The definitive version of this article was published as *Hamilton, T. and Watson Todd, R. (2016) Investigating models for second language spelling.* International Journal of Applied Linguistics *vol. 167 no. 1 pp. 16-45. doi: 10.1075/itl.167.1.02ham*

It is available at http://www.jbe-platform.com/content/journals/10.1075/itl.167.1.02ham

Investigating models for second language spelling

Abstract

In spelling research, data is easily quantifiable and offers a possible glimpse into the mind's cognitive mechanisms. Previous research has focused on two cognitive routes assumed to be used for spelling in differing situations: one route enabling spelling of words from our lexical memory, and another route facilitating sublexically constructed spellings based on a writer's rules for how phonemes map to graphemes. As the dual-route model emerged from first language alphabetic spelling data, there is a lack of research which synthesizes second language research with first language spelling models. This paper's analysis of second language spellings suggests that the traditional dual-route model of spelling may not be universally applicable to second language spellers. Instead, the data suggests that consideration of the differences between L1 and L2 writing systems may help identify directions towards developing a comprehensive model of second language spelling.

Keywords: Spelling; Dual-Route Model; Orthographic Depth Hypothesis; Grapheme-Phoneme Correspondences; Second Language Writing

Although models for first language (L1) spelling have become well established in the literature for languages using Roman alphabets, more recent theories regarding second language (L2) spelling production have yet to be synthesized into a comprehensive model. To create this model, the different possible factors influencing L2 spelling need to be identified and their influences on correct spelling measured. These factors might be similar to those found to influence L1 spellings, or they may be L2 specific. L1 models have primarily focused on the regularity of a word's spelling pattern and the frequency of exposure to that pattern, but the applicability of these factors for L2 spelling requires further investigation.

An L2 model may need to consider L1 transfer issues. Previously, researchers have had difficulties objectively measuring L1 influences on L2. One aspect of language where objective measurements are perhaps easiest is spelling (Brown & Ellis, 1994), especially when the two languages being compared are disparate. Thai and English writing systems are both alphabetic, but as they use different scripts and have distinct orthographic depths, Thai spellers of English could potentially provide clear insights into L2 spelling.

The Dual-Route Model of Spelling

Perhaps the most influential model of spelling is the dual-route model of spelling, sometimes referred to as the standard spelling model (Plaut, McClelland, Seidenberg, & Patterson, 1996, p. 56). As the name suggests, this model proposes that there are two distinct processes involved in spelling. The two extremes of these processes would explain the successful spelling of unfamiliar regularly spelled words and familiar irregularly spelled words.

Unfamiliar words or pseudowords are thought to be spelled through assembly of predicted sound to spelling patterns calculated from previous spelling experience, also referred to as Phoneme Grapheme Correspondence (PGC). However, PGC mappings would fail to predict the spellings of many words in a writing system with inconsistent PGC, like English. Therefore, the ability to correctly spell irregular PGC words is thought to be the result of memorized lexical chunks. These two approaches to spelling have commonly been described in various versions of the dual-route model of spelling (e.g. Houghton & Zorzi, 2003; Kreiner & Gough, 1990; Rapcsak, Henry, Teague, Carnahan, & Beeson, 2007).

Although spelling is sometimes described as the outcome of a single cognitive mechanism (e.g. Bullinaria, 1994; Olson & Caramazza, 1994), most spelling models predict that the key factors influencing spelling outcomes are the familiarity of a word and the regularity of its spelling (or consistency of the word's spelling patterns within the speller's idiolect). In L1 spellings, the dual-route model predicts that familiar words should be spelled correctly, regardless of regularity. As the familiarity of a word decreases, it is less likely to be remembered as a lexical chunk, and is therefore less likely to be spelled correctly. For words not remembered as chunks, spellers rely on PGC, and thus, regularly spelled unfamiliar words are more likely to be spelled correctly than irregularly spelled unfamiliar words (as long as the speller is using a valid PGC system).

Within this broad dual-route model, there are numerous variations in approach. Although some researchers have investigated the dual-route's ability to predict sublexical spelling performance (e.g. Kreiner & Gough, 1990), this research intends to only test the fundamental premise of the dual-route spelling model at the lexical level.

How a word's familiarity and regularity is operationalized is another variation in the literature. Lexical familiarity is problematic to determine, since each speaker's lexis is idiosyncratic. However, word frequency in a large reference corpus, such as the British National Corpus (BNC), provides a rough simulacrum for lexical familiarity with more frequent words in the corpus more likely to be familiar (Author, 2013). Hereafter, lexical familiarity based only on corpus frequency will be referred to as the variable of Frequency. Because spellers may have varying vocabulary breadths (especially L2 subjects), some researchers have used a receptive vocabulary test as a method for predicting a speller's familiar Frequency range (e.g. Caravolas, Hulme, & Snowling, 2001). Hereafter, a speller's individual vocabulary breadth will be referred to as the variable of Familiarity. This research investigates whether Frequency influences the correct spellings of target words, while Familiarity is used to analyze spelling performance within and beyond a speller's vocabulary breadth.

There are two main theoretical approaches to the representation of sound-to-spelling knowledge. In one view, favored by earlier researchers (e.g. Hanna, Hanna, Hodges, & Rudorf, 1966; Venezky, 1970) individual phonemes are matched to graphemes (or clusters of graphemes) to determine a word's PGC regularity. Algorithms utilizing this approach are said to accurately predict spellings for roughly 50% of English words (Spencer, 2007). Thus, words deemed to be highly regular should be spelled correctly using this system. More recently, research has begun focusing on consistency between larger lexical chunks, such as onsets and rimes (e.g. Kessler & Treiman, 1997). However, Kessler and Treiman (2001) later revised their consistency measures concluding that the most plausible spelling model "fundamentally operates on the phonemic level, but can take into account the context in which each phoneme is found" (p. 611). To make direct comparisons with previous research which compared PGC regularity and spelling

performance in L1 and L2 (e.g. Brown, 1970; Lester, 1964), the current research will calculate the predictability of a word's spelling by mapping individual phonemes to corresponding graphemes based on patterns found in a large corpus. Following Carney (1994), modifications were made to incorporate weighted frequencies of the most frequent 30,000 words found in the BNC when calculating PGCs (as found on wyrdplay.org/AlanBeale/sound-table-4.html).

However, occurrences of PGCs in a large corpora can only roughly approximate a speller's PGC rules, and thereby predict the likelihood a word will be spelled correctly which follows those rules. In line with previous dual-route spelling research, the current study does not attempt to measure all the variables that may influence a speller's PGC rules. Rather, this study attempts to minimize the possibility that these variables are skewing regularity predictions by employing various control measures. Committee selection methods are used to reduce issues of context sensitivity identified by Kessler and Treiman (2001) and semantic and etymological influences (Cummings, 1988). Randomization methods are used to control for the lexical neighborhood effects (Tainturier, Bosse, Roberts, Valdois, & Rapp, 2013). Because syntax may influence spelling patterns (Kessler & Treiman, 2003), only nouns are used as stimuli in this research. Because research suggests that a phoneme that maps to a single grapheme may be easier for a speller to process than a phoneme that maps to a cluster of graphemes (Rey, Jacobs, Schmidt-Weigand, & Ziegler, 1998), a word's length (number of letters) is measured proportionally to the number of syllables contained in that word so that grapheme clusters are evenly distributed among targeted variables.

As this research investigates lexical Frequency and Familiarity at the whole word level, a word's regularity was also measured at the whole word level. By averaging a word's individual PGC likelihoods, while considering lexical context and employing measures to avoid outlier

PGCs, a whole word's regularity can be reasonably approximated. Hereafter, the regularity of a word's spelling as determined by overall PGC likelihood will be referred to as the variable of Regularity.

Previous spelling research has generally taken two perspectives on analyzing spelling. First, some research categorizes words into two groups, such as high and low Frequency and Regularity (e.g. Jared, 2002) which allows clear assessment of the predictions of spelling models. Traditionally, this simplistic approach to dual-route spelling research (e.g. Brown, 1970) has assumed a decreasing likelihood of correct spellings following the path of a predictable slope between High-Frequency (HF) and Low-Frequency (LF), and between High-Regularity (HR) and Low-Regularity (LR). This categorization approach is also more easily applicable to analyses focused on subjects' individual idiosyncratic vocabulary range. Second, variables such as Frequency and Regularity can be treated as continuous variables giving much greater sensitivity to the analysis and allowing correlations and multiple regression analyses to be conducted (e.g. Rapcsak et al., 2007), albeit at the expense of some clarity. Treating variables as continuous is more applicable to analyses focusing on word specific factors instead of subject specific considerations. Because these different approaches may yield different results, the current study will take both perspectives. However, given that a two-group categorization of variable levels makes a possibly overly simplistic assumption of a consistent slope from high to low, we will add a third medium group to the categories so that we do not rely on the simplistic assumption while still retaining clarity.

Factors Identified in L2 Spelling

Because the dual-route model of spelling originated from research into L1 spellers, most research using the model has focused on this group. While there is some limited and dated dual-

route model research into L2 spellers (e.g. Lester, 1964; Brown, 1970), the applicability of the model to L2 spellers is less clear than it is for L1 spellers. The uncertainty about the applicability of the dual-route model to L2 spellers is largely due to the existence of various factors specific to, and with a possible greater influence on, L2 spellers. L2 spellers are likely to have a more restricted vocabulary of words memorized as lexical chunks, but the main differences are likely to concern PGC since the PGC patterns for English of L2 spellers may be influenced by their L1 and thus differ from those of native English speakers (Sun-Alperin & Wang, 2008). Hereafter, PGC will be used to specify the likely correspondences that are probably shared by native English speakers. L2 spellers may have different phoneme-grapheme correspondences since these are likely to be influenced by the correspondences in the L1. The similarities and differences between the English PGC and L2 sound-to-spelling correspondences will be generally referred to as the variable of L1-L2 Similarity (even though this variable is sometimes measured by considering the differences between a speller's L1 and L2).

Lester's hypothesis. Lester (1964) contended that spellers would not adopt the PGCs of an L2 as readily as with their L1. Second language spellers often have a limited L2 idiolect which they can use to map appropriate PGCs, and they may learn vocabulary as separate items rather than as part of a predictable PGC pattern. Lester concluded that Regularity would be less influential for L2 spellers than for L1 spellers. Therefore, Lester's hypothesis would predict the following Frequency-Regularity hierarchy of spelling for Thai spellers of English: HF-HR, **HF-LR**, **LF-HR**, LF-LR.

The orthographic depth hypothesis. Alphabetic orthographies can be classified according to the consistency (or transparency) of their PGCs, also known as a language's orthographic depth, with transparent orthographies using more predictable PGC patterns and

shallow orthographies using less predictable PGC patterns. The orthographic depth hypothesis (Katz & Frost, 1992) suggests that languages which are more phonologically transparent, e.g. Thai (Callan, Callan, & Masaki, 2005), produce spellers who rely more heavily on sound to spelling processes. On the other hand, languages that are more phonologically opaque, e.g. English (Caravolas et al., 2001), tend to rely on lexical/sublexical chunking. For this reason, Wei (2005) has attributed Thai subjects' confusion with English pronunciation to differences in orthographic depth between the two languages. The orthographic depth hypothesis would predict that Thai spellers, because of their phonologically transparent L1, would be conditioned to rely on PGC. Thus, the correct spellings of English words by Thais would be more dependent on Regularity than for native English spellers. The Frequency-Regularity Hierarchy of Spelling predicted by the orthographic depth hypothesis is HF-HR, **LF-HR**, **HF-LR**, LF-LR.

L1-L2 Similarity. Spellers of L2 words are likely to experience additional issues with PGC since processing phonological differences between the languages may be problematic, particularly for Thai spellers of English (Hamilton & Watson Todd, 2010). The Perceptual Assimilation Model (Best, 1994) predicts that cross-language speech perception will cause perceptual confusion when phonetic mappings between two languages are not equivalent. The fundamental concept of using L1-L2 phonetic mismatches to explain cross-linguistic spelling patterns has been investigated by several researchers (see Figueredo, 2006 for a review). The phonetic differences between the L1 and L2 have been associated with unique spelling patterns, especially when the L1 and L2 are both Roman script languages (e.g. Ingram & Park, 1997; Flege, MacKay, & Meador, 1999). Less common are studies that have examined cross-linguistic spelling in relation to L1-L2 differences of subjects with non-Roman L1 alphabets, (e.g. Ibrahim,

1978 for Arabic; Wang & Geva, 2003 for Cantonese; and Cook, 1997 for Japanese). Most of these studies concluded that L1-L2 differences are problematic for L2 spellers.

A corpus analysis of 125,934 Thai spellings of English (Hamilton & Watson Todd, 2010) was used to determine problematic spelling factors where the Thai L1 sound-to-spelling patterns were not similar to English PGC patterns. The data analysis suggested that L1-L2 differences were more likely to result in spelling mistakes than points of L1-L2 Similarity. Specifically identified were four types of L1-L2 differences that led to errors:

- English phonemes which are not used in Thai, i.e. /z, v, g, ʒ, tʃ, ʃ, θ, dʒ, ð, -f, -s, -tɕʰ, -tɕ/.
- English phonemes which are not always distinguishable in Thai, e.g. /r, l/ and /tʃ, ʃ/.
- Consonant stops which are realized differently in Thai and English, e.g. /-b, -pʰ, -d, -tʰ, -kʰ, -g/.
- All consonant clusters except for those which exist in Thai: /kr, kl, kw, kʰr, kʰl, kʰw, pr, pl, pʰr, pʰl, tr/.

Each of these categories of L1-L2 difference accounted for between 13.91% and 18.69% of the total number of spelling errors identified, a roughly equal proportion. Therefore, it is predicted that each category of L1-L2 phonetic differences will have a similar degree of influence on Thai spellings of English.

Finally, the physical characteristics of a writing system might differ in L1 and L2. Thai, for example, uses different graphemes than English, allows different patterns of letters, and uses very little punctuation or spacing. For instance, English words with double letters representing a single phoneme should be problematic for Thai spellers of English (Hamilton & Watson Todd,

2012) in the same way phonetic mismatches are problematic. For the purposes of this research, writing system differences between L1 and L2 will be considered an extension of L1-L2 phonetic differences. The differences and similarities in L1 and L2 writing systems can be measured in individual L2 words and then compared with correct spellings of those words to determine a relationship.

Overview of the Study

In this study, three main factors influencing spelling are considered. The dual-route model highlights the importance of Familiarity (exposure to a word's spelling pattern) and Regularity (consistency of its spelling pattern), while L2 spelling research suggests L1-L2 Similarity is important. Familiarity is operationalized in two ways: first, Frequency will be used to refer to general frequency of word use as identified in corpus analyses; second, Familiarity will examine Frequency in relation to a specific speller's vocabulary breadth. These four variables will be used to test multiple spelling hypotheses derived from the literature.

As the dual-route model of spelling suggests that a subject could switch between memorized spelling of familiar words and PGC constructions of unfamiliar words, the relationship between Familiarity and spelling output will be examined. Using a vocabulary test based on word frequency to determine each subject's idiolect, this research will investigate spelling performance for words within and beyond the familiar vocabulary range of Thai spellers of English. Additionally, Regularity and L1-L2 Similarity may influence spelling, and their influence may be different within and beyond the speller's vocabulary breadth, a point that we will investigate in this research.

In addition to examining factors possibly influencing individual speller's performance within and beyond their vocabulary breadth, we can also make predictions about the general

likelihood of a word being spelled correctly based on the variables of Frequency, Regularity, and L1-L2 Similarity. We will investigate these general predictions in two ways. First, we will compare two alternative hypotheses regarding route-preference of spellers in dual-route models (the Frequency-Regularity Hierarchy of Spelling): The orthographic depth hypothesis predicts that Thai spellers will rely heavily on the phonological route and thus prioritize Regularity; Lester's hypothesis predicts that the lexical route, and thus Frequency, will be more important. Second, we will examine the impact of the three variables on spelling but looking at the correlations between the variables and the numbers of correct spellings. Given that there may be interactions between the lexical and sublexical spelling routes (e.g. Folk, Rapp, & Goldrick, 2002), in line with previous research into L1 spelling (e.g. Lété, Peereman, & Fayol, 2008), we will conduct a regression analysis to examine how the three variables, and the interactions between them, predict spelling performance.

Specifically, the following research questions are addressed:

1. Are the proportions of correctly spelled words different within and beyond the subjects' vocabulary breadths?
2. Do Regularity and L1-L2 Similarity influence spellings within and beyond the subjects' vocabulary breadths?
3. Which of the two predicted Frequency-Regularity Hierarchies of Spelling best applies to L2 spellers?
4. What is the relationship between a word's level of each variable (Frequency, Regularity, L1-L2 Similarity) and the probability that the word is spelled correctly?
5. Which combination of these variables best predicts the correct spelling variances as calculated by multiple regression models?

Methodology

Research Design

To determine the influences of factors identified in the literature as associated with spelling, a web-based program was developed to determine vocabulary breadth and collect spelling data from Thai spellers of English. The predictions of the dual-route model of spelling were tested for applicability to L2 spellings, and the linguistic factors predicted by the literature to affect L2 spelling were analyzed in relation to the subjects' spelling of those words.

Subjects

Because Thai has many phonetic traits distinct from English and a relatively transparent orthography, data from Thai spellers of English is ideal for observation of the effects of L1-L2 Similarity and orthographic depth. The initial subject sample consisted of 198 Thai university students from a respected Thai university. The subjects were non-English major undergraduate students recruited as intact English writing classes. Common European Framework of Reference for Languages test results indicated that the subjects had A1 to B1 English proficiency levels. Demographic information regarding previous non-Thai language exposure was collected to ensure that subjects' L2s other than English were not skewing the data. Research ethics approval and full consent from the subjects were obtained before conducting the research.

As data was collected from large groups of students, it was expected there would be a portion of subjects whose data would be problematic. Steps were taken to exclude subjects who did not fully participate in the experiment. Data was excluded if the subjects were unable to accurately match Thai audio recordings with the corresponding Thai spellings (38 subjects), or if subjects' vocabulary test data was determined to be unreliable (31 subjects). An additional 32

subjects were excluded due to insufficient data for analysis. No subjects needed to be excluded based on demographic information collected in the survey (i.e. non-Thai language proficiency). After exclusions, there were 97 subjects whose data was used for analysis. Although the number of subjects excluded is high, most exclusions appear to relate to unwillingness to complete tasks seriously, despite agreeing to be research subjects, rather than competence or other issues where such high levels of exclusion might skew the results.

Instruments

Receptive vocabulary test. As the dual-route model of spelling predicts that Familiarity will influence the ability to spell irregular words, it was necessary to score each subject's vocabulary breadth so that subsequent spellings could be categorized as within or beyond the speller's familiar vocabulary range. To determine the subjects' vocabulary breadth, a receptive vocabulary test was administered to the subjects.

Subjects were asked to answer 'Yes' or 'No' as to whether they were familiar with 50 words ranging from frequent to very infrequent as determined by Nation's Range program (Nation & Heatley, 2002) based on BNC word frequencies. The program contains 1,000 word categories, or bands, of decreasing frequency. Words were selected from the Range bands by a random number generator (graphpad.com/quickcalcs/randomN1.cfm). Despite its apparent simplicity, the yes/no vocabulary test has been shown to be a reliable instrument for assessing L2 vocabulary range (Browne & Culligan, 2008; Huibregtse, Admiraal, & Meara, 2002).

Because the subjects were English language learners, a vocabulary test designed for native English speakers would be unlikely to clearly distinguish the subjects' vocabulary breadth. It was assumed that most subjects would not be familiar with many words beyond the most frequent 7,000 words found in the BNC. However, it was also assumed that subjects would

be familiar with 2,000 most frequent words found in the BNC. Therefore, the vocabulary test was weighted accordingly, with a greater word count in frequency bands which were predicted to distinguish subjects' vocabulary breadth, as shown in Table 1.

Table 1

Distribution of vocabulary test words into bands

Range Band (1,000 words)	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
No. of Words	3	5	6	7	10	4	3	3	3	3	2	1	1	1	2
Combined Vocab. Bands	1		2	3	4	5		6		7			8		
No. of Words	8		6	7	10	7		6		6			4		

Following Schmitt and Schmitt (2012) Range bands, which were designed to predict L1 lexical exposure, were merged into combined bands because BNC occurrences used to predict L1 lexical exposure are likely to be less accurate when predicting exposure for L2 subjects. To ensure that Frequency was the primary factor determining a subject's Familiarity range, the MRC Psycholinguistic Database (Coltheart, 1981) was used to verify all selected vocabulary words had high imageability and similar ages of acquisition.

Four pseudowords were also included as reliability indicators as suggested by previous researchers (e.g. Huibregtse et al., 2002). Pseudowords were derived from the Arc Nonword Database (Rastle, Harrington, & Coltheart, 2002) which allows for control of plausibility of English phoneme combinations. Choosing one pseduoword resulted in a reduction in the subject's total real word 'Yes' score. To ensure the vocabulary test indicated reliable vocabulary breadth estimations, data was excluded if a subject claimed to be familiar with more than one of

the four pseudowords as suggested by previous researchers (e.g. Browne & Culligan, 2008; Daller, Milton, & Treffers-Daller, 2007).

A subject's vocabulary range was interpreted as extending into a frequency band if they identified themselves as being familiar with a majority of words in a band. In general, subjects were categorized as being familiar with the least frequent vocabulary band where they had indicated 'Yes' to a majority of the words in that band. Although subject responses generally followed the predicted pattern, it was assumed that individual idiolects would contain some gaps in lexical familiarity instead of strictly following frequencies identified in corpora. Therefore, steps were taken to account for occasional gaps in a subject's vocabulary breadth. 'Yes' responses were categorized into vocabulary familiarity bands as follows:

1. If a subject indicated they were unfamiliar with a majority of words in a band, they were matched with the band immediately below it.
2. If a subject indicated they were unfamiliar with a majority of words in a band, but familiar with more than 40%, then they scored 70% or above on the following band, the previous low scoring band was overlooked.
3. If a subject indicated they were unfamiliar with a majority of words in a band, then they indicated familiarity with a majority of words in the next two bands (or more), the low score was overlooked.
4. If a subject chose one pseudoword, they were penalized 20% off the last band they were matched with.

The vocabulary bands that a subject was determined to be familiar with were used to categorize spelling data as being within or beyond a speller's vocabulary breadth in subsequent analyses considering the variable of Familiarity.

Spelling Production Test. To collect data on subjects' spelling performance, the web-based program also presented target spelling stimuli to the subjects. For each of 54 stimuli words, an image of a concrete noun was displayed for the subjects on each webpage. To control for imageability's potential influence on spelling (Jones, 1985), all target words were highly imageable. An audio recording of the English word corresponding to the image was played for the subject. Subjects could replay the audio file as many times as they wanted (the audio files were generated from a text-to-speech online program, oddcast.com/home/demos/tts/tts_example.php). The subjects were instructed to spell the stimulus word and press ENTER to proceed to a new page with the next stimulus word. To control for any neighborhood effect (Dijkstra, Grainger, & van Heuven, 1999), word sequences were randomly presented to each subject.

As the dual-route model is fundamental to spelling research, word selection was based on having an equal portion of words representing the hypothesized Frequency-Regularity Hierarchy of Spellings (e.g. Lester's Hypothesis: HF-HR, HF-MR, HF-LR, MF-HR, MF-MR, MF-LR, LF-HR, LF-MR, LF-LR with MF being medium Frequency and MR being medium Regularity). A committee of three linguists (British, American, and Thai nationals) selected words with equal distribution of length and L1-L2 Similarity among the combined categories of Frequency and Regularity (see Appendix A). To allow for analogy with the vocabulary test data, the spelling stimuli words were also categorized into frequency bands using similar methods as shown in Table 2.

Table 2

Distribution of Target Spellings into Frequency Bands

	High						Medium						Low	
Range Band	1	2	3	4	5	6	7	8	9	10	11	12	13-14	15+
No. of Words	7	11	8	6	4	6	2	1	1	1	1	1	1	4
Combined Band	1	2	3	4	5	6	7			8				9
No. of Words	7	11	8	6	4	6	4			4				4

Familiarity. The vocabulary test was used to classify target words into two speller-specific categories, within a speller's vocabulary breadth and beyond a speller's vocabulary breadth. Each target word's BNC frequency was compared against each subject's vocabulary band to identify which words should be considered within or beyond a speller's vocabulary breadth. Subjects were considered to be familiar with target words that fell within the speller's vocabulary breadth, and unfamiliar with target words beyond their vocabulary breadth. Familiarity is used to determine whether subjects are performing differently within and beyond their familiar vocabulary breadth.

In order to test the two Frequency-Regularity Hierarchy of Spelling predictions, data needed to be word-specific instead of subject-specific. Words were evenly distributed into HF, MF, and LF categories. The proportion of subjects' correct spellings of HF, MF, and LF were compared to proportions of subjects' correct spellings based on high, medium, and low levels of Regularity and L1-L2 Similarity.

For correlation and multiple regression analyses, the frequency of words as indicated by their combined band was used. The use of frequency bands was determined to be more plausible than the use of raw BNC occurrences to avoid an unwarranted perception of greater variable

sensitivity. Although some researchers have argued that the use of raw BNC occurrences is a better indicator of Familiarity (Crossley, Cobb, & McNamara, 2013), the use of frequency bands is a well-established measure of L2 production (Nation, 2006; Schmitt & Schmitt, 2012) where predicted lexical exposure based on corpus data is likely to be less accurate. As an additional measure to ensure the reliability of Frequency measurements used in this study, the relationship between BNC occurrences and frequency band variables was analyzed and the results indicated that the two methods yielded very similar predictions, $r(54) = .62$, $p < .0001$.

Frequency. A subject's spelling range (combined bands: see Table 2) was interpreted as extending into a frequency band if they could spell the majority of words in that band correctly. Similar to the procedure used for categorizing subjects' vocabulary range, exceptions were made to account for occasional gaps in a speller's ability. The exceptions for the spelling test were slightly different because unlike the vocabulary test where subjects had either mostly 'YES' or mostly 'NO' answers, some subjects spelled exactly half of the words correctly in a frequency band. Moreover, because subjects generally score lower on productive language tasks (like spelling) compared to receptive language tasks (like identifying familiar vocabulary) (Webb, 2008), the criteria needed to overlook a low scoring band were modified to reflect the comparative difficulty of the task:

1. If a minority of words were spelled correctly in a band, but then the subject spelled a majority correctly in two consecutive less frequent bands, the first unsuccessful band was ignored.
2. If a subject spelled exactly half of the words correctly in two or more bands, then the score was set one below the last band where half of the words were spelled correctly.

3. If a minority of words were spelled correctly in a band, but then the subject correctly spelled a majority of words in a less frequent band, the subject was categorized one band higher, regardless of their score in that band.

Regularity. As the dual-route model suggests a strong correlation between correct spellings of unfamiliar words and the regularity of those words, it was necessary to scale target words based on their level of Regularity. For a comparison with Familiarity, words were selected so that HR words contained only letters that were likely outcomes of PGC spellings, while LR words contained multiple letters that should not be the outcome of PGC spellings. For testing the Frequency-Regularity Hierarchy of Spelling, stimuli words were categorized as HR if they had a PGC average greater than 70% as MR with an average PGC of 60% to 69%, and as LR with PGC averages of less than 60%. A word's average PGC likelihood was used for correlation analyses and regression analyses. As it is arguable that successful sublexical route spelling of a word is dependent on the word's most difficult PGC (Peereman & Content, 1999), correlation analyses were run to ensure that a word's average PGC likelihood did not greatly differ from a word's least likely PGC, $r(52) = .72$, $p < .00001$.

L1-L2 Similarity. Each mismatch between L1 and L2 sound-to-spelling patterns (as identified in Hamilton & Watson Todd, 2010) was counted as one L1-L2 difference. The total L1-L2 differences for words ranged from 0 (boy) to 5 (squirrel). For a comparison of variable influence within and beyond a speller's vocabulary range, L1-L2 Similarity was divided into high and low levels. As with other independent variables, L1-L2 Similarity was also categorized into high, medium and low scores for comparison to the Frequency-Regularity Hierarchy of Spelling patterns. The number of L1-L2 differences in a word was used to determine the correlation with the number of times the word was spelled correctly by the subjects.

Scoring of spelling production test. In order to compare results with previous L2 dual-route findings (e.g. Lester, 1964), this research considers Familiarity and Regularity at the lexical level. Therefore, subjects' spellings were also scored as whole words which were either spelled correctly or incorrectly.

Results

Vocabulary Test Results Overview

To determine if a subject's vocabulary breadth predicted their spelling range, it was necessary to categorize each subject into a familiar vocabulary band which was comparable to the target spellings' Frequency bands. A YES/NO vocabulary test was administered to subjects in order to determine their vocabulary breadth. On average, subjects responded 'YES' to indicate familiarity with 44.21% of the vocabulary test words. The distribution of subjects into vocabulary Familiarity bands is shown in Table 3.

Table 3

Distribution of Vocabulary Test Frequency Bands

Range Bands (1,000 words)	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Combined Vocab. Band	1	2	3	4	5	6		7			8				
No. of Subjects	18	28	18	18	15	0		0			0				

Table 3 shows that subjects were categorized into vocabulary bands 1 to 5, suggesting that the most proficient subjects were comfortably familiar with vocabulary up to the most frequent 5,000 words appearing in the BNC.

Spelling Test Results Overview

To determine spelling performance, a spelling test was administered to subjects. On average subjects spelled 43.47% of the stimuli words correctly. The distribution of subjects into highest successful spelling bands is shown in Table 4.

Table 4

Distribution of Highest Successful Spelling Bands

	High		**Medium**				**Low**								
Range Bands (1,000 words)	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15+
Combined Spelling Bands	1	2	3	4	5	6	7			8					9
No. of Subjects	37	24	16	7	3	3	2			0					0

Table 4 shows that spelling band scores ranged from 1 to 7. Three subjects' performance meant that they could not be categorized as their scores were too low, so they are considered to be in spelling band 0. As predicted by Lester (1964), the majority of the English as an L2 subjects were only able to correctly spell higher frequency words. However, while the maximum vocabulary band to which a subject was assigned was 5, some subjects were able to spell words correctly in bands beyond their predicted vocabulary breadth. Thus, some subjects may have been correctly spelling words beyond their familiar vocabulary range using strategies other than the lexical route.

Dual-Route Prediction Results

As the lexical route of spelling should be directly related to a subject's vocabulary breadth, subjects' highest vocabulary bands and spelling bands were compared to determine if vocabulary bands accurately predicted spelling bands. The average spelling band score was 0.61 bands lower than the average vocabulary band score. The vocabulary band score matched the subject's highest successful spelling band 25.77% of the time.

To determine if subjects' vocabulary test scores predicted the variance in spelling test scores, the results of the vocabulary test and the spelling test were compared using product moment correlation analysis. The correlation between the two scores was significant if not impressive ($r(95) = .247$, $p < .01$). Fifty-one subjects scored lower on the spelling test than the vocabulary test (six of these subjects were four bands or more below the vocabulary test score). This was not unexpected as receptive scores tend to be higher than productive scores in language tests (Webb, 2008).

However, twenty-one students scored better on the spelling test than the vocabulary test (two of these subjects scored four bands or more higher than the vocabulary test). This suggests that there may be individual spelling strategies which allow some subjects to spell beyond their predicted vocabulary range. If the subjects scored lower on the spelling test than the vocabulary test, this suggests that they are relying on lexical spelling with a productive vocabulary size smaller than their receptive vocabulary size. If they scored higher on the spelling test, the traditional dual-route model suggests that they are using PGC regularity to spell unknown words.

As the dual-route model predicts that subjects will spell familiar words correctly more often than unfamiliar words, the percentage of words spelled correctly within subjects' familiar vocabulary band range was compared to the percentage of words spelled correctly beyond

subjects' familiar vocabulary band range. A paired sample t-test shows a significant difference between the proportions of correct spellings within (68.43% correct) and beyond (28.20% correct) spellers' vocabulary range (t = -34.95, p< .0001). A table of proportions of correct spellings from the perspective of high and low levels of Familiarity and Regularity is presented in Table 5.

Table 5

Proportions of correct spellings within and beyond spellers' vocabulary breadth

	Within			**Beyond**		
	No. of attempts	**No. correct**	**% correct**	**No. of attempts**	**No. correct**	**% correct**
Regular	1088	692	61.23%	1531	471	29.32%
Irregular	1126	823	75.45%	1493	291	18.40%
Similar L1-L2	1009	811	83.17%	1416	346	24.07%
Different L1-L2	1205	704	56.54%	1608	416	25.49%

The dual-route model predicts that Regularity would be the key variable allowing spellers to use PGC constructions to spell Regular words beyond their familiar vocabulary range. However, in L2 writing other variables, especially L1-L2 Similarity, may also influence correct spellings within and beyond spellers' vocabulary breadth.

Predictions Based on Regularity

To determine the influence of Regularity within and beyond each speller's predicted vocabulary range, it was first necessary to categorize each target word as within or beyond the speller's vocabulary breadth. The target words were then divided into two categories of variable influence (High Regularity and Low Regularity) in order to test whether Regularity influences correct spellings differently within and beyond a speller's vocabulary breadth. Treating the number of correct and incorrect responses for each quadrant as independent binomial samples, we can calculate the difference of proportions, standard error and odds ratio between each pair of formats to see if the differences in proportions of responses are real. Where the figures for the 99% confidence interval are either both positive, or both negative, we can conclude that there is a significant difference in the proportions (see Agresti, 2007). These comparisons are shown in Table 6 for Regularity.

Table 6

Differences in proportion of correct spellings of Regular and Irregular words

	Within Regular v. Beyond Regular	**Within Irregular v. Beyond Irregular**	**Within Regular v. Within Irregular**	**Beyond Regular v. Beyond Irregular**
Difference of proportions	0.3284	0.5360	-0.0949	0.1127
Standard error	0.0188	0.0167	0.0197	0.0156
Odds ratio	0.0188	0.1636	1.5543	0.5448
99% confidence interval (max.)	0.3767	0.5791	-0.0442	0.1530
99% confidence interval (min.)	0.2801	0.4929	-0.1456	0.0725
Interpretation	Regular words spelled correctly more often within vocabulary breadth	Irregular words spelled correctly more often within vocabulary breadth	Words within voabulary breadth spelled correctly more often if irregular	Words beyond vocabulary breadth spelled correctly more often if regular

Table 6 shows that although both regular and irregular words are more often spelled correctly within subjects' vocabulary range, there is a significant increase in the proportion of correct spellings of regular words beyond the spellers' vocabulary breadth when compared to irregular words (since the difference of proportions falls within the 99% interval, it is significant at $p < .01$). We can conclude that the tendency for regular unfamiliar words to be spelled correctly more often than irregular unfamiliar words suggests that spellers may use PGC-based spellings of words beyond their vocabulary breadth.

Interestingly, within spellers' vocabulary breadth, irregular words were spelled correctly more often than regular words (since the difference of proportions falls within the 99% interval, it is significant at $p < .01$). Thus, the role of Regularity for words familiar to spellers does not follow the predictions of previous spelling models.

Predictions Based on L1-L2 Similarity

Following the analysis procedures performed for Regularity, correct spellings of words categorized as High L1-L2 Similarity or Low L1-L2 Similarity were compared to determine if there are significant differences within and beyond a spellers' vocabulary breadth. These comparisons are shown in Table 7.

Table 7

Differences in proportion of correct spellings of similar L1-L2 and different L1-L2

	Within Similar v. Beyond Similar L1-L2	**Within Different v. Beyond Different L1-L2**	**Within Similar v. Within Different L1-L2**	**Beyond Similar v. Beyond Different L1-L2**
Difference of proportions	0.5594	0.3255	0.2195	-0.0066
Standard error	0.0169	0.0179	0.0189	0.0164
Odds ratio	0.0789	0.24836	0.3431	1.0350
99% confidence interval (max.)	0.6030	0.3717	0.2682	0.0355
99% confidence interval (min.)	0.5158	0.2794	0.1708	-0.0488
Interpretation	Similar L1-L2 words spelled correctly more often within vocabulary breadth	Different L1-L2 words spelled correctly more often within vocabulary breadth	Words within vocabulary breadth spelled correctly more often if similar L1-L2	No difference

Table 7 shows that generally, within spellers' familiar vocabulary breadth, words with high L1-L2 Similarity are spelled correctly more often than words with low L1-L2 Similarity (since the difference of proportions falls within the 99% interval, it is significant at $p < .01$). This suggests that L1-L2 differences may make a word's spelling more difficult to remember even if it is in a speller's vocabulary breadth. However, beyond a spellers' vocabulary breadth, L1-L2 Similarity is no longer a significant factor. This suggests that L1-L2 Similarity is not significantly hindering PGC constructed spellings. However, L1-L2 Similarity's high impact within a speller's vocabulary range suggests that these differences between L1 and L2 possibly have more influence than Regularity for familiar words.

The results thus far suggest that Familiarity is the major factor influencing spelling. Regularity has an unexpected role within spellers' vocabulary breadth, but its influence on spellings of unfamiliar words seems to match the predictions of the dual-route model. Because subject performance was better with similar L1-L2 words while spelling within their familiar vocabulary breadth, L1-L2 Similarity may play a role in the memorability of familiar words.

Frequency-Regularity Hierarchy Predictions

In order to compare Frequency's ability to predict spelling performance to other variables (Regularity and L1-L2 Similarity), words were no longer analyzed based on subject-specific data (within or beyond a subject's vocabulary breadth). Instead, BNC frequencies were used to categorize target spellings as HF and LF.

Lester's hypothesis, prioritizing Frequency, and the orthographic depth hypothesis, prioritizing Regularity, predict different Frequency-Regularity hierarchies of spelling with the former predicting a sequence of HF-LR then LF-HR and the latter predicting a sequence of LF-

HR then HF-LR. The proportions of correctly spelled words in the four stages of these predicted hierarchies are:

HF-HR: 65.00%

HF-LR: 64.67%

LF-HR: 17.00%

LF-LR: 5.16%

These proportions clearly support Lester's hypothesis and contradict the orthographic depth hypothesis.

To determine whether there is a consistent slope within the extremes of high and low variable levels, the proportions of correct spellings for medium levels of Regularity and Frequency were measured along with high and low levels. The addition of a medium Regularity category created unexpected results. The average correct spellings for all words in high, medium, and low levels of Frequency and Regularity are presented in Table 8.

Table 8

Proportions of correct spellings for categories in the Frequency-Regularity Hierarchy of Spelling

Category	No. of Words	No. of Attempts	Proportion Correct
HF HR	6	582	65.00%
HF MR	6	582	81.17%
HF LR	6	582	64.67%
MF HR	6	582	38.50%
MF MR	6	582	59.17%
MF LR	6	582	29.17%
LF HR	6	582	17.00%
LF MR	6	582	17.33%
LF LR	6	582	5.16%

The noticeable difference between the predicted pattern and the actual pattern is in the medium Regularity (mean = 52.56%) where corrects spellings were more common than high Regularity (mean = 40.17%) regardless of Frequency (a comparison of subjects' performance for HR and MR words indicates that this difference is significant, $t(96) = -9.32$, $p < .0001$):

General Pattern: HF> MF> LF

Within the Pattern: MR> HR> LR

When analyzing all three independent variables (including L1-L2 Similarity) in high, medium, and low categories, Regularity appears to follow the least predictable slope as shown in Figure 1.

Figure 1 shows that at high levels of all variables, subjects spelled more words correctly in comparison to low levels of those variables. However, only Frequency and L1-L2 Similarity follow the predicted slope for medium levels of these variables. Regularity did not match the predicted slope for medium levels of the variables which suggests that Regularity did not have a predictable influence on the target spellings.

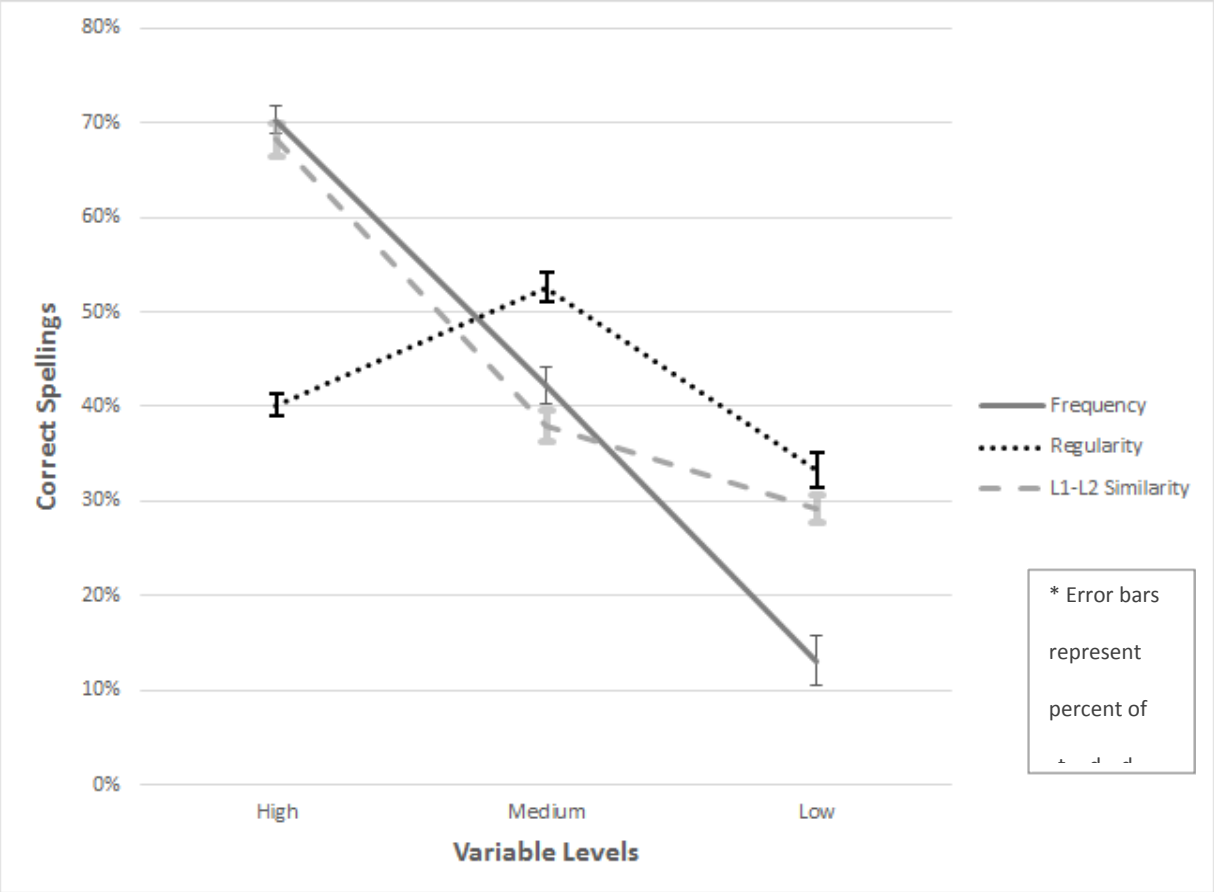

Figure 1. Slope pattern for high, medium, and low levels of variables. The percentages indicate the spelling performance for words associated with high, medium, and low levels of the variables Frequency, Regularity, and L1-L2 Similarity.

Variable Relationship with Correct Spellings

In order to determine if there are relationships between the variables, (Frequency as measured by band; Regularity as measured by percentage of overall PGC likelihood; L1-L2 Similarity as measured by number of differences) and spelling performance, the levels of the

variables in each of the 54 words were compared to the number of correct spellings for each word using product-moment correlation. The results are shown in Table 9.

Table 9

Summary of Intercorrelations, Means, and Standard Deviations

Variable	M	SD	1	2	3	4
1. Frequency	4.24	2.53	—			
2. Regularity	.655	.164	-.704	—		
3. L1-L2 Similarity	2.59	1.28	.067	-.030	—	
4. Correct Spellings	25.52	27.36	-.684**	.072	-.285*	—

Note. * $p < .05$. ** $p < .001$

Table 9 shows that Frequency had the strongest relationship to correct spellings, suggesting that less frequent words were more likely to be spelled incorrectly (the correlation is negative because frequency bands were used for this analysis, with lower-numbered bands representing more frequent words). Nearly half of the variance in correct spellings could be explained by Frequency. There was no relationship between Regularity and correct spellings, and a weak relationship with low L1-L2 Similarity words more likely to be spelled incorrectly.

Multiple Linear Regression Models

Regression analyses were run using SPSSv20 to determine which combination of variables best predicted spelling-accuracy scores. Hierarchical regression analyses indicated that the traditional dual-route prediction model (Frequency and Regularity) accounted for 46% of the variance in actual performance, $R^2 = .457$, $F(2, 51) = 22.37$; $p < .001$. All of the model's ability

to predict correct spellings came from the Frequency variable. The inclusion of Regularity actually decreased the model's effectiveness by 1%.

To determine if interaction between Frequency and Regularity improved the model's ability to predict correct spelling scores, the product of Frequency and Regularity z-scores was entered into the regression model. Based on the results of previous analyses, the L1-L2 Similarity variable was added into the regression model. To test for interaction between Frequency and L1-L2 Similarities, the product of Frequency and L1-L2 Similarity z scores was also included into the model.

In line with Lété et al., (2008), stepwise inclusion of all terms was performed to optimize the model's prediction ability. As Frequency has been described as the most influential variable for both L1 and L2 spelling (Brown, 1970), this variable was entered into the model first. The other variables were then entered stepwise to observe the subsequent impact on the model's effectiveness. The results are given in Table 10.

(Table 10 around here)

The results suggested that Regularity and interaction between Frequency and Regularity did not improve the model's ability to predict correct spelling. Therefore, these variables were removed from the optimized model. However, the inclusion of the L1-L2 Similarity variable improved the model's ability to predict correct spellings by approximately 6%. The interaction between Frequency and L1-L2 Similarity improved the model's ability to predict correct spelling scores by an additional 5%. For the optimized model (Considering the variables of Frequency, L1-L2 Similarity, and the interaction between Frequency and L1-L2 Similarity), ANOVA F statistics indicate the model is significantly better than guessing the mean ($F(1, 53) = 27.92$, $p < .0001$).

Although standardized coefficients indicated that Frequency ($r(52) = .667$, $p < .0001$) was contributing more than L1-L2 Similarity ($r(52) = .237$, $p < .05$), more than 57% of the variance in correct spelling is explained by the optimized model, which is an 11% improvement over the traditional dual-route model's prediction ability. This suggests that Regularity is not a good fit for L2 spelling models which attempt to relate the variable with overall spelling performance. In L2 models where a speller's vocabulary breadth is not considered, L1-L2 Similarity may predict overall L2 spelling performance better than Regularity.

Discussion

In this study, predictions of L1 and L2 spelling theories were tested to determine their applicability to Thai spellers of English. The results suggest that Familiarity is the dominant variable influencing subjects' spelling performance. Subjects were significantly more likely to spell words correctly within their predicted vocabulary range compared to words beyond their vocabulary range. Moreover, Frequency predicted most of the variance in subjects' spelling performance. Although Regularity and L1-L2 Similarity may have some influence on L2 spelling, Thai spellers of English appear to be highly reliant on the lexical route. This seems to be in line with Lester's (1964) prediction that the spelling of L2 words' would likely involve the memorization of each word individually because of the learners' tendency to acquire L2 vocabulary at a staggered pace and in limited amounts when compared to L1 acquisition.

Regularity appears to have some influence on spelling performance when the subjects spelled words beyond their familiar vocabulary range. This is in line with Seidenberg's (1985) prediction that the lexical route would retrieve higher frequency words' spellings before the sublexical route could construct spellings. Accordingly, sublexical constructions which rely on

Regularity should become a more competitive alternative to the lexical route when spelling lower frequency words, as observed in previous L1 spelling research (e.g. Lété et al., 2008). However, Regularity did not have an overall correlation with correct spellings which previous L1 research has identified (e.g. Kreiner & Gough, 1990). Although the results support the simple dual-route prediction (e.g. Brown, 1970) that HR words should be spelled correctly more often than LR words, irregular words were spelled correctly more often than regular words within a speller's vocabulary breadth. Moreover, MR words were spelled correctly more often than HR words.

Because MR words appear to be more memorable within a speller's vocabulary breadth than HR words, subjects may have been relying heavily on memorized spellings of L2 vocabulary instead using the sublexical route. This could indicate a general tendency in L2 spelling to rely on the lexical route. However, another explanation lies in the Thai education system. Wei (2005) reported that Thai students often utilize rote memorization techniques for the spelling of English words. Often, Thai teachers and curricula focus on irregular English spellings. Therefore, it is possible that students are able to easily recall frequent irregular words as memorized lexical items because they have been the focus of instruction.

L1 spelling researchers have also noted a trend for the spelling of HF-MR words to be recognized more quickly than HF-HR words (see discussion in Waters & Seidenberg, 1985, p. 568). The authors suggested that the unique nature of semi-irregularly spelled familiar words may allow subjects to easily distinguish these words. Therefore, regardless of whether spelling in L1 or L2, the lexical route may favor familiar words with unique characteristics that allow them to become more memorable to the speller.

Regularity's inconsistent relationship to correct spellings made it a poor predictor of the subjects' overall spelling performance. Particularly within subjects' vocabulary breadth, the results support Lester's (1964) hypothesis that L2 spellers would be unlikely to adopt the PGC of the L2 because of the learner's limited L2 vocabulary size and the potential difficulties of developing a second set of sound-to-spelling rules that may be in conflict with their previously established L1 sound-to-spelling patterns. Despite Thai's relative orthographic transparency, the subjects did not appear to be favoring the sublexical route as predicted by the orthographic depth hypothesis (Katz & Frost, 1992). Therefore, the data suggests that a speller's L1 orthographic depth is not the only factor influencing route preference.

L1-L2 Similarity had some correlation with correct spellings. The addition of L1-L2 Similarity improved the regression model's ability to predict spelling performance variance. However, it appears that L1-L2 Similarity only impacts spelling performance within a subject's vocabulary range, not beyond. This suggests that L1-L2 Similarity may play some part in a speller's ability to remember a word's spelling. L2 Spelling research (e.g. Figueredo, 2006) has noted that L1-L2 differences may create challenges for L2 spellers. This could be explained by difficulty processing unfamiliar aspects of an L2, similar to those described by the Perceptual Assimilation Model (Best, 1994) for L2 auditory processing. The unfamiliarity of PGCs unique to an L2 may make memory associations with similar lexical items less likely, and so make lexical or sublexical recall more difficult for the speller. Further investigation into the relationship between L1-L2 Similarity and lexical recall mechanisms is required.

The dual-route model of spelling considers two factors (Familiarity and Regularity) that influence spelling. However, the results suggest two different processes for words within and beyond a speller's vocabulary breadth. Within, Familiarity is very important. L1-L2 Similarity

also appears to contribute to spelling performance within a speller's vocabulary breadth, while Regularity influenced spelling performance inversely to dual-route predictions. Beyond their vocabulary breadth, subjects are more likely to utilize the sublexical route, allowing words with greater Regularity to be spelled correctly more often than irregular words. Overall, Regularity appeared to have no relationship in correlation and regression analyses because Regularity had opposite influences on spelling performance within and beyond spellers' vocabulary range.

The findings of this research can be interpreted as the dual-route model switching from Familiarity (lexical route) to Regularity (sublexical route) as words become unfamiliar. However, the impact of L1-L2 and the unexpected results for Regularity within a speller's vocabulary breadth suggest that a more complex version of the traditional dual-route model may be applicable to L2 spelling patterns.

References

Agresti, A. (2007). *An Introduction to Categorical data analysis* (2nd ed). New York, NY: Wiley.

Best, C.T. (1994) The emergence of native-language phonological influences in infants: A perceptual assimilation model. In J.C. Goodman and H.C. Nusbaum (Eds.), *The development of speech perception: The transition from speech sounds to spoken words* (pp. 167-224). Cambridge, MA: MIT Press.

Brown, G. D., & Ellis, N. C. (1994). *Handbook of spelling: Theory, process, and intervention*. Chichester, England: John Wiley & Sons Ltd.

Brown, H. D. (1970). Categories of spelling difficulty in speakers of English as a first and second language. *Journal of Verbal Learning and Verbal Behavior*, 9(2), 232-236. doi:10.1016/s0022-5371(70)80056-1

Browne, C., & Culligan, B. (2008). Combining technology and IRT testing to build student knowledge of high frequency vocabulary. *The JALT CALL Journal*, 4(2), 3-16.

Bullinaria, J. A. (1994). Connectionist modelling of spelling. In Proceedings of the sixteenth annual conference of the cognitive science society. (pp.78-83). Hillsdale, NJ: Erlbaum.

Callan, A. M., Callan, D. E., & Masaki, S. (2005). When meaningless symbols become letters: Neural activity change in learning new phonograms. *Neuroimage*, *28*(3), 553-562. doi:10.1016/j.neuroimage.2005.06.031

Caravolas, M., Hulme, C., & Snowling, M. J. (2001). The foundations of spelling ability: Evidence from a 3-year longitudinal study. *Journal of Memory and Language*, 45(4), 751-774. doi:10.1006/jmla.2000.2785

Carney, E. (1994). *A survey of English spelling*. London: Routledge.

Coltheart, M. (1981). The MRC psycholinguistic database. *The Quarterly Journal of Experimental Psychology*, 33(4), 497-505. doi:10.1080/14640748108400805

Cook, V. J. (1997). L2 users and English spelling. *Journal of Multilingual and Multicultural Development*, 18(6), 474-488. doi:10.1080/01434639708666335

Crossley, S. A., Cobb, T., & McNamara, D. S. (2013). Comparing count-based and band-based indices of word frequency: Implications for active vocabulary research and pedagogical applications. *System*, *41*(4), 965-981. doi:10.1016/j.system.2013.08.002

Cummings, D. W. (1988). *American English spelling: An informal description*. Baltimore: JHU Press.

Daller, H., Milton, J., & Treffers-Daller, J. (Eds.). (2007). *Modelling and assessing vocabulary knowledge* (p. 150). Cambridge: Cambridge University Press.

Dijkstra, T., Grainger, J., & van Heuven, W. J. (1999). Recognition of cognates and interlingual homographs: The neglected role of phonology. *Journal of Memory and Language*, 41(4), 496-518. doi:10.1006/jmla.1999.2654

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(8), 873-905. doi:10.1007/s11145-006-9014-1

Flege, J. E., MacKay, I. R. A., & Meador, D. (1999). Native Italian speakers' perception and production of English vowels. *The Journal of the Acoustical Society of America*, 106(5), 2973-2987. doi:10.1121/1.428116

Folk, J. R., Rapp, B., & Goldrick, M. (2002). The interaction of lexical and sublexical information in spelling: What's the point? *Cognitive Neuropsychology*, 19(7), 653-671. doi:10.1080/02643290244000184

Hamilton, T. G., Watson Todd, R., & Facundes, N. (2010). Unique categories of errors in Thai spellings of English. In V. Sagaravasi (Ed.), *NIDA 2nd Annual conference on Language and Communication Proceedings* (pp. 97-105). Bangkok, Thailand: National Institute of Development Administration.

Hamilton, T. G., Watson Todd, R., & Facundes, N. (2012). Reassessing traditional spelling theories from a second language perspective. In P. Subphadoongchone (Ed.), *The 32nd Thailand TESOL International Conference Proceedings 2012* (pp. 77-94). Bangkok, Thailand: Thailand TESOL.

Hanna, P. R., Hanna, J. S., Hodges, R. E., & Rudorf, E. H. (1966). Phoneme-grapheme correspondences as cues to spelling improvement. Washington, DC: U.S. Government Printing Office.

Houghton, G., & Zorzi, M. (2003). Normal and impaired spelling in a connectionist dual-route architecture. *Cognitive Neuropsychology*, 20(2), 115-162. doi:10.1080/02643290242000871

Huibregtse, I., Admiraal, W., & Meara, P. (2002). Scores on a yes-no vocabulary test: Correction for guessing and response style. *Language Testing*, 19(3), 227-245. doi:10.1191/0265532202lt229oa

Ibrahim, M. H. (1978). Patterns in spelling errors. *ELT Journal*, 32(3), 207-212. doi:10.1093/elt/xxxii.3.207

Ingram, J. C., & Park, S. G. (1997). Cross-language vowel perception and production by Japanese and Korean learners of English. *Journal of phonetics*, *25*(3), 343-370. doi:10.1006/jpho.1997.0048

Jared, D. (2002). Spelling-sound consistency and regularity effects in word naming. *Journal of Memory and Language*, 46(4), 723-750. doi:10.1006/jmla.2001.2827

Jones, G. V. (1985). Deep dyslexia, imageability, and ease of predication. *Brain and Language*, 24(1), 1-19. doi:10.1016/0093-934x(85)90094-x

Katz, L. & Frost, R. (1992). The reading process is different for different orthographies: The orthographic depth hypothesis. In Frost, R. & Katz, L., (Eds.). *Orthography, Phonology, Morphology, and Meaning* (pp. 67-84). Amsterdam: Elsevier North Holland Press.

Kessler, B., & Treiman, R. (1997). Syllable structure and the distribution of phonemes in English syllables. *Journal of Memory and language*, *37*(3), 295-311. doi:10.1006/jmla.1997.2522

Kessler, B., & Treiman, R. (2001). Relationships between sounds and letters in English monosyllables. *Journal of Memory and Language*, 44(4), 592-617. doi:10.1006/jmla.2000.2745

Kessler, B., & Treiman, R. (2003). Is English spelling chaotic? Misconceptions concerning its irregularity. *Reading Psychology*, 24, 267–289. doi:10.1080/02702710390227228

Kreiner, D. S., & Gough, P. B. (1990). Two ideas about spelling: Rules and word-specific memory. *Journal of Memory and Language*, 29(1), 103-118. doi:10.1016/0749-596x(90)90012-o

Lester, M. (1964). Graphemic-phonemic correspondences as the basis for teaching spelling. *Elementary English*, 41(7), 748-752. Retrieved from http://www.jstor.org/stable/41385716

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(4), 952-977. doi:10.1016/j.jml.2008.01.001

Nation, P., (2006). How large a vocabulary is needed for reading and listening? *Canadian Modern Language Review*, 63(1), 59-82. doi:10.3138/cmlr.63.1.59

Nation, I.S.P., & Heatley, A. (2002). Range: A program for the analysis of vocabulary in texts [software]. Downloadable from http://www.vuw.ac.nz/lals/staff/paul-nation/nation.aspx

Olson, A., & Caramazza, A. (1994). Representation and connectionist models: The NETspell experience. In G.A. Brown & N.C. Ellis (Eds.), *The handbook of normal and disturbed spelling development: Theory, process and intervention* (pp. 337-363). New York: John Wiley & Sons.

Peereman, R., & Content, A. (1999). LEXOP: A lexical database providing orthography-phonology statistics for French monosyllabic words. *Behavior Research Methods, Instruments, & Computers*, *31*(2), 376-379. doi:10.3758/bf03207735

Plaut, D. C., McClelland, J. L., Seidenberg, M. S., & Patterson, K. (1996). Understanding normal and impaired word reading: computational principles in quasi-regular domains. *Psychological Review*, 103(1), 56. doi:10.1037/0033-295x.103.1.56

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(11), 2519-2524. doi:10.1016/j.neuropsychologia.2007.03.019

Rastle, K., Harrington, J., & Coltheart, M. (2002). 358, 534 nonwords: The ARC nonword database. *The Quarterly Journal of Experimental Psychology*: Section A, 55(4), 1339-1362. doi:10.1080/02724980244000099

Rey, A., Jacobs, A. M., Schmidt-Weigand, F., & Ziegler, J. C. (1998). A phoneme effect in visual word recognition. *Cognition*, *68*(3), B71-B80. doi:10.1016/s0010-0277(98)00051-1

Schmitt, N., & Schmitt, D. (2014). A reassessment of frequency and vocabulary size in L2 vocabulary teaching. *Language Teaching*, *47*(04), 484-503. doi:10.1017/S0261444812000018

Seidenberg, M. S. (1985). The time course of information activation and utilization in visual word recognition. In D. Besner, T. G. Waller, & E. M. MacKinnon (Eds.), *Reading research: Advances in theory and practice* (Vol. 5, pp. 199–252). San Diego, CA: Academic Press.

Spencer, K. (2007). Predicting children's word-spelling difficulty for common English words from measures of orthographic transparency, phonemic and graphemic length and word frequency. *British Journal of Psychology*, *98*(2), 305-338. doi:10.1348/000712606X123002

Sun-Alperin, M. K., & Wang, M. (2008). Spanish-speaking children's spelling errors with English vowel sounds that are represented by different graphemes in English and Spanish words. *Contemporary Educational Psychology*, 33(4), 932-948. doi:10.1016/j.cedpsych.2007.12.005

Tainturier, M. J., Bosse, M. L., Roberts, D. J., Valdois, S., & Rapp, B. (2013). Lexical neighborhood effects in pseudoword spelling. *Frontiers in Psychology*, *4*, 862. http://doi.org/10.3389/fpsyg.2013.00862

Venezky, R. L. (1970). *The structure of English orthography*. The Hague: Mouton.

Wang, M., & Geva, E. (2003). Spelling performance of Chinese children using English as a second language: Lexical and visual-orthographic processes. *Applied Psycholinguistics*, 24(1), 1-26. doi:10.1017/s0142716403000018

Waters, G. S., & Seidenberg, M. S. (1985). Spelling-sound effects in reading: Time-course and decision criteria. *Memory & Cognition*, *13*(6), 557-572. doi:10.3758/bf03198326

Wei, Y. (2005). The relationship between phonological awareness and reading ability of Thai students in English and Thai in primary schools of Thailand. (unpublished doctoral dissertation). University of Maryland, College Park, MD.

Webb, S. (2008). Receptive and productive vocabulary sizes of L2 learners. *Studies in Second Language Acquisition*, *30*(01), 79-95. doi:10.1017/s0272263108080042

Appendix A

Target Spelling Words and Variable Levels

In order to observe the relationship between correct spellings of L2 words and several spelling factors identified in the literature, 54 target words were selected by a committee. The target spelling words selected and the variable levels considered for this research are shown in Table A1.

Table A1

Target Spelling Words and Variable Levels

		LETTERS	SYLLABLES	L1-L2 SIMILARITIES	FREQUENCY	REGULARITY
HF HR	VAN	3	1	1	2	99.01%
	BOX	3	1	2	1	92.84%
	BRUSH	5	1	3	2	79.60%
	PIANO	5	2	0	2	81.38%
	LIBRARY	7	3	3	2	70.14%
	SECRETARY	9	4	4	1	70.22%
	$\bar{x}$	**5.33**	**2.00**	**2.17**	**1.67**	**82.20%**
HF MR	BOY	3	1	0	1	63.71%
	HILL	4	1	2	2	63.32%
	LETTER	6	2	3	1	62.81%
	CLOTHES	7	1	3	1	65.86%
	MAGAZINE	8	3	3	2	62.56%
	TELEPHONE	9	3	1	2	60.33%
	$\bar{x}$	**6.17**	**1.83**	**2.00**	**1.50**	**63.10%**
HF LR	ICE	3	1	1	1	21.53%
	SCHOOL	6	1	2	1	45.89%
	MIRROR	6	2	3	2	39.96%
	SOLDIER	7	2	3	2	58.62%
	CIGARETTE	9	3	4	2	51.63%
	FURNITURE	10	3	3	2	47.83%
	$\bar{x}$	**6.83**	**2.00**	**2.67**	**1.67**	**44.24%**
MF HR	PET	3	1	1	3	94.44%
	BOLT	4	1	3	3	86.56%
	CRADLE	6	2	4	4	82.65%
	TRUMPET	7	2	3	4	84.30%
	INSTRUCTOR	10	3	5	3	77.02%
	THERMOMETER	11	4	3	5	78.82%
	$\bar{x}$	**6.83**	**2.17**	**3.17**	**3.67**	**83.96%**

MF MR	GOAT	4	1	2	4	63.42%
	WOLF	4	1	2	3	68.16%
	GARBAGE	7	2	3	4	65.10%
	DIAMOND	7	2	2	3	67.01%
	VEGETABLE	9	3	5	3	66.91%
	MICROSCOPE	10	3	3	4	65.43%
	$\bar{x}$	**6.83**	**2.00**	**2.83**	**3.50**	**66.00%**
MF LR	COMB	4	1	0	3	58.46%
	SWORD	5	1	4	3	59.94%
	CHOIR	5	2	2	5	6.61%
	SHIELD	6	2	4	5	54.27%
	HURRICANE	9	3	2	4	53.14%
	HANDKERCHIEF	12	3	3	5	52.28%
	$\bar{x}$	**6.83**	**2.00**	**2.50**	**4.17**	**47.45%**
LF HR	MUG	3	1	2	6	90.59%
	WIG	3	1	2	6	86.68%
	MOTH	4	1	1	7	74.88%
	SHRUB	5	1	4	7	79.60%
	APRICOT	7	3	2	9	74.57%
	FLAMINGO	8	3	3	8	70.31%
	$\bar{x}$	**5.00**	**1.67**	**2.33**	**7.17**	**79.44%**
LF MR	FAWN	4	1	0	8	64.11%
	FLUTE	5	1	3	7	66.59%
	BEEHIVE	7	2	2	8	67.44%
	SQUIRREL	8	2	5	6	68.48%
	BRACELET	8	2	5	6	62.59%
	GARDENIA	8	3	2	9	66.50%
	$\bar{x}$	**6.67**	**1.83**	**2.83**	**7.33**	**65.95%**
LF LR	CLAW	4	1	2	6	57.75%
	LYNX	4	1	3	9	61.40%
	GLACIER	7	2	4	7	59.63%
	MOCCASIN	8	3	2	9	58.35%
	ACCORDION	9	4	2	6	51.96%
	ASPARAGUS	9	4	4	8	53.96%
	$\bar{x}$	**6.83**	**2.50**	**2.83**	**7.50**	**57.17%**