1-1-2011

Speling "Successful" Sucesfuly: Statistical Learning in Spelling

Siti Syuhada Binte Faizal

WASHINGTON UNIVERSITY

Department of Psychology

Speling "Successful" Sucesfuly: Statistical Learning in Spelling

by

Siti Syuhada Binte Faizal

A thesis presented to the
Graduate School of Arts and Sciences
of Washington University in
partial fulfillment of the
requirements for the
degree of Master of Arts

December 2011

Saint Louis, Missouri

Abstract

Many spelling errors in English are doubling errors, as when people are stumped by the double ‹l› in ‹trellis›. In Study 1, we tabulated statistical patterns with regards to doubling in English. In Study 2, we collected behavioral data to see if people were sensitive to these statistical patterns in doubling and to explore other factors that might influence doubling such as context, individual differences (language background and spelling ability), and task. We gave two nonword spelling tasks to US college students (N=68) and bilingual Singaporean college students from an English-based education system but with diverse language backgrounds: Mandarin (N=54), Malay (N=44), or Tamil (N=42). In the choice task, participants heard a nonword and chose between two spelling options, e.g. dremmib/dremib. In the free task, they wrote down its best spelling. We found a vowel length effect (more doubling after short vowels than long vowels) that was moderated by spelling ability (better spellers were more influenced by vowel length) and language background. Americans had the largest vowel length effect and Tamil Singaporeans had none, as they possibly associated consonant doubling with the lengthening of doubled consonants in Tamil instead of the preceding vowel. The Mandarin group spelled nonwords least accurately, and greater knowledge of pinyin, a phoneme-based writing system, was associated with higher nonword spelling accuracy. These and other findings reflect how linguistic factors and language background moderate the role of statistical learning and context in spelling.

Keywords: spelling, context, statistical learning, language background, doubling

Acknowledgements

I would like to thank the following for all of their help and advice throughout the entire process:

My advisor, Rebecca Treiman;

Members of my Master's committee, Brett Kessler and Dave Balota;

My programming and Python consultant, Mohammed Ansari;

Nicole Rosales, for assisting in the data collection in the US;

Members of the Reading and Language lab;

My mum, for being the best mother in the world and my pillar of strength, for accompanying me during the experimental sessions, and for constantly including me in her love and prayers;

and lastly, God, for all these wonderful things could only have happened with His grace and blessings.

Table of Contents

Abstract ii

Acknowledgements iii

Table of Contents iv

List of Tables vi

List of Figures vii

Speling “Successful” Sucesfuly: Statistical Learning in Spelling 1

Doubling of Medial Consonants in English 5

Patterns in the Doubling of Medial Consonants 7

Influence of Language Background on Spelling 11

Overview of Studies 1 and 2 18

Study 1 20

Method 20

Results 21

Discussion 23

Study 2 24

Method 25

Results 30

Discussion 37

General Discussion 44

References 48

Appendix A 79

Appendix B 80

Appendix C 82

Appendix D 84

Appendix E 87

List of Tables

Table 1. *Statistical Patterns Pertaining to Proportion of Doubling of Medial Consonants in Disyllables (N = 8790)* 61

Table 2. *Statistical Patterns Pertaining to Proportion of Doubling of Medial Consonants in Disyllables (N = 8790)* 62

Table 3. *Mean Individual Characteristics across Groups in Study 2* 63

Table 4. *Bigram Frequency for the CC Bigrams in Study 2* 64

Table 5. *Proportion of Doubling of Medial Consonants as a Function of Vowel Length, Number of Letters in First Vowel Spelling, Task, and Language Group* 65

Table 6. *Nonword Spelling Accuracy across Language Groups and Spelling Ability* 66

Table 7. *Nonword Spelling Accuracy across Spelling Ability and Pinyin Accuracy (for Mandarin Group)* 67

List of Figures

Figure 1. Proportion of doubling of medial consonants for vowel length across spelling ability (for all groups). 68

Figure 2. Proportion of doubling of medial consonants for the interaction between spelling ability and vowel length across tasks. 69

Figure 3. Proportion of doubling of medial consonants for vowel length across language groups. 70

Figure 4. Proportion of doubling of medial consonants for bigram frequency across vowel length. 71

Figure 5. Proportion of doubling of medial consonants for bigram frequency across tasks. 72

Figure 6. Proportion of doubling of medial consonants for the interaction between bigram frequency and vowel length across tasks. 73

Figure 7. Proportion of doubling of medial consonants in tasks across spelling ability. . 74

Figure 8. Nonword spelling accuracy across vowels for Americans and Singaporeans. . 75

Figure 9. Proportion of doubling of medial consonants for vowel length across pinyin accuracy (for Mandarin group). 76

Figure 10. Proportion of doubling of medial consonants for the interaction between vowel length and pinyin accuracy across tasks (for Mandarin group). 77

Figure 11. Proportion of doubling of medial consonants for bigram frequency across pinyin accuracy (for Mandarin group). 78

Speling "Successful" Sucesfuly: Statistical Learning in Spelling

Spelling is an important skill. In many cultures, people need to be able to spell their names correctly in application forms, their addresses, or the things they need in a shopping list. Spelling errors could result in disastrous consequences, such as the wrong medication being given to a patient or a misunderstanding of the writer's intention. Even people who use a spell-checker may still make spelling errors. A case in point is spelling errors that are related to doubling. After all, the spell-checker cannot tell if the writer intends to spell /kænən/ as ‹cannon› or ‹canon›, both of which are plausible spellings but with different meanings. How then do people spell?

Statistical learning has been widely used as a theory to explain language acquisition and learning (e.g. Saffran, 2002; Saffran, 2003) and, more pertinently, spelling (e.g. Treiman & Kessler, 2006; Kessler, 2009; Treiman et al., 2007; Pacton, Fayol, & Perruchet, 2005; Pollo, Kessler, & Treiman, 2009; Kemp & Bryant, 2003; Deacon, Conrad, & Pacton, 2008). According to Kessler (2009), the statistical learning perspective postulates that "complex sound-spelling correspondences are learned the same way many other patterns are learned in life: by observing and internalizing the relative frequency with which objects and events occur and co-occur" (p. 20). Through repeated exposure to text, people gradually pick up sound-spelling correspondences, without necessarily formulating any conscious accounts of them, making an implicit numerical analysis where orthographic patterns are essentially observations (or computations) that some sound-spelling correspondences are more frequent than others (Kessler, 2009). Indeed, people seem to be sensitive to statistical patterns in the language when spelling and such sensitivity begins quite early in life. For example, Pollo et al. (2009) found that even the prephonological spellings of children as young as 4-years-old are not random strings of letters but reflected a sensitivity to the frequencies of letters and bigrams in the children's language.

Following this theory of statistical learning, models of spelling consider the role of learning of statistical patterns or correspondences from exposure (e.g. Barry & Seymour, 1988; Kreiner & Gough, 1990; Kreiner, 1992). Barry and Seymour (1988) postulated that there exists a set of probabilistic sound-to-spelling mappings that relate phonemes to weighted lists of alternative spelling patterns, and the mapping that ends up being selected for spelling production is the more common one. These sound-to-spelling mappings are ordered by a continuum of sound-to-spelling probabilities, i.e., the frequency with which spelling patterns represent phonemes in words. Thus, the frequency of a spelling pattern can range from high, as with the correspondence between /æ/ and ‹a› (99% of the time), to middle, as with the correspondence between /i/ and ‹ea› (40% of the time), to low, as with the correspondence between /i/ and ‹ie› (6.7% of the time). The probability with which people pick a particular letter or letter group when spelling should be related to its overall frequency. Indeed, Barry and Seymour (1988) found that, when asked to spell nonwords, spellers produced more common spelling patterns of vowels than rare spellings. Furthermore, the probability of the least typical correspondence in a word should predict spelling difficulty. Kreiner and Gough (1990) found evidence to support this—college students found it more difficult to spell /ə/, which has many alternative spellings with similar probabilities, than other vowels.

These probabilities of sound-to-spelling contingencies as proposed above imply that the weighting of spelling is context-free, and thus, does not differ by the context of the phoneme. However, it has been found that people do take context into account, weighting different spellings differently in different contexts. This has been found, at least for monosyllables in English, where both the preceding and following consonant influence the spelling of a vowel (Kessler & Treiman, 2001). For example, in Treiman and Kessler (2006), participants were sensitive to consonant context, being more likely to spell ‹ea› for /ɛd/, where the spelling ‹-ead› is more probable, than /ɛp/, where the spelling ‹-eap› is less probable. Further evidence for context effects is also found in children and adult spelling (positional and phonemic context: Bernstein & Treiman, 2001; vowel context: Hayes, Treiman, & Kessler, 2005; consonant context: Treiman & Kessler, 2006; see Kessler, 2009, for a review of these studies). Adults are also influenced by positional information, morphological status, and other constraints when spelling (Perry, Ziegler, & Coltheart, 2002). These results suggest that experienced spellers know more than just the typical or context-free correspondences; as the statistical learning hypothesis predicts, they also know something about how sounds are spelled in certain environments.

Despite these contextual patterns, it seems that even skilled readers or spellers do not use these contextual patterns to the extent that would be expected given the patterns' reliability in the English vocabulary (Treiman & Kessler, 2006). Their rate of use of these statistical patterns was lower than that found in the general vocabulary to which these readers are exposed (Treiman et al., 2007; Kessler, 2009). For example, even the most skilled spellers in Treiman and Kessler's (2006) study did not produce as many ‹ea› spellings of /ɛ/ before /d/ as expected (only about 40% of the time) given how often ‹ea› occurs in this context in real English words (at least 73% of the time). There also appear to be individual differences even among adults in the use of context, with better spellers seeming to use context more (Treiman & Kessler, 2006). This seems to be congruent with the findings by Fisher, Shankweiler, and Liberman (1985) that better spellers were more sensitive to regularities at various levels of linguistic representation (e.g. orthographic and morphology) when spelling.

The ability to use these contextual patterns also seems to develop relatively late, with the use of preceding context possibly emerging earlier than the use of following context (Treiman & Kessler, 2006; see also Bernstein & Treiman, 2001). In terms of preceding context, previous letters in a word may serve as the cue for the next letter (Ebbinghaus, 1885, as cited in Goldberg, Buchwald, Kochaniak, & Rapp, 2005). In terms of following context, for example, when spelling a vowel, people may take into account the consonants that come after it, although the influence of probabilistic context-free phoneme-to-grapheme correspondences is stronger (Perry & Ziegler, 2004). If the role of preceding context is more important than the role of following context in spelling, this could possibly hint at the idea of spelling being a serial process. Simon and Simon (1973) postulated the idea that spelling is a serial process—one takes up each phoneme in sequence and generates a list of correspondences (which can be context-sensitive) before deciding on the appropriate spelling. Thus, the spelling of a phoneme would be less likely to be influenced by what comes after, as it would not have been generated yet.

Moving forward, the role of statistical patterns (context-free or context-sensitive) in spelling needs to be further investigated so as to provide a better explanation of how people spell. Furthermore, the above studies on statistical learning and context have mostly looked at monosyllables, and this gives us the motivation to extend the work to longer words such as disyllables. It would also be worthwhile to see if individual differences (such as those related to spelling ability) exist among adults in the sensitivity to the statistical patterns in the language. The doubling of medial consonants in English is an excellent test case for the statistical learning hypothesis, as well as to test for the role of context in spelling (preceding and following). Further motivations are discussed below.

Doubling of Medial Consonants in English

Many spelling errors in English involve doubling, as when people are stumped by whether to double the ‹l› in ‹trellis›. Here, doubling is defined as two adjacent identical letters (Pollock & Zamora, 1983). Studies analyzing spelling errors in texts and in spelling production mention doubling errors, either by omission or by inclusion, as one of the common errors in English (e.g. Pollock & Zamora, 1983; Yannakoudakis & Fawthrop, 1983; Spache, 1940), by both native and non-native speakers (e.g. Bebout, 1985; Haggan, 1981; James, Scholfield, Garrett, & Griffiths, 1993; Terrebone, 1973). As analyzing spelling errors can often give us insight into the cognitive processes involved in spelling (Jorm, 1977), this gives us the motivation to study doubling and the factors influencing doubling.

On a related note, spelling errors also occur most often in the middle of words. Jensen (1962) found a serial-position effect in the relative frequencies of spelling errors as a function of letter position, administering seven-, nine-, and eleven-letter words in a spelling task. This serial-position effect was in a bow-shaped curve similar to that produced by the serial learning of other materials, i.e., errors were most frequent in the middle of a word, with the fewest errors at the beginning and at the end of the word. Similar effects were found by Wing and Baddeley (1980), Kooi, Schutz, and Baker (1965), Buchwald and Rapp (2004), and Hildreth (1955, as cited in Kooi et al., 1965), although only Kooi et al. attempted to control for the number of difficult letter combinations such as a double-letter combination or a silent letter combination occurring at the beginning, middle, and the end of the spelling words. The serial position effect of spelling errors has been interpreted as resulting from interference between neighboring letters; as medial letters of a word have more neighbors than the letters at the periphery of word, they are more prone to being misspelled, resulting in a bow-shaped error distribution (Wing & Baddeley, 1980). Relating this to doubling in English, in terms of the position of doubling in a word, positional bigram counts by Solso and Juel (1980) showed that doubling of consonants in English occur mostly in the middle of a word as medial consonants, followed by at the end of the word, and rarely at the beginning of a word (see also Cassar and Treiman, 1997, p. 632). The middle of words is thus very difficult to spell and this gives us the motivation to focus on the doubling of medial consonants in our current study, and further discussion on doubling will be focusing on medial consonants.

Patterns in the Doubling of Medial Consonants

Doubling is especially difficult in English as single and doubled consonants often sound the same (e.g. Cassar & Treiman, 1997; Deacon et al., 2011). This contrasts with languages such as Finnish and Tamil, where doubled consonants are used to mark a phonetically long phoneme (see Lehtonen & Bryant, 2005, for Finnish; Balasubramanian, 1982, for Tamil). Juul and Sigurdsson (2005) compared Danish, which has more complex and inconsistent correspondences between consonant phonemes and doubled consonants, with Icelandic, which has a relatively more consistent doubling rule and where doubling also often marks that the medial consonant itself is phonetically long, and found that Danish children made more doubling errors than Icelandic children. Fowler (1937) talked about how difficult doubling consonants in English is, lamenting that "…if a list were made of the many thousands of words whose spelling cannot be safely inferred from their sound, the doubtful point in perhaps nine tenths of them would be whether some single consonantal sound was given by a single consonant, as m or t or c, a double consonant, as mm or tt…", giving trouble examples such as ‹committee› and ‹comity›, ‹inoculate› and ‹innocuous›, ‹harass› and ‹embarrass›, and more (p. 553-555). An interesting question would thus be that when people are spelling, or more specifically, deciding whether to double a medial consonant or not, whether they are guessing randomly or whether they are being influenced by statistical patterns in the language, as per the statistical learning hypothesis. Another question would then be whether these statistical patterns in the language are context-free or context-sensitive.

Despite the difficulty in the doubling of medial consonants, it has been found to follow somewhat of a pattern. In English, according to Carney (1994), double consonant letters do not normally follow long vowels (and there tends to be more doubling after short vowels). In an experimental task to test knowledge about the relationship between medial consonant doublets and its preceding vowel, Cassar and Treiman (1997) asked participants (kindergarten, first to third grade, sixth grade, and undergraduate) to choose between nonword spellings, with a single medial consonant (e.g. tebif) or a doubled medial consonant (e.g. tebbif), as they listened to its target pronunciation. Only the sixth graders and the undergraduates chose nonwords with doubled medial consonants after a short vowel and nonwords with single medial consonants after a long vowel significantly more than chance, suggesting that knowledge of the pattern of doubling after a short vowel is learned later than elementary school (Cassar and Treiman, 1997). Deacon et al. (2011) ran a similar experiment with second to fourth graders but with real word stimuli, e.g. ‹poppy› and ‹diner›. They found that a sensitivity to the impact of vowel length on the number of medial consonants chosen (single/double) existed from third grade onwards, i.e., the third and fourth graders were more likely to choose a single medial consonant after long vowels and a doubled medial consonant after short vowels. They postulated that their finding of this sensitivity at an earlier age was likely due to the use of real words instead of Cassar and Treiman's (1997) nonwords.

Carney (1994) also stated that some consonants such as ‹b› and ‹p› double more than others. Children seem to pick this up quite early, as shown by Treiman (1993). She looked at the writings produced over the course of the school year by a group of first graders whose teacher encouraged inventive spelling, reporting that children used frequent doublets such as ‹ee› or ‹bb› more often than infrequent doublets such as ‹hh› or ‹kk›. Where these doublets occur in a word also seem to be a pattern that is picked up early on. Earlier, we pointed out that doubled consonants in English occur mostly in the middle of words and rarely in the beginning of words (Solso & Juel, 1980). Treiman (1993) found that the same children also rarely used double consonants at the beginning of a word. Similarly, Cassar and Treiman (1997) found that children by the end of kindergarten could eliminate nonword spellings with word-initial doublets, choosing ‹baff› as more word-like than ‹bbaf›.

Interestingly, there seems to be a tendency to omit an identical consonant than to insert one when making doubling errors; amongst the doubling errors, there are more instances of using a single consonant when doubling is required, e.g. ‹emba<u>r</u>ass› instead of ‹emba<u>rr</u>ass›, as compared to using a doubled consonant when a single consonant is required, e.g. ‹i<u>nn</u>oculate› instead of ‹i<u>n</u>oculate› (Pollock & Zamora, 1983; Yannakoudakis & Fawthrop, 1983; Bebout, 1985; James et al., 1993; Haggan, 1981; Deacon, Leblanc, & Sabourin, 2011). Pollock and Zamora (1983) found 227 cases of singling (where doubling is required) versus 96 cases of doubling (where singling is required) in the doubling errors that occur in normal text. As for Yannakoudakis and Fawthrop (1983), they found omission to be the most common error operation (34%), with the failure to double accounting for 7.5% of omission errors, i.e., singling where doubling is required. Insertion errors were found to be the second most important error operation (27%), with doubling resulting in 45% of insertion errors. Likewise, Bebout

(1983), Haggan (1981), Terrebone (1973), and James et al. (1993) all found consonant doubling errors to comprise more of the failure to double a consonant rather than the unnecessary doubling of one. Deacon et al. (2011) found that children between grades two to four made these same errors, having a preference for single consonants over double consonants in their incorrect as well as correct spellings. According to Cassar and Treiman (1997), this preference for single medial consonant spellings may indicate that “spellers have learned that single consonant spellings are more frequent or simpler than double consonant spellings” (p. 637).

Despite the abovementioned studies looking at the doubling patterns of vowel length, medial consonant, and position of doubling, there has been no systematic study of other patterns of doubling consonants in English, and this gives us the motivation to study more factors that influence the doubling of medial consonants in English.

To explore the statistical learning hypothesis and the use of context-sensitive patterns in the doubling of medial consonants, it would be important to see what sort of spelling patterns in relation to doubling can be found in the English language. To the best of our knowledge, no published paper has looked at the statistical patterns of the doubling of medial consonants in the English language. Carney (1994; p. 112-130) did mention several other “generalizations” in relation to consonant letter doubling other than vowel length and bigram frequency: no doubling if the preceding vowel is spelled with more than one letter; no doubling before Latinate endings, i.e., endings that are derived from Latin such as /ɪk/, /ɪd/, /ɪt/; and others. These generalizations do have exceptions. For example, ‹canon› vs. ‹cannon›, both of which are pronounced the same, /kænən/, have a preceding short vowel, but one has a single ‹n› and the other has a double ‹n›. Although Carney (1994) discussed these generalizations and exceptions in his book, he did not provide any statistics on these proportions of doubling. This gives us the motivation for Study 1, which is to tabulate the vocabulary statistics and patterns in relation to doubling in English. These statistical patterns are a first step in the test of the statistical learning hypothesis and the role of context in spelling. Using these statistical patterns, the goals of Study 2 are then to collect behavioral data on doubling to further investigate statistical learning and context in spelling as well as to explore the role of individual differences such as spelling ability. Other than spelling ability, another important individual difference that we will be exploring is language background, which will be discussed below.

Influence of Language Background on Spelling

The investigation of the influences of language background on spelling is especially important given the diversity of English-language speakers today (Share, 2008). In fact, non-native speakers of English outnumber its native speakers by a ratio of three to one (Crystal & Wang, 1997). We define language background as being immersed in an environment in two ways: 1) The use of a particular dialect of English 2) The use of a language other than English. Here, we compare American English speakers with Singaporean English speakers.

Despite English being a global language, there are still differences in the way English is spoken in many places, e.g. in terms of accent, pronunciation (e.g. rhotic: pronouncing the ‹r› before a consonant or at the end of a syllable vs. non-rhotic), spelling (e.g. ‹colour› vs. ‹color›), and more (see Crystal & Wang, 1997). These phonological differences in dialect have been shown to influence the way one spells in English (e.g. African American English: Kohler et al., 2007; Treiman, 2004; Terry, 2006; Cronnell, 1979; Appalachian: Cantrell, 2001; American and British English: Treiman, Goswami, Tincoff, & Leevers, 1997; Treiman & Barry, 2000). As it is well documented that phonological knowledge is important for spelling (e.g. Treiman, Berch, Tincoff, & Weatherstone, 1993; Treiman, 1993; Treiman, 1994), these phonological differences result in different types of spelling errors specific to that particular dialect in English. For example, comparing American (a rhotic dialect) and British English (a non-rhotic dialect), Treiman and Barry (2000) found that the percentage of misspellings of words like ‹horde› that did not include an ‹r› was almost three times larger among British college students, where ‹horde› is typically pronounced without the /ɹ/, than among American college students, where ‹horde› is typically pronounced with the /ɹ/. These errors reflecting their dialect begin from a young age of six (Treiman et al., 1997). It would thus be interesting to investigate whether these phonological differences in dialect will affect an individual's sensitivity to certain kinds of statistical patterns in the language, and thus, moderate the role of statistical learning in spelling. Here, we can conduct such an investigation through the doubling of medial consonants.

Comparing our two dialects of interest, American English is a rhotic dialect but Singaporean English, like British English, is a non-rhotic dialect. More importantly, there are some pronunciation differences in American English and Singaporean English. Many of the phonemic vowel and consonant distinctions that exist in American English do not exist in colloquial Singaporean English (Brown, 1988; Deterding, 2007; Lim, 2004). In Singaporean English, and especially in colloquial Singaporean English, there are many vowel conflations, with the main ones being /ɛ, æ/ and /i, ɪ/ (Brown, 1988). Thus, the contrast between tense-lax (or short-long) vowel pairs such as /i/ and /ɪ/, which exists in American English, is not much in standard Singaporean English and significantly indistinguishable in colloquial Singaporean English in terms of both quality and quantity (Lim, 2004). Many Singaporeans thus pronounce ‹beat› /bit/ and ‹bit› /bɪt/ similarly. In terms of doubling, as vowel length is not as differentiated in Singaporean English as in American English, Singaporean English speakers might not be as sensitive to the pattern of vowel length in doubling as compared to the American English speakers.

Furthermore, as discussed earlier, dialect differences often lead to different types of spelling errors. This gives us the motivation to study nonword spelling accuracy and whether the dialect differences between Singaporean and American English affect nonword spelling accuracy and the kinds of spelling errors they make. Nonwords make excellent stimuli as one would not have to worry about frequency or familiarity effects that might affect spelling. The common mispronunciation of certain consonants such as /d/ for /θ/ and the conflation of some vowels lead to certain consonants and vowels sounding more similar in colloquial Singaporean English than in American English. All this might have a negative impact on the Singaporeans when they are being given these specific phonemes to spell in a new word. Furthermore, the truncation of or the lack of emphasis on endings in colloquial Singaporean English might also result in Singaporeans making more errors in spelling the endings of new words. In terms of analyzing spelling errors, we choose to focus on vowel spellings due to the vowel conflations in Singaporean English.

Being exposed to a language other than English has also been shown to influence the way one spells in English (e.g. Holm & Dodd, 1996; Cook, 1997; Wang & Geva, 2003; San Francisco, Mo, Carlo, August, & Snow, 2006; see Figuerido, 2006, for a review). Similar to studies that looked into the influence of dialect in spelling, studies have shown that exposure to a particular language leads to spelling errors that are specific to that. For example, Cook (1997) found that errors in spelling ‹l›, ‹r› and epenthetic ‹e› were specific to Japanese L2 users when comparing across a variety of L2 English users from different language backgrounds. San Francisco et al. (2006) found that Spanish-English bilinguals, who underwent reading instruction in Spanish, exhibited Spanish-influenced spelling errors when asked to spell English pseudowords.

In general, the use of one writing system can affect the use of another writing system, depending on how different they are (Geva, Wade-Woolley, & Shany, 1993). Many studies have shown that exposure to a particular orthography influences how one spells, depending on whether that exposure helps in the development of knowledge of phonological processing and letter-sound correspondences in that sounds map onto letters (Sparks, Patton, Ganschow, Humbach, & Javorsky, 2008; Sparks, Patton, Ganschow, & Humbach, 2009). Being exposed to an alphabetic orthography, where there are phoneme-to-grapheme correspondences, might thus help in such development, and not exposure to non-alphabetic orthographies. For example, Wang and Geva (2003) as well as Holm and Dodd (1996) found that exposure to a logographic (non-alphabetic) orthography

(Mandarin), where there are no phoneme-to-grapheme correspondences, has a negative effect of nonword spelling accuracy in English. They also showed that English speakers with a Mandarin-speaking background who had learned pinyin were able to spell nonwords better. Pinyin is the alphabetic form of Mandarin which has one-to-one phoneme-to-grapheme correspondences.

When studying how the use of a language other than English can influence spelling, Singapore serves as an excellent subject pool for two reasons: 1) As a former British colony, English is one of the official languages in Singapore and is the de facto language for all purposes such as education, politics, law, business, and more. 2) As a multi-ethnic society, many Singaporeans have a second language that is one of the three other official languages in Singapore—Mandarin, Malay, and Tamil. From primary one to secondary four (the equivalent of first grade to tenth grade), Singaporeans are required to go through an English-based education system with English as a first language and either Mandarin, Malay, Tamil, or another approved language as their second language. Many Singaporeans thus have been through a nationally standardized way of learning English with the only difference being their second language. It is important to clarify that the term “second language” here does not necessarily refer to the second language a person learns. For many Singaporeans, their second language is their mother tongue and they are immersed in an environment where both English and their mother tongue are often spoken both at home and at school. At the very least, in terms of official instruction, many are exposed to both languages in kindergarten from the age of five. For these reasons, we do not have to be concerned about the following possible confounds: 1)

Whether our participants differ across groups in terms of English or their second language instruction; the nationally standardized curriculum for English and other languages ensures a consistent method of instruction. 2) Whether our participants differ across groups in terms of exposure and the use of English and their second language; compulsory education in Singapore ensures that they are exposed to and use both English and their second language for at least ten years.

A comparison can then be made across language groups (Mandarin, Malay, and Tamil) in relation to the doubling of medial consonants in English. Previous studies that looked at doubling errors by non-native English speakers have only discussed the difficulty of doubling in English and not the reasons behind these doubling errors. The closest that came to that was Bebout (1985) who looked at spelling errors made by Spanish-English bilinguals and postulated that their difficulty in doubling was due to the rare doubling of medial consonants in Spanish. Here, we are interested in how the knowledge of a language other than English might influence one's sensitivity to the statistical patterns in the language, and thus, the role of those patterns in influencing spelling. As there is no doubling in Mandarin and Malay, if one writing system does influence the use of another, one would thus not expect any influences from Mandarin or Malay with respect to doubling in English for Singaporean Mandarin and Malay speakers.

However, although Tamil has a different script than English, there is consonant doubling in Tamil. As in Finnish, this consonant doubling serves the purpose of lengthening the consonant sound in a particular word (Lehtonen & Bryant, 2005). Unlike in English, where single and doubled medial consonants sound the same, a doubled medial consonant in Tamil thus sounds different from a single medial consonant. Lehtonen and Bryant (2005) found that children had already internalized this pattern of doubling (to mark a longer phoneme in a word) in Finnish by third grade but no comparison was made to see if this pattern of doubling in one language influences doubling in another language where there is no such pattern, e.g. English. If it does, then Singaporean Tamil speakers might be less influenced by certain statistical patterns when doubling in English, possibly associating consonant doubling with the lengthening of doubled consonants in Tamil instead.

Just as we can use nonword spelling accuracy to study the influence of dialect differences in spelling, so can we use nonword spelling accuracy to study the influence of the use of one language on another. Even within Singaporeans, there might be differences in nonword spelling accuracy that can be attributed to language background, despite them undergoing the same English language instruction and having similar exposure to English. This would suggest that an individual’s second language can either hurt or help the individual’s ability to spell new words in one language. Singaporean Malay and Tamil speakers, who are exposed to alphabetic orthographies that have one-to-one sound-to-spelling correspondences, might perform much better in spelling new words than Singaporean Mandarin speakers, who are exposed to a logographic orthography. For Singaporean Mandarin speakers, greater knowledge of pinyin, the alphabetic form of Mandarin with sound-to-spelling correspondences, might be associated with higher nonword spelling accuracy, as shown by Holm and Dodd (1996) as well as Wang and Geva (2003).

Looking further at pinyin, there are two components of pinyin: the phonemic part (e.g. ma, xin) and the tonal part (1-4) (see Siok & Fletcher, 2001, and Ho & Bryant, 1997, for a description of the Chinese writing system and pinyin). The tonal part is unrelated to the English language as there are no tones in English. Thus, knowledge of tones might be a good measure of pinyin knowledge in bilingual English-Mandarin speakers. Any influence of tonal knowledge on spelling will not be confounded with phonemic awareness that results from the knowledge of the English language. While it was previously thought that only the phonemic part of pinyin influences word recognition and reading, studies have also shown that the tonal part (or tone processing skill) also influences word recognition and reading through being a part of phonological awareness (Wang, Perfetti & Liu, 2005; McBride-Chang, Tong, Shu, Wong, Leung, & Tardif, 2008). According to Siok and Fletcher (2001), tone awareness and pinyin knowledge become significantly correlated by the third grade; tonal information becomes more important when more characters (homophones) are encountered. With this, one could then use tonal knowledge as a valid constituent of pinyin accuracy, and thus, it gives us the motivation to focus on tonal knowledge as a measure of pinyin knowledge to see if it predicts nonword spelling accuracy.

Overview of Studies 1 and 2

In this paper, we seek to explore factors that influence spelling. Specifically, we attempt to investigate the statistical learning hypothesis and the role of context in spelling, using the doubling of medial consonants in English as a test case. We also explore the role of individual differences in spelling by looking at spelling ability and language background.

In Study 1, we seek to tabulate the statistical patterns in relation to the doubling of medial consonants in English. We focus on four patterns that will allow us to look into the role of context: preceding vowel length, number of letters spelled in preceding vowel, medial consonant, and ending type. We predict that despite the difficulty of doubling in English, there exists patterns that can be used when deciding whether to double or not in spelling.

In Study 2, we seek to collect behavioral data to show whether people are sensitive to these statistical patterns when doubling or whether they are guessing randomly. If they are influenced by statistical patterns in the language, we then investigate if the influence of these statistical patterns depend on the medial consonant itself as well as the context of these patterns, i.e., the preceding context and the following context. We also explore whether such sensitivity to statistical patterns is moderated by other factors such as spelling ability, language background, and the kind of spelling task one is asked to do. Here, we predict that people are influenced by statistical patterns when doubling and that these influences are moderated by the following: context, language background, spelling ability, and task.

In Study 2, we also examine the influence of dialect and language background on spelling through nonword spelling accuracy. We predict that differences in dialect will affect nonword spelling accuracy and the kinds of spelling errors that are produced. We also predict that if knowledge of one writing system influences the knowledge of another writing system, being exposed to certain kinds of orthography will have an effect on nonword spelling accuracy. For Mandarin speakers, pinyin knowledge is also predicted to aid in nonword spelling accuracy.

The predicted results will help us to understand the factors that underlie spelling, especially in relation to statistical learning. This is the first study that will systematically examine the influence of various statistical patterns in the doubling of medial consonants as well as the factors that moderate such influences.

Study 1

The aim of this study is to create a database and tabulate vocabulary statistics with regards to doubling in English. The lexicon used is the Unisyn lexicon, which is a master lexicon transcribed in symbols that allow the encoding of multiple accents of English. This lexicon has around 120,000 words.

Method

A General American pronunciation key was loaded onto the lexicon and a program was written in Python to extract lexical information (e.g. orthographic and phonological endings, suffixes, first vowels, medial consonants, number of morphemes, and doubling of medial consonant) from a subset of words in the Unisyn lexicon. The extracted information was then analyzed to derive statistical patterns in relation to the doubling of medial consonants.

8790 words were chosen based on the following list of criteria: 1) Disyllabic—words that can possibly be pronounced as more than 2 syllables were removed. 2)

Stressed first vowel and unstressed second vowel 3) True medial consonants phonologically—words with medial glides such as ‹power› and ‹kiwi› were removed. 4) Doubling of medial consonants involve two identical letters in a row, e.g. ‹bb›, ‹cc›, and not ‹ph›, ‹ck›. Words such as ‹hyphened› were removed. Words such as ‹burley› or ‹islet› were categorized under words with a single medial consonant where the first vowel is spelled with more than one consonant, thus they were included. Words with endings such as –le or –re, e.g. ‹treble›, ‹bubble›, and ‹acre›, have their endings coded as –le or –re and were also included. 5) Not a compound word (e.g. ‹handstand›, ‹hangman›), fused word (e.g. ‹betcha›), foreign word, or a proper noun only, i.e., words that can only be proper nouns were removed and words that can be parts of speech other than proper nouns were included. 6) Not duplicated in the database—the second or third appearance of a word in the lexicon due to ambiguous morphological status or part of speech were removed, e.g. ‹adder› (snake/someone who adds), and ‹opens› (noun/verb).

The following lexical information was then extracted from the selected 8790 disyllabic words: 1) Sound and spelling of initial consonant, first vowel, medial consonant, and ending (defined as the string of phonemes and letters after the medial consonant) 2) Doubling of medial consonant (single: C, double: CC) 3) Orthographic and phonological suffixes 4) Number of morphemes 5) Number of letters 6) Word frequency (as provided in the Unisyn lexicon).

Results

With the extracted lexical information, we were then able to derive statistical patterns based on log type frequency and token pertaining to the proportion of doubling of medial consonants in the English language (see Tables 1 and 2). Overall, medial consonants double only 37% of the time. Here, we focus on four of Carney's (1994) generalizations based on their positional context with reference to the medial consonant (preceding and following) as well as the actual medial consonant itself. All proportions are based on log type frequency so as to take into account the frequency of a particular word.

Before the medial consonant. As Carney (1994) postulated, there is little doubling if the preceding vowel is spelled with more than one letter—medial consonants double 46% of the time after one-letter vowels and only 2% of the time after vowel digraphs. There is also more doubling after short vowels and less doubling before long vowels. For all spellings of the preceding vowel, medial consonants double 65% of the time after short vowels and only 1% of the time after long vowels. For preceding vowels that are spelled with only one letter, medial consonants double 69% of the time after short vowels and only 1% of the time after long vowels.

At the medial consonant. Some medial consonants double more than others. For example, ‹b› and ‹l› double 49% of the time as compared to only 3% for ‹v›. There is no doubling for consonants ‹h›, ‹j›, and ‹q› in terms of token frequency. They do not double at all in this sample of disyllabic words, thus their proportion of doubling in log type frequency were not available.

After the medial consonant. Last, as Carney (1994) postulated, there is little doubling before Latinate endings for all disyllables—medial consonants that come before ‹-ic›, ‹-id›, and ‹-it› double 5%, 10%, and 11% of the time respectively. Latinate endings are derived from Latin whereas non-Latinate endings are not. When limited to disyllables with preceding short vowels that are spelled with a single letter, medial consonants that come before ‹-ic›, ‹-id›, and ‹-it› double 19%, 5%, and 10% of the time respectively. However, the same medial consonants that come before non-Latinate endings (and preceded by short vowels spelled with a single letter) double 91%, 60%, and 57% of the time for ‹-ow›, ‹-age›, and ‹-is› respectively.

Discussion

In order to test the statistical learning hypothesis in doubling, and thus, whether the patterns in doubling are context-free or context-sensitive, the first step is to look at the statistical patterns in the language. Study 1 thus sought to tabulate the statistical patterns concerning the doubling of medial consonants in English. Here, we focused on four of Carney's generalizations that would allow us to study the factors influencing doubling patterns and whether these patterns are context-sensitive or not, i.e., medial consonant (frequency of doublet), preceding context (number of letters spelled in the preceding vowel and preceding vowel length), and following context (type of ending).

The results supported several of Carney's generalizations about doubling medial consonants in English. The doubling patterns found also seem to be context-sensitive—there is hardly any doubling of a medial consonant if it is preceded by a vowel that is spelled with more than one letter or a long vowel, and if it is followed by a Latinate ending. Furthermore, as Carney stated, some medial consonants double more than others, e.g. ‹b› and ‹l› double much more as compared to ‹v›, which rarely doubles, or even ‹h›, ‹j›, and ‹q› which do not double at all.

The next step is to then collect behavioral data for these statistical patterns in doubling, as we do in Study 2. A question that one could then ask would be whether people have internalized these patterns when they are deciding whether to double a medial consonant or not. Another question is whether the doubling of the medial consonant is influenced by the positional context of these patterns (preceding and following) and the medial consonant itself.

Study 2

The aim of this study is to follow up on the statistical patterns in relation to doubling in English found in Study 1 and collect behavioral data on doubling. Here, we are interested in seeing if people are sensitive to the statistical patterns found in Study 1 when they are doubling and if this sensitivity depends on context. If spelling alternatives are based on probabilities weighted by overall frequency across contexts, there would be little doubling overall as singling is much more common than doubling (63% vs. 37%) based on an earlier discussion (Treiman & Kessler, 2006; Kessler, 2009). However, this does not predict any context effects; if the spelling alternatives are weighted by context, we would then find effects of context on doubling.

If spelling is a serial process, as postulated by Simon and Simon (1973), one would expect effects of preceding context but not following context on the doubling of medial consonants. Other experiments have shown that children who are better spellers took more advantage of context, using preceding and following consonants to spell vowels (Treiman and Kessler, 2006; see also Bernstein & Treiman, 2001). This also suggests the possibility of looking at spelling ability as an individual difference in the influence of these statistical patterns on doubling medial consonants. We also explore other individual differences that might influence doubling such as language background and the type of spelling task one is asked to do.

Method

Participants. Native English-speaking participants in the United States were recruited through a subject pool in the psychology department of a Midwestern university (N = 68). Participants comprised 27 Whites, 14 African Americans, 22 from Asian origin, 2 of mixed race, and 3 unspecified. There were 41 females and 27 males, and 66 were students and 2 were working. Participants were either given course credit or compensated USD10 for their participation.

Participants in Singapore (N = 140; 54 Mandarin-speaking, 44 Malay-speaking, 42 Tamil-speaking) were recruited through e-mail invitations to the respective language societies in a Singaporean university as well as through word of mouth. To ensure proficiency in their second language, participants were required to have at least a B grade in the national second language exam taken usually at the age of 16. 40% of the Singaporean participants reported that they were native English speakers; among the self-reported non-native English speakers, 62% of them reported that English was the first language they learned to write. In terms of ethnicity, Mandarin-speaking participants comprised 53 Chinese and 1 Eurasian; Malay-speaking participants comprised 35 Malays, 2 Indians, 1 Chinese, 1 Eurasian, 3 of mixed ethnicity, and 1 unspecified; Tamil-speaking participants comprised 42 Indians. There were 74 females and 66 males, and

101 were students and 39 were working. Participants were compensated the equivalent of around USD10 for their participation.

Table 3 shows the mean ages and their standard deviations of each language group. All participants reported that they had no current diagnoses of reading, hearing, speech, or spelling disorders. The experiment lasted for an hour in both experimental venues.

Measures (Individual-level predictors).

Language background and demographics questionnaire. A questionnaire (adapted from Rickard Liow & Poon, 1998; see Appendices C & D) was administered to determine the participants' language background, proficiency, and pinyin background. Participants were then grouped according to their self-reported language dominance, e.g. American English, Mandarin, Malay, and Tamil.

WRAT spelling test. A spelling test of 56 words was adapted from the list of words used in the spelling subtest of the Wide Range Achievement Test (WRAT). The first 14 items from both the Blue and Green forms were removed as they were determined to be too easy for adults, e.g. ‹go›, ‹cat›, and ‹boy›; it was assumed that participants would have gotten these items correct. Participants' raw scores were adjusted according to the norms of their respective age groups. The standardized WRAT score is used to operationalize *spelling ability* as an individual difference variable**.** Table 3 shows the mean standardized WRAT score for the four groups.

Shipley vocabulary test. The Shipley (1940) Vocabulary Test is a 40-item multiple-choice standardized test to measure proficiency in English. Here, we used it to ensure that the Singaporean and American groups have similar proficiency in English. Items were presented in the test in order of increasing difficulty. Participants were scored out of 40 and the resulting score corresponded to a vocabulary age**.** Table 3 shows the mean vocabulary age for the four groups.

Chinese character transcription (pinyin) test. This was adapted from a language screening test used by Tham et al. (2005) involving pinyin transcription with tone assignment. The Chinese characters used had a wide range of difficulty that was derived from Singaporean norms based on work by Rickard Liow, Tng, and Lee (1999). Appendix B shows the full list of characters used in the test. Participants were required to write out the pinyin and the tone (e.g. xin1) of a Chinese character. Error rates were based on two components: pinyin and tone. Participants were graded on whether they transcribed the correct pinyin of the character. If so, they were then graded on whether they assigned the correct tone. High error rates in the latter component reflect low tonal knowledge that is separate from Chinese character knowledge. Tonal knowledge was used to operationalize *pinyin accuracy* of the Chinese participants. Overall accuracy on the test was used to operationalize their Chinese character knowledge, i.e., Mandarin vocabulary score.

Task type. This refers to the nonword spelling task participants were assigned to: choice versus free spelling. Both tasks are further described below.

Choice task. Participants listened to a nonword three times and chose what they thought was the best spelling of that nonword from two options, e.g. blemic | blemmic.

Free spelling task. Participants listened to a nonword three times and wrote down what they thought was the best spelling of that nonword. Participants' responses were scored based on a list of plausible spellings for the initial consonant, first vowel, medial consonant, and ending consonants of that particular nonword based on their dialect (American English or Singaporean English). There was a list of plausible spellings of phonemes for each dialect of English and the two were different in some respects. This list can be found in Appendix E. An example of a difference would be the inclusion of a silent ‹r› in a vowel, e.g. ‹or› for /ɒ/, for Singaporean English, whereas such an inclusion of ‹r› would be inaccurate for American English. A nonword was considered to have been accurate only if all phonemes were plausibly spelled.

Measures (Item-level predictors).

Vowel length. This refers to the quality of the vowel that comes before the medial consonant of the nonword, either long (e.g. /aɪ/, /o/, /i/, /e/) or short (e.g. /ɑ/, /æ/, /ɪ/, /ɛ/).

Number of letters in the first vowel. This refers to the number of letters used to spell the vowel that comes before the medial consonant of the nonword stimulus. In the free task, this is coded as either one, e.g. ‹gribit› for /graɪbɪt/, or more than one, e.g. ‹graibit› for /graɪbɪt/. In the choice task, all nonword stimuli have a single vowel letter before the medial consonant, so they were all coded as one.

Log (bigram frequency). This refers to the log of the averaged frequency (per million) of the unconstrained bigrams for consonants such as ‹mm›, ‹vv›, and ‹ll›. An unconstrained bigram is defined as a specific two letter combination (bigram) within a word, regardless of its position, or the word length. For example, the 'ba' in 'bat' is considered the same as the 'ba' in 'tabasco'. The frequency norms were retrieved from the MCWord database (Medler & Binder, 2005), which was based on the CELEX efw.cd file. This file includes all the English word forms from a COBUILD (Collins Birmingham University International Language Database) corpus of both written and spoken text, which contains approximately 17,900,000 instances of word use. Table 4 shows the values for the specific CC bigrams used in the experiment.

Ending type. This refers to the type of ending of the nonword stimulus: Latinate versus non-Latinate. Latinate endings are derived from Latin, e.g. /ɪk/, /ɪd/, /ɪt/, whereas non-Latinate endings are not, e.g. /ɪb/, /ɪp/, /ɪg/. These non-Latinate endings were chosen because they are orthographically and phonologically similar to Latinate endings, differing by just one phoneme or letter. Furthermore, they could not be interpreted as a separate morpheme like /ɪng/.

Stimuli. Ninety-six disyllabic nonwords were created (four groups of 24 nonwords each) with different combinations of vowel length (short: /ɑ/, /æ/, /ɪ/, /ɛ/ vs. long: /aɪ/, /o/, /i/, /e/) and ending type (Latinate: /ɪk/, /ɪd/, /ɪt/ vs. Non-Latinate: /ɪb/, /ɪp/, /ɪg/). The medial consonants were ‹m›, ‹l›, ‹n›, ‹p›, ‹b›, and ‹v›. A full list of the stimuli appears in Appendix A. Each group of stimuli was matched on the following characteristics: number of letters, phonemes, medial consonants, onsets or initial consonant clusters, vowel, and ending. Care was taken to ensure that the nonword stimuli had neither rhyming nor Coltheart neighbors to real words. Coltheart neighbors were defined both orthographically and phonologically, i.e., neighbors with one letter or phoneme replaced. Five sets of stimuli were created with different randomized orders of the 96 nonwords so as to prevent any order effects that may be caused by presenting all participants with the same order of stimuli.

Design. The study is a mixed-models design looking at whether the proportion of doubling of medial consonants in the nonword stimuli was predicted by the following variables: Individual differences (spelling ability, language group, pinyin accuracy for Mandarin group); linguistic characteristics (vowel length, ending type, number of letters in the first vowel, bigram frequency of medial consonants); and task (choice vs. free).

Procedure. In both the United States and Singapore, participants were tested in groups ranging from one to 10. They were first given the WRAT spelling test and the Shipley vocabulary test. Groups were randomly assigned their spelling task (free/choice) as well as the set of nonword stimuli that they were told to spell during the experimental session. Mandarin-speaking participants in Singapore were then given the *pinyin* transcription test, where they were asked to write the *pinyin* transcription of Chinese characters. Last, participants were asked to complete a demographics and language background questionnaire. After the experiment, participants received remuneration for their participation and were then debriefed.

Results

Data from the free and choice task were combined and analyzed using the lme4 software package (Bates, 2011) via a multilevel model with participant and stimulus as random factors. We looked at two main dependent variables separately: proportion of doubling of medial consonants and nonword spelling accuracy.

Main analysis: All language groups. All language groups were included in this analysis.

Proportion of doubling: All endings. We asked whether the proportion of doubling of medial consonants was predicted by the following variables: Individual differences (spelling ability, language group); linguistic characteristics (vowel length, ending type, number of letters in the spelling of the first vowel, bigram frequency of medial consonants); and task (choice vs. free). The continuous variables, spelling ability and bigram frequency, were centered in all analyses. A preliminary analysis revealed significant main effects of spelling ability ($p < .001$), bigram frequency ($p < .001$), vowel length ($p < .001$), number of letters in the spelling of the first vowel ($p < .001$), and task ($p < .001$).

In our next preliminary analysis, we asked if the linguistic characteristics that were significant in predicting the proportion of doubling interact with individual differences (spelling ability, language group) and task. Thus, in addition to the main effect of each variable, the second model included these two-way interactions. An interaction between bigram frequency and vowel length was also included to see if the preliminary interaction found in raw proportions was significant. The interaction between language group and number of letters in the spelling of first vowel was not included as its inclusion resulted in an unusually high standard error due to a much lower number of first vowels spelled with more than one letter for the Mandarin group (7.6%) as compared to the other groups (all > 9%). The analysis revealed significant two-way interactions between bigram frequency and vowel length ($p < .001$), bigram frequency and task ($p <$

.001), language group and vowel length ($p < .001$), and spelling ability and vowel length ($p < .001$).

Last, we were interested to see if the above significant two-way interactions interact with task, so we decided to test for these possible three-way interactions. Thus, in addition to the main effects and two-way interactions from the first two models, the final model included these three-way interactions. Results from the final multilevel model are organized based on the location of the statistical patterns in reference to the medial consonant.

Before the medial consonant. Results from the final multilevel model showed a significant main effect of the number of letters in the spelling of the first vowel (p = .001). Similar to the statistical patterns, if the vowel before the medial consonant was spelled with more than one letter, participants were unlikely to double the medial consonant (see Table 5 for proportion of doubling).

There was also a significant main effect of vowel length ($p < .001$). Similar to the statistical patterns, participants were more likely to double medial consonants after a short vowel than after a long vowel (see Table 5 for proportion of doubling). This effect was moderated by spelling ability as well as language background. The spelling ability and vowel length interaction ($p < .001$) reflects that better spellers were more influenced by vowel length. That is, the better the speller, the more likely he or she was to double a medial consonant after a short vowel than a long vowel (see Figure 1). Following this interaction, there was a significant three-way interaction with spelling ability, vowel length, and task ($p < .001$). The interaction between spelling ability and vowel length, where better spellers were more likely to double a medial consonant that comes after a short vowel than a long vowel, was more pronounced in the choice task than in the free spelling task (see Figure 2).

The language background and vowel length interaction ($p < .001$) reflected the fact that the American group was the most influenced by vowel length when doubling. As Figure 3 shows, they were the most likely to double medial consonants after a short vowel than a long vowel. The Mandarin and Malay groups were also influenced by vowel length in the same direction, although to a lesser extent. The Tamil group was not significantly influenced by vowel length. There was no significant three-way interaction of language background, vowel length, and task.

At the medial consonant. Participants doubled more with medial consonants of higher bigram frequency (High: 28% vs. Low: 23%) but this initially significant main effect was no longer significant in the final model. However, there was a significant bigram frequency and vowel length interaction ($p < .01$). There was a stronger effect of bigram frequency (greater doubling of medial consonants with higher bigram frequency) for medial consonants after short vowels than after long vowels (see Figure 4). There was also the bigram frequency and task interaction ($p < .001$), where there was a larger effect of bigram frequency in the free spelling task than in the choice task (see Figure 5). Following these two interactions, there was a significant three-way interaction with bigram frequency, vowel length, and task ($p < .05$). The interaction between bigram frequency and vowel length was more pronounced in the free task than in the choice task (see Figure 6).

After the medial consonant. Participants did not show a significant effect of ending type when doubling (Latinate: 25% vs. Non-Latinate: 26%).

Task. Participants were more likely to double medial consonants in the choice task than in the free spelling task, as shown by the significant main effect of task ($p < .001$; see Table 5 for proportion of doubling). There was also a significant spelling ability and task interaction ($p < .001$), such that better spellers showed more doubling than poorer spellers in the choice task. Not only was there a smaller effect of spelling ability in the free task, the interaction could also have reflected a floor effect in the free task (see Figure 7).

Language background. There was a significant main effect of language background ($p = .01$). The Mandarin group doubled the most, followed by the American English and Malay groups, and the Tamil group doubled the least (see Table 5 for proportion of doubling).

Proportion of Doubling: Only endings as defined by experimenter. To see whether the lack of an ending effect was due to participants not spelling the endings as intended by the experimenter, we removed endings that were not spelled as ‹ib›, ‹ip›, ‹ig›, ‹ic›, ‹id›, and ‹it› in the free task. A total of 1057 trials were removed and 16190 trials remained in both tasks. We then ran the same multilevel model as in the previous analysis and found similar effects but still no effect of ending.

Nonword Spelling Accuracy. We asked whether the accuracy of nonword spelling in the free spelling task was predicted by the following variables: age, sex, spelling ability, and language group. Age and sex did not predict nonword spelling accuracy but language group ($p < .001$) and spelling ability ($p < .01$) did. The effect of language group on nonword spelling accuracy is interesting given the groups' comparable standardized WRAT scores. Table 6 shows the nonword spelling accuracy of each language group as well as for participants at high and low WRAT scores. On average, the Mandarin group were good spellers, yet they were poorer at nonword spelling than the Malay and Tamil groups. The American group was the most accurate in nonword spelling. The effect of spelling ability confirms the value of the WRAT spelling task in that the better the spellers performed on the WRAT, the better they were at spelling nonwords.

The Singaporean group spelled nonwords less accurately (65%) than the American group (94%). Looking at spelling accuracy across vowels, the Singaporean group spelled all vowels almost as accurately as the American group except for two vowels: /æ/ (47% vs. 86%), and /ɪ/ (86% vs. 98%). Figure 8 shows spelling accuracy across vowels for both groups.

Supplemental analysis: Mandarin group. Supplemental multilevel model analyses were carried out specifically on the Mandarin group using a similar general statistical approach as the main analysis, but with the inclusion of the individual difference variable of pinyin accuracy as a main effect and interactions with the significant main effects of linguistic characteristics and task. Number of letters in the spelling of the first vowel had to be removed in all analyses as its inclusion in the model caused an unusually high standard error due to a much lower number of first vowels spelled with more than one letter (7.6%) than first vowels spelled with one letter.

Proportion of Doubling: All endings.

Pinyin accuracy. Those with higher pinyin accuracy tend to double medial consonants more overall, but this effect was only marginal (p = .056). A new finding of interest is the significant interaction between vowel length and pinyin accuracy (p < .05), which showed that those who were better at pinyin were more influenced by vowel length when doubling medial consonants. That is, the better he or she was at pinyin, the more likely he or she was to double a medial consonant that comes after a short vowel than a long vowel (see Figure 9). There was a significant three-way interaction with pinyin accuracy, vowel length, and task (p < .05). The interaction between pinyin accuracy and vowel length, where those who were better at pinyin were more likely to double a medial consonant that comes after a short vowel than a long vowel, was more pronounced in the choice task than the free task (see Figure 10).

Another new finding of interest is the significant interaction between pinyin accuracy and bigram frequency (p < .001). Those who had higher pinyin accuracy had a stronger effect of bigram frequency than those with lower pinyin accuracy (see Figure 11). The three-way interaction between bigram frequency, pinyin accuracy, and task was not significant.

It is important to note that Mandarin vocabulary score replaced pinyin accuracy in a separate analysis and did not significantly predict proportion of doubling or significantly interact with any linguistic characteristic or task.

Proportion of Doubling: Only endings as defined by experimenter. To see whether the lack of an ending effect was due to participants not spelling the endings as intended by the experimenter, we removed endings that were not spelled as ‹ib›, ‹ip›, ‹ig›, ‹ic›, ‹id›, and ‹it› in the free task. A total of 286 trials were removed and 3686 trials remained in both tasks. We then ran the same multilevel model as in the previous analysis and found similar effects but still no effect of ending.

Nonword Spelling Accuracy. Here, we asked whether the accuracy of nonword spelling in the free spelling task was predicted by age, sex, spelling ability, and pinyin accuracy. In addition to the main effect of each variable, the model included a two-way interaction between the spelling ability and pinyin accuracy. There were significant main effects of spelling ability ($p < .001$) and pinyin accuracy ($p < .01$). Table 7 shows the nonword spelling accuracy for Mandarin-speaking participants at high and low WRAT scores as well as at high and low pinyin accuracy. As previously mentioned, the effect of spelling ability was expected in that the better the spellers performed on the WRAT, the better they were at spelling nonwords. The effect of pinyin accuracy was also expected in that being better at pinyin also predicted higher accuracy in spelling nonwords. Interestingly, there was no significant interaction between spelling ability and pinyin accuracy in predicting nonword spelling accuracy, thus indicating that the main effects of spelling ability and pinyin accuracy are in the same direction.

When Mandarin vocabulary score replaced pinyin accuracy in a separate analysis, it did not significantly predict nonword spelling accuracy.

Discussion

The aim of Study 2 was to follow up on the statistical patterns in relation to doubling in English found in Study 1 and collect behavioral data on doubling. We wanted to see if people were sensitive to the statistical patterns found in Study 1 and if this sensitivity depended on context. We also explored other individual differences that might influence spelling such as language background, spelling ability, and the type of spelling task one is asked to do.

People seemed to be influenced by statistical patterns when doubling but these influences were moderated by other factors such as context, individual differences such as spelling ability and language background, and task. First, people were influenced by doubling patterns involving the medial consonant itself. People doubled more with medial consonants of higher bigram frequency, which is similar to the statistical pattern that we found in the vocabulary. They did so especially after short vowels in the free spelling task. The bigram frequency effect was also moderated by pinyin accuracy in that those who were better at pinyin were more influenced by bigram frequency in doubling medial consonants. It is unclear as to why such an effect was found. More can be said only after a replication of such an interaction as well as further understanding about the characteristics of pinyin that can make one more sensitive to the statistical patterns of doubling in medial consonants.

Second, to investigate whether people are influenced by context-sensitive or context-free probabilities in doubling, we looked at whether people were influenced by preceding and following context. Indeed, people were influenced by preceding context such as the number of letters that are spelled in the preceding vowel and preceding vowel length. Similar to the statistical pattern we found in the vocabulary, people hardly ever doubled if they spelled the preceding vowel with more than one letter, such as ‹graivit› for /graɪvɪt/ instead of ‹grivit›. Looking at vowel length, people tended to double more after a short vowel, which is similar to the statistical pattern we found in the vocabulary. This effect was moderated by spelling ability, pinyin accuracy (better spellers and those who were better at pinyin were more influenced by vowel length), and language background. With regards to the interaction of the vowel length effect with spelling ability, this was similar to what was found by Treiman and Kessler (2006) in that better spellers were more sensitive to context in spelling. With regards to the interaction of the vowel length effect with pinyin accuracy, this was an unexpected finding, but it might be that pinyin knowledge allows for the development of phonological knowledge in spelling, which has been found to correlate with spelling ability in previous studies, and thus, making one more sensitive to phonological patterns in doubling.

With regards to the interaction of the vowel length effect and language background, we found that both dialect and the language one was exposed to other than English affected one's sensitivity to the pattern of vowel length in spelling. In terms of dialect differences, Americans had a larger vowel length effect than Singaporeans, possibly because, as mentioned earlier, vowel length is not as differentiated in Singaporean English, therefore the Singaporeans were not as sensitive to the pattern of vowel length when doubling. Looking at the language one is exposed to other than English, as we expected, the Tamils had the smallest vowel length effect, as they possibly associated consonant doubling with the lengthening of doubled consonants in Tamil instead of the preceding vowel. This is further corroborated by the finding that the Tamils double medial consonants the least across all language groups. As there is no doubling of consonants in Malay and Mandarin that could influence one's doubling in new words, it is not surprising that both the Malay and Mandarin groups have similar vowel length effects that are larger than the Tamils'. These findings augment our understanding of the role of language background in spelling and support the idea that the use of one writing system can affect the use of another writing system, depending on their linguistic characteristics.

Although people seemed to be influenced by preceding context, people were not sensitive to the ending when doubling. This seems to suggest that people were more affected by the preceding context and the consonant itself. There are two possible interpretations of this lack of ending effect. First, the results may support a left to right (serial) view of spelling as postulated by Simon and Simon (1973), where people's choice of a single or doubled medial consonant is affected by what comes before but not by what is to come after. Second, it might be possible that there is actually an effect of ending (people double less with Latinate endings) but the effect was not detected due to the rare non-Latinate endings that were used for comparison. These endings were chosen mainly because they could not be interpreted as a separate morpheme and that they were similar to Latinate endings. However, as these endings are rare in English, people might not be sure whether they should double or not double the medial consonant before these endings, and they ended up doing the simplest and statistically most frequent action, i.e., not doubling. In this view, experience with a particular ending is important in determining what one does when one gets a new item with that ending, but if one has not had much experience with a particular ending, one resorts to something simple and basic.

This possible interpretation suggests that one would perhaps see more doubling of the medial consonant if one is presented with a more common ending that usually has doubling before it and after a short vowel. This would speak against a serial view of spelling in showing that what is to come does influence the spelling of the consonant and support previous findings that showed an effect of following context on spelling (Treiman & Kessler, 2006). What is clear, however, is that people have not internalized a doubling pattern that will generalize across all endings.

Despite the effects that we found showing the influence of preceding context on doubling, there is overall a low rate of doubling in the experiment in both tasks, and especially in the free spelling task. People have a strong preference for single medial consonants when they have to produce their own spellings, corroborating the findings of Pollock and Zamora (1983), Yannakoudakis & Fawthrop (1983), Bebout (1985), James et al. (1993), Haggan (1981), and Deacon et al. (2011). Even when they were presented with a doubled medial consonant in the choice task, they preferred single medial consonants more than 50% of the time. Results speak against the idea that people rely strongly on the kind of rule they are sometimes taught in spelling instruction: double a consonant after a short vowel. These results also support previous findings that people are not applying these statistical patterns to the same extent as to what was found in the language (Treiman & Kessler, 2006; Kessler, 2009); despite being experienced readers and spellers, their behavioral patterns in doubling do not match or even come close to the context-sensitive statistical patterns found in the vocabulary. Rather, people's use of context is constrained by their knowledge of typical context-free spellings (Perry &

Ziegler, 2004; Treiman Bowey, & Bourassa, 2002), which is in this case, the more frequent single medial consonant (63% of the time in disyllables). This supports theories that postulate the use of context-free probabilities in spelling. Nonetheless, better spellers were more sensitive to context as their behavioral patterns in doubling came closer to the context-sensitive statistical patterns found in the vocabulary; this supports previous findings by Treiman and Kessler (2006). More clearly needs to be done to understand the role of context in spelling.

We also explored individual differences in spelling in terms of nonword spelling accuracy. Results showed that both spelling ability and language background have an effect on being able to spell new words. First, better spellers can spell new words more accurately perhaps because they are able to take advantage of contexts when spelling (Treiman & Kessler, 2006), but also because they are better at applying phoneme-to-grapheme correspondences, an important skill in spelling new words, as found in previous studies.

Second, dialect differences in English resulted in Singaporeans spelling nonwords less accurately than the Americans despite having comparable spelling ability. The common mispronunciation of certain consonants such as /d/ for /θ/ and the conflation of vowels lead to certain consonants and vowels sounding more similar in Singaporean English than in American English such as /i/ and as well as /ɛ/ and /æ/. All this seemed to have a negative impact on the Singaporeans when they were given these specific phonemes to spell in a new word. The finding that Singaporeans found it the hardest to spell the vowel /æ/ and /ɪ/ as compared to other vowels replicated Brown's (1988) finding. These errors speak to the important role of phonology in spelling and how dialect differences can impact spelling accuracy, resulting in certain kinds of errors specific to that dialect, as found in previous studies.

Third, even within the Singaporeans, there were differences in nonword spelling accuracy that can be attributed to language background, despite them undergoing the same English language instruction and having similar exposure to English. The exposure to a logographic orthography with no phoneme-to-grapheme correspondences has a negative effect on nonword spelling accuracy, as shown by the Mandarin group which spelled nonwords least accurately, and greater knowledge of *pinyin*, the alphabetic form of Mandarin with phoneme-to-grapheme correspondences, was associated with higher nonword spelling accuracy. It is important to note that Mandarin vocabulary score did not predict nonword spelling accuracy. The Malay and Tamil groups, which are exposed to orthographies that have a one-to-one phoneme-to-grapheme correspondences, performed much better than the Mandarin group in spelling new words. It is noteworthy that the Tamil group had a much lower average score than the Mandarin group on the standardized WRAT spelling task but they still had a higher nonword spelling accuracy. This supports the idea that the orthography of an individual's second language can either hurt or help the individual's ability to spell new words. Being exposed to an alphabetic orthography, where there are direct phoneme-to-grapheme correspondences like Malay and Tamil, helps in the development of knowledge of phonological processing and letter-sound correspondences in that sounds map onto letters. This is not the case for non-alphabetic orthographies with no phoneme-to-grapheme correspondences like Mandarin.

Similar to what we found on the influences of language background in doubling, these findings further augment our understanding of the role of language background in spelling and support the idea that the use of one writing system can affect the use of another writing system, depending on their linguistic characteristics.

General Discussion

The general issues addressed by these two studies were sensitivity to statistical patterns in the English language and the role of that sensitivity in spelling. Despite English having a reputation for being chaotic and extremely difficult to spell, there are still statistical regularities in English spelling (e.g. Kessler, 2003), and these context-sensitive statistics do affect decisions in spelling (e.g. Kessler, 2009). We have proposed that although the doubling of medial consonants in English is difficult, it follows statistical patterns and that sensitivity to these patterns plays a role in doubling. We have also proposed that these patterns are not only context-sensitive, but also that the impact of these patterns on doubling are moderated by the position of the context, spelling ability, and language background. To begin to investigate this proposal, we looked at patterns pertaining to the doubling of medial consonants in the English language.

In Study 1, we tabulated the statistical patterns in relation to the doubling of medial consonants in English. We focused on four patterns: preceding vowel length, number of letters spelled in preceding vowel, medial consonant, and ending type. In Study 2, we collected behavioral data to see whether people were sensitive to these statistical patterns when doubling or whether they were guessing randomly. If they were being influenced by statistical patterns in the language, we then investigated if the influence of these statistical patterns depend on the positional context of these patterns, i.e., preceding or following. We also explored whether such sensitivity to statistical patterns was moderated by other factors such as spelling ability, language background, and the kind of spelling task one was asked to do. Together, the results of both studies suggest that people are sensitive to the statistical patterns in the language and these statistical patterns influence the way they spell. Various factors moderate the sensitivity to these statistical patterns—spelling ability, language background, and task.

Our findings on the influence of language background on spelling (above and beyond spelling ability) have important implications, given the diversity of English-language speakers and learners in the world. Educators must appreciate the linguistic profile of the environment their students are in as well as the languages that they are exposed to so as to tailor their teaching methods accordingly. For example, students who are exposed to a non-alphabetic orthography may need more help in learning phoneme-to-grapheme correspondences and how to cope with spelling a word that they have never seen before. On another note, researchers need to appreciate the idea that the influence of certain processes in spelling is not universal; rather, it depends on one's language background.

Although we found some evidence for the influence of context in doubling, it is still inconclusive as we also found evidence for the influence of context-free probabilities in doubling. Either way, these experienced readers and spellers are still not applying statistical patterns to their spellings in proportions that are close to what was found in the vocabulary. Perhaps humans are not optimal statistical learners (which explains our abundant spelling errors) or perhaps spelling is a much more complex process that cannot be fully explained through the theory of statistical learning. More needs to be done to provide a holistic understanding of the factors influencing the processes that underlie spelling, so that people can be better spellers.

There are still further improvements that could have been made to both studies. In Study 1, the statistical patterns derived from the database were based on General American pronunciation. To accommodate dialect differences, Singaporean pronunciation and frequency counts could be used to tabulate the statistical patterns more accurately for the Singaporeans. In Study 2, a listening pinyin task in which participants listen to Chinese characters and write down their pinyin might be a more valid test of pinyin ability. Not only that, the non-Latinate endings that were used in this study were rare in English. Future studies should seek to include a listening pinyin task and include more common endings to test for an effect of ending in doubling medial consonants. Furthermore, having studied the influence of these patterns and their moderating factors on adult spelling, it would be important to look at these influences on spelling in children so as to understand the development of these influences.

The present studies highlight the need for a fuller understanding of the factors contributing to spelling. Here, we only looked at some factors influencing doubling in the spelling of nonwords. There is still much more to learn about the development and influence of statistical learning in the spelling of real words. For example, we need to look at other factors in doubling or spelling such as morphology. Tracking the development of these other factors is necessary to create an integrated picture of the competent speller. The first step in this process is appreciating the interactive nature of these factors that underlie spelling.

References

Albrow, K. H. (1972). *The English writing system: Notes towards a description*. Longman for the Schools Council.

Balasubramanian, T. (1982). Intervocalic double nasal and lateral consonant articulations in Tamil. *Journal of Phonetics, 10*, 99-104.

Barry, C. (1992). Interactions between lexical and assembled spelling (in English, Italian and Welsh). *Psychology, Spelling and Education* (pp. 71–86). Multilingual Matters Ltd.

Barry, C., & Seymour, P. H. K. (1988). Lexical priming and sound-to-spelling contingency effects in nonword spelling. *The Quarterly Journal of Experimental Psychology*, *40*(1), 5-40.

Bates, D., Maechler, M., & Bolker, B. (2011). lme4: Linear mixed-effects models using S4 classes. R package version 0.999375-39. http://CRAN.R-project.org/package=lme4

Bebout, L. (1985). An error analysis of misspellings made by learners of English as a first and as a second language. *Journal of Psycholinguistic Research*, *14*(6), 569-593.

Bernstein, S. E., & Treiman, R. (2001). Learning a novel grapheme: Effects of positional and phonemic context on children's spelling. *Journal of Experimental Child Psychology*, *79*(1), 56-77.

Bourassa, D., & Treiman, R. (2007). Linguistic factors in spelling development. *Encyclopedia of Language and Literacy Development* (pp. 1-8). London, ON: Canadian Language and Literacy Research Network. Retrieved August, 2011 from http://www.literacyencyclopedia.ca/pdfs/topic.php?topId=228.

Brown, A. (1988). A Singaporean corpus of misspellings: Analysis and implications. *Journal of the Simplified Spelling Society, 3,* 4-10.

Bruck, M., & Waters, G. (1988). An analysis of the spelling errors of children who differ in their reading and spelling skills. *Applied Psycholinguistics, 9*(1), 77-92.

Buchwald, A., & Rapp, B. (2004). Rethinking the graphemic buffer? *Brain and Language, 91*(1), 100-101.

Bullinaria, J. A. (1997). Modeling reading, spelling, and past tense learning with artificial neural networks. *Brain and Language, 59*(2), 236-266.

Burt, J. S., & Shrubsole, C. S. (2000). Processing of phonological representations and adult spelling proficiency. *Australian Journal of Psychology, 52*(2), 100-109.

Cahen, L. S., Craun, M. J., & Johnson, S. K. (1971). Spelling difficulty. A survey of the research. *Review of Educational Research, 41*(4), 281-301.

Cantrell, R. J. (2001). Exploring the relationship between dialect and spelling for specific vocalic features in Appalachian first-grade children. *Linguistics and Education, 12*(1), 1-23.

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

Carlisle, J. F. (1988). Knowledge of derivational morphology and spelling ability in fourth, sixth, and eighth graders. *Applied Psycholinguistics, 9*(3), 247-266.

Carney, E. (1997). *English spelling*: Psychology Press.

Cassar, M., & Treiman, R. (1997). The beginnings of orthographic knowledge: Children's knowledge of double letters in words. *Journal of Educational Psychology, 89*(4), 631.

Chen, X., Ku, Y. M., Koyama, E., Anderson, R. C., & Li, W. (2008). Development of phonological awareness in bilingual Chinese children. *Journal of Psycholinguistic Research, 37*(6), 405-418.

Cheung, H., & Chen, H. C. (2004). Early orthographic experience modifies both phonological awareness and on-line speech processing. *Language and Cognitive Processes, 19*(1), 1-28.

Cheung, H., Chen, H. C., Lai, C. Y., Wong, O. C., & Hills, M. (2001). The development of phonological awareness: Effects of spoken language experience and orthography. *Cognition, 81*(3), 227-241.

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

Cronnell, B. (1979). Black English and spelling. *Research in the Teaching of English, 13*(1), 81-90.

Crystal, D., & Wang, F. (1997). *English as a global language* (Vol. 2). Cambridge: Cambridge University Press.

Cummins, J. (1979). Linguistic interdependence and the educational development of bilingual children. *Review of Educational Research, 49*(2), 222.

Deacon, S., Leblanc, D., & Sabourin, C. (2011). When cues collide: children's sensitivity to letter-and meaning-patterns in spelling words in English. *Journal of Child Language, 38*(4), 809-827.

Deacon, S. H., Conrad, N., & Pacton, S. (2008). A statistical learning perspective on children's learning about graphotactic and morphological regularities in spelling. *Canadian Psychology/Psychologie Canadienne, 49*(2), 118.

Deterding, D. (2007). *Singapore English*. Edinburgh: Edinburgh University Press.

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.

Fischer, F. W., Shankweiler, D., & Liberman, I. Y. (1985). Spelling proficiency and sensitivity to word structure. *Journal of Memory and Language, 24*(4), 423-441.

Fowler, H. W. (1937). *A Dictionary of Modern English Usage*. USA: Oxford University Press.

Frith, U. (Ed.). (1980). *Cognitive processes in spelling*. NY: Academic Press Inc.

Geva, E., Wade-Woolley, L., & Shany, M. (1993). The concurrent development of spelling and decoding in two different orthographies. *Journal of Literacy Research, 25*(4), 383-406.

Goldberg, A. M., Buchwald, A., Kochaniak, J., & Rapp, B. (2005). Investigating serial order and graphemic representations in spelling: A simple recurrent network Simulation. Retrieved October, 2011 from http://www.psych.unito.it/csc/cogsci05/frame/poster/1/ma332-goldberg.pdf

Haggan, M. (1991). Spelling errors in native Arabic-speaking English majors: A comparison between remedial students and fourth year students. *System*, *19*(1-2), 45-61.

Hayes, H., Treiman, R., & Kessler, B. (2006). Children use vowels to help them spell consonants. *Journal of Experimental Child Psychology, 94*(1), 27-42.

Ho, C. S. H., & Bryant, P. (1997). Development of phonological awareness of Chinese children in Hong Kong. *Journal of Psycholinguistic Research, 26*(1), 109-126.

Ho, C. S. H., & Bryant, P. (1997). Phonological skills are important in learning to read Chinese. *Developmental Psychology*, *33*(6), 946.

Holm, A., & Dodd, B. (1996). The effect of first written language on the acquisition of English literacy. *Cognition*, *59*(2), 119-147.

Hong, S. C., & Chen, S. H. (2011). Roles of position, stress, and proficiency in L2 children’s spelling: A developmental perspective. *Reading and Writing*, *24*(3), 355-385.

Ibrahim, M. H. (1978). Patterns in spelling errors. *ELT Journal, 32*(3), 207-212.

James, C., & Klein, K. (1994). Foreign language learners' spelling and proofreading strategies. *Papers and Studies in Contrastive Linguistics, 29*, 31-46.

James, C., Scholfield, P., Garrett, P., & Griffiths, Y. (1993). Welsh bilinguals' English spelling: An error analysis. *Journal of Multilingual and Multicultural Development, 14*(4), 287-306.

Jensen, A. R. (1962). Spelling errors and the serial-position effect. *Journal of Educational Psychology, 53*(3), 105.

Jorm, A. (1977). Children's spelling processes revealed by transitional error probabilities. *Australian Journal of Psychology, 29*(2), 125-130.

Juul, H., & Sigurdsson, B. (2005). Orthography as a handicap? A direct comparison of spelling acquisition in Danish and Icelandic. *Scandinavian Journal of Psychology, 46*(3), 263-272.

Kandel, S., Alvarez, C. J., & Vallée, N. (2006). Syllables as processing units in handwriting production. *Journal of Experimental Psychology: Human Perception and Performance, 32*(1), 18.

Kemp, N., & Bryant, P. (2003). Do Beez Buzz? Rule–Based and Frequency–Based Knowledge in Learning to Spell Plural–s. *Child Development, 74*(1), 63-74.

Kessler, B. (2003). Is English spelling chaotic? Misconceptions concerning its irregularity. *Reading Psychology, 24*(3-4), 267-289.

Kessler, B. (2009). Statistical learning of conditional orthographic correspondences. *Writing Systems Research, 1*(1), 19-34.

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

Kohler, C. T., Bahr, R. H., Silliman, E. R., Bryant, J. B., Apel, K., & Wilkinson, L. C. (2007). African American English dialect and performance on nonword spelling and phonemic awareness tasks. *American Journal of Speech-Language Pathology, 16*(2), 157.

Kooi, B. Y., Schutz, R. E., & Baker, R. L. (1965). Spelling errors and the serial-position effect. *Journal of Educational Psychology, 56*(6), 334.

Kreiner, D. S. (1992). Reaction time measures of spelling: Testing a two-strategy model of skilled spelling. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18*(4), 765.

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.

Kreiner, D. S., Price, R. Z., & Gross, A. M. (2008). Linguistic boundaries as predictors of the time between letters in oral and typed spellings. *The Journal of General Psychology, 135*(2), 117-132.

Lecours, A. R. (1966). Serial order in writing—A study of misspelled words in "developmental dysgraphia". *Neuropsychologia*, *4*(3), 221-241.

Lefly, D. L., & Pennington, B. F. (1991). Spelling errors and reading fluency in compensated adult dyslexics. *Annals of Dyslexia, 41*(1), 141-162.

Lehtonen, A., & Bryant, P. (2005). Doublet challenge: Form comes before function in children's understanding of their orthography. *Developmental Science*, *8*(3), 211-217.

Lian, H. M. (1997). This is Chickenfits: The effect of pronunciation on writing. *RELC Journal*, *28*(2), 22-34.

Lim, L. (2004). *Singapore English: A grammatical description* (Vol. 33). Amsterdam: John Benjamins Publishing Company.

Liow, S. J. R., & Lau, L. H. S. (2006). The development of bilingual children's early spelling in English. *Journal of Educational Psychology, 98*(4), 868.

Marcus, G. F., & Berent, I. (2003). Are there limits to statistical learning? *Science*, *300*(5616), 53.

McBride-Chang, C., Bialystok, E., Chong, K. K. Y., & Li, Y. (2004). Levels of phonological awareness in three cultures. *Journal of Experimental Child*

Psychology, 89(2), 93-111.

McBride-Chang, C., Tong, X., Shu, H., Wong, A. M. Y., Leung, K., & Tardif, T. (2008). Syllable, phoneme, and tone: Psycholinguistic units in early Chinese and English word recognition. *Scientific Studies of Reading, 12*(2), 171-194.

Medler, D. A., & Binder, J. R. (2005). MCWord: An on-line orthographic database of the English language. Retrieved October, 2010, from www.neuro.mcw.edu/mcword/

Mitton, R. (1987). Computer-Readable Corpora of Spelling Errors. *TESOL Quarterly, 21*(1), 153-155.

Pacton, S., Fayol, M., & Perruchet, P. (2005). Children's implicit learning of graphotactic and morphological regularities. *Child Development, 76*(2), 324-339.

Pacton, S., Perruchet, P., Fayol, M., & Cleeremans, A. (2001). Implicit learning out of the lab: The case of orthographic regularities. *Journal of Experimental Psychology: General, 130*(3), 401.

Perry, C., & Ziegler, J. C. (2004). Beyond the two-strategy model of skilled spelling: Effects of consistency, grain size, and orthographic redundancy. *Quarterly Journal of Experimental Psychology: Section A, 57*(2), 325-356.

Perry, C., Ziegler, J. C., & Coltheart, M. (2002). How predictable is spelling? Developing and testing metrics of phoneme-grapheme contingency. *The Quarterly Journal of Experimental Psychology: Section A, 55*(3), 897-915.

Pollo, T. C., Treiman, R., & Kessler, B. (2008). Three perspectives on spelling development. In E. L. Grigorenko, & A. J. Naples (Eds.), *Single-word Reading: Behavioral and Biological Perspectives* (pp. 175-189). New York: Lawrence

Erlbaum Associates.

Pollock, J. J., & Zamora, A. (1983). Collection and characterization of spelling errors in scientific and scholarly text. *Journal of the American Society for Information Science, 34*(1), 51-58.

Read, C., Zhang, Y., Nie, H., & Ding, B. (1987). The ability to manipulate speech sounds depends on knowing alphabetic writing. *The Onset of Literacy: Cognitive Processes in Reading Acquisition*, 31-44.

Rickard Liow, S. J., & Poon, K. K. L. (1998). Phonological awareness in multilingual Chinese children. *Applied Psycholinguistics*, *19*(03), 339-362.

Rickard Liow, S. J., Tng, S. K., & Lee, C. L. (1999). Chinese characters: Semantic and phonetic regularity norms for China, Singapore, and Taiwan. *Behavior Research Methods, 31*(1), 155-177.

Saffran, J. R. (2002). Constraints on statistical language learning. *Journal of Memory and Language*, *47*(1), 172-196.

Saffran, J. R. (2003). Statistical language learning: Mechanisms and constraints. *Current Directions in Psychological Science, 12*(4), 110-114.

San Francisco, A. R., Mo, E., Carlo, M., August, D., & Snow, C. (2006). The influences of language of literacy instruction and vocabulary on the spelling of Spanish–English bilinguals. *Reading and Writing, 19*(6), 627-642.

Schiffman, H. F. (2007). Tamil language policy in Singapore: The role of implementation. *Language, Capital, Culture: Critical Studies of Language in Education in Singapore*, 209–226.

Schiller, N. O., Greenhall, J. A., Shelton, J. R., & Caramazza, A. (2001). Serial order effects in spelling errors: Evidence from two dysgraphic patients. *Neurocase, 7*(1), 1-14.

Seidenberg, M. S. (2005). Connectionist models of word reading. *Current Directions in Psychological Science, 14*(5), 238-242.

Seidenberg, M. S., MacDonald, M. C., & Saffran, J. R. (2002). Does grammar start where statistics stop? *Science, 298*(5593), 553.

Share, D. L. (2008). On the Anglocentricities of current reading research and practice: The perils of overreliance on an" outlier" orthography. *Psychological Bulletin, 134*(4), 584.

Shu, H., Peng, H., & McBride Chang, C. (2008). Phonological awareness in young Chinese children. *Developmental Science, 11*(1), 171-181.

Simon, D. P., & Simon, H. A. (1973). Alternative uses of phonemic information in spelling. *Review of Educational Research, 43*(1), 115-137.

Siok, W. T., & Fletcher, P. (2001). The role of phonological awareness and visual-orthographic skills in Chinese reading acquisition. *Developmental Psychology, 37*(6), 886.

Solso, R. L., & Juel, C. L. (1980). Positional frequency and versatility of bigrams for two-through nine-letter English words. *Behavior Research Methods & Instrumentation, 12*(3), 297-343.

Spache, G. (1940). Validity and reliability of the proposed classification of spelling errors. II. *Journal of Educational Psychology, 31*(3), 204.

Sparks, R., Patton, J., Ganschow, L., & Humbach, N. (2009). Long-Term Crosslinguistic Transfer of Skills From L1 to L2. *Language Learning, 59*(1), 203-243.

Sparks, R. L., Patton, J., Ganschow, L., Humbach, N., & Javorsky, J. (2008). Early first-language reading and spelling skills predict later second-language reading and spelling skills. *Journal of Educational Psychology, 100*(1), 162.

Steffler, D. J. (2004). An investigation of Grade 5 children's knowledge of the doubling rule in spelling. *Journal of Research in Reading, 27*(3), 248-264.

Terry, N. P. (2006). Relations between dialect variation, grammar, and early spelling skills. *Reading and Writing, 19*(9), 907-931.

Tesdell, L. S. (1984). ESL spelling errors. *TESOL Quarterly, 18*(2), 333-334.

Tham, W. W. P., Rickard Liow, S. J., Rajapakse, J. C., Choong Leong, T., Ng, S. E. S., Lim, W. E. H., et al. (2005). Phonological processing in Chinese-English bilingual biscriptals: An fMRI study. *NeuroImage, 28*(3), 579-587.

Thompson, G. B., Connelly, V., Fletcher-Flinn, C. M., & Hodson, S. J. (2009). The nature of skilled adult reading varies with type of instruction in childhood. *Memory and Cognition, 37*(2), 223-234.

Treiman, R. (1993). *Beginning to spell: A study of first-grade children*. USA: Oxford University Press.

Treiman, R. (1994). Use of consonant letter names in beginning spelling. *Developmental Psychology, 30*(4), 567.

Treiman, R. (2004). Spelling and dialect: Comparisons between speakers of African American vernacular English and White speakers. *Psychonomic Bulletin &*

Review, 11(2), 338-342.

Treiman, R., & Barry, C. (2000). Dialect and authography: Some differences between American and British spellers. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26*(6), 1423.

Treiman, R., Berch, D., Tincoff, R., & Weatherston, S. (1993). Phonology and spelling: The case of syllabic consonants. *Journal of Experimental Child Psychology*.

Treiman, R., Berch, D., & Weatherston, S. (1993). Children's use of phoneme-grapheme correspondences in spelling: Roles of position and stress. *Journal of Educational Psychology, 85*(3), 466.

Treiman, R., Bowey, J. A., & Bourassa, D. (2002). Segmentation of spoken words into syllables by English-speaking children as compared to adults. *Journal of Experimental Child Psychology, 83*(3), 213-238.

Treiman, R., & Danis, C. (1988). Syllabification of intervocalic consonants. *Journal of Memory and Language, 27*(1), 87-104.

Treiman, R., Goswami, U., Tincoff, R., & Leevers, H. (1997). Effects of dialect on American and British children's spelling. *Child Development, 68*(2), 229-245.

Treiman, R., & Kessler, B. (2006). Spelling as statistical learning: Using consonantal context to spell vowels. *Journal of Educational Psychology, 98*(3), 642.

Treiman, R., Kessler, B., & Evans, R. (2007). Anticipatory conditioning of spelling-to-sound translation. *Journal of Memory and Language, 56*(2), 229-245.

Treiman, R., Kessler, B., Knewasser, S., Tincoff, R., & Bowman, M. (2000). English speakers' sensitivity to phonotactic patterns. In M. B. Broe & J. B. Pierrehumbert

(Eds.), *Papers in Laboratory Phonology V: Acquisition and the Lexicon* (pp. 269-282). Cambridge, England: Cambridge University Press

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, 9*(5), 387-406.

Wang, M., & Geva, E. (2003). Spelling performance of Chinese children using English as a second language: Lexical and visual–orthographic processes. *Applied Psycholinguistics*, *24*(01), 1-25.

Wang, M., Perfetti, C. A., & Liu, Y. (2005). Chinese-English biliteracy acquisition: Cross-language and writing system transfer. *Cognition, 97*(1), 67-88.

Waters, G. S., Bruck, M., & Malus-Abramowitz, M. (1988). The role of linguistic and visual information in spelling: A developmental study. *Journal of Experimental Child Psychology, 45*(3), 400-421.

Wing, A. M., & Baddeley, A. D. (1980). Spelling errors in handwriting: A corpus and a distributional analysis. *Cognitive Processes in Spelling*, 251-285.

Worthy, M. J., & Invernizzi, M. (1990). Spelling errors of normal and disabled students on achievement levels one through four: Instructional implications. *Annals of Dyslexia, 40*(1), 138-151.

Yannakoudakis, E., & Fawthrop, D. (1983). The rules of spelling errors. *Information Processing & Management*, *19*(2), 87-99.

Table 1

Statistical Patterns Pertaining to Proportion of Doubling of Medial Consonants in Disyllables (N = 8790)

	Proportion of Doubling	
	By lg(type frequency)	By token
Number of letters in first vowel spelling		
One	0.46	0.49
More than one	0.02	0.03
Vowel length		
Short: /ɑ/, /æ/, /ɛ/, /ɪ/	0.65	0.69
Long: /o/, /e/, /aɪ/, /i/	0.01	0.01
*Latinate endings**		
‹-ic›	0.19	0.21
‹-id›	0.05	0.09
‹-it›	0.10	0.18
*Non-Latinate endings**		
‹-ow›	0.91	0.93
‹-age›	0.60	0.63
‹-is›	0.57	0.47

*Medial consonant preceded by short vowel spelled with one letter

Table 2

Statistical Patterns Pertaining to Proportion of Doubling of Medial Consonants in Disyllables (N = 8790)

	Proportion of Doubling	
Medial Consonant	By lg(type frequency)	By token
b	0.49	0.55
c	0.06	0.09
d	0.32	0.35
f	0.61	0.63
g	0.52	0.56
k	0.02	0.03
l	0.49	0.45
m	0.32	0.36
n	0.31	0.30
p	0.47	0.48
r	0.29	0.26
s	0.39	0.41
t	0.41	0.42
v	0.03	0.04
z	0.32	0.36

Table 3

Mean Individual Characteristics across Groups in Study 2

	Group							
			Singaporean					
	American (*N* = 68)		Mandarin (*N* = 54)		Malay (*N* = 44)		Tamil (*N* = 42)	
Characteristic	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Age	20.66	1.58	22.91	2.22	24.64	4.51	22.48	5.94
WRAT Score	124.43	10.87	121.28	10.91	123.59	9.35	113.02	12.87
Shipley Vocabulary Age	18.54	1.19	17.68	1.38	17.81	1.21	16.55	1.77

Table 4

Bigram Frequency for the CC Bigrams in Study 2

Bigram	Bigram Frequency	lg(Bigram Frequency)
bb	523.30	3.72
ll	21368.91	5.33
mm	2480.73	4.39
nn	2263.76	4.35
pp	4527.34	4.66
vv	2.44	1.39

Table 5

Proportion of Doubling of Medial Consonants as a Function of Vowel Length, Number of Letters in First Vowel Spelling, Task, and Language Group

	Proportion of Doubling
Vowel Length	
Short	0.38
Long	0.13
Number of Letters in First Vowel Spelling	
One	0.28
More Than One	0.00
Task	
Choice	0.41
Free	0.06
Language Group	
American English	0.25
Mandarin	0.29
Malay	0.25
Tamil	0.22

Table 6

Nonword Spelling Accuracy across Language Groups and Spelling Ability

	Accuracy
Language Group	
American English	0.94
Mandarin	0.56
Malay	0.71
Tamil	0.69
Standardized WRAT Score	
High (> 121)	0.80
Low (≤ 121)	0.69

Table 7

Nonword Spelling Accuracy across Spelling Ability and Pinyin Accuracy (for Mandarin Group)

	Accuracy
WRAT Score	
High (> 119.5)	0.63
Low (≤ 119.5)	0.48
Pinyin Accuracy	
High	0.61
Low	0.49

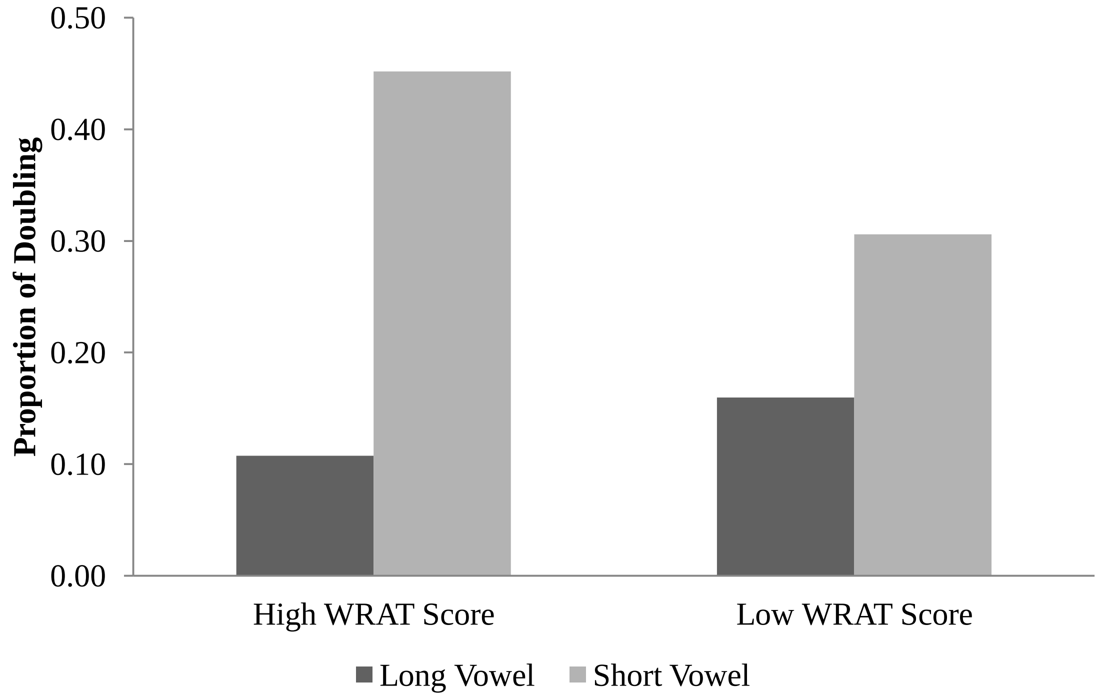

Figure 1. Proportion of doubling of medial consonants for vowel length across spelling ability (for all groups).

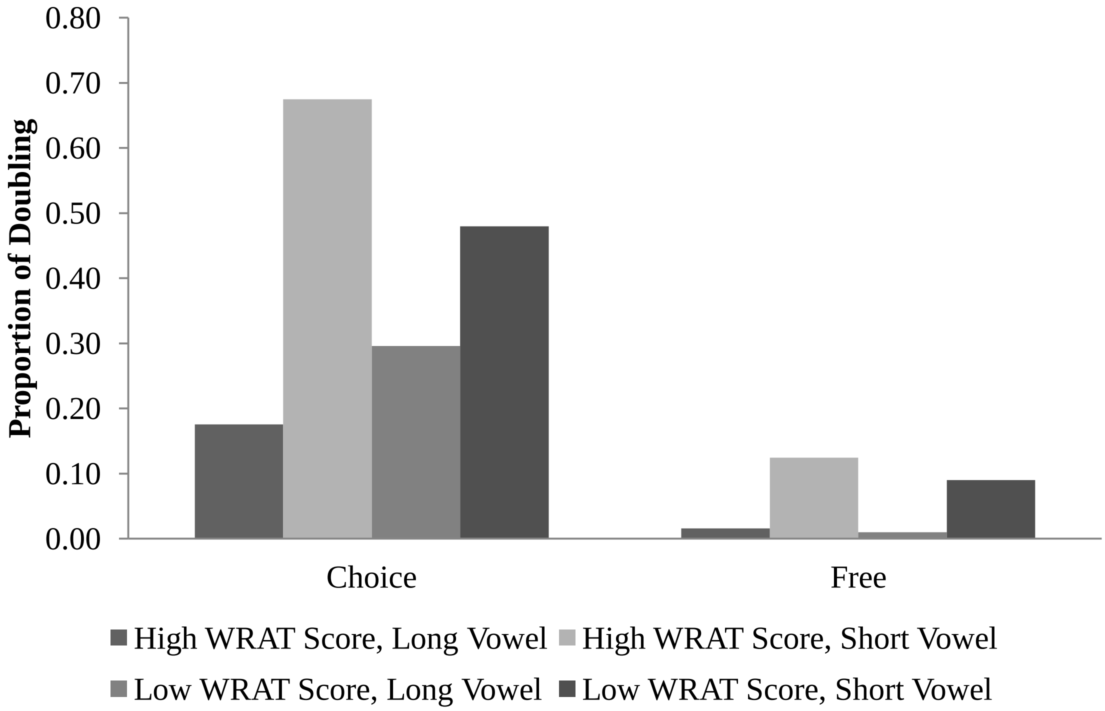

Figure 2. Proportion of doubling of medial consonants for the interaction between spelling ability and vowel length across tasks.

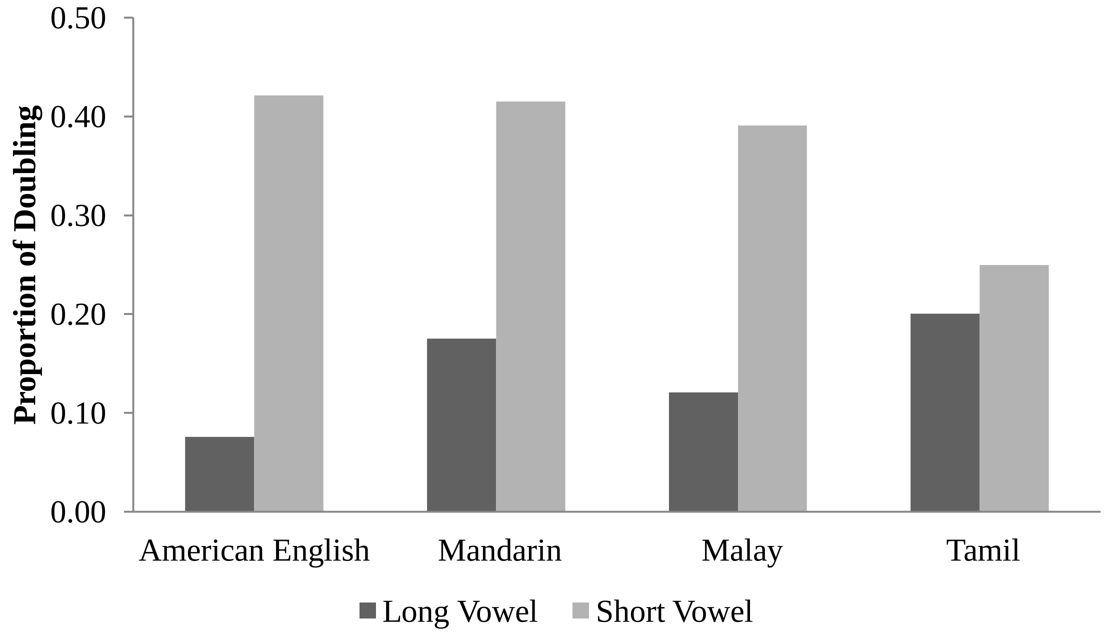

Figure 3. Proportion of doubling of medial consonants for vowel length across language groups.

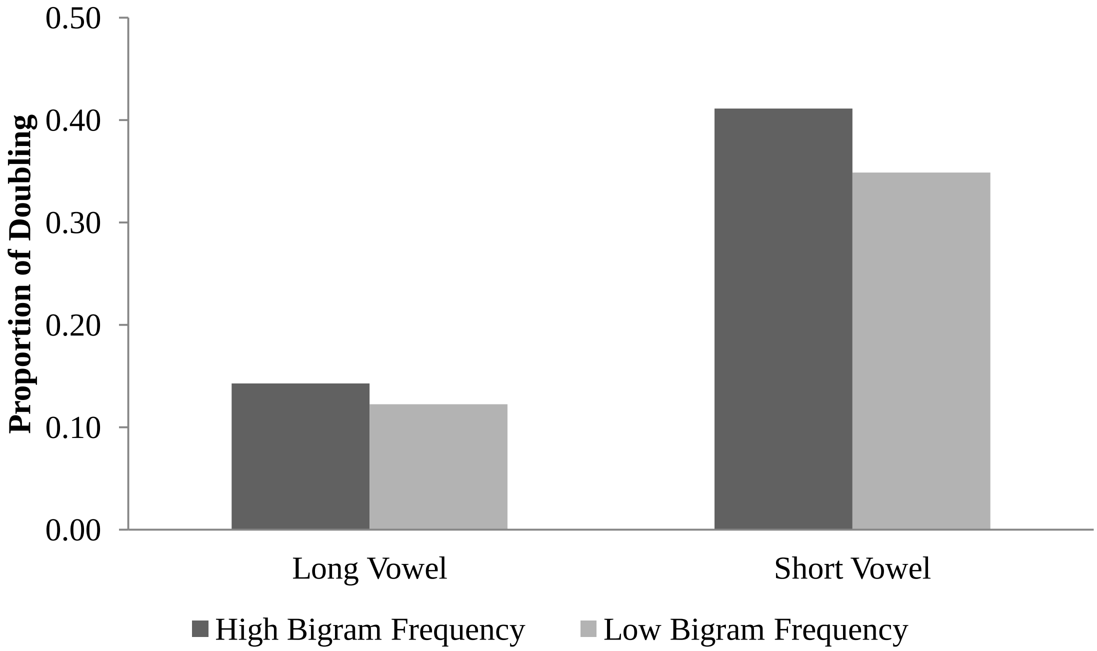

Figure 4. Proportion of doubling of medial consonants for bigram frequency across vowel length.

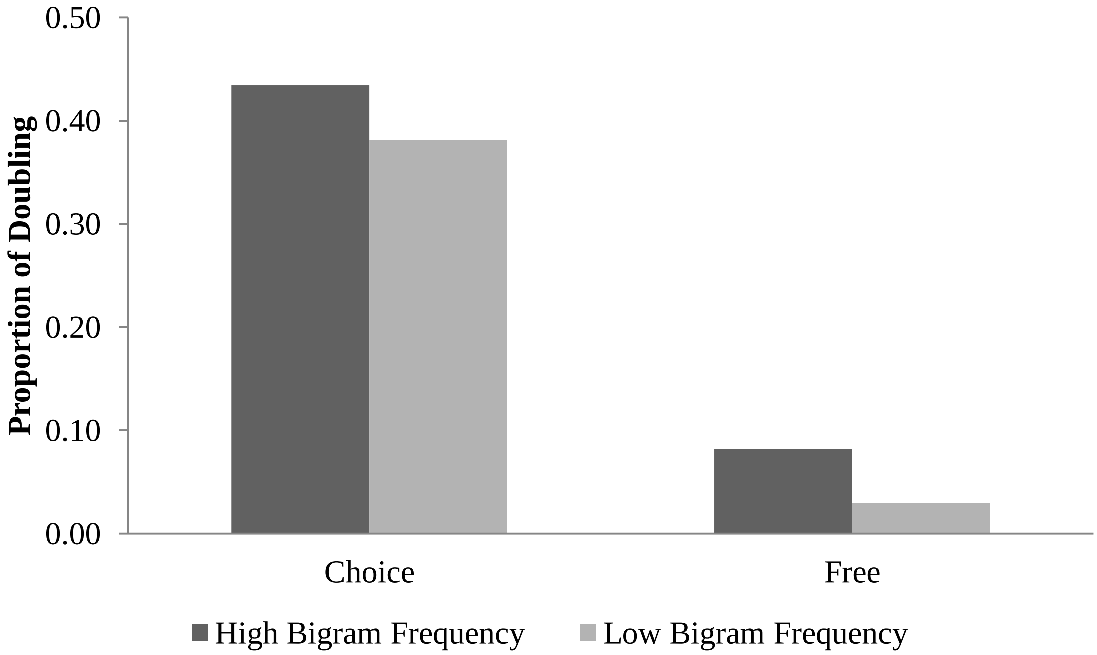

Figure 5. Proportion of doubling of medial consonants for bigram frequency across tasks.

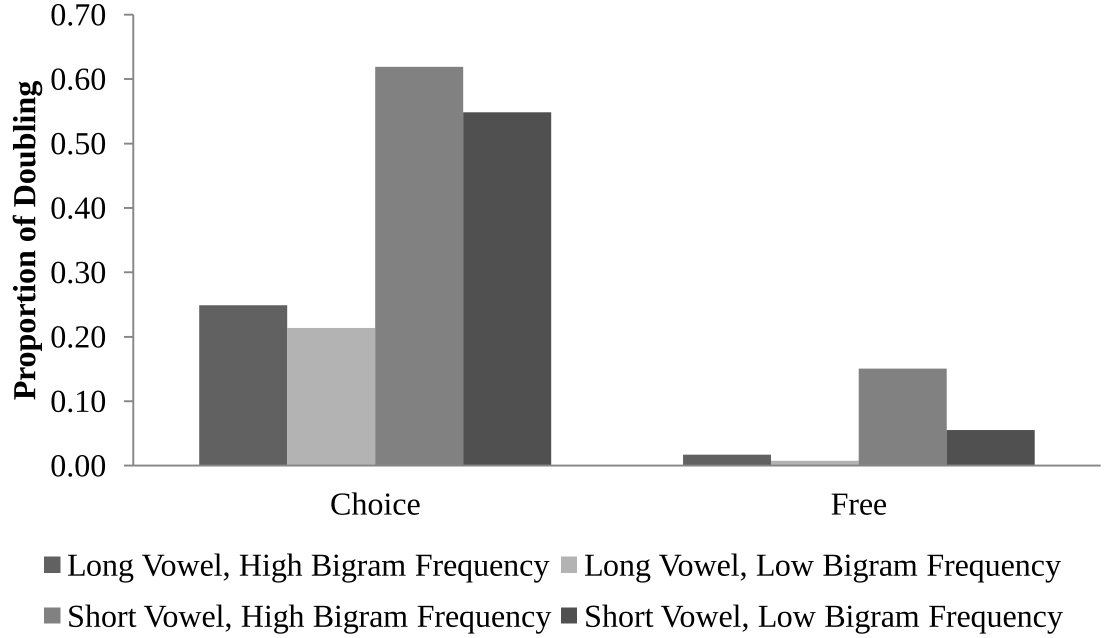

Figure 6. Proportion of doubling of medial consonants for the interaction between bigram frequency and vowel length across tasks.

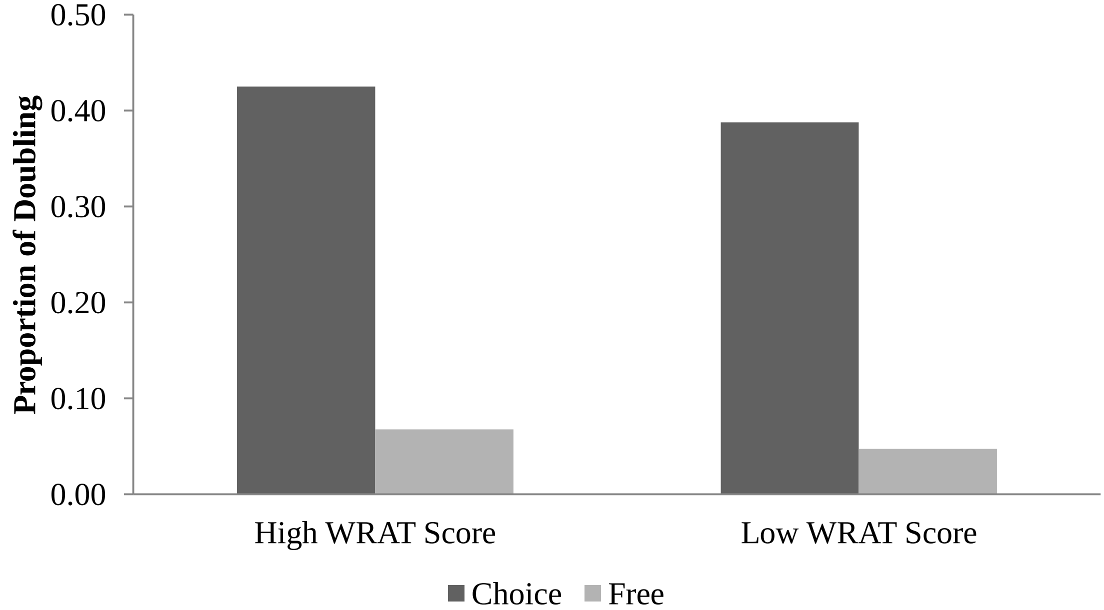

Figure 7. Proportion of doubling of medial consonants in tasks across spelling ability.

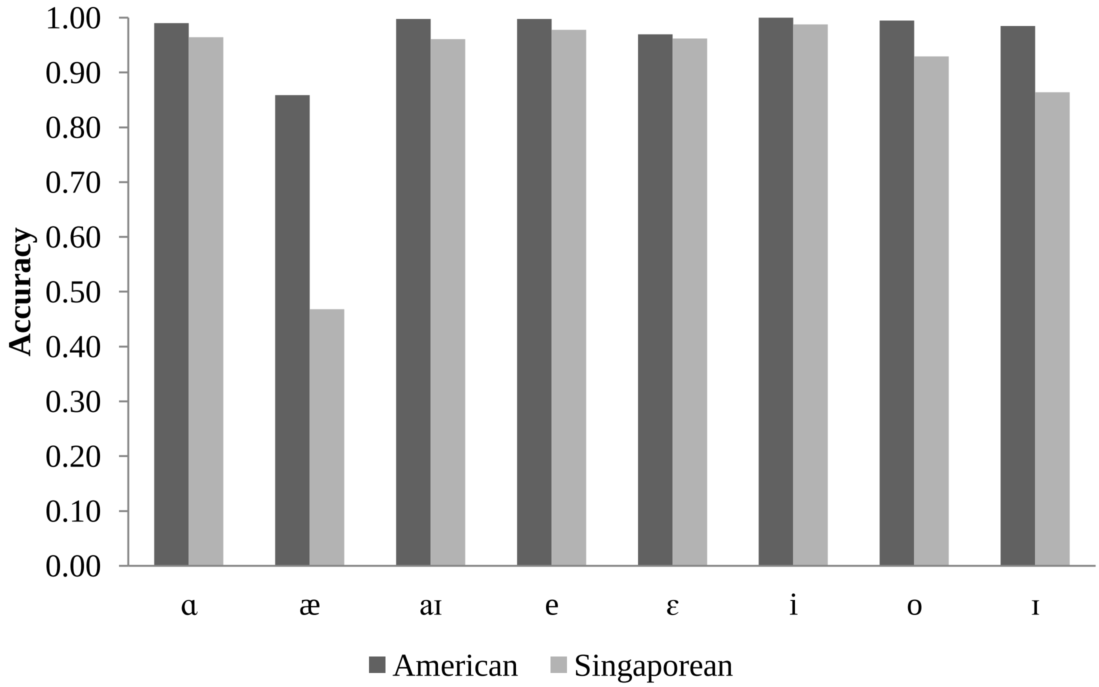

Figure 8. Nonword spelling accuracy across vowels for Americans and Singaporeans.

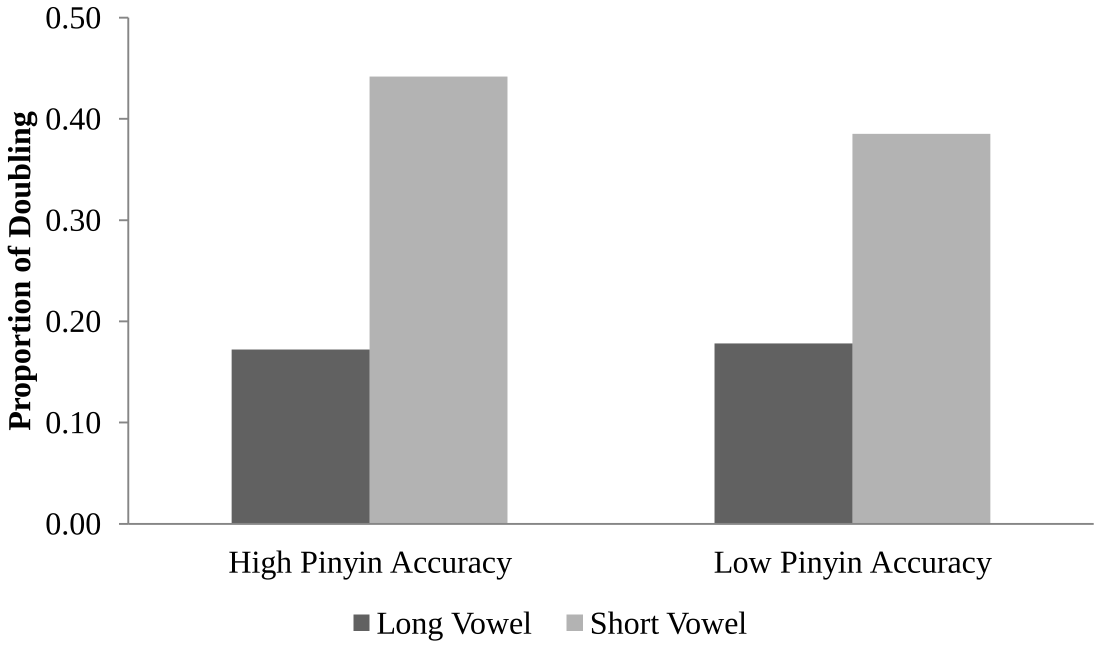

Figure 9. Proportion of doubling of medial consonants for vowel length across pinyin accuracy (for Mandarin group).

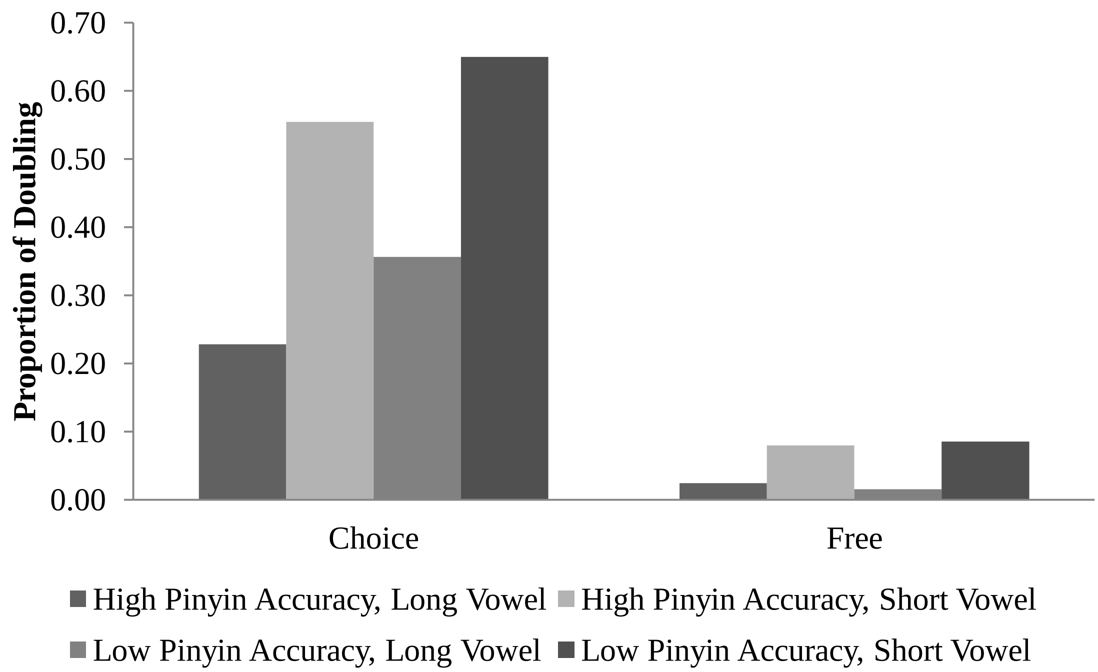

Figure 10. Proportion of doubling of medial consonants for the interaction between vowel length and pinyin accuracy across tasks (for Mandarin group).

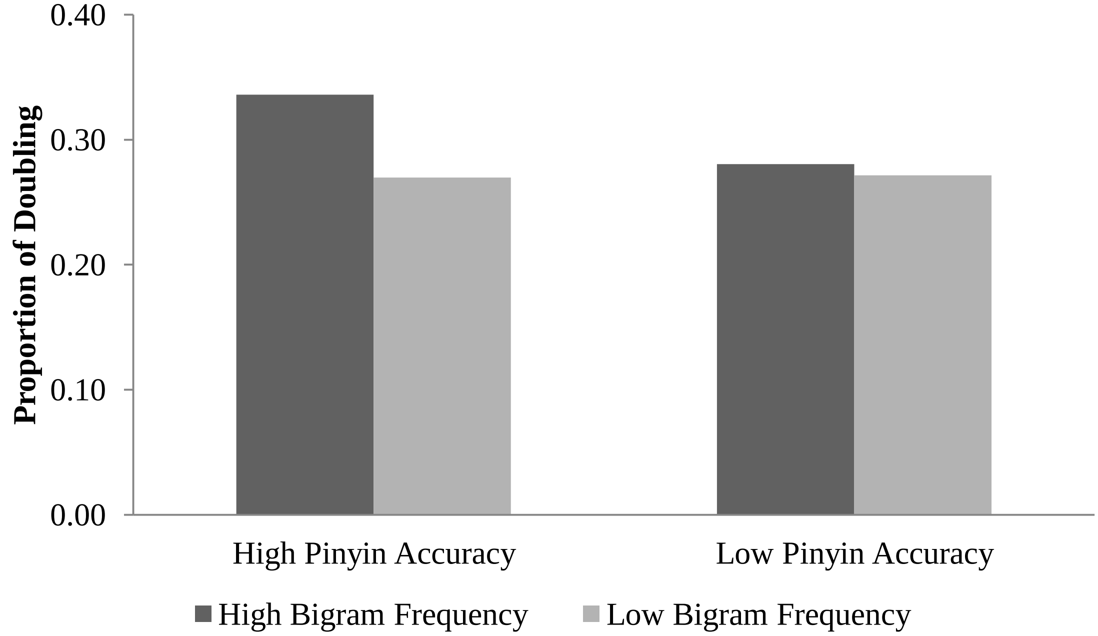

Figure 11. Proportion of doubling of medial consonants for bigram frequency across pinyin accuracy (for Mandarin group).

Appendix A

Stimuli for Choice/Free Spelling Task

Group 1 Short Vowel, Non-Latinate Ending	Group 2 Short Vowel, Latinate Ending	Group 3 Short Vowel, Non-Latinate Ending	Group 4 Long Vowel, Latinate Ending
spræbɪb	spræbɪk	spraɪbɪb	spraɪbɪk
vɑbɪg	vɑbɪd	vobɪg	vobɪd
zɪbɪp	zɪbɪd	gribɪp	gribɪt
zæbɪg	zæbɪk	zibɪb	zibɪk
trɑmɪg	trɑmɪd	blemɪg	blemɪd
drɛmɪp	drɛmɪt	zemɪp	zemɪd
snɪlɪb	snɪlɪk	snaɪlɪb	snaɪlɪk
grɛlɪg	grɛlɪd	zaɪlɪg	zaɪlɪk
snɑlɪp	snɑlɪt	snelɪp	snelɪt
plɛmɪb	plɛmɪk	dromɪb	dromɪk
tʃæmɪg	tʃæmɪt	zemɪg	zemɪt
zɪlɪp	zɪlɪk	tʃelɪp	tʃelɪd
zɪnɪb	zɪnɪd	plonɪb	plonɪk
blænɪg	blænɪd	slinɪg	slinɪt
zænɪp	zænɪd	vonɪp	vonɪt
glɛnɪb	glɛnɪd	tronɪb	tronɪd
fɪpɪg	fɪpɪt	zipɪg	zipɪd
slɛpɪp	slɛpɪt	glaɪpɪp	glaɪpɪt
vɪpɪb	vɪpɪk	snipɪb	snipɪk
zɑpɪg	zɑpɪd	smaɪpɪg	smaɪpɪd
snævɪp	snævɪk	graɪvɪp	graɪvɪt
tʃɑvɪb	tʃɑvɪk	tʃevɪb	tʃevɪk
smɛvɪb	smɛvɪt	fovɪp	fovɪd
grɑvɪp	grɑvɪt	blivɪb	blivɪt

Appendix B

Stimuli for Chinese Character Transcription (Pinyin) Test

1	胜	**2**	记	**3**	姑
4	缔	**5**	订	**6**	坟
7	板	**8**	蚊	**9**	论
10	饼	**11**	粽	**12**	踌
13	抽	**14**	浩	**15**	胎
16	旺	**17**	扯	**18**	妃
19	扰	**20**	袖	**21**	消
22	魂	**23**	庞	**24**	村
25	庇	**26**	持	**27**	枯
28	拙	**29**	抹	**30**	汗
31	碑	**32**	姓	**33**	拱
34	饱	**35**	特	**36**	牲
37	锈	**38**	供	**39**	脍
40	柚	**41**	讨	**42**	打
43	犹	**44**	猬	**45**	疼
46	拢	**47**	批	**48**	酷
49	优	**50**	精	**51**	肝

52	唱	**53**	始	**54**	仙
55	脾	**56**	宠	**57**	洪
58	瑰	**59**	清	**60**	追
61	猜	**62**	沫	**63**	珑
64	讪	**65**	社	**66**	绘
67	跑	**68**	叮	**69**	拼
70	抬	**71**	愧	**72**	晴
73	肚	**74**	皈	**75**	涛
76	厅	**77**	伦	**78**	嫖
79	诱	**80**	终	**81**	咚
82	埋	**83**	淙	**84**	理
85	趾	**86**	灿	**87**	杞
88	悄	**89**	油	**90**	奸
91	饭	**92**	较	**93**	返
94	胞	**95**	踪	**96**	咬

Appendix C

Language Background and Demographics Questionnaire (United States)

1. What is your age? _____

2. Are you: Male ___ Female___

3. Are you a student? Yes ___ No ____

4. Are you: Hispanic ____ Not Hispanic ____ I do not wish to report this information ___

5. Please place a check beside one or more of the following racial categories that apply to you:

   _____ American Indian / Alaskan Native

   _____Asian (Please specify: ________________________________)

   _____Native Hawaiian or Other Pacific Islander

   _____Black / African American

   _____White / Caucasian

   _____More Than One Race (Please specify: ____________________________)

   _____ I do not wish to report this information

6. Is English your native language? Yes ___ No ___
   If no, please answer the following:

   a. What is your native language? _________________

   b. At what age did you learn English? ______

   c. Was English the first language in which you learned to read? Yes ___ No ___

   d. Was English the first language in which you learned to write? Yes ___ No___

7. Do you speak or know any language(s) aside from English? Yes ___ No ___
If yes, please answer the following:

	Language 1	Language 2	Language 3	Others
List the languages you know:				
Age you learned the language:				
How/Where you learned the language:				
Fluency (Reading): Excellent/Fair/Poor				
Fluency (Writing): Excellent/Fair/Poor				
Fluency (Speaking): Excellent/Fair/Poor				

8. Have you ever been diagnosed with any disorder(s) that have affected in the past or might currently affect your ability to read or spell? Yes___ No___
If so, please explain. ____
____.

9. Have you ever been diagnosed with any hearing or speech disorder(s)? Yes ___ No ___
If so, please explain. ____

10. Are you aware of any rules for doubling a medial consonant in spelling?
If yes, please explain. ____

Appendix D

Language Background and Demographics Questionnaire (Singapore)

1. What is your age? _____

2. Are you: Male ___ Female___

3. Are you a student? Yes ___ No ____

4. Are you: Hispanic ____ Not Hispanic ____ I do not wish to report this information ___

5. Please place a check beside one or more of the following racial categories that apply to you:

   _____ American Indian / Alaskan Native

   _____Asian (Please specify: _________________________________)

   _____Native Hawaiian or Other Pacific Islander

   _____Black / African American

   _____White / Caucasian

   _____More Than One Race (Please specify: _____________________________)

   _____ I do not wish to report this information

6. Is English your native language? Yes ___ No ___
   If no, please answer the following:

   e. What is your native language? _________________

   f. At what age did you learn English? ______

   g. Was English the first language in which you learned to read? Yes ___ No ___

   h. Was English the first language in which you learned to write? Yes ___ No___

7. Do you speak or know any language(s) aside from English? Yes ___ No ___
 If yes, please answer the following:

	Language 1	Language 2	Language 3	Others
List the languages you know:				
Age you learned the language:				
How/Where you learned the language:				
Fluency (Reading): Excellent/Fair/Poor				
Fluency (Writing): Excellent/Fair/Poor				
Fluency (Speaking): Excellent/Fair/Poor				

8. Please report examination grades for ALL languages:

Language	**'O' Level**	**'AO' Level (GP)/ Higher MT**	**Other Certificates: (please specify)**
English			
Mother Tongue (MT) (Chinese/Malay/Tamil)			
Others: (please specify)			

9. Have you previously taken Higher Chinese/Malay/Tamil (delete where applicable)?
Yes ___ No ___

10. Have you previously taken Chinese/Malay/Tamil Literature (delete where applicable) at the "A-level"? Yes ___ No ___

11. Were you taught *Hanyu Pinyin* formally? Yes ___ No ___
    If yes, please answer the following:
    a. At what age did you learn Hanyu Pinyin? ____________

    b. How were you taught Hanyu Pinyin?
       ____

    c. Where were you taught Hanyu Pinyin?
       ____

    d. Duration of formal instruction of Hanyu Pinyin: _________________

12. Have you ever been diagnosed with any disorder(s) that have affected in the past or might currently affect your ability to read or spell? Yes___ No___
    If so, please explain. ____

13. Have you ever been diagnosed with any hearing or speech disorder(s)? Yes ___ No ___
    If so, please explain. ____

14. Are you aware of any rules for doubling a medial consonant in spelling?
    If yes, please explain. ____

Appendix E

List of Plausible Spellings per Phoneme for American and Singaporean English

	Plausible Spellings	
Phoneme	American	Singaporean
Vowel		
/aɪ/	ae, ai, aie, aille, ais, ay, aye, ei, eigh, ey, eye, i, ia, ie, igh, is, oi, ui, uy, uye, y, ye	ae, ai, aie, aille, ais, ay, aye, ei, eigh, ey, eye, i, ia, ie, igh, is, oi, ui, uy, uye, y, ye
/o/	o, oa, ow, ou, oe, oo, eau, oh, ew, au, aoh, ough, eo, aux	o, oa, ow, ou, oe, oo, eau, oh, ew, au, aoh, ough, eo, aux
/i/	e, ea, ee, ae, ie, eo, oe, ay, ey, i, y, oi, ue, ey	e, ea, ee, ae, ie, eo, oe, ay, ey, i, y, oi, ue, ey
/e/	a, aa, ae, ai, aigh, al, ao, au, ay, e, ea, ei, eigh, ee, eh, es, et, ey, ez, ie, oeh	a, aa, ae, ai, aigh, al, ao, au, ay, e, ea, ei, eigh, ee, eh, es, et, ey, ez, ie, oeh
/ɑ/	a, au, aw, ough, augh, o, oa, oo, al, uo, u, ao, eau, ach, au, ou	a, au, aw, ough, augh, o, oa, oo, al, uo, u, ao, eau, ach, au, ou, or, ore, our, oar
/æ/	a, ai, al, au, i	a, ai, al, au, i, air, aire, are
/ɪ/	i, y, ui, e, ie, o, u, a, ei, ia, ea, ai, ey, oe	i, y, ui, e, ie, o, u, a, ei, ia, ea, ai, ey, oe
/ɛ/	e, ea, a, ae, ai, ay, ei, eo, ie, ieu, u, ue, oe, eh	e, ea, a, ae, ai, ay, ei, eo, ie, ieu, u, ue, oe, eh, ear, eir, ere, err, erre, eh
Consonant		
/p/	p, pp, ph	
/b/	b, bb, bh	
/t/	t, tt, ed, pt, th, ct	
/d/	d, dd, ed, dh	
/g/	g, gg, gue, gh, go	
/k/	c, k, ck, ch, cc, qu, q, cq, cu, que, kk, kh	
/m/	m, mm, mb, mn, mh, gm, chm	
/n/	n, nn, kn, gn, pn, nh, cn, mn	
/ŋ/	ng, n, ngue, ngh	
/r/	r, rr, wr, rh, rrh	
/f/	f, ph, ff, gh, pph	
/v/	v, vv, f, ph	
/θ/	th, chth, phth, tth	

/ð/	th	
/z/	s, z, x, zz, ss, ze	
/ʃ/	sh, ti, ci, ssi, si, ss, ch, s, sci, ce, sch, sc	
/ʒ/	si, s, g, z, j, zh, ti	
/tʃ/	ch, t, tch, ti, c, cz, tsch	
/dʒ/	g, j, dg, dge, d, di, gi, ge, dj, gg	
/h/	h, wh, j, ch	
/j/	y, i, j, ll	
/l/	l, ll, lh	
/w/	w, u, o, ou, wh	
/s/	s, c, ss, sc, st, ps, cc, se, ce	s, c, ss, sc, st, ps, cc, se, ce, sz