Do Skilled Readers Use Orthographic Regularities in Learning to Read and Spell?

By: Hannah Wade

A Thesis Submitted to
Saint Mary's University Halifax, Nova Scotia
in Partial Fulfillment of the Requirements for
the B.A. Honours Degree of Psychology

April 20th, 2018, Halifax, Nova Scotia

Examining Committee Member: Dr. Nicole Conrad
Honours Thesis Supervisor

Examining Committee Member: Dr. Jason Ivanoff
Second Reader

Do Skilled Readers Use Orthographic Regularities in Learning to Read and Spell?

Hannah Wade

April 20th 2018

Abstract

The English language has many different spelling patterns. For example, the letters ‘al’ occur in the words practic**al** and **al**low, but within different positions. This study examines whether skilled readers use this type of orthographic knowledge when learning to spell and read new words. This is a two-part study. Part 1 manipulates the frequency of letter patterns using a statistical learning task. Part 2 examines whether skilled readers use knowledge of newly learned letter patterns when reading and spelling new words. Participants were 7 undergraduate students (5 females, 2 males) between the ages of 18 to 27 ($M = 21.7$). Participants read five stories, each containing four repetitions of two target words differing in the frequency of the letter pattern. A spelling and reading test followed, which contained practiced and unpracticed nonwords, made with high and low frequency letter patterns. Results indicated participants were more accurate at reading and spelling practiced words, in comparison to the unpracticed words $F(1,6) = 9.51$, $p < 0.05$, $\eta^2 = .61$. Findings suggest that adults benefit from setting up word representations.

Table of Contents

Do Skilled Readers Use Orthographic Regularities in Learning to Read and Spell?

Introduction

The 20th volume of the Oxford English Dictionary 2nd Edition holds 171,476 words which contain immense amounts of orthographic regularities (Oxford University Press, 2017). Orthographic regularities are the patterns of order and arrangement of letters within a word, including the statistical frequencies with which letter combinations occur (Treiman, 2017). For example, the letters 'al' can be used in the words s*al*t, *al*phabet, and practic*al*, but in different positions within each word. In this case, the letters 'al' would also be considered a bigram, which is one type of orthographic regularity. Orthographic regularities are important to study as it may contribute to children's spelling and reading development and understanding orthographic regularities can provide potential implications for elementary school teachers in teaching children how to read and spell. This study is novel as no one has used the English alphabet to expose children to different combinations of regularities to determine whether children are able to implicitly use knowledge from previous exposure and apply it in novel reading and spelling. The current study examines whether knowledge of orthographic regularities aids with learning to read and spell new words for elementary school-aged children.

Written English is full of regularities that may aid in reading and spelling. For example, children may use rime units to aid their reading and spelling skills. Rime units are combinations of letters that follow the initial consonant onset in a word (e.g., the "ake" in the word "snake"). These phonological units occur with varying regularity across words, and children use knowledge of these regularities to read and spell new words by analogy; that is, reading the word "flake" because they already know the word "snake"

(Goswami, 2000). As well as rime units, children may also use morphological units which is understanding a unit of language in its smallest form, where the word can no longer be broken down any smaller while still having a meaning (Ontario Ministry of Education, 2017). For an example, the word ‘play’ cannot be broken down further without a meaning but can be used in other words such as replay and playpen (Ontario Ministry of Education, 2017). Although children may use rime and morphological units, children also may use orthographic regularities to help read and spell. Children may use past knowledge of the different letter combinations that appear together when attempting to spell a new, but similar word. For example, the letters *ea* appear in the word “h*ea*t” and when given the task to spell the word “b*ea*n,” children may notice that the letters *ea* occur together frequently and apply that knowledge to correctly spell “bean.” The letters *ea* are only one example, of the thousands of different regularities.

There are currently three notable theories in the literature that support the role of orthographic regularities in reading and spelling; Ehri’s (2005 & 2014) orthographic mapping theory, word-specific and generalized knowledge of orthographic regularities in developing readers from stage/phase theory (Ehri, 2014), and lastly the integration of multiple patterns theory (Treiman & Kessler, 2014). These theories combined suggest that children use orthographic knowledge in reading and spelling and demonstrate the importance of studying the development of orthographic regularities in children.

Orthographic Mapping

Being sensitive to orthographic regularities aids the development of orthographic mapping and sight words (Ehri, 2014). Orthographic mapping is the ability to form connections in memory between written units (e.g., graphemes), and spoken units (e.g., phonemes; Ehri, 2014). Sight words are simply words that are already stored in memory and while looking at the word, you immediately have the activation and ability to pronounce and understand the word, which you are currently doing now as you read this sentence (Ehri, 2014). Sight words require setting up word representations, and when those sight words are stored into memory, the orthographic regularities that make up the word are also being stored (Ehri, 2014). While setting up word specific representations, children are sensitive to the letter patterns that create the word and may be able to use this information in novel situations by extracting orthographic regularities from sight words. Therefore, orthographic mapping and sensitivity to orthographic regularities are key to setting up the word specific representations in memory that enable sight word reading (Ehri, 2014).

With orthographic mapping, children can apply strategies to enhance their ability to spell and read such as through analogy (Ehri, 2014). The use of analogy is related to the awareness of setting up correct word representations of orthographic regularities in memory. The word representations set up in memory can be used by extracting the knowledge of the orthographic regularities already stored in memory when reading and spelling new words alike. Orthographic mapping is important for reading and spelling as past research demonstrates that children can make connections from memory based on analogies and sight words. Sensitivity to orthographic regularities is needed to create these connections (Ehri, 2014).

Ehri (2005) and other studies suggests that children can use information of orthographic regularities in large units to remember how to spell and read specific words (Conrad, Harris & Williams, 2013; Treiman & Kessler, 2006). For example, when children are given an option between spellings of two words with the vowel doublets ‘aa’ or ‘ee’, most children would choose ‘ee’, because that bigram is more frequent in the English language (Cassar & Treiman, 1997). This indicates that children have more word representations set up in memory with more frequent letters. Using stored word representations and orthographic regularities indicates that children develop an awareness of orthographic regularities very early in literacy development and use orthographic knowledge to aid spelling and reading.

Word-Specific and Generalized Orthographic Knowledge

There are two important ideas proposed by Ehri (2005) in her stage/phase theory, suggesting that orthographic regularities may be helpful for children learning to read and spell. Firstly, Ehri (2005) found that with practice and increased alphabetic knowledge (of spelling patterns and regularities), children may use this knowledge to remember how to spell specific words. The idea of increased alphabetic knowledge aiding children’s ability to spell specific words is equivalent to orthographic mapping because the practice and awareness of different letter patterns increases children’s knowledge of orthographic regularities in words while simultaneously increasing the connections being formed into memory. The second idea Ehri (2005) proposes is children’s knowledge of recurring letter patterns. Recurring letter patterns may help children learn rules about how a word can be read or spelt. Overtime, as children learn about the constraints of words regarding how words may be written, or read, connections are reinforced in memory (Ehri, 2005). Overall, Ehri’s (2005) theory strongly relates to orthographic regularities and the influence orthographic regularities may have in reading and spelling.

Increased alphabetic knowledge suggests that exposure of learning the alphabet influences children's choices of letters when spelling words because of the connections formed through orthographic mapping (Treiman & Kessler, 2006). Treiman and Kessler (2006) demonstrated that younger pre-phonological (children who cannot identify syllables or rimes) spellers were more likely to use letters to spell words in alphabetical order. Children are learning and storing the alphabet into memory in a specific order, as well as applying their knowledge of the alphabetic letter patterns in novel words. The same effect was found when comparing the letters children used to spell words, to the letters in the younger pre-phonological spellers' names. Younger pre-phonological children were more likely to use the letters in their name to spell words indicating that the exposure to learning their name influences their choice of letters in spelling words (Treiman and Kessler, 2006). Once again, with this evidence, it supports Ehri's (2014) theory of children being sensitive to orthographic regularities as letter patterns from children's names are stored into memory and used to aid spellings of novel words.

Children have an awareness of reoccurring letter patterns and writing rules (Pacton, 2013). When children are exposed to non-words with double consonants through reading and asked to spell, commonly, full-phase children (children who are developed readers/spellers) knew that they had to include a double consonant somewhere, and usually used it in a legal position, even though it was exposed to them in an illegal position (Wright & Ehri, 2007). This supports Ehri's (2005) theory as children are using awareness of orthographic regularities and constraints from memory to accurately spell. The same was found in Pacton (2013) where French children spellings used correct rules, even when taught incorrect spelling rules (Pacton et al., 2013). This again suggests that previous knowledge of orthographic regularities and constraints learnt through reading are used to accurately spell.

Integration of Multiple Patterns

Most importantly, integration of multiple patterns theory focuses on patterns children use to spell words (Treiman & Kessler, 2014). There are two types of patterns, the first one being the patterns of the outer form of the word which is whether a word ‘looks correct.’ For example, the word “mom,” children know that it cannot be “mmm” because no word in the English writing system has three identical letters corresponding one another (Treiman & Kessler, 2014). The second pattern focuses on repeated exposure of orthographic regularities in the inner function of a word and how it influences reading and spelling (Treiman & Kessler, 2014). It is thought that this knowledge is acquired through statistical learning. Statistical learning allows children to identify patterns in print and the likelihood of the patterns to reoccur (Treiman, 2017). Combing the theory altogether, when children spell, they are motivated by two classes of patterns; visual representations and repeated exposure of orthographic regularities.

Children are more likely to spell a word in a specific way when previous multiple patterns support that spelling in a new word (Treiman & Kessler, 2014). It is thought that when children are asked to read words with infrequent letter patterns, they are more likely to find it difficult to spell the word because children do not have word representations in memory to support the unique spelling. Similarly, Treiman (2017) found that the spellings of the older pre-phonological spellers looked more like English words, indicating that they knew more about letter patterns and lengths of words because of the increased exposure received in comparison to younger pre-phonological spellers (Treiman, 2017). These findings show that statistical frequency of letter patterns and an increased exposure of words and letters influences children's choice of letter patterns when spelling (Treiman, 2017).

Skilled Readers

Although the theoretical models reviewed are developmental models, adults use knowledge of orthographic regularities and common writing rules when spelling. For example, Sobaco and colleagues (2014) found that French university students were accurate at spelling words with more common orthographic regularities and found it harder to spell words with unique letter patterns. Similarly, Treiman and Boland (2017) found adults with higher pre-measure spelling skill used explicit rules to spell non-words with common orthographic regularities. This suggests more exposure of common orthographic regularities influences spelling. Therefore, adults rely on orthographic knowledge and legitimate rules when learning to spell new words.

Adults also rely on rules and statistical regularities when spelling new words. Chetail (2017) used artificial scripts and manipulated the exposure French university students received to determine whether manipulated exposure of artificial scripts influences students' performance on word-likeness and letter detection tasks. Results indicated that participants were more accurate at detecting the high frequency artificial scripts, in comparison to the low frequency scripts. This supports the use of statistical learning as the mechanism through which we acquire awareness of orthographic regularities. Together, these studies suggest that adults use their awareness of orthographic regularities when spelling new words. The current study also examines the use of orthographic regularities in adults' spelling attempts, but here we focus on a more subtle regularity - the statistical frequency of bigrams within words. Treiman and Boland (2017) examined the use of regularities that can also be expressed as explicit rules (e.g., double consonants only follow single vowels), whereas bigram frequencies are often considered to be implicit awareness of orthographic probabilities (Chetail, 2017). In addition, our work extends past work in several other important ways. First, the bigrams will be low frequency bigrams from the English language instead of artificial scripts. As well as using English bigrams, bigrams will be incorporated into non-words and participants will be tested on a reading and spelling outcome after a letter detection task.

The Current Study

Although our initial intent was to examine the use of orthographic regularities in children's spelling and reading, due to unforeseen circumstances, the current study examines this with university students. As demonstrated in the preceding section, our research question with skilled readers is also theoretically motivated. We made this switch to skilled readers in mid-March, which accounts for our relatively small sample size to date.

The current study extends past research in a number of ways. The study focused on the English language by using discrete orthographic regularities instead of focusing on double consonants or explicit writing rules. This is novel because past studies have focused on university students' spellings of legitimate or illegitimate letter patterns. As well, the study will manipulate exposure of English low frequency bigrams and incorporate them into non-words. To my knowledge, no previous study has manipulated exposure adults receive of subtle English regularities that cannot be explained through explicit rules and formal instruction. Through incorporating bigrams into non-words, it will help determine whether once adults acquire orthographic knowledge learnt from exposure, whether adults use this knowledge to aid their reading and spelling of new words. Once again, this is novel as past research focused on letter detection task outcomes but did not examine the effect manipulating exposure has on reading and spelling outcomes.

Ten low frequency bigrams (it is expected participants have limited exposure to these bigrams) are selected based on perceived likelihood of being able to pronounce the letters in non-words (Solso & Barbuto, 1979). Bigrams were then manipulated in a statistical learning task. The manipulation of bigrams in the statistical learning task enabled us to create low or high frequency bigrams in the study. The study then focused on whether exposure of the bigrams would affect adults' ability to set up word representations in memory and affect subsequent reading and spelling. Participants were assessed on their accuracy of reading non-words with high or low frequency bigrams through short stories. Participants were then assessed reading and spelling practiced words and new words with the practiced bigrams to determine whether adults are more accurate at reading and spelling non-words with high or low frequency bigrams. Lastly, participants were assessed on their generalization of reading and spelling new non-words, with the same exposed bigrams. Specifically, we hypothesized that adults will learn to read and spell non-words with high-frequency bigrams better than non-words with low-frequency bigrams. The second hypothesis is adults will generalize their knowledge of orthographic regularities when spelling and reading new non-words because of the opportunity to set up word representations.

Method

Participants

Twenty university students were tested however 6 participants did not speak English as their first language. An additional 6 participants' data had to be excluded due to a technical error in the Statistical Learning Phase. Lastly, 1 participant's data could not be used because of their low pre-test measures. Therefore, this study includes 7 (5 females and 2 males) university students between the ages of 18-27 with a mean age of 21.7. Students who participated gave informed consent. Students were recruited through Saint Mary's University SONA system and compensated with bonus points towards a psychology course.

Materials and Design

This study is considered as a second portion of a two-part study. Part 1 focused on statistical learning of orthographic regularities. Part 2 focused on whether awareness of those manipulated orthographic regularities affects learning to read and spell new words. Both parts of the studies were administered on the same day. This study was an experimental within-subject design where all participants completed seven phases (Part 1 having three phases, while Part 2 has four phases) of the experiment. Part one of the study included pre-test measures, and the Statistical Learning (SL) Phase, and the Statistical Learning Outcome Phase. Part two included the story reading phase (learning phase), non-verbal reasoning, word spelling test, and lastly the reading test phase.

Pre- existing Knowledge

Wide Range Achievement Test. Participants were assessed with the Wide Range Achievement Test- 4th Edition (WRAT-4) which tests students' spelling. The WRAT-4 measured spelling accuracy (Pearson Education, 2017). Participants were asked to start at number six and heard the words that had to be spelled and wrote their answers on paper next to the corresponding numbers. There was a total of 56 items, and participants continued spelling until they had ten consecutive errors.

As well, the WRAT-4 was used to assess the participants' current reading ability (Pearson Education, 2017). Participants were given a list of words and asked to read the words out loud clearly, and slowly. Participants started on the word ‘cat.’ There was a total of 70 items and participants read until the end of the list or until they made ten consecutive errors.

Word Likeness Test. Participants were also measured of their current knowledge of orthographic regularities through a test created by Conrad, Harris, and Williams (2013) (see Appendix A). The assessment of pre-existing knowledge of orthographic regularities had 32 questions, and students had two options for each question. Students were asked to identify which non-word looked like a real word. For an example, students had the option between ‘plew’ and ploo’ and had to circle which word seemed more word-like. This tests pre-existing orthographic knowledge as students identify which word is most word-like based on their understanding of which letter arrangements are more likely to occur in the English language.

Non-Verbal Reasoning Task. As a distractor, participants were assessed on the Matrix Reasoning task from the WASI-II. The task examines participants’ intelligence, spatial ability, perceptual organization and simultaneous processing (Pearson Education, 2011). There is a total of 30 items, and participants are asked to indicate which object corresponds to the incomplete matrix shown. Participants were tested until they completed all 30 items, or had three consecutives wrong answers.

Part One

Statistical Learning (SL) Phase. The SL phase was designed to expose participants to bigrams that were manipulated with different frequencies. Bigrams can be defined as two letters that always appear together during the SL phase. Bigrams were selected from a pool of bigrams generated by Solso and Barbuto (1979). All selected bigrams were ‘low-frequency’, as in not as likely to occur in the English writing system, to ensure participants are not already familiar with the common letter patterns and cannot rely on memory/previous knowledge when performing this task. From this pool of 577 low frequency bigrams, ten bigrams were then selected based on perceived likelihood of being able to pronounce the bigrams once embedded into non-words. As well as picking bigrams based on ease of pronunciation, bigrams were also selected to ensure there was balance of consonant-consonant (DP and WF) and consonant-vowel (IH and EJ) bigrams across both sets (see Appendix B). The ten bigrams selected were HI, YC, DP, RZ, KM, EJ, AQ, WF, GB, SN. Although all bigrams are considered low-frequency based on the frequency of occurrence in the English writing system, for this study the terms 'high-frequency' and 'low-frequency' bigrams refer to the manipulated exposure participants received of the bigrams during the statistical learning task. To put into simpler words, all bigrams used in this study were low-frequency, but manipulation of the exposure participants received made the difference between high, and low-frequency bigrams. High-frequency bigrams were seen by the participants 12 times, whereas low-frequency bigrams were only seen 6 times. For an example, bigrams IH, KM, DP, UX, GB were considered high-frequency because they were seen 12 times on the computer screen whereas bigrams RZ, EJ, SN, AQ, WF were considered low-frequency because they were only exposed 6 times (see Appendix B).

The SL phase, including the exposure of bigrams, included 20 letters of the alphabet, which were each exposed 16 times each, for 373 trials (see Appendix C). All letters were presented an equal number of times so participants would receive the same amount of exposure to all letters. For the letters used in the high frequency bigrams, the letters were each presented twelve times within a bigram, and four times alone. For the letters used in the low frequency bigrams, the letters were each presented six times within a bigram, and 10 times alone. Each letter was exposed individually in black, lower case font, on a white background computer screen. A happy face emoji appeared on the screen a total of 53 times and was used as a distracter task to encourage participant's attention during the task (see Appendix C). As the letters appeared one-by-one on the computer screen, participants were asked to say the letters out loud to ensure the letters were being attended to and asked to hit the space-bar when the emoji appeared on the screen. The letters exposed individually included the presence of manipulated frequency bigrams which served as the purpose of exposing the orthographic regularities, without participants' knowledge. As well, the presence of the distractor task was to hide the reason behind the experiment, and to ensure participants were paying attention while being exposed to the letters one-by-one. Statistical learning was assessed following the learning phase, by a learning outcome phase where participants had the options of two sets on bigrams and had to indicate which bigram seemed most familiar. Although the results of this task will not be directly reported in this thesis, this phase provided the bigram exposure manipulation that will be further explored in the current thesis.

Part Two

Learning phase. Participants each read five stories containing non-words that were made with the same ten bigrams that were used in the statistical learning phase (see Appendix E). After the learning phase, participants were assessed in the learning outcome phase for word specific learning. Word specific learning means participants were measured based on spelling and reading accuracy of the non-words from the stories, thus providing a measure of whether nonwords with high frequency bigrams were read and spelled more accurately than nonwords with low frequency bigrams. Participants were also assessed for generalization of bigram learning by spelling and reading new non-words that were not in the stories, but contained the same bigrams from the statistical learning part of the study. Measuring word specific learning, and generalization learning uses three sets of non-words (see Appendix D); one set for exposure in the story, one set to assess generalization of spelling, and one set to assess generalization of reading. Each set of words will be counterbalanced across tasks, so participants are not all seeing the same set of words in the stories. For an example, Set A may be in the story for Participant 1, but Set B could be incorporated into the story for Participant 2.

Thirty non-words were generated by a "Fake Word Generator" and divided into three sets (A, B, C). The target bigrams were adjusted into the non-words accordingly. Each set of words had constraints to control for extraneous variables. Each set contained two 4-letter words, six 5 letter words, and two 6 letter words. As well, each set of words had constraints on the position of the bigram within the word: the bigram appeared in the second-third position four times (e.g. E*dp*ow), the third-fourth position four times (e.g. Ur*km*as), and in the fourth-fifth position twice (e.g. Cor*ej*). These constraints on the non-words that were used during testing attempts to control confounding factors that could sway the results of children's reading and spelling accuracy, eliminating the effect of different letter-lengths, and bigram position within the word may have. Each non-word was checked for a potential meaning through google searching. In some cases, words were last names, recipe names, or geographical locations, and had to be eliminated since participants could not be exposed to existing words.

Five stories were created to meet the expected grade four reading level (as previously the experiment was designed for Grade 4 children) and each contained 122 words. Two of the stories were modified to the grade four reading level from samples used by Pacton et al. (2014), one of the stories was modified from Pacton, Foulin, Casalis, & Treiman (2013), and the last two stories were created using the previous three as examples. In each story, one non-word with a high-frequency bigram, and one non-word with a low-frequency bigram were used as nouns in the story. Each non-word with the high-frequency, and low-frequency bigram included were repeated four times in the story (see Appendix E). In total, five stories, each containing one high-frequency non-word repeated four times, and one-low frequency non-word repeated four times was used. The manipulation of bigram frequency was completed during the SL phase, therefore participants received equal exposure of non-words, no matter the bigram frequency incorporated into the non-words, while reading the story. Sets A, B, and C were used equally often within the story, and as generalization for reading and spelling outcomes across participants. Questions based on the stories were created for participants to answer after each story was read to encourage attention while reading. For an example, "What was the teacher's last name in the story?"

Spelling Test Phase. Participants' ability to remember the spellings of the five non-words with high-frequency bigrams and the five non-words with low-frequency bigrams was assessed with an oral spelling production task. Participants heard the pre-recorded word that was to be spelt twice, and responded by writing the spelling of the non-word on paper. The non-words were recorded to ensure consistency in pronunciations. After the ten non-words were spelt, participants were also asked to spell ten novel non-words (see Appendix D, set B) that were not included in the story. These novel non-words contain the same 10 bigrams. This part of the spelling test was used to assess whether additional exposure to the bigrams while reading words containing these bigrams in the stories would generalize to spelling new words with the same bigrams. No corrective feedback was given while participants were writing down their answers. Spelling accuracy was scored with the following criteria: one point for using the correct bigram, one point for the bigram being in the correct position within the word, and another point for spelling the whole word correctly. Therefore, each non-word spelt was marked out of 3 points, and overall each participant received a mark out of 60 accordingly. In addition to the 20 non-words that were asked to be spelt, participants were also given an additional 10 non-words that did not include any of the bigrams manipulated in the statistical learning phase to draw comparison between the spelling of the 10 non-words with exposed bigrams, to 10 non-words without any previously exposed bigrams. The 10 additional non-words (without bigrams) will be helpful in determining whether participants were able to spell non-words with bigrams better than non-words without bigrams. Therefore, five non-words (without exposed bigrams) were added to each set of non-words with bigrams, totaling 15 words in each set meaning participants were asked to spell 30 items (see Appendix D), 10 nonwords they read in the stories, 10 new nonwords with the same bigrams, and 10 new nonwords with new bigrams.

Reading Test Phase. This outcome measures focused on whether participants had learnt to read based on the exposure they received during the statistical learning phase, and after reading the five stories that included the 10 non-words containing five high-frequency bigrams and five low-frequency bigrams during the Story Reading (or exposure) Phase. This Reading Test measured speed and accuracy of participants' pronunciations of the non-words. Participants were asked to read the 10 non-words out loud and were recorded. Participants were then asked to read another set of 10 non-words that included the same 5 high-frequency and the same 5 low-frequency bigrams as the original 10 non-words, but the bigrams were embedded into new words that have never been seen before (see Appendix D). The purpose of the new 10 non-words was to see whether generalization occurred in reading new non-words, but with the same exposed high and low-frequency bigrams. Once again, no corrective feedback was given for the pronunciations of the non-words. As mentioned previously, an additional 10 non-words without manipulated bigrams were also administered for comparison between non-words with bigrams, and non-words without bigrams. Once again, the comparison between the non-words with and without bigrams will be helpful in determining whether participants found it easier to read the non-words with previous bigram exposure. Thus, each participant read 30 nonwords – the 10 nonwords from the stories, 10 new nonwords with the same bigrams, and 10 new nonwords with new bigrams. The total number of correct pronunciations of non-words was marked for accuracy. Response time was defined as the time between when a nonword first appeared on the computer screen and a vocal response was made.

Procedure

The three tests assessing word reading, spelling, and pre-existing knowledge of orthographic regularities were first administered before the experiment commenced. After completing the pre-existing measure of orthographic knowledge, spelling, and reading, participants then completed the exposure phase called the 'Statistical Learning (SL) Phase'. Participants were tested individually. Each participant came to the Literacy Lab at the university and completed the SL phase. Each participant was told that the purpose of the task was for them to hit the space bar when they saw an emoji appear on the computer screen. This task took 8 minutes. After the SL phase, participants were tested to see if they acquired knowledge of the manipulated bigrams through statistical learning. These results will not be reported here (see Craig Isnor's thesis for further details).

Participants were then asked to read five stories out loud. After each of the short stories were read out loud participants were asked to answer questions which were based on the stories they had just read. As a distractor task, participants were then administered a non-verbal reasoning task and asked to answer the matrix questions out loud. Participants then completed the spelling production and reading tests. Participants were provided with no corrective feedback on the test phase, but were encouraged to "try their best" or "make their best guess". At the end of the session, participants were provided an opportunity to ask questions, or inquire about the purpose of the study.

Results

Hypothesis 1

To address the first hypothesis that adults read and spell words with high frequency bigrams more accurately than words with low frequency bigrams, we conducted two different analyses. First, we explored reading accuracy during the story reading phase by conducting a within subjects t-test to compare accuracy means of reading non-words with high frequency (HF) bigrams to means of reading non-words with low frequency (LF) bigrams during this exposure phase. There was a significance difference in reading accuracy, $t\ (6) = -4.510$, $p < 0.05$. As illustrated in Table 1, reading accuracy for non-words with HF bigrams was less accurate than reading performance in the non-words with LF bigrams, contrary to the first hypothesis.

Table 1
Proportion of Non-Words Read Correctly During the Exposure Phase

	Reading non-words with high frequency bigrams	Reading non-words with low frequency bigrams
Readers (n = 7)		
M	.62	.76
SD	.14	.12

The second analysis addressing the first hypothesis was a 2 (reading/spelling) x 2 (HF nonwords/LF nonwords) x 2 (practiced words/generalization words) within-subjects ANOVA. The main effects of modality and frequency, and their potential interaction, would enable us to determine whether the practiced nonwords with HF bigrams were read or spelled more accurately than practiced nonwords with LF bigrams (among other things addressed in the next section). This analysis revealed a significant main effect for reading and spelling F (1,6) = 25.58, $p < 0.05$, $\eta^2 = .81$ revealing that overall participants were more accurate at reading. The main effect for frequency and the interaction between frequency and modality were not significant.

Hypothesis 2

The same 2 x 2 x 2 within-subjects ANOVA addressed the second hypothesis, that adults use orthographic regularities when reading and spelling new non-words for which they have no previous exposure. In this analysis, there was a main effect for word type F (1,6) = 9.51, $p < 0.05$, $\eta^2 = .61$. As expected, participants were more accurate at reading and spelling the practiced words compared to the new words with the same bigrams (see Table 2). However, there was no significant main effect for frequency, and no significant interactions.

Table 2

Proportion of Exposed and Generalized Words Read and Spelt Correctly

Outcome	Exposure words with HF bigrams	Exposure words with LF bigrams	New words with HF bigrams	New words with LF bigrams
Readers (n = 7)				
M	.77	.80	.59	.54
SD	.21	.26	.36	.30
Spellers (n = 7)				
M	.46	.29	.14	.23
SD	.28	.25	.15	.08

To further probe whether adults use the practiced bigrams, whether HF or LF, when reading and spelling new nonwords, a second 2 x 3 ANOVA was conducted in which modality (reading/spelling) and word type (practiced nonword/new nonword with practiced bigrams/control nonwords with no practiced bigrams) were within-subject factors. Analysis revealed a main effect for modality F (1,6) = 25..01, $p < 0.05$, $\eta^2 = .81$, again illustrating that adults were more accurate reading than spelling, and a main effect for word type F (1,6) = 8.02, $\eta^2 = .57$. Post hoc tests using Bonferroni revealed participants spelt and read practiced words more accurately than generalized words, $p <$ 0.05 and control words more accurately than generalized words, $p < 0.05$ (see Table 3).

Table 3

Proportion of Exposed, Generalized Words and Control Words Spelt Correctly

Outcome	Exposure words with both HF and LF bigrams	Generalized words with both HF and LF bigrams	Control words with no bigrams
Reading outcome (n = 7)			
M	.79	.57	.79
SD	.21	.23	.17
Spelling outcome (n = 7)			
M	.37	.24	.47
SD	.23	.18	.18

Overall, the manipulated frequency of the bigrams did not affect accuracy of reading or spelling in the expected direction. With the practice data, we see that adults are more accurate reading nonwords with LF bigrams than HF bigrams. With the outcome measures, frequency did not affect reading or spelling at all.

Discussion

The current study had two research questions. The first question was do adults read words with high frequency bigrams more accurately than words with low frequency bigrams? The results indicated that during the reading practice phase, adults were more accurate at reading non-words with low-frequency bigrams, which was not expected. An explanation for these results could be the non-words that were created with low frequency bigrams were easier to read. To overcome this, it would be beneficial to counter balance all non-words across both high and low frequency conditions. Although this was our original intention, we have not yet completed data collection and have not yet done this. In addition, we unfortunately discovered during the study that the bigram *SN is* more frequent in the English language than previously identified. For example, the bigram *SN is* used in the words snooze, snake, snap, snow, snail, sneak, snake and many more. In the current study, *SN* was used as a low frequency bigram but the high frequency in the English language could have been a factor influencing the higher accuracy of reading scores for non-words with low frequency bigrams. To determine if the bigram *SN* was an extraneous variable, future studies should replace this bigram with a low frequency bigram from our writing system.

In addition, frequency was not a factor influencing accuracy on the reading and spelling outcome measures. An explanation for these results could be the manipulation during the statistical learning phase did not work. If that were the case, all non-words in the stories would have been considered of equal exposure. Another explanation could be adults do not attend to regularities at this level. However, other research suggests that adults do. To determine if frequency is a factor influencing accuracy of reading and spelling, future research should focus on ensuring the statistical learning phase is effective in implicitly exposing adults to different frequencies of bigrams.

The second research question was do adults use orthographic regularities when reading and spelling new non-words? Once again data analysis was broken into two parts. For part one, results revealed that there was a main effect for word type. Adults were more accurate at spelling and reading the non-words from the practice phase than the new non-words containing the practiced bigrams. This result was expected as participants were given the opportunity to set up word specific representations in their memory before being asked to spell and read the same non-words in the spelling and reading outcome. However, we could not include the control words in the within-subjects ANOVA because we cannot compare the control words to the non-words with high or low frequency bigrams. Thus, the second within-subjects ANOVA focused on modality and word type, with the inclusion of control words. As reported above, a main effect was found in modality and word type. Post hoc tests identified where the differences in accuracy lie. Adults spelt and read the non-words from the practice phase more accurately than the generalization words. As well, post hoc revealed that control words were spelt and read more accurately than generalized words. It became evident during testing that the control words were considerably easier to spell and read than the other non-words that were created, and will be edited for further studies. However, that the practiced words ended up being read with equal ease as the control words is interesting and suggests that four practices reading these new, difficult to pronounce words was sufficient to set up a good quality word representation in memory that enabled efficient reading and spelling of those words.

Implications for Theories of Reading Development and Practice

The main effect found in word type supports Ehri's theory of orthographic mapping. Ehri (2014) says that sight words are words that children can look at and read instantly, and through orthographic mapping it allows children to apply strategies to enhance their reading and spelling. Likewise, in the current study adults were better at reading and spelling words they were exposed to in the short stories because participants had the opportunity to set up word representations and use the orthographic regularities from those representations stored in memory. Therefore, an implication that can be made is that adults benefit from setting up word representations and to increase their spelling accuracy, adults should receive exposure of whole words before learning to spell. In practice, this would indicate that adults should read more often to have a general knowledge of different word representations and orthographic regularities.

Limitations and Future Directions

The current study has a few limitations such as the very small sample size. To gain a boarder understanding of whether adults are using orthographic regularities when reading and spelling, a larger sample is needed. As well, with a larger sample size future studies should counter balance the non-words with bigrams to determine whether some words are easier to spell than others. As well as counter balancing, it would be beneficial to record participants' responses when reading the non-words during the exposure phase. Recording responses would allow scoring to be more precise as other researchers have the opportunity to deem responses as correct or incorrect.

Another limitation is that the results of this study depended on the manipulation in the statistical learning phase. Results of the statistical learning outcome did not find significance (Isnor, 2017). This means that adults did not gain knowledge of orthographic regularities from the statistical learning phase (Isnor, 2017). This also indicates that the manipulation was not effective, so the non-words that were considered high or low frequency in part two of the study may instead be of equal frequency since participants saw each word four times in stories.

Ecological validity was weak as children and adults cannot sit in front of a computer screen to learn different letter patterns in the English language. It is boring. The idea behind the statistical learning task was to implicitly expose participants to bigrams to manipulate frequency, to test whether statistical learning is the mechanism through which children acquire knowledge of these regularities. Theory suggests that children are exposed to orthographic regularities before they can read and are implicitly picking up letter patterns from words before developing the skill to create whole word representations into memory. Our study was to explore this mechanism without adding in reading. However, to increase ecological validity, future studies could manipulate the exposure of bigrams in words, just as children and adults experience in their everyday lives. As well, future studies could compare the accuracy of spelling and reading by manipulating the type of exposure participants receive, such as exposing bigrams through words or through the statistical learning task by showing letters individually. Lastly, the original intention of this study was to focus on children as it is important to understand how children develop as readers and spellers, which is valuable for this field of research. Therefore, further research should focus on how children use orthographic regularities when learning to read and spell.

Conclusion

The current study set out to answer two research questions; (1) Do adults read words with high frequency bigrams more accurately than words with low frequency bigrams? (2) Do adults use orthographic regularities when reading and spelling new non-words? The hypothesis was not supported for the first research question as non-words with low frequency bigrams were spelt more accurately. However, for the second research question it was found that adults are more accurate at spelling the non-words from the exposure phase which supports Ehri (2014) theory of orthographic mapping and sight words. The current study is valuable as it provides a direction and methodology for further study on this topic. Future research would benefit from using this study’s design with some adaptions.

References

Apel, K, Brimo, D, Wilson-Fowler, E B., Vorstius, C, & Radach, R. (2013). Children develop initial orthographic knowledge during storybook reading. *Scientific Studies of Reading, 17*(4), 286-302.

Cassar, M., Treiman, R., & Pressley, G. M. (1997). The Beginnings of Orthographic Knowledge: Children's Knowledge of Double Letters in Words. *Journal of Educational Psychology, 89*(4), 631-644.

Chetail, F. (2017). What do we do with what we learn? Statistical learning of orthographic regularities impacts written word processing. *Cognition, 163*, 103-120.

Conrad, N., & Harris, Karen R. (2008). From Reading to Spelling and Spelling to Reading: Transfer Goes Both Ways. *Journal of Educational Psychology, 100*(4), 869-878.

Conrad, Nicole J., Harris, Nicholas, & Williams, Jennifer. (2013). Individual Differences in Children's Literacy Development: The Contribution of Orthographic Knowledge. *Reading and Writing: An Interdisciplinary Journal, 26*(8), 1223-1239.

Isnor, C. (2017). Statistical Learning of Orthographic Regularities in Adults.

Liennea C. Ehri (2014) Orthographic Mapping in the Acquisition of Sight Word Reading, Spelling Memory, and Vocabulary Learning, Scientific Studies of Reading, 18:1, 5-21, DOI:10.1080/10888438.2013.819356

Ontario, G. O. (2017). Early Reading Strategy. Retrieved December 15, 2017, from http://www.edu.gov.on.ca/eng/document/reports/reading/effective.html

Oxford University Press. (2017). How many words are there in the English language?

Oxford Dictionaries. Retrieved October 24, 2017, from https://en.oxforddictionaries.com/explore/how-many-words-are-there-in-the-english-language

Samara, & Caravolas. (2014). Statistical learning of novel graphotactic constraints in children and adults. *Journal of Experimental Child Psychology*, *121*, 137-155.

Sebastien Pacton, Amélie Sobaco, Michel Fayol, & Rebecca Treiman. (2013). How does graphotactic knowledge influence children's learning of new spellings? *Frontiers in Psychology, 4*, Frontiers in Psychology, 01 October 2013, Vol.4.

Share, D., & Cooper, Harris. (2008). On the Anglocentricities of Current Reading Research and Practice: The Perils of Overreliance on an "Outlier" Orthography. *Psychological Bulletin, 134*(4), 584-615.

Sobaco, A., Treiman, R., Borchardt, G., & Pacton, S. (2014). The influence of graphotactic knowledge on adults' learning of spelling. *Memory & Cognition,* 43(4), 593-604. Doi:10.3758/s13421-014-0494-y

Treiman, & Boland. (2017). Graphotactics and spelling: Evidence from consonant doubling. *Journal of Memory and Language, 92*, 254-264.

Treiman, R., Kessler, B., & Harris, Karen R. (2006). Spelling as Statistical Learning: Using Consonantal Context to Spell Vowels. *Journal of Educational Psychology, 98*(3), 642-652.

Treiman, Rebecca, & Others. (1993). Children's Use of Phoneme-Grapheme Correspondences in Spelling: Roles of Position and Stress. *Journal of Educational Psychology, 85*(3), 466-77.

Treiman, R. (2017). Learning to Spell Words: Findings, Theories, and Issues. *Scientific Studies of Reading, 21*(4), 265-276.

Wright, Donna-Marie, & Ehri, Linnea C. (2007). Beginners Remember Orthography when They Learn to Read Words: The Case of Doubled Letters. *Applied Psycholinguistics*, *28*(1), 115-133.

Appendix A

Assessment of pre-existing knowledge of orthographic regularities

From: Conrad, Harris, & Williams (2013)

Children select which of each pair "looks most like a real word".

Frequent letter pattern/Infrequent letter pattern
siff siph
vime vyme
moin moyn
hool hewl
poaf pofe
kade kayd
tave taiv
gilk gilc
murn mirn
waff waph
tays tayz
zame zaym
deef defe
nide nyde
hife hyfe
jick jikk
sive syve
bope boep
goom gewm
rork rorc
tupe tuep
plew ploo

Appendix B

High and Low-Frequency Bigrams

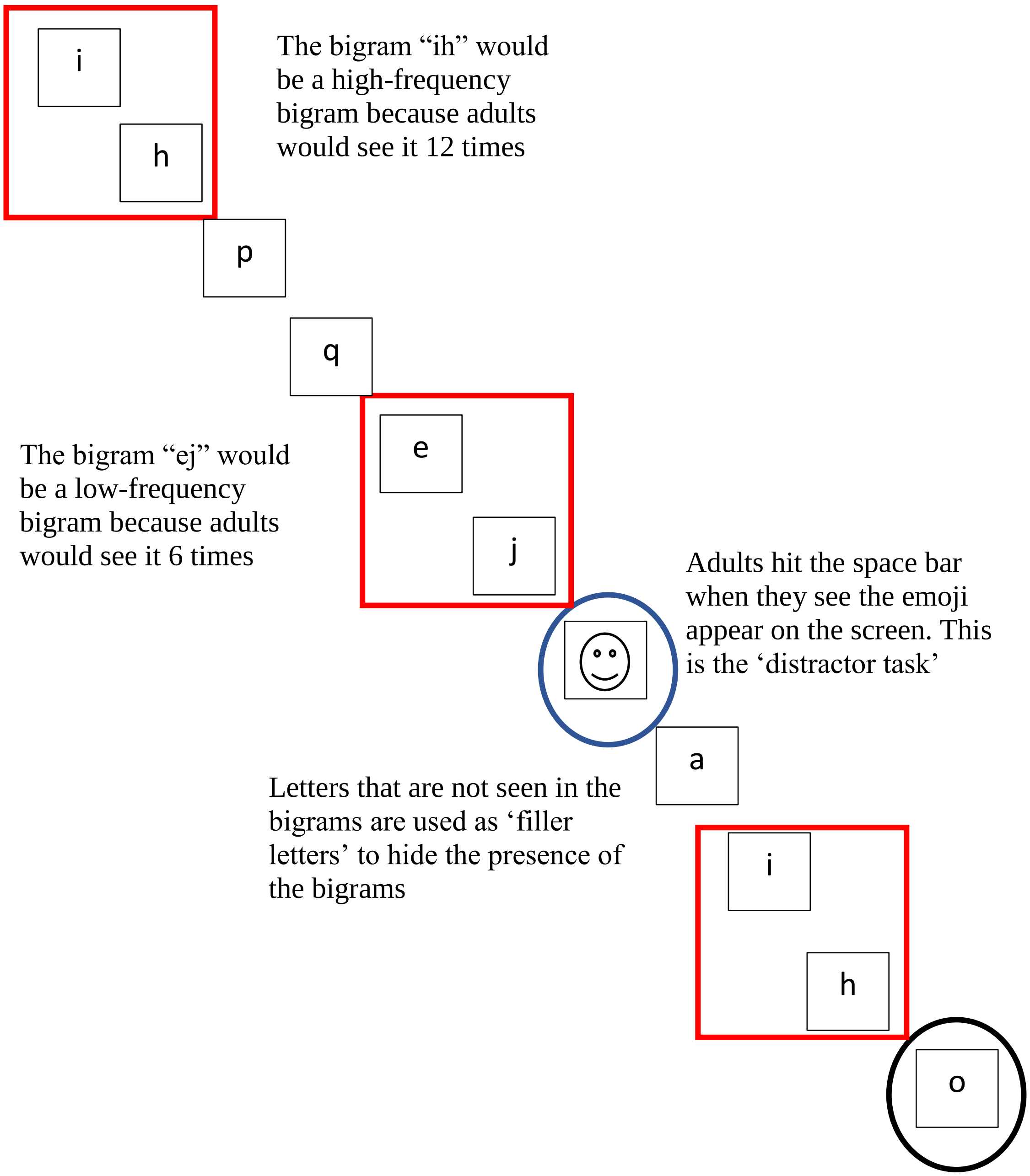
10 low-frequency bigrams were selected from Solso and Barbuto (1979). Bigrams were selected on perceived likelihood of pronunciation. Each row was balanced by consonant-vowel, or consonant-consonant. High-frequency bigrams were seen twelve times, and low-frequency bigrams were seen six times by participants in the statistical learning phase.

High-Frequency	Low-Frequency
IH	EJ
DP	WF
GB	RZ
UX	AQ
KM	SN
= 12	= 6

Appendix C

Statistical Learning Phase

Each box represents a computer screen. Adults are exposed to 20 letters, 16 times individually with the inclusion of the ten bigrams that were selected from Solso and Barbuto (1979)



Appendix D

Non-Words

These non-words were created through using "Fake Word Generator" for ideas. Three sets of non-words were created. Each set had letter length constraints; 2-four letter words, 6-five letter words, and 2-six letter words. As well, each set had bigram position constraints; position 2-four times, position 3-four times, and position 4 twice. (Letter-length, bigram position)

	IH	**DP**	**GB**	**UX**	**KM**	**EJ**	**WF**	**RZ**	**AQ**	**SN**
A	Do**ih**y (5,3)	E**dp**ow (5,2)	Ro**gb**ot (6,3)	Pis**ux** (5,4)	E**km**i (4,2)	Dw**ej**it (6,3)	O**wf**el (5,2)	Fi**rz**o (5,3)	Nes**aq** (5,4)	E**sn**o (4,2)
B	B**ih**ow (5,2)	Sco**dp** (5,4)	Ra**gb**o (5,3)	T**ux**a (4,2)	Ur**km**as (6,3)	S**ej**er (5,2)	Eno**wf** (5,4)	Wa**rz**i (5,3)	T**aq**i (4,2)	Hi**sn**ov (6,3)
C	Jer**ih** (5,4)	El**dp**i (5,3)	A**gb**i (4,2)	Ap**ux**es (6,3)	o**km**et (5,2)	Cor**ej** (5,4)	To**wf** (4,3)	O**rz**up (5,2)	Fr**aq**y (5,3)	A**sn**int (6,2)

Non-words without manipulated bigrams:

1. Purg
2. Ludat
3. Lashu
4. Togra
5. Offlo
6. Edgra
7. Rivis
8. Nurne
9. Duvi
10. Arno
11. Conow
12. Uppat
13. Supin
14. Mutro
15. Lasti

Source for creating non-words:

Fake Word Generator . (n.d.). Retrieved November 28, 2017, from https://www.feldarkrealms.com/words/#.Wh4iF0qnHIV

Appendix E

Reading Phase- Sample Stories

Non-words are bolded. The manipulated "bigram-frequency" incorporated into the non-words was done during the statistical learning phase.

The **sejer** is one of the hardest musical instruments to learn to play. The **sejer** is a very special instrument. It looks like a bass with more than fifty strings. One must use a **pisux** to manipulate the strings. The **pisux** is a little pincer made of wood. It is very flexible. One pulls the strings of the **sejer** with the **pisux** to make them vibrate. The **pisux** twists the strings. This gives a unique sound that is more or less long depending on the position on the string. Some strings are thick and others are thin, while some strings are very tight and others are not tight. A nice melody can also be played on the **sejer** with only one's fingers.
122 words Grade 4.5
Q: What is the pisux made out of?

People in the village often gathered for a **towf**. This was a party for all the villagers. It was a chance to sing and dance. At the **towf**, people would share a big cake made of **agbi**. **Agbi** is a delicious fruit. It is found in forests surrounding the village. Laura wanted to taste the cake before the end of the **towf.** Unfortunately, the cake tumbled off the table and fell over. Laura spent the rest of the night looking for the plant that grows **agbi.** Finally, Laura found the plant with the treasured fruit behind a waterfall. Laura made a new cake. In the end, nobody noticed anything and the **towf** finished as planned. The **agbi** cake made many people happy.
122 words Grade 4.4
Q: What happened to the cake?

The **doihy** was a mean sea monster that lived in the bottom of the ocean. It gobbled up the seawater with a terrible noise. It swallowed everything in its path. Once, the **doihy** swallowed up large wooden ship called an **orzup.** The **orzup** was close to a rock where the monster was hiding. Stretching its head up from the beneath the rock, the **doihy** revealed its three ugly heads. The sailors tried to climb up the masts of the **orzup.** Not fast enough! The monster devoured each of the sailors on board with one quick gulp. The only thing remaining of the **orzup** was a few wooden boards from the deck. To this day, the **doihy** is greatly feared by sailors everywhere.
122 words Grade 4.5
Q: Where does the sea monster live?

The **eldpi** is the newest tool in writing technology. It can help fix all of your tricky homework needs. The **eldpi** is a pen. It can correct your spelling. It has special ink, called **asnint**, which makes this possible. The **asnint** chamber contains a tiny chip. The chip works just like a computer to spell check your writing. Teacher Smith thinks the **eldpi** is a great tool for students. She says they will be able to spell better. She thinks the **asnint** inside the **eldpi** shows how new technology has made learning easier for students. You can even buy this pen with different scented **asnint**. All your papers can each have a different smell! Your homework will be correct and smell nice!

122 words Grade 4.5
Q: What was the teacher's last name in the story?

An **urkmas** is a tiny animal found in the jungles by the Amazon River. It can live for over 100 years if provided with a rich food source. The **urkmas** likes to eat **nesaq**. **Nesaq** is a type of grass that grows along the water's edge. This grass grows over four feet tall. The **urkmas** has to fight off many other animals to reach the sweet grass. Crocodiles and wild boars like to eat the grass. Once in the water there is no danger. The **urkmas** is the same colour as the **nesaq.** Both are a bluish colour. The fish cannot see the tiny animal because it blends into the **nesaq.** Therefore, these unique animals can eat in peace and live forever!
122 words Grade 4.6
Q: What do crocodiles and wild boars like to eat?