined	8	2	7	50.2	2372
country	7	2	6	65.9	31401
defeat	6	2	5	51.7	3578
differentiate	13	5	11	42.9	503
diversity	9	4	8	47.8	1394
embarrass	9	3	7	40	202
enough	6	2	4	66.7	31149
estimation	10	4	9	39.2	342
famous	6	2	5	59.6	6407
flourish	8	2	6	44.9	663
friendly	8	2	7	56.9	3949
government	10	3	8	65.6	61987
guarantee	9	3	7	49.6	3006
infant	6	2	6	50.5	1672
infer	5	2	4	43.2	319
initial	7	3	6	52.6	6549
manipulates	11	4	10	32.8	34
neutral	7	2	7	52	1565
opportunity	11	5	10	57	10096
prepare	7	2	5	56.4	2965
restaurant	10	3	9	52.4	3438
sustainable	11	4	9	28.3	677
temperature	11	4	9	60.9	4340
unfortunately	13	5	11	53.4	4550

Table 9. Intervention #1, audio dictation pre-test/post-test control items

ITEM	Letters	Syllables	Phonemes	Zeno Frequency	BNC Frequency
barbecue	8	3	7	37.8	348
because	7	2	5	70.3	100509
calendar	8	3	7	50.2	1086
contradiction	13	4	12	41.8	779
council	7	2	5	56	31394
expansion	9	3	9	53.8	3532
intelligence	12	4	10	54.6	3431
opacity	7	4	7	31	58
something	9	2	6	67.8	50060
stereotype	10	4	9	42.5	277

Table 10. Intervention #1, lexical decision pre-test/post-test control items

ITEM (“yes” responses)	Letters	Syllables	Phonemes	Zeno Frequency	BNC Frequency
cemetery	8	4	7	46.7	732
immigration	11	4	9	43.7	1081
laundry	7	2	6	49.3	510
phenomenon	10	4	9	50.4	2174
tolerance	9	3	8	47.9	716

ITEM (“no” responses)	Letters
abandun	7
Amirican	8
conseption	10
repproduce	10
visble	6

APPENDIX B

STIMULI FROM INTERVENTION STUDY #2

Number of syllables and length in phonemes were taken from the MRC Database (http://www.psy.uwa.edu.au/mrcdatabase/uwa_mrc.htm). The Standard Frequency Index (SFI) comes from *The Educator's Word Frequency Guide* (Zeno, Ivens, Millard, & Duvvuri, 1995). This value is a logarithmic transformation of a dispersion-weighted frequency of type per million tokens. BNC values represent the combined written and spoken frequency taken from the British National Corpus (http://www.natcorp.ox.ac.uk/). Thorndike-Lorge (TL) and Kučera and Francis (KF) frequency values were also taken from the MRC database (http://www.psy.uwa.edu.au/mrcdatabase/uwa_mrc.htm), when available. AWL refers to the sublist number of the Academic Word List (Coxhead, 2000), when applicable.

Table 11. Intervention #2 training items, set 1

Item	Letters	Syllables	Phonemes	SFI	BNC	TL	KF	AWL
access	6	2	5	51	10789	34	24	4
achieve	7	2	5	54.5	6713	249	51	2
assess	6	2	4	46.4	2662	2	6	1
bias	4	2	4	46	1393	17	8	8
conceive	8	2	6	44.7	450	79	14	10
consist	7	2	7	52.3	1223	200	17	1
contact	7	2	7	57.2	10655	229	63	5
context	7	2	8	52.2	9248	7	35	1
emerge	6	2	4	49.3	2035	159	18	4
impose	6	2	5	46	1878	98	9	4
occur	5	2	3	58.4	5540	400	43	1
perceive	8	2	5	49	890	83	13	2
process	7	2	6	62.7	22483	293	196	1
region	6	2	5	59.8	9851	149	76	2
reveal	6	2	5	51.5	2601	243	30	6

Table 12. Intervention #2 training items, set 2

ITEM	Letters	Syllables	Phonemes	SFI	BNC	TL(L)	KF	AWL
dispel	6	2	6	34.1	219	19	3	--
displace	8	2	7	38.9	177	21	3	8
display	7	2	6	53.6	6150	232	41	6
function	8	2	7	57.8	8591	165	113	1
general	7	3	7	63.9	38313	770	497	--
journal	7	2	4	52.2	2445	1008	42	2
journey	7	2	4	56.5	4700	190	28	--
label	5	2	4	52.7	2044	90	19	4
labor	5	2	4	59.4	182	628	4	1
lapel	5	2	5	34.8	98	21	1	--
minor	5	2	4	53.5	4895	83	58	3
normal	6	2	5	58.7	12179	335	136	2
survey	6	2	4	52.1	8113	193	37	2
theory	6	2	4	59.4	12875	220	129	1
transfer	8	2	7	53.6	6833	118	38	2

Table 13. Intervention #2, audio dictation pre-test/post-test control items

ITEM	Letters	Syllables	Phonemes	SFI	BNC	TL	KF	AWL
approach	8	2	6	57	16005	460	123	1
conflict	8	2	8	55.7	5868	85	52	5
data	4	2	4	57.4	22179	26	173	1
distinct	8	2	8	52.6	3158	76	42	2
equate	6	2	5	34.9	260	--	8	2
margin	6	2	5	51.6	1443	71	10	5
persist	7	2	6	44.8	535	124	6	10
pursue	6	2	5	49.1	1934	108	20	5
schedule	8	2	6	52.8	2460	124	36	8

Table 14. Intervention #2, sentence completion pre-test/post-test items

Correct Response	Distractor Item	Sentence
achieve	conceive	I want to ___ my goal of graduating next year.
acquire	require	He likes to ___ a new car every year.
adapt	adopt	The teenager tried to ___ to his new school.
complex	context	The directions to get to his house are very ___.
contact	context	It is good to keep in ___ with your friends.
dispel	display	She wants to ___ bad influences from her life.
display	displace	The store had many beautiful clothes on ___.
emerge	merge	We hope that the sun will ___ from behind the clouds.
evolve	involve	Your opinions about people may ___ over time.
journey	journal	The doctor planned a long ___ to Australia.
label	labor	It is important to put a ___ on your graphs.
lapel	label	The reporter had a small microphone on his ___.
license	incense	The man has a special ___ to drive a bus.
migrate	migraine	Many birds ___ to warmer places in the winter.
minor	miner	She made a ___ change to the homework assignment.
obtain	sustain	I am going to ___ the new version of my favorite video game.
occur	occlude	The party will ___ at noon tomorrow.
perceive	conceive	Dogs can ___ sounds that people cannot hear.
process	access	Getting into college can be a difficult ___.
region	legion	Many kinds of fruit are grown in the southern ___ of this country.
require	acquire	That job might ___ a college education.
survey	survive	The researchers gave a ___ to everyone in the room.
sustain	obtain	The runner tried to ___ her steady rate during the race.
theory	thereby	I have a ___ about why English spelling is so unusual.

APPENDIX C

STIMULI FROM COGNITIVE EXPERIMENTS

AUDIO DICTATION TASK

Number of syllables, number of phonemes, and age of acquisition were taken from the MRC Database (http://websites.psychology.uwa.edu.au/school/MRCDatabase/uwa_mrc.htm).

Frequency values were the word frequencies per million according to Brysbaert and New's (2009) Subtlex-US database for United States English (http://subtlexus.lexique.org/moteur2/).

Table 15. Lexical characteristics of the audio dictation task items

Item	Letters	Syllables	Phonemes	Age of Acquisition	Frequency
acre	4	2	3	411	1.82
bedroom	7	2	6	206	36.71
bequest	7	2	7	600	0.20
blanket	7	2	7	211	12.98
booth	5	1	3	508	20.37
buffer	6	2	4	553	1.67
carnage	7	2	5	628	1.12
circle	6	2	4	214	21.51
coat	4	1	3	197	42.08
darkness	8	2	6	242	17.49
degree	6	2	5	508	14.88
disquiet	8	3	8	617	0.08
doll	4	1	3	161	24.76
elephant	8	3	7	222	11.37
fashion	7	2	5	467	18.76
finance	7	2	6	522	5.35
golf	4	1	4	364	25.53
knuckle	7	2	4	356	1.29
nutrient	8	3	9	611	0.37
oven	4	2	4	236	8.88
pencil	6	2	5	225	9.86
picture	7	2	6	219	138.45
pillow	6	2	4	217	11.39
pressure	8	2	5	444	53.12
puck	4	1	3	572	2.88
sequel	6	2	6	556	1.76
shoulder	8	2	5	264	26.20
siege	5	1	3	503	2.31
sock	4	1	3	172	8.98
spoon	5	1	4	186	7.61
strut	5	1	5	511	1.57
sunshine	8	2	6	206	11.84
theory	6	2	4	557	28.61
thumb	5	1	3	183	11.82
transfer	8	2	7	489	20.55
uncle	5	2	4	192	124.06
union	5	2	6	503	21.78
vote	4	1	3	486	34.33
water	5	2	4	153	225.06
winter	6	2	5	236	26.22

APPENDIX D

STIMULI FROM THE CROSS-MODAL MATCHING TASK

Table 16. Identity match items in the cross-modal matching task

adequate	purpose
allergy	quarter
apply	quilt
brace	radical
correction	relevant
courage	rhythm
danger	science
delicate	scream
earth	service
face	simple
grammar	stampede
gutter	stoop
height	stoves
imitate	suit
injure	super
interpret	talon
journal	tariff
label	therapy
lord	tragic
lucrative	vigil
mansion	violet
orchard	waiter
parcel	wash
proposition	widow
pull	zipper

Table 17. Foil items in the cross-modal matching task

Orthographic Foils		Phonological Foils	
Auditory Stimulus	Visual Stimulus	Auditory Stimulus	Visual Stimulus
align	alien	anchor	anger
along	among	arrange	arraign
beard	heard	attack	attach
blood	brood	baked	ached
body	bode	boast	most
bone	done	cancer	answer
cameo	camel	chore	door
chief	chef	clothing	closing
college	collage	comb	home
cross	gross	cough	off
daughter	laughter	crowd	proud
debit	debt	daisy	daze
demon	lemon	dare	chair
fever	never	drama	trauma
floor	flood	fashion	passion
freak	break	good	could
gave	have	his	fizz
give	five	home	foam
nature	mature	hutch	much
notice	novice	loose	deuce
perpetrate	perpetuate	love	of
previous	precious	many	penny
quest	guest	most	coast
ratio	patio	plaid	glad
said	raid	rarely	fairly
this	his	wallow	hollow
tour	hour	warm	form
tower	lower	watch	notch
wager	lager	worse	curse
work	fork	youth	tooth

Orthographic and Phonological Foils		Transposition Foils	
Auditory Stimulus	Visual Stimulus	Auditory Stimulus	Visual Stimulus
adapt	adopt	angel	angle
bath	path	broad	board
caused	paused	clam	calm
comma	coma	cloud	could
computer	commuter	complaint	compliant
conceal	congeal	diary	dairy
curious	furious	dose	does
decree	degree	expect	except
devote	denote	fear	fare
dread	thread	fiend	fined
feather	father	from	form
field	yield	quite	quiet
fist	list	reverse	reveres
flown	frown	sacred	scared
foot	soot	sauce	cause
gold	cold	slave	salve
gravel	travel	sliver	silver
hand	sand	trail	trial
lease	cease	tried	tired
meditate	medicate	until	unlit
mild	wild		
mouse	house		
pity	city		
procession	profession		
property	properly		
really	realty		
region	legion		
short	shout		
touch	torch		
weary	leery		

APPENDIX E

STIMULI FROM THE SPELL CHECKING TASK

Table 18. Items in the spell checking task

Version A – Correctly Spelled Items				
afraid	damage	history	ninety	rhythm
ambulance	debacle	horizon	northern	royal
annual	delicate	human	obstacle	sarcasm
appendix	destroy	hurricane	official	schedule
arena	devotion	image	opposite	scream
audience	document	immune	orange	seizure
balance	elbow	invent	paradise	skeleton
ballerina	eligible	junction	peculiar	souvenir
blame	embrace	jungle	perplex	station
blossom	errand	language	plastic	success
breakfast	establish	lantern	pocket	surgery
budget	exactly	lawyer	porch	surprise
business	excellent	leather	possible	tendency
career	exhale	library	probably	thirsty
ceremony	factor	liquid	prominent	thousand
clarity	familiar	lottery	protect	tragedy
coffee	famous	maiden	pyramid	umbrella
colony	fanatic	mechanic	quality	vacant
comfort	fatigue	medicine	reckless	velvet
competent	garbage	military	recommend	village
confirm	garden	mimicked	regular	voyage
courteous	genius	minimum	religious	wardrobe
creature	glory	murmur	represent	welcome
currency	harass	museum	return	yacht

Version A – Phonology Altered Items		Version A – Phonology Preserved Items	
alairm	intecept	absense	influince
anoether	kidnep	accesory	kichen
apertment	lameint	anatamy	legicy
aupron	maimmoth	arguement	licence
bailcony	manace	atention	loyel
bauchelor	meloday	baloon	magizine
cafetaria	negatuve	bargan	markit
catious	nughty	benifit	milage
centiry	papular	campis	moter
chacolate	peibble	catagory	necesary
cheldren	perheps	channal	ocasion
claenser	plaesure	colomn	originel
corruopt	profassor	compair	paralel
coutton	pronunce	constent	pardan
cuorage	qualifay	custamer	pasttime
entertan	quartit	delivary	peeple
escepe	remerk	dependant	procede
etarnal	remimber	dirtey	pusition
exest	reveval	disgise	rainge
exploide	riebbon	emergancy	refrence
furtune	smoike	encurage	relivant
garege	speiral	excede	responce
gratful	staendard	forcast	silvur
haizard	strutegy	fourty	skaite
hasband	suppart	furnature	squaire
haybrid	thrishold	gallary	steem
heisitate	torniedo	govenor	tommorrow
ievory	tweunty	helth	tradetion
impetient	vacotion	icecle	truble
incume	zippar	indistry	wagun

Version B – Correctly Spelled Items				
absence	compare	governor	menace	relevant
accessory	constant	grateful	mileage	remark
alarm	corrupt	hazard	motor	remember
anatomy	cotton	health	naughty	response
another	courage	hesitate	necessary	revival
apartment	customer	husband	negative	ribbon
apron	delivery	hybrid	occasion	silver
argument	dependent	icicle	original	skate
attention	dirty	impatient	parallel	smoke
bachelor	disguise	income	pardon	spiral
balcony	emergency	industry	pastime	square
balloon	encourage	influence	pebble	standard
bargain	entertain	intercept	people	steam
benefit	escape	ivory	perhaps	strategy
cafeteria	eternal	kidnap	pleasure	support
campus	exceed	kitchen	popular	threshold
category	exist	lament	position	tomorrow
cautious	explode	legacy	proceed	tornado
century	forecast	license	professor	tradition
channel	fortune	loyal	pronounce	trouble
children	forty	magazine	qualify	twenty
chocolate	furniture	mammoth	quartet	vacation
cleanser	gallery	market	range	wagon
column	garage	melody	reference	zipper

Version B – Phonology Altered Items		Version B – Phonology Preserved Items	
areuna	lequid	affraid	lenguage
ballerana	libary	ambulence	lether
blassom	mechenic	annuel	maidun
brakfast	midicine	apendix	milatary
buisiness	mimiced	audiance	minamum
camfort	muesum	ballance	murmer
ceremany	ninty	blaime	northurn
claority	paradase	budgit	obsticle
coeffee	payramid	carreer	oficial
craeture	perplax	colany	oposite
debecle	plestic	compatent	orenge
deilicate	probally	conferm	peculier
demage	protict	currancy	pockit
devoition	regalar	curteous	portch
documnet	reickless	distroy	possable
elbaw	retarn	eligable	prominant
embroce	saizure	errend	qualaty
exatly	sercasm	establush	recomend
fanaitic	skeileton	excelent	religous
gairden	souvener	exhail	reprisent
gerbage	stetion	facter	royel
hestory	teandency	familliar	rythm
horezon	tharsty	famos	schedual
huiman	thuosand	fategue	screem
hurricene	umbrulla	genious	sergery
invint	vayage	glorey	sucess
juingle	vealvet	harrass	suprise
juntion	vecant	imege	tradgedy
laentern	waelcome	imune	vilage
lattery	werdrobe	lauyer	yaght

BIBLIOGRAPHY

Abu-Rabia, S. (1997). Reading in Arabic orthography: The effect of vowels and context on reading accuracy of poor and skilled native Arab readers. *Reading and Writing: An Interdisciplinary Journal, 9*, 65-78.

Abu-Rabia, S., & Siegel, L. S. (2002). Reading, syntactic, orthographic, and working memory skills of bilingual Arabic-English speaking Canadian children. *Journal of Psycholinguistic Research, 31*(6), 661-678.

Al-Hamouri, F., Maestú, F., del Río, D., Fernández, S., Campo, P., Capilla, A., García, E., González-Marqués, J., & Ortiz, T. (2005). Brain dynamics of Arabic reading: A magnetoencephalographic study. *NeuroReport, 16*(7), 1861-1864.

Andrews, S., & Lo, S. (2011). Not all skilled readers have cracked the code: Individual differences in masked form priming. *Journal of Experimental Psychology: Learning, Memory and Cognition, 38*(1), 152-163.

Balass, M. (2004). Individual differences in representations of newly learned words: An event-related investigation. Unpublished Master's thesis, University of Pittsburgh.

Balass, M. (2011). *Learning words in context: An ERP investigation of word experience effects on familiarity and meaning acquisition*. Unpublished doctoral dissertation, University of Pittsburgh.

Balass, M., Nelson, J. R., & Perfetti, C.A. (2009). *Word learning: An ERP investigation of word* *experience effects on recognition and word processing.* Manuscript submitted for publication.

Balota, D. A., & Ferraro, F. R. (1993). A dissociation of frequency and regularity effects in pronunciation performance across young adults, older adults, and individuals with senile dementia of the Alzheimer type. *Journal of Memory and Language, 32,* 573-592.

Bates, T. C., Castles, A., Luciano, M., Wright, M. J., Coltheart, M., & Martin, N. G. (2007). Genetic and environmental bases of reading and spelling: A unified genetic dual route model. *Reading and Writing, 20,* 147-171.

Bialystok, E., & Hakuta, K. (1994). *In other words: The psychology and science of second language acquisition.* New York: Basic Books.

Bosman, A. M. T. (1994). *Reading and spelling in children and adults: Evidence for a single-route model*. Unpublished doctoral dissertation, Universiteit van Amsterdam.

Bruder, M. N., & Henderson, R. T. (1985). Beginning reading in English as a second language. *Language in Education: Theory and Practice, 64,* 1-76.

Brysbaert, M., & New, B. (2009). Moving beyond Kučera and Francis: A critical evaluation of current word frequency norms and the introduction of a new and improved word frequency measure for American English. *Behavior Research Methods, 41*(4), 977-990.

Caramazza, A., Miceli, G., Villa, G., & Romani, C. (1987). The role of the graphemic buffer in spelling: Evidence from a case of acquired dysgraphia. *Cognition, 26* 59-85.

Caravolas, M. (2004). Spelling development in alphabetic writing systems: A cross-linguistic perspective. *European Psychologist, 9*(1), 3-14.

Chall, J. S. (1979). The great debate: Ten years later, with a modest proposal for reading stages. In L. B. Resnick & P. A. Weaver (Eds.), *Theory and practice of early reading*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Chen, B. G., Zhou, H. X., Dunlap, S., & Perfetti, C. A. (2007). Age of acquisition effect in reading Chinese: Evidence in favor of the arbitrary mapping hypothesis. *British Journal of Psychology, 98*, 499-516.

Chikamatsu, N. (1996). The effects of L2 orthography on L2 word recognition: A study of American and Chinese learners of Japanese. *Studies in Second Language Acquisition, 18,* 403-432.

Chikamatsu, N. (2006). Developmental word recognition: A study of L1 English readers of L2 Japanese. *The Modern Language Journal, 90*, 67-85.

Cocking, R. R., & McHale, S. (1981). A comparative study of the use of pictures and objects in assessing children's receptive and productive language. *Journal of Child Language, 8,* 1-13. doi:10.1017/S030500090000297X

Conrad, M., Carreirras, M., Tamm, S., & Jacobs, A. M. (2009). Syllables and bigrams: Orthographic redundancy and syllabic units affect visual word recognition at different processing levels. *Journal of Experimental Psychology: Human Perception and Performance, 35*(2), 461-479.

Coxhead, A. (2000). A new academic word list. *TESOL Quarterly, 34*(2), 213-238.

Craik, F. I. M., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior, 11*, 671-682.

Dijkstra, T., & van Heuven, W. J. B. (2002). The architecture of the bilingual word recognition system: From identification to decision. *Bilingualism: Language and Cognition, 5*(3), 175-197. doi:10.1017/S1366728902003012

Ehri, L. C. (1991). Learning to read and spell words. In L. Rieben & C. A. Perfetti (Eds.), *Learning to read: Basic research and its implications*. Hillsdale, NJ: Lawrence Erlbaum Associates, pp. 57-73.

Ehri, L. C., & Wilce, L. S. (1985). Movement into reading: Is the first stage of printed word learning visual or phonetic? *Reading Research Quarterly, 20*, 163-179.

Ellis, A. W., & Lambon Ralph, M. A. (2000). Age of acquisition effects in adult lexical processing reflects loss of plasticity in maturing systems: Insights from connectionist networks. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26*, 1103-1123.

Fender, M. (2001). A review of L1 and L2/ESL word integration skills and the nature of L2/ESL word integration development involved in lower-level text processing. *Language Learning, 51*, 319-396.

Fender, M. (2003). English word recognition and word integration skills of native Arabic- and Japanese-speaking learners of English as a second language. *Applied Psycholinguistics, 24*, 289-315.

Fender, M. (2008). Spelling knowledge and reading development: Insights from Arab ESL learners. *Reading in a Foreign Language, 20*(1), 19-42.

Fiez, J., Balota, D., Raichle, M. E., & Petersen, S. E. (1999). Effect of lexicality, frequency, and spelling-to-sound consistency on the functional anatomy of reading. *Neuron, 24*, 205-218.

Figueredo, L. (2006). Using the known to chart the unknown: A review of first language influence on the development of English-as-a-second-language spelling skill. *Reading and Writing, 19*, 873-905. doi:10.1007/s11145-006-9014-1

Frith, U. (1985). Beneath the surface of developmental dyslexia. In K. E. Patterson, J. C. Marshall, & M. Coltheart (Eds.), *Surface dyslexia*. London: Routledge and Kegan Paul.

Frost, R., Katz, L., & Bentin, S. (1987). Strategies for visual word recognition and orthographical depth: A multilingual comparison. *Journal of Experimental Psychology: Human Perception and Performance, 13*, 104-115.

Gentry, J. R. (1982). An analysis of developmental spelling in "*GNYS AT WRK*." *The Reading Teacher, 36*(2), 192-200.

Goswami, U., Ziegler, J. C., & Richardson, U. (2005). The effects of spelling consistency on phonological awareness: A comparison of English and German. *Journal of Experimental Child Psychology*. doi:10.1016/j.jecp.2005.06.002.

Gough, P. B., & Juel, C. (1991). The first stages of word recognition. In L. Rieben & C. A. Perfetti (Eds.), *Learning to read: Basic research and its implications*. Hillsdale, NJ: Lawrence Erlbaum Associates, pp. 47-56.

Harris, L. N. (2012). Error-related negativities during spelling judgments expose orthographic knowledge. Unpublished Master's thesis, University of Pittsburgh.

Hayes-Harb, R. (2006). Native speakers of Arabic and ESL texts: Evidence for the transfer of written word identification processes. *TESOL Quarterly, 40*(2), 321-339.

Henderson, E. (1981). *Learning to read and spell: The child's knowledge of words.* Dekalb, IL: Northern Illinois University Press.

Henderson, E. H., & Beers, J. W. (1980). *Developmental and cognitive aspects of learning to spell: A reflection of word knowledge.* Newark, DE: International Reading Association.

Henderson, E. H., & Templeton, S. (1986). A developmental perspective of formal spelling instruction through alphabet, pattern, and meaning. *The Elementary School Journal, 86*(3), 304-316.

Hino, Y., & Lupker, S. J. (2000). Effects of word frequency and spelling-to-sound regularity in naming with and without preceding lexical decision. *Journal of Experimental Psychology: Human Perception and Performance, 26*(1), 166-183.

Hofman, J. E., & Habib-Allah, M. (1982). The cloze technique in Arabic: Words or semantic units? *Journal of Psycholinguistic Research*, *11*(4), 275-282.

Krashen, S. (1989). We acquire vocabulary and spelling by reading: Additional evidence for the input hypothesis. *Modern Language Journal, 73*(4), 440-464.

Kučera, H., & Francis, W. N. (1967). *Computational analysis of present-day American English*. Providence, RI: Brown University Press.

Landi, N., Perfetti, C. A., Bolger, D. J., Dunlap, S., & Foorman, B. R. (2006). The role of discourse context in developing word form representations: A paradoxical relationship between reading and learning. *Journal of Experimental Child Psychology, 94,* 114-133.

Lange, M. (1997/2001). Quantitative description of grapheme–phoneme relations in English.

Lange, M. (2002). Activation of multiple phoneme associates of graphemes in visual word recognition. *Brain and Language, 81*, 610-620. doi:10.1006/brln.2001.2551

Lee, S. H., & Muncie, J. (2006). From receptive to productive: Improving ESL learners' use of vocabulary in a postreading composition task. *TESOL Quarterly, 40*(2)295-320.

Lété, B., Peereman, R., & Fayol, M. (2008). Consistency and word-frequency effects on spelling among first- to fifth-grade French children: A regression-based study. *Journal of Memory and Language, 58*, 952-977. doi:10.1016/j.jml.2008.01.001

Liu, Y., Dunlap, S., Fiez, J., & Perfetti, C. A. (2003). Evidence for neural accommodation to a writings system following learning. *Human Brain Mapping*, *28*, 1223-1234.

Macizo, P., & Van Petten, C. (2006). Syllable frequency in lexical decision and naming of English word. *Reading and Writing.*

MacWhinney, B. (2005a). A unified model of language acquisition. In J. F. Kroll & A. M. B. de Groot (Eds.), *Handbook of bilingualism: Psycholinguistic approaches* (pp. 49-67). New York: Oxford University Press.

MacWhinney, B. (2005b). Extending the competition model. *International Journal of Bilingualism*, 9(1), 69-84.

Margolin, D. I. (1984). The neuropsychology of writing and spelling: Semantic, phonological, motor, and perceptual processes. *The Quarterly Journal of Experimental Psychology*, *36A*, 459-489.

Martin, K. I. (2011). *Reading in English: A comparison of native Arabic and native English speaker*. Unpublished Master's thesis, University of Pittsburgh.

Martinet, C., Valdois, S., & Fayol, M. (2004). Lexical orthographic knowledge develops from the beginning of literacy acquisition. *Cognition*, *91*(2), B12-B22.

Medler, D. A., & Binder, J. R. (2005). MCWord: An on-line orthographic database of the English language. http://www.neuro.mcw.edu/mcword/.

Miceli, G., & Capasso, R. (2006). Spelling and dysgraphia. *Cognitive Neuropsychology, 23*(1), 110-134.

Monaghan, J., & Ellis, A. W. (2002a). What exactly interacts with spelling-sound consistency in word naming? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 28*,183–206.

Monaghan, J., & Ellis, A. W. (2002b). Age of acquisition and the completeness of phonological representations. *Reading and Writing*, *15*, 759–788.

Morrison, C. M., & Ellis, A. W. (1995). The roles of word frequency and age of acquisition in word naming and lexical decision. *Journal of Experimental Psychology: Learning, Memory and Cognition, 21*, 116–133.

Nassaji, H. (2005). The development of spelling and orthographic knowledge in English as an L2: A longitudinal case study. *Canadian Journal of Applied Linguistics, 10*(1), 77-98.

Nation, I. S. P. (2001). *Learning vocabulary in another language*. Cambridge: Cambridge University Press.

Nelson, J. R. (2010). *Reading skill and components of word knowledge affect eye movements during reading*. Unpublished doctoral dissertation, University of Pittsburgh.

Norris, J., & Ortega, L. (2000). Effectiveness of L2 instruction: A research synthesis and quantitative meta-analysis. *Language Learning, 50*, 417-528.

Perfetti, C. A., & Dunlap, S. (2008). Learning to read: General principles and writing system variations. In K. Koda & A. M. Zehler (Eds.), *Learning to read across languages: Cross-linguistic relationships in first- and second-language literacy development* (pp. 13-38). New York: Routledge.

Perfetti, C. A., & Hart, L. A. (2001). The lexical bases of comprehension skill. In D. Gorfien (Ed.), *On the consequences of meaning selection* (pp. 67–86). Washington, DC: American Psychological Association.

Perfetti, C. A., Liu, Y., Fiez, J., Nelson, J, Bolger, D. J., & Tan, L. H. (2007). Reading in two writing systems: Accommodation and assimilation of the brain's reading network. *Bilingualism: Language and Cognition, 10*(2), 131-146.

Perfetti, C. A., Rieben, L., & Fayol, M. (1997). *Learning to spell: Research, theory, and practice across languages*. Lawrence Erlbaum Associates: Mahwah, NJ.

Polak, J., & Krashen, S. (1988). Do we need to teach spelling? The relationship between spelling and voluntary reading among college ESL students. *TESOL Quarterly, 22*(1), 141-146. doi:10.2307/3587067

Rapcsak, S. Z., Henry, M. L., Teague, S. L., Carnahan, S. D., & Beeson, P. M. (2007). Do dual-route models accurately predict reading and spelling performance in individuals with acquired alexia and agraphia? *Neuropsychologia*, *45*, 2519-2524.

Reber, A. S. (1989). Implicit and tacit knowledge. *Journal of Experimental Psychology: General, 118*, 219-235.

Red, D. L. (1999). Adults learning to read in a different script: What we've learned. In Alatis, J. E., & Tan, A. H. (Eds.), *Language in our time: Bilingual education and official English, Ebonics, and Standard English, immigration, and the Unz initiative*. Georgetown University Round Table on Languages and Linguistics. Washington, DC: Georgetown University Press.

Rice, G. A., & Robinson, D. O. (1975). The role of bigram frequency in the perception of words and nonwords. *Memory and Cognition*, *3*(5), 513-518.

Ryan, A., & Meara, P. (1991). The case of the invisible vowels: Arabic speakers reading English words. *Reading in a Foreign Language, 7*(2), 531-540.

Saz, O., Lin, Y., & Eskenazi, M. (2010). *What REAP teaches us about vocabulary learning*. Presentation at the Second Language Acquisition Research Symposium, University of Pittsburgh.

Schmidt, R. W. (1990). The role of consciousness in second language learning. *Applied Linguistics, 11*(2), 129-158.

Schneider, W., Eschman, A., & Zuccolotto, A. (2000). *E-Prime User's Guide*. Pittsburgh: Psychology Software Tools, Inc.

Seymour, P. H. K., Aro, M., & Erskine, J. M. (2003). Foundation literacy acquisition in European orthographies. *British Journal of Psychology, 94*(2), 143-174.

Share, D. L. (2008). On the Anglocentricities of current reading research and practice: The perils of over-reliance on an "outlier" orthography. *Psychological Bulletin, 134*, 584-615.

Spencer, K. (2007). Predicting children's word-spelling difficulty in common English words from measures of orthographic transparency, phonemic and graphemic length, and word frequency. *British Journal of Psychology, 98*(2), 305-338.

Swain, M. (1985). Communicative competence: Some roles of comprehensible input and comprehensible output in its development. In S. Gass & C. Madden (Eds.), *Input in second language acquisition* (pp. 235–253). Rowley, MA: Newbury House.

Swain, M. (1993). The output hypothesis: Just speaking and writing aren't enough. *Canadian Modern Language Review, 50*, 158-164.

Thompson-Panos, K., & Thomas-Ruzic, M. (1983). The least you should know about Arabic: Implications for the ESL writing instructor. *TESOL Quarterly, 17*,609-623.

Thorndike, E. L., & Lorge, I. (1968). *The teacher's word book of 30,000 words*. New York: Teachers College Press.

Tokowicz, N., & MacWhinney, B. (2005). Implicit and explicit measures of sensitivity to violations in second language grammar: An event-related potential investigation. *Studies in Second Language Acquisition, 27*, 173-204.

Treisman, R., & Bourassa, D. C. (2000). The development of spelling skill. *Topics in Language Disorders, 20*, 1-18.

van Hell, J. G., Bosman, A. M. T., & Bartelings, M. C. G. (2003). Visual dictation improves the spelling performance of three groups of Dutch students with spelling disabilities. *Learning Disability Quarterly, 26*(4), 239-255.

VanPatten, B., Farmer, J. L., & Clardy, C. L. (2009). Processing instruction and meaning-based output instruction: A response to Keating and Farley (2008). *Hispania, 92*(1), 116-126.

Verhoeven, L. (2000). Components in early second language reading and spelling. *Scientific Studies of Reading, 4*(4), 313-330.

Wade-Woolley, L. (1999). First language influences on second language word reading: All roads lead to Rome. *Language Learning, 49*, 447-471.

Wade-Woolley, L., & Siegel, L. S. (1997). The spelling performance of ESL and native speakers of English as a function of reading skill. *Reading and Writing: An Interdisciplinary Journal, 9*, 387-406.

West, M. (1953). *A general service list of English words with semantic frequencies and a supplementary word-list for the writing of popular science and technology*. London: Longman.

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

Zevin, J. D., & Seidenberg, M. S. (2002). Age of acquisition effects in word reading and other tasks. *Journal of Memory and Language, 47*, 1-29.

Zevin, J., & Seidenberg, M. (2004). Age-of-acquisition effects in reading aloud: Tests of cumulative frequency and frequency trajectory. *Memory and Cognition, 32*, 31-38.

Zhou, X. L., Ye, Z., Cheung, H., & Chen, H. C. (2009). Processing the Chinese language: An introduction. *Language and Cognitive Processes*, *24*, 929-946.