Effects of Frequency and Regularity on New Learning in Preference Readers

Jessica R. Wise

Thesis completed in partial fulfillment

of the requirements of the

Honors Program in the Psychological Sciences

Under the Direction of Prof. Bruce D. McCandliss

Vanderbilt University

April, 2012

**Abstract**

Learning to read English requires both knowledge of grapheme-phoneme (GP) rules and rapid recognition for familiar words, which can be mediated by whole-word (WW) mappings. Previous research has suggested that readers differ in their preferences for GP vs. WW strategies of word reading. In an artificial script training study with literate adults, we tested whether preference learners are affected by aspects of the learning demands relevant to English - GP regularity and word frequency. WW preference learners suffered higher costs when reading irregular words, but showed no frequency effects. Further, WW preference learners were more likely to apply GP rules to regular words. These findings suggest that preference may be a measure of ability to use GP strategy appropriately.

**Introduction**

When learning to read, teachers and parents alike want to ensure that their children are being taught using the best teaching strategies available. Copious research has been done about which reading strategies will lead to the greatest achievement in reading ability. In English, this research has centered around two main strategies. The first is the grapheme-phoneme (GP) rules learning method, in which students are explicitly taught to map certain letters (or graphemes) to sounds (phonemes), which can then be put together to make the sound of a whole word. The second is a whole-language approach, in which learners are not specifically taught the grapheme-phoneme rules of phonics, but are allowed to explore and discover reading on their own through their environment. Such learners start with learning on the whole level and later progresses to the part level. Beginning readers, therefore, tend to start out with looking at words as a whole and are not aware of the individual grapheme-phoneme relationships. At this stage, they are using a whole word (WW) strategy for learning. Proponents of the GP strategy believe students should be taught about phonemic relationships so they are able to decode both familiar and unfamiliar words (Farstrup & Samules, 2002). Proponents of WW strategy believe this approach is too rigid, especially when the large number of irregular words that do not follow these basic GP rules is considered. They believe that an overemphasis on these “artificial” rules can hinder word comprehension, the ultimate goal of reading (Tierney & Readence, 2005). In the past, teachers and even governments have fluctuated on their support for either of these two methods. As more research became available, however, extreme positions have become more moderate. Many teachers now incorporate strategies from both methods, teaching both phonemic decoding and requiring students to learn familiar sight words.

However, when considering the debate over the optimal strategy for learning, one aspect of learning that should not be lost sight of is the learner himself. Each person comes to the classroom with his own set of preferences that will affect how well a teaching method actually works for that particular learner. One such learning preference was proposed by Baron and Strawson in 1976. They found two groups of adult readers that differed significantly on their list reading speed for two different lists. One group of readers read a list of irregular words (those that violate traditional grapheme-phoneme rules such as ‘phlegm’) faster than a list of pseudowords (pronounceable nonwords). These readers were deemed Chinese readers, referring to the Chinese writing system of ideographs representing individual words. A second group of readers showed the opposite pattern: reading a list of pseudowords faster than a list of irregular words. They were deemed Phoenician readers in reference to the Phoenician’s consistent alphabetic writing system. The relative skill on one type of words over the others was seen as a greater familiarity with the strategy used on the superior list due to a preference for that strategy. Previous research examining the effects of such strategy preferences in children (Freebody and Byrne1988, Bowey 2008) has shown that two preference groups differed on reading abilities, but that these differences disappeared after the third grade.

Additionally, we have shown that these reading strategy preferences affect new learning in adult readers trained to read an artificial orthography (Wise, Yoncheva, & McCandliss, 2011). In this study the script was a novel symbol system composed of letter symbols that are then joined to create a single three letter word. This design allowed subjects who are aware of the letter symbols to read words using a GP strategy, but the symbols could be viewed as a whole as well, allowing the possibility of a WW strategy. In the 2011 Wise et al study, participants trained on two sets of words. One was a list of regular words there were instructed to use a GP strategy, and the other a list of irregular words that were instructed to use a WW strategy on. When their performances were compared, we found that although each group was able to achieve a similar accuracy on all trained words (regardless of strategy used), each preference group exhibited a unique weakness not shown by the other group. Specifically, Phoenician preference readers suffered from interference effects: performance on words learned using a WW strategy suffered as they were required to learn an increasing number of words. Chinese readers, on the other hand, suffered in transfer ability: they were slower to read novel words (or transfer words) that followed the grapheme-phoneme rules learned during training compared to trained words. Meanwhile Phoenician readers read these words at the same speed as trained words. These results indicated that preference affects learning ability even when the optimal strategy for each type of word is explicitly instructed.

The Wise et al (2011) study showed that preference readers suffered a deficit in their learning when forced to use their unpreferred strategy. This then begs the question of whether the differences found in the previous study would remain if strategy is not specifically instructed and is left to the decision of the learner and regular and irregular words are intermixed. These conditions will make the learning conditions more similar to those in the classroom, where students are presented with different techniques they can use, but are not forced to use one specific strategy for a given set of words. This will allow the study to better address how preference readers deal with the challenges the English orthography presents. The current study therefore aims to determine if transfer and interference effects are still present when strategy is not forced and both regular words and irregular words are mixed together.

Another factor that will be addressed to make the study more generalizable to classroom learning is the ratio of regular words to irregular words. In the previous study, participants were trained on an equal number of regular and irregular words and were only trained on 24 total words in each strategy. However, in practice, children must learn large sets of words, of which there are more regular than irregular words. Previous research with artificial orthography has shown that, as set size increases, the GP strategy shows an advantage relative to the WW strategy (McCandliss, Schneider, & Smith, 1997). This was not found in the previous study, possibly due to the small set size. In the current study, set size will be increased and will include more regular words than irregular words. This advantage for the GP strategy should magnify differences between Phoenician and Chinese readers. Phoenician readers should perform worse on irregular words since they are more infrequent, and Chinese readers should perform worse overall, especially when learning a second set of words when set size increases.

The last factor that should be considered in this research is the method used to determine preference. There is some conflicting research over whether the classification method used to classify readers in the Wise et al (2011) study actually creates meaningfully different groups. Many studies that examine preference use a list reading design (Freebody & Byrne1988, Bowey 2008, Bowey & Rutherford 2007) similar to the one used in Wise et al 2011 study. This design is modeled after the results found by Baron and Strawon (1976) and compares list reading time on a list of nonwords to a list of irregular words. Previous research has used this method to identify groups of readers in elementary school that did perform differently on this task (Freebody & Byrne, 1988), but the results with adults is not so conclusive. Brown, Lupker, and Colombo (1996) attempted to replicate Baron and Strawson’s 1976 study, but found no significant differences between the groups on the list reading tasks. Given this body of research, the list reading method may not be appropriate for adults, or a modified set of stimuli may be needed.

This list reading method, however, was not the original classification method used by Baron and Strawson in 1976. The original method is not timed, but measures the same type of information. In this method, participants’ scores on two tasks are compared. The first is a spelling test of difficult to spell words. Two phonetically plausible spellings of a word are presented, and participants must choose the correctly spelled version. Because both options are phonetically plausible (assistant vs asisstant), Phoenician readers should perform poorly, while Chinese readers, who should have memorized the correct spelling, will not have difficulty. The second task tests the subject’s knowledge of the orthographic rules of English. Pseudowords are presented, and participants must decide whether the word is pronounced like an English word or if it is a nonsense word (i.e. ‘caik’ would be pronounced ‘cake’). Chinese readers would have difficulty with this task as these words are unfamiliar and their pronunciation is not known, while Phoenician readers would have little difficulty sounding out the words. This method focuses on reader knowledge and eliminates individual differences in reading time, and therefore might be a clearer indicator of preference than the previously relied upon list reading paradigm. Given these potential advantages of the original 1976 classification, this method will be used in the current study.

Additionally in order to examine other group differences hypothesized about preference readers, we will also examine the effects of frequency and regularity on test performance. If Brown et al’s (1996) hypotheses about preference groups are correct, Phoenician reader should suffer from regularity effects, but not frequency. Because Phoenician readers rely on a GP strategy for all words, the number of exposures of a word should not affect performance and they should perform similarly on high and low frequency words. However, they should have difficulty with irregular words since they are not able to apply the GP strategy to these words. Conversely, Chinese readers should suffer from frequency effects, but not necessarily regularity effects. Because they rely on memorizing the word as a whole, they should be highly affected by the number of exposures and perform better on high frequency words than low frequency words. They should not however, be affected by regularity since they apply a WW strategy to all words. This additional factor should reveal important differences between preference groups not previously seen in training as well as lend support to the idea that these factors affect Phoenician and Chinese readers differently, as postulated by Brown et al (1996).

An artificial orthography training paradigm can add some additional information to this issue of whether preference exists in adults. Because it uses unique symbols not found in any other writing system, an artificial orthography allows researchers to mimic the early reading process when preferences are most likely to manifest, in a manner that controls for various factors including previous experience and the orthographic design. This training paradigm may be able to demonstrate differences between the two groups that were not exhibited when reading lists of words in English, a language they are already experts at reading.

**Current Study**

In the current study, we seek to answer the previously described questions and determine if the results from the previous study occur as well as extend it to answer some new ones. Specifically, we will investigate whether Phoenician readers will still suffer from interference effects and Chinese readers will suffer from transfer costs if the word sets are constructed more similar to a classroom learning environment and strategy is left to the discretion of the learner. We will also investigate how the different preference groups are affected by frequency and regularity. We hypothesize that the previous transfer and interference results will be even greater now that preference readers are relying more exclusively on their preferred strategy. In line with previous research, we predict Phoenician readers and Chinese readers will show no differences in their overall performance, but that Phoenician readers will suffer from regularity effects and Chinese readers will suffer from frequency effects.

**Methods**

**Participants**

**Preference Assessment.** In order to qualify for participation in the learning assessments, participants were required to complete an online survey. This survey included a spelling test and orthographic rules task using the stimuli from the Brown et al (1996) modeled after those used by Baron and Strawson (1976). The stimuli used by Baron and Starwson were not published, so it was not possible to use the same stimuli as in the original study. See Appendix B for full list of stimuli. Survey data from a total of 131 participants was collected. Errors from the orthographic rules task were subtracted from the spelling task such that positive scores indicated a Chinese reader and a negative score indicated a Phoenician reader. Errors were compared in this manner to ensure that a participant’s preference was determined by his own scores, not that of the overall group. Therefore high or low scores on either test did not determine preference. Scores ranged from -7 to 7 with a mean difference score was 0.1 and a standard deviation of 2.9.

Native English speakers with a score of at least two (positive or negative) were invited to complete the learning study in the lab. This cut off score was designed to ensure that all participants in the main experiment were at least a half of a standard deviation away from the mean and allowed provided a larger subject pool to draw from than if a higher cut off score was used. In all, 47 Phoenicians and 48 Chinese readers were invited to participate. A total of 36 participants agreed to participate in the learning study. For these participants, the average Phoenician reader score was -3.5 and the average Chinese reader score was 3.5.

**Learning Task.** Participants were excluded if scores on the combination of reading tests were below the 16th percentile, data were not collected from all testing sessions (1), they reported a learning disability (2), or were native speakers of another language (1). Final analysis included 32 participants, 16 of each preference group.

The groups did not differ in terms of either mean age or years of education. The final group did not exhibit any reading difficulties as measured by scores above the 16th percentile on a combination of the Nelson Denny Reading Rate and the Nonword Repetition Task from the Comprehensive Test of Phonological Processing. Demographics for each group are reported in Chart 1.

**Stimuli**

Participants were trained on an artificial script similar to that used in Yoncheva, Blau, Maurer, and McCandliss (2010). The script is composed of letter like characters that are combined to make words that can also be viewed as a whole. Words were composed of 10 consonants (b, d, k, m, n, r, s, t, p, and g) and four vowels (a, e, i, and u). See Appendix A for an example. Each participant learned two sets of 30 words each. Each set had two frequency groups, one group that was presented seven times during a training session, and another group that was presented three times. Each frequency group had 10 regular words and five irregular words. This resulted in four types of words: High Frequency Regular Words (HighReg), High Frequency Irregular Words (LowIrreg), Low Frequency Regular Words (LowReg), and Low Frequency Irregular Words (LowIrreg). In the regular word list, each consonant was used in each position, and vowels were as evenly distributed as possible. Irregular words were created so that one letter in any of the three positions differed from its paired sound (for example, ‘bed’ was paired with ‘bad’). Rules were created to account for the irregular pronunciation of words. For example, if the letters ‘ad’ are preceded by a ‘b’, then the ‘a’ is actually pronounced as ‘e’. Some of these rules had greater evidence than others (i.e. one rule is applied to two irregular words while another is applied to only one).

**Procedure**

Participants took part in a two day study during which they were trained on an artificial script as described above. Each day began with a standardized assessments session. Nelson Denny, WTAR, and TOWRE were completed on the first day and NWR and MD on the second day. The first day continued with letter training during which participants learned the sounds the individual letter symbols make. Participants trained on these letters until they reached an accuracy criterion of 85% (two errors) on a test following training. Each letter was presented three times during training before testing. The average number of cycles to reach accuracy criterion was 1.5.

Following letter training, participants began training on a set of word symbols. This was followed by a test of words they learned as well as a test of six ‘transfer’ words they were not trained on in order to test their ability to transfer their knowledge to novel words. After the first training session, participants did another task as part of a different experiment before training for a second time. The same procedures and tests were used for this second training session. On the second day, participants followed the same training procedures previously described with a second set of word symbols. However, in place of letter training, a retention test was completed to measure how much was remembered from the previous day’s training. After all training sessions were completed on the second day, participants took final tests on all words learned as well as a set of transfer words.

Participants were instructed in both learning strategies prior to the first training. They were told that they could use whichever strategy they preferred, but once they decided which strategy they like, they should stick to it for the rest of training. They were also informed that they would be tested on words they were trained on as well as some new words they were not trained on in order to allow participants to make an informed decision about their strategy.

During the course of a training session, high frequency words were seen seven times each and low frequency words were seen three times each. In order to ensure that presentations of each word were equally distributed throughout the course of training (i.e. all three presentations of a low frequency word were not presented in close succession), words were broken up into 10 smaller lists of ten words. Each word was then randomly assigned to either seven or three of these lists (depending on frequency). Word symbols w symbilsyere presented for a total of 4000ms. After 2500ms, the auditory word for that symbol was played, which lasted approximately 600ms. During the test, a word symbol was presented for 250ms followed by a spoken word. Participants were instructed to decide whether the spoken word matched the symbol based on what they learned in training and had 3000ms to respond. Each regular word was paired with a foil from the list of trained words. Each foil shared either 1 or 2 letters in common with the target word. Each target word was presented once, either as a ‘yes’ response or a ‘no’ response. In this design, participants saw each word once, but heard half of the words twice and did not hear the other half at all. However, the ‘yes’ and ‘no’ words were switched during the second training session so that over the full course of training, each word was presented once as a ‘no’ trial and once as a ‘yes’ trial. Each irregular word was presented twice, once as a ‘yes’ trial and once as a ‘no’ trial. The foils for irregular words were the regular pronunciation of the word symbol. Final tests were constructed in the same manner.

**Results**

**Standardized Assessments**

There were no significant differences between the two preference groups on any of the standardized tests, all *t’s* < 1.5 (30), *p’s* >.2. In addition to comparing performance on each test a z-score was determined for each test for each participant in order to compare relative performance on each task. These z-scores were then averaged together to examine overall test performance. Again, there were no significant differences between the two groups, all *t’s* <1.4 (30), *p’s* >.15. Lastly, a correlation between these averaged z-scores and preference score was performed to determine whether preference affected reading ability. This was performed using the preference score to determine whether one preference has an advantage in reading ability as well as the absolute value of the preference score to determine whether a strong preference in either direction is related to reading ability. However, there was no significant correlation for either of these factors, all *r’s*(30) <0, *p’s* <.5.

**Learning Tasks**

Accuracy and reaction times for each condition (HighReg, LowReg, HighIrreg, and LowIrreg) was compared at each time point. Additionally, tests were collapsed across regularity, frequency, session, and day in order to determine larger learning patterns. Analysis of reaction times revealed no significant differences not revealed by accuracy, so the remaining analyses were done on accuracy data. Overall there was a main effect of regularity with accuracy for regular words being greater than that of irregular words regardless of preference, all *t’s*(30) >6, *p’s* <.01. There was a significant main effect of frequency at Day 2 Session 1 words and Day 2 Final Test*,* all *t’s*(30) >4*, p’s* <.01. Although the effect was not significant at other test points, the data do trend towards this all *t’s* (30) <1.5, *p* <.1. Additionally, there was a main effect of session and day. All participants performed better on Session Two compared to Session One, and Day 2 compared to Day 1, all *t’s* (30) > 2, *p’s* <.04.

At only one individual time point was there a significant difference between Phoenician and Chinese readers. Phoenician readers were significantly more accurate on Day 1 Session 1 HighIrreg words, *t*(30) = 2.4, *p* = .02. When data were collapsed across frequency or regularity, no significant differences were found all *t’s*(30) <1.5, *p’s* >.1. This indicates that all groups were able to learn words of each type to a similar degree as well as apply the decoding method as indicated by high performance on the transfer words. See Figure 1.

**Final Tests Analysis**

In order to examine the most robust effects, further analysis on regularity, frequency, interference, and transfer affects were assessed at the final test sessions collapsed across words learned on each day. This should show the most lasting effects since all training is complete.

**Regularity Effect.** In order to examine the effects of regularity on different preference groups, data were collapsed across frequency, and regular word performance was compared with irregular word performance. There was a significant interaction between preference and regularity at the final test session *F*(1,15) = 4.91, *p* = .043. To examine this further, a regularity cost score was determined to examine how regularity affected each preference group. This regularity cost score was determined by subtracting irregular word performance from regular word performance. Chinese readers showed a greater cost for irregular words than Phoenician readers *t*(30) = .027, *p* = .02. See Figure 2.

This result was inconsistent with the hypothesis that Chinese readers perform better on irregular words than Phoenician readers. To examine this inconsistency further, error rates for different test trial types were compared. Irregular words were tested twice, once as a ‘yes’ match in which the sound and symbol were consistent training and once as a ‘no’ match in which the sound and symbol were consistent with regular rules, but not with training. By comparing the error rates for these two types of trials, we can determine how much one group incorrectly applies the regular rules to irregular words. Chinese readers showed more errors on ‘no’ trials relative to Phoenician readers (*t*(30) = 2.8, *p*<.01), indicating that they are over generalizing the grapheme-phoneme rules they learned during training. There was no such difference between the two groups on ‘yes’ trials, *t*(30) = .8, *p* = .4. See Figure 3.

**Frequency Effect.** Similarly, in order to examine frequency effects, data were collapsed across regularity, and high frequency words were compared with low frequency words. However, in this case there was not a significant interaction between preference and frequency *F*(1,15) = 2.19, *p* = .16.

**Transfer Effect**. In order to determine transfer ability, transfer words were compared with regular words collapsed across frequency. There was no significant interaction at the final test point *F*(1,15) = .09, *p* = .76, but throughout training, Chinese readers performed significantly better at transfer words compared to regular words, all *t’s*(15) >2, *p’s* <.05. Phoenician readers showed no significant differences between transfer words and regular words all *t’s* (15)<1.3, *p’s* >.19. See Figure 4. In order to examine this further, a transfer effect score was calculated by subtracting transfer word performance from regular word performance when data were collapsed across all training sessions. This difference, however, was not significant *t*(30) = 1.5, *p* = .13. See Figure 5.

**Interference Effect**. Because interference effects are characterized by a difference in performance after additional learning, this comparison differs from the previous ones. The Retention Test of Day 1 words served as a pre-test and the Final Test of Day 1 words served as a post-test after Day 2 words were learned. Both tests took place on the same day, and without additional training on Day 1 words. Both Phoenician and Chinese readers performed better on the retention test than the final test *t*(15) = 3.4, *p* <.01. See Figure 6. However, there was no interaction between preference and test time point *F*(1,15) = .19, *p* = .67.

**Session and Day Effects.** Although there was a main effect of session and day, these effects were compared within each group to determine if one group was driving the main effect. Data was collapsed across frequency and regularity in order to assess overall learning patterns. There was no significant interaction between preference and session or preference and day, indicating both groups improved with more training, *F*(1,15) <1.5, *p* > .2.

**Discussion**

This study aimed to determine whether preference group differed on frequency and regularity effects, as well as to determine whether transfer and interference effects were still present when strategy was not forced in the context of interspersed regular and irregular words. Overall, there were little significant differences between the two groups on frequency, regularity, and interference under these testing conditions. However, the low number of significant effects could be due to high variability in the samples, as indicated by several trends in the data that did not reach significance. The large nominal differences that were not significant may become so if a larger sample size were used.

Although we results for transfer effects were not as expected, this may be because of the learning conditions in this experiment. All participants were instructed on both possible strategies and informed them that they would have to take a test on words they were not trained on before training began. This may have influenced participants to choose the GP strategy despite their natural inclination so that they might accomplish the test of untrained words. Participants were not asked which strategy they intended to use or if they stuck to that strategy the whole time (though they were asked to). Another factor that may have influenced Chinese readers to perform better on transfer words than regular words is the testing conditions. Transfer words were done in a completely separate test and participants were warned that they would be dealing with untrained words. Regular words, however, were mixed with other trained words. Because they knew they would have to rely on decoding, Chinese readers may have switched strategy or concentrated more on decoding, rather than relying on a WW strategy they may have applied to at least some regular words.

The most salient result of this study was the types of errors on irregular words. Chinese readers were more likely to ‘regularize’ irregular words and agree that the regular pronunciation of an irregular word is correct, despite having learned differently in training. This systematic error brought their performance below chance level. This may be because of the 2:1 ratio of regular words to irregular words. This meant the likelihood of encountering an irregular word was lower than a regular word. This may mean that Chinese readers were more likely to forget their training on irregular words and rely on the decoding skills they were able to learn through training on regular words, which is indicated by high performance on transfer words. This also fits in with the hypothesis that Chinese readers would perform poorly on less frequent (or improbable) words.

These results could mean that ‘preference’ is really a measure of ability to incorporate more complex GP rules that are seen in irregular words. Because irregular words in this study were not completely arbitrary, but only one letter was changed under certain conditions, these words can be viewed as complex regular words. Although this complexity gave both groups trouble, it seems to have affected Chinese readers more since they were more likely to agree with the regular pronunciation than Phoenician readers.

When learning English, Phoenician readers’ grasp of complex GP rules may allow them to make relatively few errors when relying on their decoding skills for both irregular and regular words. They may therefore rely on this strategy more, and when completing tests of irregular words make mistakes when reading highly irregular words. This also manifests in the Baron and Strawson classification spelling test when they are presented with a forced choice with a distractor that is equally likely to be correct, given their knowledge of GP rules. Therefore, they are more likely to make mistakes on trials where they must decide between words like putrify and putrefy but not on words like whistle and whisle. ‘Putrify’ is highly probable phonetically, so they are more likely to make a mistake, but ‘whisle’ is much less probable and they do not make a mistake.

Chinese readers, on the other hand, may be able to apply decoding strategies to only highly regular words. In a classroom setting where they receive feedback on their pronunciation, they are aware that their GP knowledge does not work on words with complex GP rules and may be forced to rely on a WW strategy to learn them correctly. Their awareness would then lead them to study the spelling of these words rather than rely on GP rules. Therefore, they were not affected by equally probable spellings because they had previously learned the specific correct spelling of that word.

If this theory of preferences is correct, the key difference between the two groups may not be skill at applying one strategy over another, but ability of inferential learning of pronunciation of more complex combinations of letters. This notion is consistent with Beech’s 2002 study examining how decoding ability affects reading abilities. Further research should be done to examine performance between the two groups on words that have varying amounts of evidence on a training paradigm (for example, there is a large amount of evidence that the ‘a’ in bad is pronounced the same as the ‘a’ in cat). These groups should also be compared on their orthographic knowledge in a more explicit manner in order to determine what types GP knowledge they differ on. This could better characterize the differences between the two preference groups, which could then lead to classification tests that better differentiate between readers.

Additionally, further research should be done to determine how feedback affects strategy choice. In this study, subjects were not given any feedback on their learning progress. It could be that if feedback were given, participants would have been able to recognize the weakness in applying the GP strategy to irregular words and adapt by switching to a WW strategy. Chinese subjects may be more likely to switch strategies than Phoenician readers, which would then allow them to perform better on irregular words, as they have been shown to do with tests in English. This could provide important insight into how learners acquire a preference, as well as what skills should be focused on to improve reading skills in preference readers.

**References**

Balota, D. A., Cortese, M. J., Sergent-Marshall, S. D., Spieler, D. H., & Yap, M. (2004). Visual word recognition of single-syllable words. *Journal of Experimental Psychology: General, 133*(2), 283-316.

Baron, J., & Strawson, C. (1976). Use of Orthographic and Word Specific Knowledge in Reading Words Aloud. *Journal of Experimental Psychology-Human Perception and Performance, 2*(3), 386-393.

Beech, J. (2002). Individual differences in mature readers in reading, spelling, and grapheme-phoneme conversion. *Current Psychology: Developmental, Learning, Personality, Social*, *21*(2),

Bishop, C. (1964). Tramsfer Effects of Word and Letter Training in Reading. *Journal of Verbal Learning and Verbal Behavior, 3*, 215-221.

Bitan, T., & Karni, A. (2003). Alphabetical knowledge from whole words training: effects of explicit instruction and implicit experience on learning script segmentation. *Cognitive Brain Research, 16*(3), 323-337.

Bitan, T., & Karni, A. (2004). Procedural and declarative knowledge of word recognition and letter decoding in reading an artificial script. *Cognitive Brain Research, 19*(3), 229-243.

Bitan, T., Manor, D., Morocz, I. A., & Karni, A. (2005). Effects of alphabeticality, practice and type of instruction on reading an artificial script: an fMRI study. *Cognitive Brain Research, 25*(1), 90-106.

Bowey, J. A. (2008). Is a ‘‘Phoenician’’ reading style superior to a ‘‘Chinese’’ reading style? Evidence from fourth graders. *Journal of Experimental Child Psychology, 100*, 186-214.

Bowey, J. A., & Rutherford, J. (2007). Imbalanced word-reading profiles in eighth-grade students. *Journal of Experimental Child Psychology*, *96,* 169–196.

Brown, P., Lupker, S. J., & Colombo, L. (1994). Interacting sources of information in word naming - a study of individual-differences. *Journal of Experimental Psychology-Human Perception and Performance, 20*(3), 537-554.

Desroches, A. S., Cone, N. E., Bolger, D. J., Bitan, T., Burman, D. D., & Booth, J. R. (2010). Children with reading difficulties show differences in brain regions associated with orthographic processing during spoken language processing. *Brain Research, 1356*(14), 73-84.

Farstrup, A. E., & Samuels, S. J. (2002). *What research has to say about reading instruction*. Newark, DE: International Reading Assoc.

Freebody, P., & Byrne, B. (1988). Word-Reading strategies in elementary school children: relations to comprehension, reading time, and phonemic awareness. *Reading Research Quarterly, 23*(4), 441-453.

McCandliss, B. D., Posner, M. I., & Givon, T. (1997). Brain plasticity in learning visual words. *Cognitive Psychology, 33*(1), 88-110.

McCandliss, B. D., Schneider,W., & Smith, T. (1997). Learning to read new visual symbols as integrated wholes or component parts. Paper presented at the 38th Annual Meeting of the Psychonomic Society.

Muter, P., & Johns, E. (1985). Learning Logographies and Alphabetic Codes. *Human Learning, 4*, 105-125.

Seghier, M. L., Lee, H. L., Schofield, T., Ellis, C. L., & Price, C. J. (2008). Inter-subject variability in the use of two different neuronal networks for reading aloud familiar words. *Neuroimage*, 42, 1226–1236.

Tierney, R. J., & Readence, J. E. (2006). *Reading strategies and practices*. (6th ed.). Boston: Pearson.

Treiman, R. (1984). Individual differences among children in spelling and reading styles. *Journal of Experimental Child Psychology, 37*, 463–477.

Wise, J., Yoncheva, Y., & McCandliss, B. (2011). Effects of preference and strategy on learning to read an artificial script. *Indiana University Undergraduate Journal of Cognitive Science*, *6*,

Yoncheva, Y. N., Blau, V. C., Maurer, U., & McCandliss, B. D. (2010). Attentional focus during learning impacts N170 ERP responses to an artificial script. *Dev Neuropsychol, 35*(4), 423-445.

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

**Table 1**

*Demographic Profiles and Mean Standardized Test Performance Scores of Participants by Preference*

|  |  |  |
| --- | --- | --- |
|  | **Phoenician** | **Chinese** |
| N | 16.0 | 16.0 |
| Males | 5.0 | 4.0 |
| Preference Rank | -3.6 | 3.6 |
| Age | 20.8 | 21.0 |
| Education | 14.3 | 14.1 |
| Nelson Denny RR | 62.0 | 71.6 |
| ND Qs | 0.4 | 0.1 |
| WTAR | 43.9 | 43.6 |
| TOWRE | 69.5 | 68.4 |
| MD | 80.3 | 80.5 |
| NWR | 45.3 | 40.1 |

*Note.* RS = raw score; PR = percentile ranking.

**Figure Captions**

*Figure 1.* Accuracyfor all word types collapsed across frequency and regularity for each testing session. No significant differences indicate that both groups performed similarly for all types of words.

*Figure 2.* Size of regularity effect score for each preference. This difference indicates Chinese readers had a higher cost of accuracy for irregular words compared to Phoenicians.

*Figure 3.* Number of errors for each trial type for irregular words. Chinese readers made more errors on ‘no’ trials indicating over application of GP rules.

*Figure 4.* Accuracy for transfer and regular words for each preference. Chinese readers show significant improvement for transfer words over regular words while Phoenicians show no difference.

*Figure 5.* Size of transfer effect score for each preference.

*Figure 6.* Accuracy for irregular words on retention test and final test of Day 1 words. Both preferences exhibit interference effects.

**Figure 1**

**Figure 2**

\*

\*

**Figure 3**

\*

**Figure 4**

\*

\*

\*

\*

\*

**Figure 5**

**Figure 6**

\*

\*

Appendix A: Sample Stimuli

“b” “a” “d”

**Appendix B**

**Classification Task Stimuli**

**Spelling Test Stimuli**

|  |  |
| --- | --- |
| Incorrect | Correct |
| Asisstant | Assistant |
| Atempt | Attempt |
| Beseige | Besiege |
| Boistrous | Boisterous |
| Comission | Commission |
| Curteous | Courteous |
| Crape | Crepe |
| Currancy | Currency |
| Despare | Despair |
| Elagent | Elegant |
| Facinate | Fascinate |
| Gorgous | Gorgeous |
| Hemesphere | Hemisphere |
| Honorible | Honrable |
| Insistant | Insistent |
| Interpilate | Interpolate |
| Invoise | Invoice |
| Incorrect | Correct |
| Lascerate | Lacerate |
| Ligiment | Ligament |
| Loveable | Lovable |
| Militery | Military |
| Narative | Narrative |
| Oscilate | Oscillate |
| Overwelm | Overwhelm |
| Perenial | Perennial |
| Primative | Primitive |
| Putrify | Putrefy |
| Quarentine | Quarantine |
| Renevate | Renovate |
| Sesession | Secession |
| Testiment | Testament |
| Terpitude | Turpitude |
| Unanamous | Unanimous |
| Whisle | Whistle |

**Appendix B**

**Orthographic Rules Test Stimuli**

|  |  |
| --- | --- |
| Psuedohomophone | Nonword |
| mait | freest |
| turse | hane |
| groop | preet |
| fownd | gurst |
| gess | tunce |
| muel | feap |
| hokes | turth |
| phocks | bleef |
| gide | klite |
| feal | wirch |
| berd | mong |
| hoap | vawl |
| plaie | pern |
| sope | neech |
| coph | jaul |
| coad | turle |
| grean | bruve |