INSTRUCTIONAL METHODS FOR PROMOTING THE DEVELOPMENT OF ORTHOGRAPHIC AND PHONOLOGICAL KNOWLEDGE IN SECOND LANGUAGE LEARNERS OF INDIC LANGUAGES

by

Adeetee Bhide

Bachelor of Arts, Washington University in St. Louis, 2011

Masters of Philosophy, University of Cambridge, 2012

Submitted to the Graduate Faculty of the

Kenneth P. Dietrich School of Arts & Sciences in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

University of Pittsburgh

2017

UNIVERSITY OF PITTSBURGH

DIETRICH SCHOOL OF ARTS & SCIENCES

This dissertation was presented

by

Adeetee Bhide

It was defended on

April 6, 2017

and approved by

Julie Fiez, Professor, Department of Psychology

Marta Ortega-Llebaria, Assistant Professor, Department of Linguistics

Natasha Tokowicz, Associate Professor, Department of Psychology

Dissertation Advisor: Charles Perfetti, Distinguished Professor, Department of Psychology

INSTRUCTIONAL METHODS FOR PROMOTING THE DEVELOPMENT OF ORTHOGRAPHIC AND PHONOLOGICAL KNOWLEDGE IN SECOND LANGUAGE LEARNERS OF INDIC LANGUAGES

Adeetee Bhide, PhD

University of Pittsburgh, 2017

In three experiments, I test whether the application of particular instructional principles improves the teaching of the orthographic and phonological systems of Indic languages to second language learners. In Experiment 1, I developed a mobile game that teaches 4th grade children Hindi decoding skills, with an emphasis on complex akshara. There were two versions of the game that varied in terms of stimuli spacing (narrow and wide). I found that the game improved participants' akshara recognition and their ability to read and spell words that contain complex akshara. Both versions of the game yielded equivalent levels of improvement, but participants played the narrow spacing version faster. Analysis of the game data revealed interesting patterns of common mistakes. Children struggled with akshara that were non-linear and opaque. When spelling words, children struggled when the complex akshara crossed a syllabic boundary and they often made phonological errors. In Experiment 2, I examined whether motor encoding and testing benefit orthographic learning. I found that motor encoding benefits orthographic learning when tasks require pure orthographic knowledge or the production of an orthographic form when given a phonological form. Testing does not benefit beginning learners. In Experiment 3, I tested whether pedagogical differences or individual differences affect the learning of non-native phonemic contrasts. I found that learning of the difficult dental/retroflex contrast can be improved by increasing the voice onset times of the dental sounds. Both English phonological skills and rise time discrimination positively predict learning the non-native contrasts. Furthermore, pairing phonemes with English transliterations impairs discrimination learning, likely because of

interference from the English pronunciation. Orthographic support helps people remember which phonemes are in words. Therefore, the use of akshara can benefit second language learners because the graphs are not already associated with phonological referents and the graphs help people remember which phonemes are in vocabulary words. When considered together, these three experiments suggest that multisensory encoding and reducing interference benefit second language learners.

TABLE OF CONTENTS

PREFACE..XVI

1.0 GENERAL INTRODUCTION 1

1.1 THE PHONOLOGICAL AND ORTHOGRAPHIC SYSTEMS OF HINDI AND MARATHI 2

1.2 INSTRUCTIONAL PRINCIPLES 5

1.3 SUMMARY 7

2.0 EXPERIMENT 1: MOBILE GAME THAT TEACHES HINDI DECODING 9

2.1 INTRODUCTION 9

2.1.1 Spacing manipulation 10

2.1.2 Overview 12

2.2 METHODS 12

2.2.1 Pre and post-tests 12

2.2.1.1 Non-verbal IQ 13

2.2.1.2 Akshara recognition 13

2.2.1.3 Word reading 13

2.2.1.4 Spelling 14

2.2.1.5 Akshara construction 14

2.2.1.6 Vocabulary 15

2.2.1.7 Math 15

2.2.2 Game design 15

2.2.2.1 Akshara decomposition 17

2.2.2.2 Spelling 18

2.2.2.3 Game play 18

2.2.3 Game development 20

2.2.4 Participants 20

2.2.5 Schedule 21

2.3 RESULTS 22

2.3.1 Pre and post-tests 22

2.3.1.1 Non-verbal IQ 22

2.3.1.2 Akshara recognition 22

2.3.1.3 Word reading 24

2.3.1.4 Spelling 27

2.3.1.5 Akshara construction 30

2.3.1.6 Vocabulary 31

2.3.1.7 Math 32

2.3.2 Game play 33

2.3.2.1 By-item analyses 34

2.3.2.2 By-subject analyses 42

2.4 DISCUSSION 43

3.0 EXPERIMENT 2: COMPARISON OF PEDAGOGICAL METHODS FOR BUILDING ORTHOGRAPHIC KNOWLEDGE 48

3.1 INTRODUCTION 48

3.1.1 Transparency 48

3.1.2 Benefit of motor encoding 49

3.1.3 Benefit of testing 54

3.1.4 Study overview and hypotheses 55

3.2 METHODS 56

3.2.1 Participants 56

3.2.2 Materials 57

3.2.3 Learning methods 58

3.2.3.1 Decompose complex akshara MC 59

3.2.3.2 Compose complex akshara MC 59

3.2.3.3 Copy 59

3.2.3.4 Write 60

3.2.4 Tests 60

3.2.4.1 Hear and choose 60

3.2.4.2 Orthographic legality 61

3.2.4.3 Reading 61

3.2.4.4 Writing 61

3.2.5 Procedure 62

3.2.5.1 Day 1 62

3.2.5.2 Day 2 63

3.2.5.3 Day 3 64

3.3 RESULTS 66

3.3.1 Simple akshara learning 66

3.3.2 Data analysis for tests of complex akshara learning 67

3.3.2.1 Day 2 hear and choose 69

3.3.2.2 Day 3 hear and choose 69
3.3.2.3 Day 3 orthographic legality 70
3.3.2.4 Day 3 reading accuracy 70
3.3.2.5 Day 3 writing accuracy 71
3.3.3 Time on task 71
3.4 DISCUSSION 72
3.4.1 Transparency 72
3.4.2 Benefit of motor encoding 73
3.4.3 Benefit of testing 74
3.4.4 Time on task 76
3.4.5 Discrepancies in the prior literature 77
3.4.6 Use of artificial orthographies 80
3.4.7 Educational implications 81
3.5 CONCLUSION 82
4.0 EXPERIMENT 3: DISCRIMINATION OF PHONEMIC CONTRASTS 83
4.1 INTRODUCTION 83
4.1.1 Hindi and Marathi phonemic contrasts 83
4.1.2 Orthographic support 86
4.1.3 Individual differences in learning phonemic contrasts 88
4.1.3.1 Rise time discrimination 89
4.1.3.2 Phonological awareness and decoding 91
4.1.4 Pilot study 93
4.2 METHODS 94

4.2.1 Participants 94

4.2.2 Overview 94

4.2.3 Procedure 95

4.2.3.1 Repetition 96

4.2.3.2 Introduce phonemes 96

4.2.3.3 Forced-choice identification phonemes 97

4.2.3.4 Forced-choice identification words 98

4.2.3.5 Learning words 98

4.2.3.6 Choose correct word 99

4.2.3.7 Individual difference measures 99

4.2.4 Materials 101

4.3 RESULTS 105

4.3.1 Language History Questionnaire 105

4.3.2 Descriptive statistics 108

4.3.3 Pronunciation 111

4.3.3.1 Pronunciation (all participants) 112

4.3.3.2 Pronunciation with individual differences 113

4.3.4 Forced-choice identification analyses 114

4.3.4.1 Forced-choice identification phonemes 114

4.3.4.2 Forced-choice identification words 117

4.3.5 Choose correct word 119

4.4 DISCUSSION 121

4.4.1 Manipulated utterances 121

4.4.2 Orthography 124
4.4.3 Individual differences 127
4.4.3.1 Rise time discrimination 128
4.4.3.2 Phonological awareness and decoding 129
4.5 CONCLUSION 131
5.0 GENERAL DISCUSSION 132
5.1 SUMMARY OF EXPERIMENT 1 132
5.2 SUMMARY OF EXPERIMENT 2 133
5.3 SUMMARY OF EXPERIMENT 3 134
5.4 THEORETICAL DISCUSSION 136
5.4.1 Multisensory encoding 136
5.4.1.1 Motor encoding 136
5.4.1.2 Orthographic support 137
5.4.2 Reducing interference 138
5.4.2.1 Reducing interference from prior knowledge 138
5.4.2.2 Reducing interference from incorrect responses 138
5.5 EDUCATIONAL IMPLICATIONS 139
5.5.1 Multisensory encoding 139
5.5.1.1 Motor encoding 139
5.5.1.2 Orthographic support 140
5.5.2 Reducing interference 140
5.5.2.1 Reducing interference from prior knowledge 140
5.5.2.2 Reducing interference from incorrect responses 141

5.6 CONCLUSION 141

BIBLIOGRAPHY 143

LIST OF TABLES

Table 1: Examples of different akshara types 3

Table 2: Mean (standard deviation) raw scores on the matrix reasoning assessment. 22

Table 3: Mean (standard deviation) accuracy on the akshara recognition test (in percents). 23

Table 4: Mean (standard deviation) accuracy on the word reading test (out of 8). 25

Table 5: Mean (standard deviation) accuracy on the spelling test (out of 6). 29

Table 6: Mean (standard deviation) accuracy on the akshara construction test (out of 7). 31

Table 7: Mean (standard deviation) accuracy on the vocabulary test (out of 24). 32

Table 8: Average (standard deviation) raw scores on the math assessment. 33

Table 9: Difficult akshara problems. 36

Table 10: Easy akshara problems 37

Table 11: Difficult word problems. 38

Table 12: Easy word problems. 40

Table 13: Simple and complex akshara used in the experiment. 58

Table 14: Procedure overview. 62

Table 15: Distribution of stimuli across tests. 65

Table 16: Means (standard deviations) on the complex akshara tests, by learning method. 68

Table 17: Summary of expected and actual results. 68

Table 18: The graphs in the two orthography conditions 95

Table 19: The tasks included in the pre-test, learning, and post-test phases of the experiment. .. 95

Table 20: The voice onset time (in s) of the K stimuli. 103

Table 21: The voice onset time (in s) of the T stimuli 103

Table 22: The voice onset time and pre-voicing duration (in s) of the D stimuli 104

Table 23: One heritage Marathi speaker's performance on the tasks. 105

Table 24: Accuracy on the sound contrast tasks. 109

Table 25: Performance on the individual difference measures. 109

Table 26: Frequency of response types (in percents) when participants were attempting to say the D and T contrasts 112

LIST OF FIGURES

Figure 1: A screenshot of an akshara decomposition problem. 16

Figure 2: A screenshot of a spelling problem. 17

Figure 3: Examples from the orthographic legality test.. 61

Figure 4: Examples of Hindi movie posters. 88

Figure 5: Examples of the reference cards participants could use during the experiment. 97

Figure 6: Example sound waveforms of K phoneme recordings from one speaker. 102

Figure 7: Example sound waveforms of T phoneme recordings from one speaker. 102

Figure 8: Example sound waveforms of D phoneme recordings from one speaker. 102

PREFACE

This dissertation would not have been possible without the generosity of many people. First, I would like to thank my parents for teaching me Marathi. Without them, I would not be able to speak this beautiful language and would not have become curious about reading development in alphasyllabaries.

Second, I would like to thank my advisor, Dr. Chuck Perfetti, for all of his assistance and support, as well as for giving me the freedom to explore topics that interested me.

Third, I would like to thank Dr. Sonali Nag for inviting me to India to collaborate with her on a project. My time in India was a truly incredible experience and it inspired the rest of this dissertation.

I would also like to thank the rest of my dissertation committee, Dr. Natasha Tokowicz, Dr. Julie Fiez, and Dr. Marta Ortega-Llebaria, for all their help throughout the process. Marta was especially helpful with Experiment 3; it would not have been completed without her.

Dr. Jingtao Wang and Wencan Luo programmed the game used in Experiment 1 and helped me extract the data about the game play. Adrian Maries also helped me with the programming.

Dr. Scott Fraundorf was immensely helpful for all of my data analysis. Dr. Melinda Fricke helped me write my Praat scripts. Dr. Tessa Warren gave me valuable guidance during my mentoring meetings and feedback on my practice talks. Dr. Lori Holt and Casey Roark helped me find literature relevant to Experiment 3 and assisted me with the linguistic descriptions. Dr. Usha Goswami and Nichola Daily provided me with the Rise Time task.

There were many research assistants who directly contributed to this project. Nivita Vijay, Sajma Aravind, Abhineeta Rangunathan, and the other members of the Promise Foundation helped develop and test the mobile game. Nora Kazour and Tara Nair helped me score the data from Experiment 1. Austin Marcus, Sonali Agrawal, and Aarya Wadke all helped me run participants. Austin Marcus, Sonali Agrawal, Aarya Wadke, Elena Cimino, Shalaka Fadnis, and Paula Pascual all helped with the Experiment 3 stimuli.

Finally, I would like to thank everyone whose voice recordings I used in Experiment 3: Dnyanada and Rajeev Bhide, Sudheer Apte, Shirish and Uma Kher, Soniya Gadgil, and Punit Marathe.

I am especially thankful to all the wonderful colleagues I work with for the supportive environment they create. Every member of the Reading and Language Group, both past and present, has helped me in some way with this research. This includes Evelyn Milburn, Michelle Colvin, Katherine Martin, Alba Tuninetti, Chelsea Eddington, Caitlin Rice, Gabby Terrazas, Gina Calloway, Anne Helder, Xiaoping Fang, Erika Laing, Travis Alvarez, Joseph Stafura, Elizabeth Hirshorn, Wendy Chang, Juan Zhang, Benjamin Rickles, Lindsay Harris, Nora Presson, Hannah Legerwood, and Carrie Epp. Kimberly Muth, our fabulous lab manager, helped my write the international IRB and with innumerable other tasks.

Of course, this dissertation would not have been possible without my participants. I would like to thank the Baldwin Girls School for collaborating with me on Experiment 1 and all of my Experiment 2 and 3 participants who came in for my multi-day studies.

Finally, this study would not have been possible without all the friends and family who allowed me to stay with them. While I was in India, I lived with Rajeev Bhide and Vrunda, Vijay, Mihir, and Manasi Gadgil. In Pittsburgh, Soniya Gadgil and Vinod Sharma were infinitely helpful while I was writing my dissertation.

1.0 GENERAL INTRODUCTION

Successful text comprehension is predicated on the ability to recognize words (Gough & Tunmer, 1986). The lexical quality hypothesis identifies three components of word knowledge: orthography, phonology, and semantics (Perfetti & Hart, 2002). Here I focus on two of those components, orthography and phonology, and identify effective methods for building orthographic and phonological knowledge.

Which methods are most effective for building orthographic and phonological knowledge may depend on the properties of the orthography being taught. For example, because Finnish has consistent grapheme-phoneme correspondences, reading instruction that explicitly teaches those correspondences is effective. In contrast, English is very inconsistent at the grapheme level but more consistent at the rime level. Thus, for English, instruction that focuses on both the phoneme and rime level is effective (Kyle, Kujala, Richardson, Lyytinen, & Goswami, 2013). Similarly, both Japanese kana and alphasyllabaries have graphs that represent multiple phonemes. However, in kana, the phonemes represented by one graph are always part of the same syllable. In contrast, in alphasyllabaries, the phonemes represented by one graph can cross a syllabic boundary (Nag, 2014). Thus, for Japanese kana, syllable-level instruction is sufficient. In contrast, for alphasyllabaries, instruction should draw attention to the resyllabification process and, at least partly, focus on the phoneme level.

I study two Indic languages that use alphasyllabic writing systems, Hindi and Marathi. Both languages use the Devanagari script and the mapping between orthography and phonology is nearly identical (see Bhide & Perfetti, 2017). Thus, because the focus of this dissertation is on decoding, both languages can be used interchangeably.

1.1 THE PHONOLOGICAL AND ORTHOGRAPHIC SYSTEMS OF HINDI AND MARATHI

Hindi and Marathi present unique phonological challenges. First, they contain many similar phonemes that only contrast in terms of place of articulation, aspiration, or length (e.g., द /d̪/ and ड /ɖ/ vary by place of articulation; क /k/ and ख /kʰ/ vary by aspiration; इ /i/ and ई /i:/ vary by duration). Second, although the name "alphasyllabary" suggests that graphs map onto syllables, in reality the relationship between graphs and phonological units is not that straightforward.

In addition to being phonologically challenging, Hindi and Marathi are also orthographically challenging. They use an alphasyllabic writing system, in which phonemes combine in a non-linear manner to form open syllabic graphs called *akshara*. Because there are more syllables than phonemes in language, alphasyllabic writing systems have very large grapheme sets (Nag, 2011). A large grapheme set also results in high graphic complexity (Chang, 2015).

The number of phonemes the different akshara represent varies greatly. *Simple* akshara either represent a vowel phoneme or a consonantal phoneme and an inherent schwa vowel. *Consonant-vowel (CV)* akshara have consonant and vowel subcomponents. *Complex akshara* contain two or more consonants and may also have a vowel subcomponent (see Table 1).

Table 1: Examples of different akshara types

Akshara Type	Akshara	Phonology
Simple Vowel	उ	/u/
Simple Consonant	स	/sə/
CV	सु	/su/
Complex	स्तु	/st̪u/

Even simple akshara can be relatively difficult to learn. There are many orthographically similar simple akshara that are easy to confuse (e.g., म /m/ and भ /b^h/). The presence of many orthographically similar pairs can be challenging to novice learners.

Complex akshara present an additional layer of difficulty. First, the rules for concatenating consonants are very complex. The most common way to concatenate consonants is to remove the right-most portion of the first consonant and physically attach it to the second consonant (e.g., स + त = स्त ; ण + ट = ण्ट). However, there are other ways of joining consonants (e.g., ट + ट = ट्ट ; द + व = द्व). All of these methods are relatively transparent; both consonantal forms are easily visible. Other complex akshara are more opaque; their components are not easily visible. For example, whenever र /r/ is the first consonant in a complex akshara, it is represented by a curved line over the second consonant (e.g., र + द = र्द ; र + त = र्त). Whenever r /r/ is the second consonant, it is depicted as either one or two diagonal lines (प + र = प्र ; क + र = क्र ; ट + र = ट्र ; ड + र = ड्र). Finally, some complex akshara are very opaque and need to be memorized (e.g., क + ष = क्ष ; त + त = त्त ; त + र = त्र; द + द = द्द).

Second, complex akshara are often nonlinear. One of the largest sources of non-linearity is the इ /i/ vowel, whose diacritic occurs to the left of the consonants although it is pronounced after the consonants (e.g., प्र + इ = प्रि). The र /r/ is another source of non-linearity. For example, in the akshara र्ता (र + त + आ), the first phoneme (/r/) is found in the upper-right, the second phoneme (/t̪/) is found in the left, and the third phoneme (/a/) is found in the lower right.

The mapping between complex akshara and phonology is also quite challenging. Complex akshara can either represent consonantal blends or two adjacent complex akshara that cross a syllabic boundary. For example, in the word व्यायाम /wja.yam/ (exercise), the complex akshara व्या /wja/ represents a consonantal blend. In contrast, in the word बर्तन /bər.t̪ən/ (pan; kitchen utensil), the complex akshara र्त /r. t̪ə/ represents two consonants that cross a syllabic boundary. Complex akshara that cross a syllabic boundary are more difficult to learn than complex akshara that represent a blend (Nag, 2014).

Another difficulty with learning complex akshara is that, although complex akshara as a type is common, an individual complex akshara is rare. For example, Patel, Bapi, and Nag (2013) identified 702 different akshara in texts for children in grades 1-5. Although 285 of those akshara were complex akshara (40%), only sixty of those complex akshara occurred more than ten times. Of the fifty most common akshara, only three were complex akshara (Nag, 2014). Therefore, although complex akshara recognition is very important for text comprehension, texts may not provide enough examples of a given complex akshara for a child to be able to easily recognize it. Furthermore, many instructors report that a very small percentage of complex akshara are explicitly taught (Nag, 2014; Nag & Sircar, 2008; P. G. Patel, 2004). Thus, instruction that explicitly teaches complex akshara may be beneficial.[1]

It is important to note that Hindi and Marathi's orthographic properties can present learning challenges to both first and second language learners. In contrast, although some of the phonological distinctions may be difficult for first language learners, they are particularly difficult for second language learners. Specifically, Hindi and Marathi have consonants that sound very similar to each other, and only vary in terms of aspiration or place of articulation. People who were not exposed to these close phonemic pairs as children lose their ability to discriminate them, and it can be very difficult to learn these discriminations at an older age (Tees & Werker, 1984). Here, I focus on second language learners of Hindi and Marathi, so they should struggle to learn both the orthographic and phonological properties.

1.2 INSTRUCTIONAL PRINCIPLES

This dissertation compares pedagogical methods for teaching second language learners the orthographic and phonological systems of two Indic languages, Hindi and Marathi. The pedagogical methods incorporate different instructional principles known to benefit learning. The benefit of these instructional principles has been demonstrated in other learning situations, but not for teaching the orthographic and phonological systems of Indic languages. Thus, comparison of

[1] The research cited here was not done with Hindi and Marathi, but rather with other languages that use alphasyllabic orthographies. However, to the best of my knowledge, a similar analysis has not been done with Hindi and Marathi. I am presuming that distributional properties are relatively similar across alphasyllabic orthographies and that instructional principles are similar across India.

the pedagogical methods tests the applicability of those instructional principles in the acquisition of Hindi and Marathi as a second language.

The first study tests an intervention for teaching Indian children Hindi's phonological and orthographic systems. Although the children live in India, they are not native speakers of Hindi. There are two versions of the game, one in which similar stimuli are presented in a grouped manner (narrow spacing) and one in which similar stimuli are presented in a distributed manner (wide spacing). Thus, in addition to testing the effectiveness of the game, this experiment examines whether spacing affects learning. According to the spacing effect, the wide spacing version of the game should be more effective (Cepeda, Pashler, Vul, Wixted, & Rohrer, 2006; Underwood, 1961).

The second study compares four pedagogical methods for teaching native English-speaking adults an artificial orthography strongly modeled on Devanagari. Two of the methods (copying and writing) incorporate motor encoding (Wollscheid, Sjaastad, & Tømte, 2016). Furthermore, writing incorporates testing but copying does not (Roediger & Karpicke, 2006). Thus, I test whether motor encoding and testing benefit orthographic learning.

The third study tests two instructional principles for improving phonemic perception. First, it tests whether manipulating sounds to emphasize their differences benefits learning (e.g., McCandliss, Fiez, Protopapas, Conway, & McClelland, 2002). Second, it tests whether orthographic support benefits learning (Steele, 2005) and if different orthographies vary in their efficacy. In addition to examining these two instructional principles, individual differences in phonological learning are also measured.

Taken together, the three experiments examine how multisensory encoding (Shams & Seitz, 2008) and interference influence learning. The effect of multisensory encoding is explored in Experiments 2 and 3. In Experiment 2, I examine whether motor encoding benefits learning. Motor encoding is believed to be beneficial because it combines motor and visual information. In Experiment 3, I examine whether orthographic support aids phonological learning. Orthographic support is believed to be helpful because it combines visual and phonological information.

The effect of interference is explored in all three experiments. In Experiment 3, I examine how knowledge of English grapheme-phoneme correspondences may interfere with phonological learning when English letters are paired with Marathi phonemes. In Experiments 1 and 2, I examine how interference from errors during learning affect final performance. In Experiment 1, I vary spacing within the game. The wide spacing version should be more difficult, and thus induce more errors. In Experiment 2, I compare copying with writing (from memory). The copying condition is much easier and should elicit almost no errors, whereas the writing condition should elicit many errors. I examine whether the additional difficulty of the wide spacing version and of writing is desirable (McDaniel & Butler, 2011), or if the additional errors create too much interference.

1.3 SUMMARY

The pace of orthographic learning in Indic languages is slow due to their phonological and orthographic intricacies (Nag, 2007). More research on effective methods for teaching people to read Indic languages is needed. This dissertation seeks to identify which instructional principles are useful for teaching people to read Indic languages. In Experiment 1, I test a mobile game that teaches children Hindi's orthographic and phonological systems. Then, in Experiments 2 and 3, I test specific hypotheses regarding how the game could better teach orthographic and phonological properties, respectively.

2.0 EXPERIMENT 1: MOBILE GAME THAT TEACHES HINDI DECODING

2.1 INTRODUCTION

As described in the General Introduction, Devanagari has three main akshara types: simple, CV, and complex. Complex akshara are particularly difficult to learn, a result demonstrated in many alphasyllabaries (Telugu: Vasanta, 2004; Bengali: Nag & Sircar, 2008; Malayalam: Tiwari, 2011; Kannada: Nag, 2007, Nag, Treiman, & Snowling, 2010, and Joshi, 2013). For example, Nag (2007) tested children between grades one and four learning Kannada on their knowledge of 20 akshara: eight simple consonantal akshara, one simple vowel akshara, five CV akshara, and six complex akshara. Children in grade 1 were 71.8% correct at naming simple consonantal akshara, but were near-zero on the other akshara types. By grade 2, children were nearly perfect at naming consonantal simple akshara, but continued to struggle with the other akshara types. Children in grade 4 were only 80% correct at naming the akshara overall. More specifically, although they were nearly perfect at naming consonantal simple akshara, they were on average only 72.5% correct at naming CV akshara, and 55.2% correct at naming complex akshara. Similarly, Tiwari (2011) studied children in grade 3 learning Malayalam and found that though they had a strong grasp on simple akshara, they found CV akshara more difficult, and complex akshara extremely difficult. Out of the six complex akshara they were tested on, children in the 25th percentile got all of them wrong, the median score was three, and children in the 75th percentile got only 4.75 correct. Not only is akshara recognition difficult, but so is production. Both good and poor spellers in grades 4-5 learning Kannada had more difficulty spelling words containing complex akshara than words containing CV akshara, which in turn were more difficult than words containing only simple akshara. Furthermore, the difference between the good and poor spellers was largest on words containing CV and complex akshara (Nag et al., 2010).

Because recognizing complex akshara is both difficult and necessary for proficient reading, we need to know if mobile games can increase akshara knowledge. In this study, I test a mobile game that teaches children these difficult complex akshara. The game teaches students to recognize complex akshara both in isolation and in word contexts. The game also seeks to improve orthographic and phonological knowledge more generally by including close orthographic and phonological foils.

The mobile game format has several benefits for use in educational interventions. First, it allows students to progress at their own pace so that every student is appropriately challenged. Second, it allows for extensive data collection because the game logs every button press the participants make. Thus, in addition to pre- and post-test data, I have detailed data about performance on the intervention itself. Third, the format is very engaging and motivating for students.

2.1.1 Spacing manipulation

There were two game versions that tested whether spacing and desirable difficulties increase learning in the game. In one version of the game, problems about the same akshara were grouped together (narrow spacing). In the other version, problems about the same akshara were presented in a more distributed manner (wide spacing). Previous research has shown that spacing stimuli leads to slower initial learning, but better long-term retention (Cepeda et al., 2006; Underwood, 1961). The experiment realizes the spacing manipulation slightly differently than has been done previously. Most studies utilizing a spacing manipulation repeat the same or very similar problems at different intervals. In contrast, in the present study, the same akshara are presented in different contexts. Specifically, in the narrow spacing version, students learn an isolated akshara and then immediately practice it in a word context. In the wide spacing version, students learn 10 isolated akshara and then practice all of them in word contexts.

In the narrow spacing version, students are exposed to the same akshara twice in a row. In contrast, in the wide spacing version, students are exposed to a given akshara in a more distributed manner. Thus, the schedule of akshara exposure is similar to the traditional spacing manipulation. Furthermore, the narrow spacing version of the game is easier because students know that the akshara they just learned will be in the word. Therefore, students do not need to rely solely on phonology-graph correspondences to spell the word; they already know what one of the correct answers is and they need to only fill in the remaining akshara. Furthermore, at least in the words with only one complex akshara (which is true of all the words in the earlier levels), students can automatically eliminate all of the complex akshara foils. Normally when children are spelling words, they do not know which complex akshara is in the word, and have to rely purely on phonology-graph correspondences. Thus, the wide spacing version is more authentic and is better training students to use phonology-graph correspondences. Because the wide spacing version requires students to spell words from scratch, it may take longer to play this version but lead to better learning outcomes. This prediction is in line with the literature on desirable difficulties (McDaniel & Butler, 2011).

However, the narrow spacing version does have one advantage over the wide spacing version. Because the isolated akshara and the akshara-in-word-context are shown consecutively, students can better understand how the akshara they learned functions in a word context. In the wide spacing version, this relationship is harder to discern. Thus, it is possible that the additional difficulties created by the wide spacing version will not be desirable in this instance (McDaniel & Butler, 2011).

2.1.2 Overview

The efficacy of a mobile game that teaches Hindi decoding skills is measured using a pre-test, post-test format. There are two versions of the game that vary in the order of their stimulus presentation. The two groups that play the game are compared to an unseen control group. In addition to analyzing the data from the pre and post-tests, the data from the game play itself is also analyzed. The pre and post-test data is primarily used to test the efficacy of the game. The data from the game play is primarily used to better understand the nature of akshara learning and to identify orthographic and phonological aspects of Hindi that are challenging for learners.

2.2 METHODS

2.2.1 Pre and post-tests

Pre and post-tests included measures of the children's Hindi akshara recognition, reading, spelling, and math abilities and their knowledge of vocabulary and ligaturing rules. Non-verbal IQ was measured at pre-test only. The testing took approximately one hour per child.

2.2.1.1 Non-verbal IQ

The matrix reasoning subtest from the Wechsler abbreviated scale of intelligence (WASI) was administered at pre-test only (Wechsler, 1999).

2.2.1.2 Akshara recognition

Children were presented with 20 akshara (9 simple, 6 CV, 5 complex) and were asked to read them aloud as quickly as possible. The children were given 6 akshara as practice items before the test began. The number of correct responses, error type, and time to complete the task were documented. The task was recorded to allow a second experimenter to check the scoring.

2.2.1.3 Word reading

The children were asked to read 48 words, presented in 6 lists of 8 words each. The first two lists consisted of words that did not contain complex akshara. The words in the first list were composed of simple akshara and were not taught in the game (i.e., simple list). The words in the second list all had two CV akshara and were not taught in the game (i.e., CV list). The words in the final four lists all had one complex akshara and tested differing levels of transfer. The words in the third list were taught between levels 4 and 20 in the game (i.e., learned list). The words in the fourth list were not taught in the game, but the complex akshara within them was taught between levels 4 and 20 in the game (e.g., नष्ट, wherein ष्ट was taught in the game; i.e., near transfer list). The words in the fifth list were not taught in the game and their full complex akshara was also not taught in the game. However, the consonants within the complex akshara were taught between levels 4 and 20 in the game, but paired with a different vowel (e.g., चन्दा, wherein although न्दा was never taught, न्द was taught; i.e., medium transfer list). The words in the sixth list were not taught in the game, and the consonants within the complex akshara were never taught in the game (e.g., अग्नि, wherein ग्न was never taught; i.e., far transfer list). The children were given unlimited time to read the words. If a child got a score of zero on five consecutive words within one list, the testing of that list was discontinued. Three practice items were given before the testing began. The task was recorded to allow a second experimenter to check the scoring.

2.2.1.4 Spelling

The children were asked to spell 30 words. Six types of words were presented from the same categories described in the word reading section (i.e., simple, CV, learned, near transfer, medium transfer, and far transfer). The lists were presented in an interspersed manner; the first word from each list was presented, then the second word from each list, etc. If the child got a score of zero on three consecutive words from one list, the remaining words from that list were not administered.

2.2.1.5 Akshara construction

The akshara construction assessment measured students' knowledge of ligaturing rules and their ability to apply those rules in novel contexts. The children were introduced to a "made-up" akshara and were told to "pretend that it makes the /l/ sound". They were also told that "it can be combined with diacritics and other Hindi akshara that you know". The children then practiced drawing it in isolation, combining it with diacritics, and combining it with other akshara to make complex akshara. After the practice session, they were given six akshara construction problems: two of them were in the CV pattern (/la/, /lu:/), three were in the CCə pattern (/klə/, /blə/, /plə/), and one was in the CCV pattern (/gle/).

2.2.1.6 Vocabulary

The children were asked to define 24 Hindi words. 1/3 of the words were taught during the first 20 levels of the game, 1/3 of the words were morphologically related to words taught during the first 20 levels of the game, and 1/3 of words were not taught during the game. Definitions given in either Hindi or English were accepted, as were English translations. Before testing began, the experimenter demonstrated one item and then had the child practice with two items. The task was recorded to allow a second experimenter to check the scoring.

2.2.1.7 Math

The Math Fluency subtest from the Woodcock-Johnson was administered (Woodcock, McGrew, & Mather, 2001). The children had three minutes to complete as many simple arithmetic problems (i.e., addition, subtraction, multiplication) as possible. The score consisted of the number of correct answers. This test was included to see whether the gains from the intervention were specific to Hindi literacy skills. If the intervention group improved more than the control group on the math assessment (a skill the game was not teaching), that would suggest that the improvements resulted from greater interaction with the experimenters. If not, that would suggest that the other gains resulted from the game itself.

2.2.2 Game design

The game consisted of two types of problems. In the first type of problem, a complex akshara was shown and the children had to select the simple akshara of which it was composed (i.e., akshara decomposition, see Figure 1). In the second type of problem, the children would hear a word and have to spell it using the akshara provided (i.e., spelling, see Figure 2). Each level consisted of 10 akshara decomposition problems and 10 spelling problems. The spelling problems contained the akshara taught through akshara decomposition (e.g., after learning the complex akshara ग्रा /gra/, children had to spell the word ग्राम /gram/ (village)). To pass a level, the child had to finish all 20 problems before time ran out and earn a sufficient number of points.

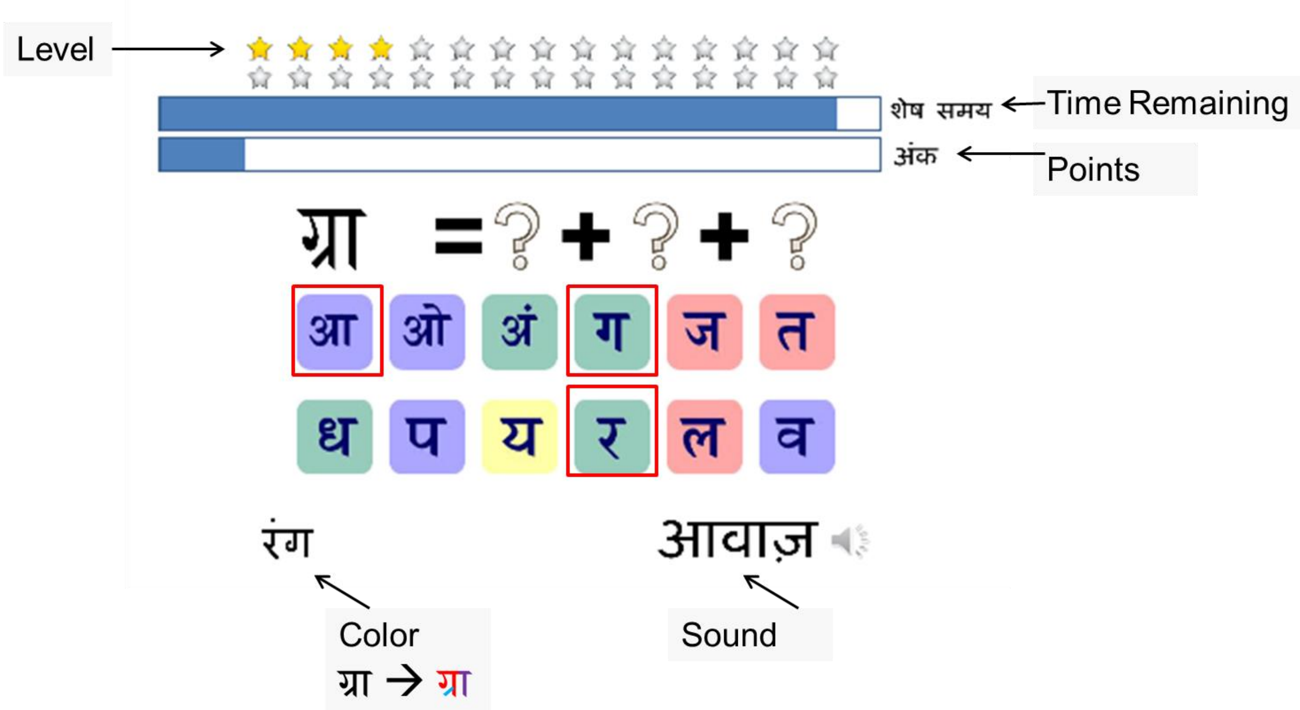

Figure 1: A screenshot of an akshara decomposition problem. The stars at the top show how many levels have been successfully completed. The stars were highly motivating for the children. The bars at the top display how much time is left in the level and how many points have been earned so far in the level. The complex akshara ग्रा /gra/ is shown. The simple akshara options are below. The three correct ones have red boxes around them. The two hint buttons, color and sound, are at the bottom.

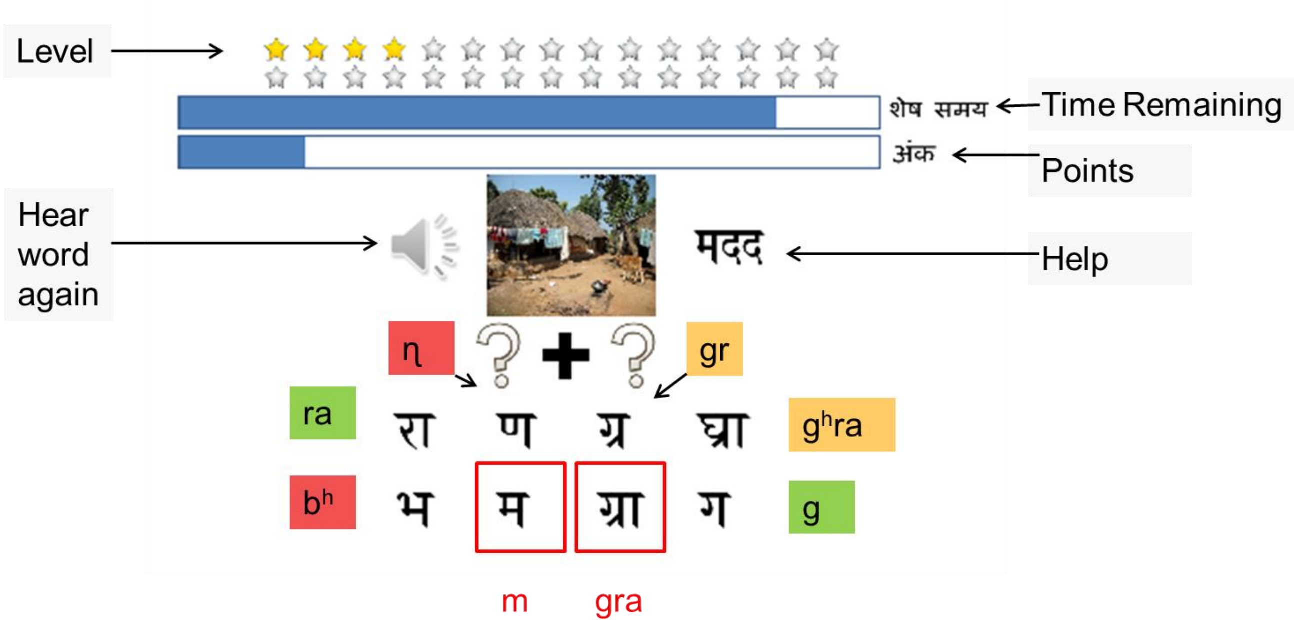

Figure 2: A screenshot of a spelling problem. Children would hear the word /gram/ (village). They could press the speaker to hear it again. The picture depicts the meaning of the word. If a child presses the "help" button, two foils are automatically removed. Children have to choose the correct akshara (the ones in red boxes) to spell the word. Note that the complex akshara they just learned (/gra/) is in this word. The orthographic foils are shown with red backgrounds, the phonological foils are shown with yellow backgrounds, and the split foils (one of the consonants in the complex akshara) are shown with green backgrounds. A few words also had combination foils (combining consonants from two different akshara into one akshara), but the problem displayed here did not.

2.2.2.1 Akshara decomposition

A complex akshara was displayed and the child had to choose the simple akshara of which it was composed from the options at the bottom of the screen. The number of simple akshara options ranged from 7 to 16. Typically, the number of options increased as the child progressed through the game. The simple akshara options remained constant within a level and were displayed in alphabetical order. If the child was having trouble, there were two hint buttons that she could use. The first one color-coded the different parts of the akshara, making it easier to decompose. The second hint button would pronounce the akshara, so it was possible to hear which simple akshara it was composed of. The pronunciations were recorded by a native Hindi speaker. Three points were given for correct answers obtained without using any hint buttons, two points if one hint was used, and one point if both hints were used.

2.2.2.2 Spelling

The game played an audio recording of a word and displayed a picture that represented the meaning of the word. The words' pronunciations were recorded by a native Hindi speaker. The child had to spell the word using the akshara choices at the bottom of the screen. In the first level, there were 6-7 akshara choices for each word. After that, there were 8-10 akshara choices. The foils were designed to be orthographically and phonologically similar to the correct answers (see Figure 2). The child could listen to the word as many times as necessary without any penalties. If a child was struggling, she could press the help button, which would delete two of the foils at random. Two points were given for correct answers obtained without using a hint button and one point was given if the hint button was used.

All the words were chosen to be age-appropriate for the children. There are two popular curricula in India, the Central Board of Secondary Education (CBSE) and the Indian Certificate of Secondary Education (ICSE). All of the words were chosen from either CBSE or ICSE 3rd-5th standard textbooks.

2.2.2.3 Game play

At the top of the screen were two bars, one that counted down the amount of time remaining and another that counted how many points had been earned. The children were given 4 minutes to complete levels 1-15 and 3.5 minutes to complete levels 16-30. The children needed to earn 40 points to pass levels 1-15 and 45 points to pass levels 16-30. If the child ran out of time during a level, a message would show on the screen stating that time has run out and the level would restart. If the child finished the level within the time limit but did not earn enough points to move onto the next level, a message would be displayed explaining this and then the level would restart. If the child successfully completed a level, a congratulatory message would be shown and then the next level would begin. The top of the screen contained 30 stars. One star was colored in for every level completed.

The first level consisted of very high frequency words that the children should be familiar with. All of the words contained CV akshara but did not contain complex akshara. In the spelling problems, there were only 6-7 akshara choices. Levels two and three also contained CV akshara and no complex akshara. They were designed such that each of the following diacritics (/a/, /i/, /i:/, /u/, /u:/, /e/, /ai/, /o/, /ou/, and nasal) was taught within one problem. In levels 4-28, every word contained one complex akshara. In levels 29 and 30, every word contained two complex akshara. Only one of those complex akshara was practiced in the akshara decomposition problem, but the unpracticed complex akshara was taught in a previous level. The game kept a log of all activity that could be analyzed at a later time.

There were two versions of the game. In the narrow spacing version, the akshara decomposition and spelling problems would alternate in every level. Therefore, right after the children learned to decompose an akshara, they would practice spelling a word that contained that akshara. In the wide spacing version, the children would do all 10 akshara decomposition problems and then all 10 spelling problems. Other than the order of the stimulus presentation, the two games were identical.

2.2.3 Game development

The mobile game was programmed using e-Chimera, a visual end-use programming environment that can design experiments for mobile devices (Luo, Head, Schneider, & Wang, 2014).

2.2.4 Participants

The participants were all in the 4th standard at a large all-girls private school in Bangalore, Karnataka, India. The state language of Karnataka is Kannda. However, Bangalore is a large, cosmopolitan city with immigrants from all over India, so the students speak a wide variety of languages at home. The content areas are all taught in English so all of the students are fluent in English. The students have to select two additional languages to study. For their first additional language, they can choose between Hindi and Kannada. For their second additional language, they can choose between Hindi, Kannada, and Sanskrit. More time is dedicated to their first additional language than to their second additional language. All of the children in my sample had chosen Hindi as their first additional language. Children who spoke Hindi or a language that was similar to Hindi (e.g., Marathi, Urdu) at home were excluded.

A total of 122 children were pre-tested. If a child showed very low Hindi literacy skills (did not know even the simple akshara) pre-testing was discontinued because the game would be too challenging for her. 108 children had high enough pre-test scores to continue with the study. 36 of them played the narrow spacing version of the game, 36 played the wide spacing version, and 36 were in an unseen control group. The three groups were selected to match as closely as possible on spelling pre-test scores, with no significant differences on the other pre-test measures.

Five children were dropped because they went on vacation during the study or elected to discontinue with the study.

2.2.5 Schedule

The pre-testing was conducted over a period of 2.5 weeks. The intervention was conducted over the next 4 weeks. Children were seen in groups of nine and each group was seen for a total of 12 sessions, 25 minutes long each. 28 children finished the entire game within the 12 sessions; the minimum number of sessions it took to finish the game was six. Of the 39 children who did not finish the game, 1 student was seen for nine session, 8 students were seen for ten sessions, and 12 students were seen for eleven sessions due to absences; the remaining 18 students were seen for all twelve sessions. The fewest levels successfully completed by a student was 14 levels. The post-testing was done approximately 2 weeks after the intervention. The children who finished the game early were post-tested first (along with the children from the unseen control group with the highest scores) to keep the time between completing the intervention and the post-test approximately constant across all children.

2.3 RESULTS

2.3.1 Pre and post-tests

2.3.1.1 Non-verbal IQ

The means and standard deviations for the three groups are displayed in Table 2. There were no statistically significant differences across the three groups in terms of non-verbal IQ, all *p*s > .5.

Table 2: Mean (standard deviation) raw scores on the matrix reasoning assessment. The students in the present study were 8.34 – 10.18 years old, with an average age of 9.26 years. The average scores for 8.34, 9.26, and 10.18 year old children are 15.5, 18, and 21, respectively (normed to a US population). Thus, the participants' non-verbal IQs were average for their age.

Experimental Group	Score
Control	17.583 (6.447)
Narrow	16.667 (7.119)
Wide	17.111 (6.968)

2.3.1.2 Akshara recognition

The descriptive statistics from the akshara recognition test are displayed in Table 3. The akshara recognition data were analyzed using a binomial linear mixed effects model[2]. Although ideally the model would have included a three-way interaction among akshara type, experimental group, and time (pre-test/post-test), this model did not converge, likely because participants were nearly at ceiling on the simple akshara at post-test. Therefore, it included akshara type and the interaction between experimental group and time. It also included the following random effects: 1) random intercept for subjects; 2) the effect of akshara type (simple as compared to CV and complex) to vary by subjects; and 3) random intercept for items.

[2] Binomial linear mixed effects models yield z and p-values.

Orthogonal contrasts were used to compare akshara type. The first contrast compared simple and CV akshara to complex akshara and the second contrast compared simple and CV akshara to each other. Orthogonal contrasts were also used to assess the effect of experimental group. The first contrast compared the control group to the average of both experimental groups. The second contrast compared the two experimental groups (narrow and wide spacing) to each other.

Participants performed better on the simple akshara than on the CV and complex akshara, $z = 7.513$, $p < .001$. The odds of answering a simple akshara problem correctly were 83.263 times higher than the odds of answering a CV or complex akshara problem correctly.

Participants improved from pre-test to post-test, $z = 6.750$, $p < .001$. The experimental groups performed better than the control group, $z = 2.060$, $p = .039$. Both of these main effects were qualified by an interaction between time and experimental group, $z = 3.030$, $p = .002$. This interaction is driven by the fact that participants in the control group did not improve from pre-test to post-test, $z = 1.624$, $p = .104$, but participants in the experimental groups did, $z = 8.849$, $p < .001$. For participants in the experimental groups, the odds of answering a post-test question correctly were 2.433 times higher than the odds of answering a pre-test question correctly.

Table 3: Mean (standard deviation) accuracy on the akshara recognition test (in percents). Note that participants that played the game improved to a greater degree than did control participants.

	Control		Narrow		Wide	
	Pre	Post	Pre	Post	Pre	Post
Simple	95.37 (9.71)	95.68 (9.31)	94.44 (9.78)	98.96 (3.29)	94.29 (15.33)	98.69 (3.63)
CV	46.3 (21.50)	50.93 (24.86)	46.88 (21.77)	57.29 (20.71)	55.71 (22.49)	59.8 (21.37)
Complex	31.11 (28.86)	35.56 (26.67)	36.25 28.93)	53.13 (29.89)	34.29 (27.26)	51.18 (27.05)

2.3.1.3 Word reading

The descriptive statistics from the word reading data are displayed in Table 4. The reading data were analyzed using binomial linear mixed effects models. The models included a three-way interaction between word type, experimental group, and time (pre-test/post-test) as well as the following random effects: 1) random intercept for subjects; 2) the effect of word type to vary by subjects; and 3) random intercept for items.

To restrict the number of contrasts to a manageable amount, the model collapsed across some of the word types. Specifically, the "simple" and "CV" categories were collapsed into a "no complex akshara" category, the "learned" and "near transfer" categories were collapsed into a "learned complex akshara" category, and the "medium transfer" and "far transfer" categories were collapsed into a "transfer complex akshara" category. Orthogonal contrasts were used to assess the effect of word type. The first contrast compared the "no complex akshara" category to the average of the "learned complex akshara" and "transfer complex akshara" categories. Thus, this contrast compared words that did and did not contain complex akshara. The second contrast compared the "learned complex akshara" and "transfer complex akshara" categories, thus measuring the degree of transfer.

Orthogonal contrasts were also used to assess the effect of experimental group. The first contrast compared the control group to the average of both experimental groups. The second contrast compared the two experimental groups (narrow and wide spacing) to each other.

The data were scored in the following way: if the participant read the word correctly, she received a score of 1. For words containing complex akshara, if she read the complex akshara correctly but made a mistake elsewhere in the word, she received a score of 0.5. For words in the CV category, if she read both CV akshara correctly but made a mistake elsewhere in the word, she received a score of 0.5. Scores of 0.5 were not possible for words in the simple category. All other readings received a score of 0. Because binomial models only accept scores of 1 or 0, two scoring criteria were used. In the lenient scoring criterion, scores of 0.5 were scored as 1. In the strict scoring criterion, scores of 0.5 were scored as 0.

Table 4: Mean (standard deviation) accuracy on the word reading test (out of 8). D = Difference between the pre-test and post-test scores. Note that participants that played the game improved to a greater degree than did control participants on words containing complex akshara.

	Control			Narrow			Wide		
	Pre	Post	D	Pre	Post	D	Pre	Post	D
Simple	5.06 (1.74)	6.35 (1.66)	1.29	5.58 (1.46)	6.16 (1.49)	0.58	5.51 (1.63)	6.17 (1.52)	0.66
CV	3.89 (1.96)	4.96 (2.04)	1.07	4.18 (2.08)	5.48 (1.71)	1.31	4.09 (1.94)	4.99 (1.99)	0.90
Learned	3.14 (1.53)	4.18 (2.01)	1.04	3.42 (1.49)	4.85 (1.52)	1.44	3.46 (1.59)	5.26 (1.15)	1.80
Near	2.11 (1.64)	2.85 (1.74)	0.74	2.23 (1.61)	3.15 (1.84)	0.92	2.43 (1.88)	3.49 (1.86)	1.06
Medium	3.08 (2.46)	4.07 (2.13)	0.99	3.85 (2.18)	4.10 (2.50)	0.24	3.70 (2.23)	4.63 (1.79)	0.93
Far	2.93 (2.00)	3.17 (1.86)	0.24	2.81 (1.77)	3.65 (1.86)	0.84	3.17 (1.83)	3.93 (1.83)	0.76

Lenient scoring criteria. Participants improved from pre-test to post-test, z = 14.198, $p < .001$. Participants were more accurate on words that did not contain complex akshara than on words that did contain complex akshara, z = 3.226, $p = .001$. These main effects were qualified by two interactions. First, there was an interaction between word type (contains complex akshara/does not contain complex akshara) and experimental group (narrow/wide), z = -2.023, $p = .043$. The difference between words with and without complex akshara was larger for the narrow spacing group. There was also an interaction between word type (learned complex akshara/transfer complex akshara) and time (pre/post), z = 3.098, $p = .002$. There was more improvement on the learned complex akshara than on the transfer complex akshara.

Importantly, there was three-way interaction among experimental group (control/experimental), word type (no complex akshara/complex akshara), and time, z = -2.034, p =0.042. Post-hoc[3] analyses were conducted to better understand this three-way interaction. At pre-test, there was no difference between the control and experimental groups on words without complex akshara, z = 1.263, p = .207, nor on words with complex akshara, z = 0.879, p = .379. At post-test, there was still no difference between the groups on words without complex akshara, z = 0.096, p = .923. But, the experimental groups marginally outperformed the control group on words with complex akshara, z = 1.688, p = .092. At post-test, the odds of the experimental groups correctly reading a word with a complex akshara were 1.491 times higher than the odds of the control group correctly reading a word with a complex akshara. To summarize, by post-test, the experimental groups outperformed the control group on words containing complex akshara, the words the game was training them on.

Strict scoring criteria. The results using the strict scoring criteria were very similar. Participants improved from pre-test to post-test, z = 13.800, $p < .001$. Participants were more accurate on words that did not contain complex akshara than on words that did contain complex akshara, z = 4.359, $p < .001$. These main effects were qualified an interaction: There was an interaction between word type (learned complex akshara/transfer complex akshara) and time (pre/post), z = 3.167, p = .001. There was more improvement on the learned complex akshara than on the transfer complex akshara.

[3] All post-hoc analyses were conducted using the same linear mixed effects model but with fewer variables and a subset of the data. For example, here the time and word type variables were removed and separate analyses were conducted for pre-test words with complex akshara, pre-test words without complex akshara, post-test words with complex akshara, and post-test words without complex akshara.

The three-way interaction among word type (contains complex akshara/does not contain complex akshara), time (pre/post), and group (control/experimental) that was significant with the lenient scoring criteria, was not significant with the strict scoring criteria, although the effects were in the same direction, z = -1.599, p = .110. However, because this interaction was significant with the lenient scoring criteria, I performed the post-hoc analyses with the strict scoring criteria. The results of the post-hoc analyses closely mirror those done with the lenient scoring criteria. At pre-test, there was no difference between the control and experimental groups on words without complex akshara, z = 1.122, p = .262, nor on words with complex akshara, z = 1.261, p = .207. At post-test, there was still no difference between the groups on words without complex akshara, z = 0.032, p = .975. But, the experimental groups marginally outperformed the control group on words with complex akshara, z = 1.709, p = .087.

Summary. In summary, the control and experimental groups were matched at pre-test. However, at post-test, the experimental groups outperformed the control group on words that contained complex akshara. This was the word type I was expecting to see the most gains in because words with complex akshara are more difficult and the game specifically focuses on them. This result is more clearly seen with the lenient scoring criteria than the strict scoring criteria. Because the lenient scoring criteria gives points for pronouncing the complex akshara correctly, this finding suggests that participants in the experimental groups made large gains in pronouncing the complex akshara correctly, but may have continued to make mistakes elsewhere in the word.

2.3.1.4 Spelling

The descriptive statistics for the spelling data are displayed in Table 5. The spelling data were analyzed using a binomial linear mixed effects model. The model included a three-way interaction among word type, experimental group, and time (pre-test/post-test) as well as the following random effects: 1) random intercept for participants and 2) random intercept for items.

To restrict the number of contrasts to a manageable amount, the model collapsed across some of the word types. Specifically, the “simple” and “CV” categories were collapsed into a “no complex akshara” category, the “learned” and “near transfer” categories were collapsed into a “learned complex akshara” category, and the “medium transfer” and “far transfer” categories were collapsed into a “transfer complex akshara” category. Orthogonal contrasts were used to assess the effect of word type. The first contrast compared the “no complex akshara” category to the average of the “learned complex akshara” and “transfer complex akshara” categories. Thus, this contrast compared words that did and did not contain complex akshara. The second contrast compared the “learned complex akshara” and “transfer complex akshara” categories, thus measuring the degree of transfer.

Orthogonal contrasts were also used to assess the effect of experimental group. The first contrast compared the control group to the average of both experimental groups. The second contrast compared the two experimental groups (narrow and wide spacing) to each other.

The data were scored in the following way: if the participant spelled the word correctly, she received a score of 1. For words containing complex akshara, if she wrote the complex akshara correctly but made a mistake elsewhere in the word, she received a score of 0.5. For words in the CV category, if she wrote both CV akshara correctly but made a mistake elsewhere in the word, she received a score of 0.5. Scores of 0.5 were not possible for words in the simple category. All other spellings received a score of 0. Because binomial models only accept scores of 1 or 0, two scoring criteria were used. With the lenient scoring criterion, scores of 0.5 were scored as 1. With the strict scoring criterion, scores of 0.5 were scored as 0.

Table 5: Mean (standard deviation) accuracy on the spelling test (out of 6). D = Difference between the pre-test and post-test scores. Note that participants that played the game improved to a greater degree than did control participants on words containing complex akshara.

	Control			Narrow			Wide		
	Pre	Post	D	Pre	Post	D	Pre	Post	D
Simple	3.25 (1.34)	4.19 (0.98)	0.94	3.59 (1.48)	4.03 (1.36)	0.44	3.34 (1.37)	4.11 (0.83)	0.77
CV	1.13 (1.54)	1.94 (1.40)	0.82	1.34 (1.46)	2.14 (1.52)	0.80	1.36 (1.44)	2.24 (1.26)	0.89
Learned	0.63 (0.81)	0.94 (0.99)	0.32	0.61 (0.92)	1.11 (1.01)	0.50	0.77 (0.92)	1.60 (1.37)	0.83
Near	0.50 (0.73)	0.57 (0.90)	0.07	0.44 (0.82)	0.70 (0.85)	0.27	0.36 (0.71)	0.71 (1.09)	0.36
Medium	0.33 (0.76)	0.65 (1.02)	0.32	0.33 (0.69)	1.03 (1.26)	0.70	0.37 (0.81)	0.81 (1.14)	0.44
Far	0.06 (0.20)	0.11 (0.34)	0.06	0 (0)	0.17 (0.49)	0.17	0 (0)	0.20 (0.62)	0.20

Lenient scoring criteria. The[4] results showed that there was improvement from pre-test to post-test, z = 12.546, p < .001. Accuracy was affected by word type: participants performed better on the words that did not contain complex akshara than on words that did contain complex akshara (z = 6.568, p < .001) and participants performed better on learned complex akshara items than on transfer complex akshara items (z = 2.343, p = .019). There was also an interaction between test and word type (learned complex akshara/transfer complex akshara), z = -2.470, p = .014. Participants showed more improvement on the transfer complex akshara items than on the learned complex akshara items.

Importantly, the overall degree of improvement was greater for participants in the experimental groups than for participants in the control group, z = 2.312, p = .021. This interaction was qualified by a three-way interaction among time, experimental group (control/experimental),

[4] Although the model did not converge, the relative gradient was equal to 0.001. A relative gradient of 0.001 or less indicates sufficient convergence.

and word type (no complex akshara/complex akshara), z = -2.800, *p* = .005. Post-hoc analyses were conducted to better understand this interaction. At pre-test, there was no difference between the control and experimental groups on words without complex akshara, z = 1.128, *p* = .259, nor on words with complex akshara, z = -0.811, *p* = .417. At post-test, there was still no difference between the groups on words without complex akshara, z = 0.423, *p* = .672. But, the experimental groups marginally outperformed the control group on words with complex akshara, z = 1.668, *p* = .095. At-post, the odds of the experimental groups correctly spelling a word with a complex akshara were 1.992 times higher than the odds of the control group correctly spelling a word with a complex akshara. To summarize, by post-test, the experimental groups outperformed the control group on words containing complex akshara, the words the game was training them on.

Strict scoring criteria. The model with the strict scoring criteria did not converge because of nearly floor effects on the words with transfer complex akshara (mean = 0.062).

2.3.1.5 Akshara construction

The descriptive statistics for the akshara construction data are displayed in Table 6.The akshara construction data were analyzed using a binomial linear mixed effects model. The model included an interaction between experimental group and test (pre-test/post-test) as well as random intercepts for both subjects and items. Orthogonal contrasts were used to assess the effect of experimental group. The first contrast compared the control group to the average of both experimental groups. The second contrast compared the two experimental groups (narrow and wide spacing) to each other. There was a main effect of test, with participants improving from pre-test to post-test, z = 7.965, *p* < .001. There were no effects of group nor group by test interactions, all *p*s > .38. Therefore, no benefits of the game were seen on the akshara construction test.

Table 6: Mean (standard deviation) accuracy on the akshara construction test (out of 7). Note that there were no statistically significant differences across the three experimental groups.

	Pre	Post
Control	4.89 (2.03)	6.09 (1.22)
Narrow	4.97 (1.73)	6.06 (1.16)
Wide	5.06 (1.82)	6.09 (1.26)

2.3.1.6 Vocabulary

The vocabulary data were scored on scale of 0-3. Cross-rater agreement on scoring was high, kappa = 0.98. The descriptive statistics are in Table 7. Because binomial linear mixed effects models require items to be scored as correct/incorrect, these data were transformed using strict and lenient scoring criteria. For the strict criteria, only items scored as a '3' were marked correct. For the lenient criteria, items scored as '1-3' were marked correct.

The vocabulary data were analyzed using two binomial linear mixed effects models (one for strict coding, one for lenient coding). The models included three-way interactions among word type, experimental group, and test (pre-test/post-test) as well as the following random effects: 1) random intercept for subjects; 2) the effect of word type to vary by subjects[5]; and 3) random intercept for items. Orthogonal contrasts were used to assess the effect of word type. The first contrast compared the "words not in game" category to the average of the "words in game" and "words morphologically related to those in game" categories. Thus, this contrast compared words related and unrelated to those in the game. The second contrast compared the "words in game" and "words morphologically related to those in game" categories, thus measuring the degree of transfer. Orthogonal contrasts were also used to assess the effect of experimental group. The first

[5] The strict scoring criteria analysis did not allow for the words morphologically related to those in the game/words in game contrast to vary by subjects because that random slope explained very little variance.

contrast compared the control group to the average of both experimental groups. The second contrast compared the two experimental groups (narrow and wide spacing) to each other.

Table 7: Mean (standard deviation) accuracy on the vocabulary test (out of 24). Note that there were no statistically significant differences across the three experimental groups.

	Control		Narrow		Wide	
	Pre	Post	Pre	Post	Pre	Post
Unrelated to game	15.06 (5.66)	16.61 (5.93)	14.00 (5.83)	15.34 (5.71)	14.40 (6.36)	16.57 (5.39)
Morphologically related to words in game	7.78 (5.00)	9.61 (5.84)	7.75 (5.41)	8.75 (6.54)	8.83 (6.06)	9.54 (6.38)
Learned in game	6.36 (4.98)	7.47 (4.63)	5.03 (4.59)	6.50 (5.56)	6.03 (5.81)	7.66 (6.01)

The data showed that participants improved from pre-test to post-test (strict: z = 6.254, $p < .001$; lenient: z = 6.592, $p < .001$). The participants performed more poorly on words that were related to the game than words unrelated to the game (strict: z = -2.313, $p = .021$; lenient: z = -2.669; $p = .008$). This effect suggests that the words related to the game were more difficult than words unrelated to the game. Surprisingly, the participants also showed more improvement on words unrelated to the game than those related to the game (strict: z = -2.459, $p = .014$; lenient: z = -2.272, $p = .023$). There was no effect of, nor interaction with, experimental group. Therefore, there is no evidence that the game improved students' vocabulary knowledge.

2.3.1.7 Math

The descriptive statistics for the math data are displayed in Table 8. Because the math assessment required students to complete as many math problems as possible in a set amount of time, the most appropriate model to analyze these data is the Poisson distribution. The data were analyzed using a Poisson general linear model that included an interaction between experimental group and time (pre-test/post-test). Orthogonal contrasts were used to assess the effect of experimental group. The first contrast compared the control group to the average of both experimental groups. The second contrast compared the two experimental groups (narrow and wide spacing) to each other.

Table 8: Average (standard deviation) raw scores on the math assessment. Scores of 41 and 57 correspond to age equivalents of 8 years 6 months and 9 years 10 months, respectively (normed to a US population). The participants' average age was 9.26 years, so they were performing appropriately for their age group. Note that participants in all three groups showed equivalent levels of improvement.

	Pre	Post
Control	40.78 (10.09)	54.39 (9.98)
Narrow	42.91 (10.76)	56.8 (10.21)
Wide	42.15 (14.26)	55.38 (18.11)

Participants improved from pre-test to post-test, z = 13.842, $p < .001$. Participants in the experimental groups performed marginally better than did participants in the control group, z = 1.707, $p = .088$. Importantly, there were no interactions between time and experimental group, *p*s > .79. As expected, the intervention groups did not improve on the math assessment to a greater degree than the control group did, suggesting that other improvements made by the intervention groups are due to the game itself, and not due to their greater interaction with the experimenters.

2.3.2 Game play

There is evidence that the wide spacing version of the game was more difficult than the narrow spacing version of the game because it took longer to play. On average, students playing the narrow spacing version were able to complete significantly more levels in one session (M = 3.9) than students playing the wide spacing version (M = 3.5), t (65) = 2.5, $p = .015$. Furthermore, students playing the narrow spacing version spent marginally less time on each level (M = 326.4s) than students playing the wide spacing version (M = 352.2s), t (65) = 1.9, $p = .056$.

2.3.2.1 By-item analyses

Akshara problems. The game tracked every click that the participants made. These data were used to quantify the difficulty of various items and extract general principles about akshara learning. First, participants' performance on the "akshara decomposition" problems was analyzed. For every akshara, the average number of clicks it took participants to arrive at the correct answer was calculated. First the data were cleaned to remove times in which participants quit in the middle of the problem (perhaps they quit the game because they had to go to class or the game quit in the middle of a problem because they had run out of time and had to re-start the level). Thus, the problems the participants had solved correctly remained. Then, the number of button clicks it took participants to arrive at the correct answer was counted. Button clicks included selecting an akshara, de-selecting an akshara, and asking for a hint. If a given participant did the same problem multiple times, all attempts were included. It is important to note that all participants completed the problems in the earlier levels, whereas fewer participants completed the problems in the higher levels. In the highest levels, the highest performing participants and participants in the narrow spacing group are over-represented. It is also important to note that this analysis cannot account for instances in which the students asked the experimenters for help. Then, the average number of button clicks it took to correctly answer each problem was calculated separately for the narrow and wide spacing groups. There was no significant difference between the narrow (mean = 4.365 clicks, SD = 5.566) and wide (mean = 4.472 clicks, SD = 6.458) spacing groups, $t(299) = 0.425$, $p = .671$. There were 13 problems for which participants, on average, took more than 10 clicks to respond to correctly (the threshold of 10 was chosen because it was approximately equal to the mean + 1 SD). Therefore, these problems were particularly difficult and analyzing the mistakes made while attempting these problems can shed light on common akshara recognition difficulties.

On the whole, participants had difficulties with irregular and non-linear akshara. There was also one instance in which participants were confused by orthographic similarity. Table 9 summarizes the difficult akshara. For the sake of comparison, Table 10 summarizes the 13 easiest akshara.

As you can see from Table 9, participants had the most difficulty with complex akshara in which the components are not easily visible (i.e., opaque complex akshara). They also had trouble with non-linear akshara. For example, the इ /i/ vowel was problematic, likely because its diacritic is to the left of the consonant even though the vowel is pronounced after the consonant. The /r/ was also problematic when it occurred at the beginning of the consonant cluster because that created non-linearity.

It is interesting to contrast the easy and difficult akshara. All of the easy akshara are composed of only two consonants, the minimum number of components possible. This pattern was likely seen because it reduces the minimum number of clicks it takes to get the problem correct. Thus, this pattern is an artifact of the analytical method and is not particularly informative. However, what is informative is the fact that all of the easy akshara are transparent. In fact, 12/13 of the easiest akshara are formed using the most common method for concatenating consonants, dropping the right-hand portion of the first consonant and physically attaching it to the second consonant. Only one complex akshara (21.5, ट्ठ) is not formed using that method. Nevertheless, it is still transparent because both components are easily visible. The fact that 11/13 of the most difficult akshara were opaque but none of the easy akshara were opaque tells us that opaque akshara are particularly difficult.

Table 9: Difficult akshara problems. The 3rd column shows the simple akshara that comprise the complex akshara. The color-coding scheme is: green = vowel; red = र /r/, this akshara always changes shape when a part of a complex akshara; purple = opaque akshara, black/blue = consonant akshara that is easily visible in the complex akshara. In the last column, out of order means that the students selected to correct components, but in the wrong order.

Level. Problem	Complex akshara	Components	Why was this akshara difficult?	What common errors did students make?
4.7	त्त	त त	Opaque akshara	Participants knew one akshara was त; randomly guessed the second akshara
5.7	र्ता	र त आ	Opaque akshara with र Non-linear with र Orthographic similarity	-Out of order -This akshara looks similar to the CV akshara तो (त + ओ); many participants chose the vowel ओ
6.9	स्त्र	स त र	Opaque akshara	Many participants chose स and र but missed त
7.2	क्षी	क ष ई	Opaque akshara	Some participants chose क and ई; others chose ष and ई
7.4	न्ति	न त इ	Non-linear with इ	Out of order
7.9	स्त्रि	स त र इ	Opaque akshara Non-linear with इ	Out of order
9.2	ण्टि	ण ट इ	Non-linear with इ	Out of order
9.3	र्णि	र ण इ	Opaque akshara with र Non-linear with र Non-linear with इ	Out of order
9.6	ष्ट्री	ष ट र ई	Opaque akshara with र	Out of order
14.3	र्ती	र त ई	Opaque akshara with र Non-linear with र	Out of order
19.7	द्दी	द द ई	Opaque akshara	Participants did not seem to know that द + द = द्द; they used the hint sound and chose consonants that were phonologically similar
22.1	क्षि	क ष इ	Non-linear with इ Opaque akshara	Both श and ष make the /ʃ/ sound. Many participants chose श by mistake.[6]
22.6	र्षा	र ष आ	Opaque akshara with र Non-linear with र	Out of order

[6] Although traditionally श and ष are pronounced as /ʃ/ and / ʂ/ respectively, the difference is diminishing in modern Hindi (Kachru, 2006).

Table 10: Easy akshara problems

Level.Problem	Complex akshara	Components
6.3	क्त	क त
10.1	ज्य	ज य
11.7	ल्ल	ल ल
13.1	त्म	त म
13.7	न्न	न न
16.3	न्न	न न
19.8	न्त	न त
21.5	ट्ठ	ट ठ
26.9	स्व	स व
28.10	स्व	स व
29.3	व्य	व य
29.5	प्त	प त
29.7	ष्ट	ष ट

The way in which the most common errors (right hand column in Table 9) were determined was by looking at the trials in which participants took the most button clicks to get the correct answer. Then, patterns in the first few button clicks of those trials were noted. For example, for akshara 22.6 (र्षा = र + ष + आ), the 30 trials in which people took the most button clicks to respond were examined. Of those 30 trials, in 16 trials the first akshara chosen were ष आ र. In another 5 trials, the first akshara chosen were ष र आ. Therefore, for this problem, the primary error type was coded as out of order.

It is important to note that there was large variation in the number of clicks it took participants to get a problem correct; the large number of clicks for some problems were driven by a few people. For example, for akshara 22.6, it took people an average of 10.3 clicks to get the answer correct, but the standard deviation was 11.2 clicks. Getting the problem correct requires a minimum of three clicks. On 13 trials, people got it in 3 clicks. On 40 trials, people took 7 clicks or less. On 29 trials, people took between 9 and 43 clicks. On one trial, a student took 79 clicks!

There was improvement on a given akshara across the game (i.e., learning). For example, participants struggled with त्त in level 4, they took, on average, 10.2 clicks to get it correct. In levels 6 and 7, they encountered the akshara त्ती, and took an average of 7.5 and 5.2 clicks to get it correct, respectively. This reduction is especially impressive because the akshara in level 4 does not have a vowel diacritic, so it takes a minimum of 2 clicks to get correct. The akshara in levels 6 and 7 does have a vowel diacritic, so it takes a minimum of 3 clicks to get correct. Another example is that participants struggled with the स्त्र akshara in level 6; on average they took 11.8 clicks to get it correct. When they encountered it again in level 16, they took an average of 3.9 clicks to get it correct.

Word problems. A similar analysis was performed for the word problems. Participants in the wide spacing group (mean = 11.327 clicks, SD = 12.475) required, on average, significantly more clicks to get the word problems correct than participants in the narrow spacing group (mean = 8.989 clicks, SD = 9.954), $t(299) = 12.043$, $p < .001$. Thus, the fact that students in the narrow spacing group played faster seems to be primarily driven by the word problems, not the akshara problems.

Table 11: Difficult word problems. The first row has the level number.problem number, difficult word, its pronunciation, and its meaning. The next row shows the options the students could choose from. The correct responses are on the left and the foils are on the right. The correct responses are listed in the order they appear in the word. The foils are listed from most to least selected. The next row shows the pronunciations of the akshara. The complex akshara in the word is bolded. The next row shows the foil type. P = phonological, S = splitting the complex akshara, O = orthographic, C = combining two akshara. For example, in the word बर्तन /bər.t̪ən/, the foil भ /b^h/ is phonologically similar to ब /b/. Therefore, भ is marked as ‘P’ and both akshara are coded blue to show their connection. व is orthographically similar to ब. Therefore, व is marked as ‘O’ and both akshara are coded blue. The complex akshara is र्त /rt̪ə/. The foils र /r/ and त /t̪/ are examples of incorrectly splitting the complex akshara. Therefore, they are marked as ‘S’ and all three akshara are color-coded red. Finally, this word has the sounds /t̪/ and /n/ in different akshara. The तं /t̪ən/ foil combines both of those sounds and is therefore marked as ‘C’. It is not color coded because it is not associated with one particular akshara. Finally, the number of clicks is listed in the last row. Note that for correct akshara, schwas are shown as appropriate. For the foils, no schwas are shown because schwa placement depends on the word context.

	Correct					Foils					
6.6 चिकित्सक / t͡ʃi.kit̪.sək/ (doctor)											
Akshara	चि	कि	त्स	क		की	ची	स	त	ख	
Phonology	t͡ʃi	ki	**t̪.sə**	k		ki:	t͡ʃi:	s	t̪	k^{h}	
Foil Type						P	P	S	S	P	
# clicks	243	262	312	273		136	135	133	115	95	
6.7 बर्तन /bər.t̪ən/ (pan; kitchen utensil)											
Akshara	ब	र्त	न			र	तं	भ	र्ट	त	व
Phonology	bə	**r.t̪ə**	n			r	t̪ən	b^{h}	rʈ	t̪	w
Foil Type						S	C	P	P	S	O
# clicks	224	372	180			168	155	93	42	37	20
8.2 गणतन्त्र /gəɳ.t̪ən.t̪rə/ (Republic)											
Akshara	ग	ण	त	न्त्र		न	त्र	न्त	घ	र	ट
Phonology	gə	ɳ	t̪ə	**n.t̪rə**		n	t̪r	nt̪	g^{h}	r	ʈ
Foil Type						S/P	S	S	P	S	P
# clicks	197	165	199	300		142	130	128	83	42	24
9.8 दर्द /d̪ərd̪/ (pain)											
Akshara	द	र्द				र्ढ	द्र	र्ध	ढ	ड	र
Phonology	d̪ə	**rd̪**				rɖ^{h}	d̪r	rd̪^{h}	ɖ^{h}	ɖ	r
Foil Type						P/O	C	P	P/O	P	S
# clicks	250	737				253	170	148	132	86	62
13.6 सुगन्ध /su.gən.d̪^{h}ə/ (fragrance)											
Akshara	सु	ग	न्ध			न्द	घ	सू	न्घ	ध	न
Phonology	su	gə	**n.d̪^{h}**			nd̪	g^{h}	su:	ng^{h}	d̪^{h}	n
Foil Type						P	P	P/O	O	S	S
# clicks	134	177	173			153	109	93	92	67	24
14.5 निर्दयी /nir.d̪ə.ji:/ (merciless)											
Akshara	नि	र्द	यी			र	यि	नी	र्ड	णि	द
Phonology	ni	**r.d̪ə**	ji:			r	ji	ni:	rɖ	ɳi	d̪
Foil Type						S	P	P	P	P	S
# clicks	176	265	165			103	97	85	65	46	42
15.5 परिवर्तित /pə.ri.wər.t̪i.t̪ə/ (change)											
Akshara	प	रि	व	र्ति	त	र्ती	री	र	ति	फ	
Phonology	pə	ri	wə	**r.t̪i**	t̪ə	rt̪i:	ri:	r	ti	p^{h}	
Foil Type						P	P	S	S	P	
# clicks	170	177	216	331	210	140	109	102	74	14	
19.1 आनन्दित /a.nən.d̪it̪/ (rejoice)											
Akshara	आ	न	न्दि	त		न्दी	अ	थ	दि	ण्दि	
Phonology	a	nə	**n.d̪i**	t̪		nd̪i:	ə	t̪^{h}	d̪i	ɳd̪i	
Foil Type						P	P	P	S	P	
# clicks	182	226	280	261		145	115	105	76	72	
23.1 व्यायाम /wja.jam/ (exercise)											
Akshara	व्या	या	म			य	यां	व	व्य	भ	ब्या
Phonology	**wja**	ja	m			j	jam	w	wj	b^{h}	bja
Foil Type						P	C	S	P	O	O
# clicks	166	216	100			81	76	60	43	17	14

Table 12: Easy word problems. The complex akshara are color-coded in red. Note: Easy words in levels 28-30 were excluded from this list because they contained two complex akshara.

Level.Problem	Word	Pronunciation	Meaning
15.2	मात्रा	ma.t̪ra	quantity
16.6	वस्त्र	wə.st̪rə	cloth, textile
17.8	रस्ता	rə.st̪a	road
20.4	लम्बा	ləm.ba	long
22.4	प्रेम	prem	love
22.6	कोष्ठक	ko.ʂʈʰək	nationalistic
25.2	कच्चा	kət͡ʃ.t͡ʃa	raw
27.5	तुम्हें	t̪u.mhõ	you
27.9	कक्षा	kək.ʂa	classroom

There were 9 word problems in which participants required, on average, more than 21 clicks to get correct (The threshold of 21 was chosen because it is approximately equal to the mean + 1 SD). Each of those problems is described in Table 11. For the sake of comparison, the 9 easiest words are shown in Table 12.

As seen from Table 11, there were a few common sources of confusion. The first was phonological confusion with vowels. Students had particular trouble with vowel length. For example, the word चिकित्सक /t͡ʃi.kit̪.sək/ has two short /i/ sounds. Students often chose the long /i:/ by mistake. Similar mistakes were seen in आनन्दित /a.nən.d̪it̪/, परिवर्तित /pə.ri.wər.t̪i.t̪ə/, निर्दयी /nir.d̪ə.ji:/, and सुगन्ध /su.gən.d̪ʰə/. It is interesting that none of the easy words contain the vowels /i/ and /i:/ and only one of the easy words contains the vowels /u/ and /u:/. This reinforces the fact that vowel length confusion is particularly problematic. Another problem commonly seen within the difficult words was vowel reduction (choosing the schwa vowel rather than the /a/ sound); it was seen in both आनन्दित /a.nən.d̪it̪/ and व्यायाम /wja.jam/.

There were also some examples of orthographic confusion. For example, in the word सुगन्ध, the orthographic foil न्घ was chosen fairly often. However, there seemed to be more phonological difficulties than orthographic difficulties. For example, in the words बर्तन /bər.t̪ən/ and व्यायाम /wja.jam/, the orthographic foils are chosen very infrequently.

It is possible for complex akshara to represent a consonantal blend or to cross a syllabic boundary. It is interesting to note that in 7/9 of the hardest words, the complex akshara crosses a syllabic boundary. In contrast, in only 3/9 of the easiest words does the complex akshara cross the syllabic boundary. Thus, it seems that complex akshara that cross syllabic boundaries were particularly challenging. Furthermore, it appears that split-type errors were more common when the complex akshara crosses a syllabic boundary. In both दर्द /d̪ərd̪/ and व्यायाम /wja.jam/, the split-type foils were not chosen very often (all <65 clicks). In all the other words, at least one split-type foil had more than 65 clicks. गणतन्त्र /gəɳ.t̪ən.t̪rə/ is a particularly good example of a splitting error. The complex akshara न्त्र is composed of three consonants, न /n/, त /t̪/, and र /r/. The त and र combine to form the opaque complex akshara त्र. Then, the न is added to form the complete form न्त्र. There is a syllabic break between /n/ and /t̪r/. The most common split-type errors were choosing न and त्र, which fall along the syllabic boundary. The third most common split-type error was choosing न्त. Thus, it seemed that the three consonant cluster was especially challenging and students tended to choose different combinations of the three consonants.

There were two words with particularly interesting response patterns that I want to discuss in more detail. First, for the word बर्तन /bər.t̪ən/, students commonly chose the combination, ब /b/, र /r/, तं /t̪ən/ (this combination was chosen in 48 out of 93 trials at some point and on 26 trials as the first three akshara). This is interesting because this combination is phonologically correct; both बर्तन and बरतं would be pronounced as /bər.t̪ən/.

Second, for the word दर्द /d̪ərd̪/, on 33 out of 108 trials र्द /rd̪/ was chosen first. This suggests that participants may have known that र्द has the phonemes /r/ and /d̪/, but were confused about their order. This order confusion mirrors some of the common mistakes students made on the akshara problems.

It is important to note that there was a high level of variability. For example, for the word आनन्दित /a.nən.d̪it̪/, on six trials participants got the word correct with 4 clicks (the minimum number of clicks required). The first quartile was 10.75 clicks, the median was 17.5 clicks, the 3rd quartile was 25 clicks, and the maximum was 103 clicks.

2.3.2.2 By-subject analyses

The data were filtered to only include the first time a subject successfully completed a given item. The average number of clicks it took each subject to correctly answer the akshara and word problems in each level was then calculated. The average number of clicks was then correlated with the level number[7]. If the correlation is negative, this indicates that participants improved as they played the game. The average correlations for both akshara and word problems were significantly less than zero, indicating that participants did improve throughout the game (akshara: average r = -0.289, t (69) = -12.759, $p < .001$; words: average r = -0.073, t (69) = -2.010, $p = .048$). However, the magnitude of improvement was much larger for the akshara problems than the word problems, in fact the improvement on the word problems was negligible. This discrepancy could result from the fact that some akshara repeated throughout the game, whereas words never repeated. Furthermore, akshara are more similar to each other than are words, so insights gleaned from one akshara problem are more easily applicable to a future problem.

[7] Only levels 4-30 were included because levels 1-3 did not contain complex akshara.

2.4 DISCUSSION

Overall, the game improved students' akshara knowledge. This improvement was apparent in both the game data and the pre/post-test data. In terms of pre/post-tests, the benefit of the game was most apparent on the akshara recognition assessment. Students in the control group did not improve on this assessment from pre-test to post-test whereas students in the experimental groups showed significant improvement. There were also gains on the reading and spelling tasks, although they were not as robust as those with akshara recognition. Specifically, the game helped students read and spell words that contained complex akshara. On the reading assessment, improvement was more apparent when using the lenient scoring criteria than the strict scoring criteria, suggesting that participants were getting the complex akshara correct but making mistakes elsewhere in the word.

The experimental design used an unseen control group, rather than an active group that does a non-reading related activity. The experimental groups may perform better due to Hawthorne effects (Cook, 1962), specifically the experimental groups may do better because they feel more comfortable with the experimenters and they know they received the treatment and are expected to do better. The mathematical assessment, an assessment the game is not expected to improve, was included to look for the presence of Hawthorne effects. The experimental groups did not improve to a greater degree than the control group did on the mathematical assessment, suggesting that the improvements in Hindi are due to the game itself, and not due to Hawthorne effects.

Although the goal of the game was not to teach vocabulary, a vocabulary assessment was included to test for incidental vocabulary learning in the game. There were no intervention effects on the vocabulary assessment. There also were not any intervention effects on the akshara construction assessment, which was surprising because I did expect to see effects on this assessment. It is possible that there was not enough power to see intervention effects on the akshara construction assessment because there were only seven items. Furthermore, ceiling level performance could have made it difficult to see significant effects.

It was interesting that, although the wide spacing version of the game was more difficult, it did not yield more gains on the post-tests. I expected the wide spacing version to be more beneficial because previous research has shown that spaced practiced is generally more beneficial than massed practice (see Cepeda, Pashler, Vul, Wixted, & Rohrer, 2006; Underwood, 1961 for review), including for learning spelling (Fishman, Keller, & Atkinson, 1968). Typically, in massed practice, the to-be-learned item is shown repeatedly, whereas in spaced practice, the to-be-learned item is shown in a more distributed fashion. The narrow/wide spacing distinction in the present study is slightly different in that the game did not present the same item repeatedly, rather once the akshara was shown in isolation and later the akshara was shown in a word context. Nevertheless, I expected the wide spacing to be beneficial. Furthermore, the narrow spacing is theoretically easier because once you have learned an akshara, you know that akshara will be in the following word. Thus, you already know one of the correct answers. In the wide spacing version, you do not have this advantage. This narrow spacing advantage was demonstrated in the game data; students playing the narrow spacing version required fewer button clicks to correctly answer only the word problems. From the point of view of desirable difficulties (McDaniel & Butler, 2011), the wide spacing version should be more beneficial. Furthermore, students in the wide spacing version were less likely to finish all 30 levels in less than 12 sessions, so overall, they spent more time playing the game. Although, from a theoretical standpoint, I expected the wide spacing version to be more beneficial, this advantage was not borne out in the pre and post-test data.

It is possible that the wide spacing version was not beneficial because, although it was difficult, this difficulty was not desirable. McDaniel and Butler (2011) suggest that whether or not a difficulty is desirable depends on properties of the subjects, materials, and criterial tasks. I believed that the difficulty created by the wide spacing version would be desirable because one of the criterial tasks (spelling), required students to spell words based solely on phonology-graph correspondences. Therefore, the type of processing required by the wide spacing version was more in line with the type of processing required on the spelling post-test. However, it is possible that my subjects had trouble seeing the relationship between isolated complex akshara and how they function in word contexts. The fact that the wide spacing version obscured this relationship could have induced a difficulty that was not desirable. Thus, it seems that the narrow spacing version is both efficient and effective and should be used going forward.

One particularly beneficial aspect of analyzing the game data is that I was able to collect data on the difficulty level and common mistakes on 300 words, more than would be able to be studied in most experiments. This rich data set is an invaluable resource to researchers interested in Hindi decoding skills.

The akshara-level game data showed that students had the most trouble with opaque akshara and non-linear akshara (e.g., akshara containing the /i/ vowel). The /i/ vowel is challenging because the diacritic occurs to the left of the consonant, but is pronounced after the consonant. Similar difficulty with the /i/ vowel has been reported in other studies (Nag, 2011; Vaid & Gupta, 2002). These results suggest that educators need to spend more helping students with these two sources of difficulty.

For the words, there were three very common errors. The first was with vowel length, students commonly confused /i/ and /i:/ as well as /u/ and /u:/. Vowel length can be difficult because the pronunciation differences are slight. In fact, in the word final position, both vowel lengths are pronounced the same except for in careful, educated speech (Kachru, 2006). Thus, educators may need to help students hear vowel length differences, and in some instances, the correct spellings may need to be memorized. The second common error was also phonological in nature; students tended to confuse aspirated and unaspirated consonant pairs. Again, educators may need to help students phonologically distinguish these pairs.

The third common error was splitting the complex akshara, especially when the complex akshara straddled a syllabic boundary. It is also interesting to note that, in 7/9 of the most difficult words, the complex akshara does straddle a syllabic boundary. This difficulty with akshara that cross syllabic boundaries has been noted in previous literature (Nag, 2014).

Although the game was helpful, the pre and post-test data indicated that there were three areas in which students continued to struggle. First, students struggled with some of the vowels. For example, only 14% of students pronounced /tʃɛ/ correctly on the akshara recognition post-test. The /ɛ/ vowel may have been difficult because it has two possible pronunciations; it is pronounced as /ɛ/ in most contexts but as /əi/ when preceding a /j/ (Kachru, 2006).

Second, Hindi has many similar phonemes that vary only in aspiration or place of articulation. I expected the game to help students distinguish these similar-sounding phonemes because I used many close phonological foils in the word level of the game. However, the post-tests showed that phonological errors were still rampant. Specifically, students tended to use dental/alveolar forms when they should have used retroflex forms and they tended to use unaspirated forms when they should use aspirated forms. For example, on the spelling post-test, 88% of students spelled a word containing the /ɳ/ phoneme with /n/ and 66% of students spelled a word containing the /tʃʰ/ phoneme with /tʃ/. On the akshara recognition post-test, 67% of students pronounced /b^hi/ as /bi/. This pattern may reflect the frequencies of certain phonemes in Hindi; dental/alveolar form are more frequent than retroflex forms and unaspirated forms are more frequent than aspirated forms (Khan et al., 1991). Helping students better distinguish similar-sounding phonemes will be explored in Experiment 3.

Third, the improvement on the reading and spelling tests was relatively modest. Experiment 2 will explore if a different game design could increase reading and spelling gains.

3.0 EXPERIMENT 2: COMPARISON OF PEDAGOGICAL METHODS FOR BUILDING ORTHOGRAPHIC KNOWLEDGE

3.1 INTRODUCTION

In Experiment 1, children were taught complex akshara both in isolation and in word contexts. On the akshara problems, students were shown a complex akshara and had to choose its components. The present study tests whether or not that was the most effective method for teaching complex akshara. The method used in the game is compared to three other methods. The four methods test three specific hypotheses: 1) does emphasis on part-whole versus whole-part relationships matter? (i.e., learning to build up simple akshara to form complex akshara as opposed to decomposing complex akshara); 2) does motor encoding benefit learning?; and 3) does testing benefit learning? I predicted that motor encoding and testing would benefit learning, but that the emphasis on part-whole versus whole-part relationships would not matter. Finally, in Hindi, some complex akshara are transparent (i.e., the subcomponents are easily visible) whereas others are opaque (i.e., the subcomponents are not easily visible). I also tested whether the efficacy of the learning method varied by transparency.

3.1.1 Transparency

As explained in the General Introduction, some complex akshara are transparent (i.e., their components are easily visible) whereas others are opaque (one or both components is not easily visible). It is reasonable to expect that opaque complex akshara are more difficult to learn than transparent complex akshara because they must be memorized and cannot be derived using rules. It is also possible that learning strategies that build orthographic knowledge, such as copying/writing, are more important for opaque complex akshara. Therefore, in this study, the opacity of the graphs is varied. Transparent graphs were all formed using the most common rule: drop the right portion of the first graph and physically connect it to the second graph (e.g., क + य = क्य). Opaque graphs did not contain one or both of their component graphs (e.g., complex graph that does not contain one component graph: क + स = के; complex graph that does not contain both component graphs: ग + य = ए)[8].

3.1.2 Benefit of motor encoding

Physically copying/writing graphs/words has been shown to be more helpful for building orthographic knowledge than viewing, tracing, typing, and manipulating tiles (here copying is defined as writing while viewing a model, writing is defined as writing from memory) (see Wollscheid, Sjaastad, & Tømte, 2016 for review). Furthermore, the benefit of copying/writing has been demonstrated with children and adults learning English letters, Chinese characters, Bengali and Gujarati akshara, Arabic graphs, and pseudoletters/characters. Ouellette (2010) found that second grade children could better spell English non-words if they learned via writing than viewing. Cunningham and Stanovich (1990) taught first graders to spell several English words in three conditions: copying, typing, and physically re-ordering letter tiles. Students learned best in the copying condition and this result held whether the post-test required them to spell via writing,

[8] Note: These are not the mappings between simple and complex akshara in Hindi; they are the mappings used in the present experiment.

typing, or manipulating tiles. Longcamp, Zerbato-Poudou, and Velay (2005) found a similar result: four year-old children learned English letters better when they copied the letters than when they typed them. Another study examined first, third, and fifth grade Japanese students learning Chinese pseudocharacters. The children learned the characters better when they copied them rather than just viewed them. However, the benefit of copying decreased when children traced the characters or copied them with the other side of their pen (thus producing no physical mark). Therefore, it is not only the motor action that is important, but also momentarily holding the orthographic form in memory and seeing the physical form you produce (Naka, 1998).

The benefit of motor encoding has also been demonstrated with children and adults learning a second language. Naka (1998) demonstrated that Japanese children learned Arabic graphs better when they copied the graphs than when they traced or viewed them. In another study (Longcamp et al., 2008), adults learned the orthographic forms of Bengali and Gujarati akshara; the akshara were simple akshara and were not paired with their phonological forms. The participants practiced one set of akshara via copying and the other via typing. Participants were better at determining whether an akshara was in its correct orientation or was a mirror-image when they learned via copying than via typing, and participants remembered handwritten akshara for a longer period of time. In another study, English-speaking adults learned Chinese characters by either writing or viewing them. Participants were shown the character, and they heard its pronunciation and saw its English translation. Once the character was off the screen, participants studied it either by mentally recalling it or by writing it. Participants learned the meaning and orthographic form better when they learned via writing than via viewing (Guan, Liu, Chan, Ye, & Perfetti, 2011). Another study with a similar population found that writing was more beneficial than passive viewing for orthographic learning, but less beneficial for retrieving the sound and meaning of the characters. In the study, adults were shown a character and they heard its pronunciation/saw its pinyin transcription, and saw its English translation. In the passive viewing condition, the character remained on the screen and participants were asked to study it. In the writing condition, the screen went blank and participants were asked to write it from memory. The writing condition produced superior results for the three tasks requiring detailed orthographic knowledge: a lexical decision task, a task in which participants wrote the character from a pronunciation prompt, and a task in which participants wrote a character from a meaning prompt. In contrast, the viewing condition produced superior results when participants were shown a character and asked to produce its pronunciation or meaning (Xu, Chang, Zhang, & Perfetti, 2013).

Why is copying/writing beneficial? There is evidence that motor experience with objects alters neural activation patterns; the motor system is active when viewing objects that we have had motor experiences with. Therefore, it is possible that the physical action of copying/writing helps engage the motor system when viewing graphs (James, 2010). Multiple neuroimaging studies have demonstrated that learning graphs via copying/writing changes the neural response to the graphs. For example, children show more activity in the inferior frontal gyrus, posterior parietal cortex, anterior cingulate, and left fusiform gyrus when they learn letters via copying as opposed to various other learning mechanisms including tracing, typing, and viewing (James, 2010; James & Engelhardt, 2012). Because the posterior parietal cortex is associated with motor representation and the left fusiform gyrus is associated with orthographic representations, these results suggest that copying helps engage the motor system and improves the visual representation for the graphs (James & Engelhardt, 2012). Similarly, adults show more activity in the left posterior fusiform gyrus and left dorsal precentral gyrus when they learn pseudoletters via copying than via typing or passively viewing (James & Atwood, 2009). One study that examined adults learning Chinese characters replicated the behavioral results summarized above, that writing leads to better performance on a lexical decision task than does thinking about the character's pronunciation. Furthermore, it found that writing led to more activation in the bilateral superior parietal lobules and bilateral lingual gyri. Activation in the right homologues of these areas was correlated with performance (Cao et al., 2013). Finally, another study (Longcamp et al., 2008) demonstrated that writing leads to better akshara learning than does typing and also found neural differences between the two learning conditions. Writing was associated with more brain activation in several areas: the cerebellum which is critical for motor memory consolidation, the posterior part of the middle temporal gyrus which is important for learning associations between visual stimuli and motor responses, and the somatosensory cortex, which is associated with motor execution and imagery.

Studies with English letters, Chinese characters, Arabic graphs, and simple akshara have shown that copying/writing is a highly effective way to develop orthographic knowledge. However, this effect is not seen consistently (Chang, Xu, Perfetti, Zhang, & Chen, 2014; Naka & Naoi, 1995; Ouellette & Tims, 2014; Vaughn, Schumm, & Gordon, 1992, 1993). Given these inconsistencies, I wanted to see if I could replicate the beneficial effect of copying/writing and extend the effect to different materials than those used in previous studies. I used complex akshara as the to-be-learned materials, which are composites of simple akshara. Chang et al.'s (2014) study suggests that copying/writing may not be beneficial for complex akshara. They found that drawing beginning Chinese learners' attention to the chunks comprising characters was more beneficial than writing. Similarly, in the present study, identifying components draws learners' attention to the components in transparent complex akshara. Therefore, it is possible that copying/writing will not be beneficial for learning complex akshara. Furthermore, the benefit of copying/writing may vary for transparent and opaque complex akshara.

I chose to work with novice learners rather than Hindi[9] students because Hindi students may vary significantly in their akshara exposure and knowledge of ligaturing rules; working with novice learners minimized the confound of previous knowledge. Also, motor effects seem especially beneficial for novice learners, although this relationship is not completely straightforward. Cunningham and Stanovich (1990) and Longcamp et al. (2005) found a benefit of copying with first grade and pre-school children, respectively. However, a study with older children (2^{nd} – 3^{rd} grades), failed to replicate this effect (Vaughn et al., 1992). I believe this difference is experience-based rather than developmental because studies with adults beginning to learn a foreign language have shown a benefit of writing. Specifically, two studies demonstrating the benefit of writing for learning Chinese characters used relatively novice learners; Xu et al.'s (2013) participants had 20 weeks of Chinese instruction and 57.2% of Guan et al.'s (2011) participants had less than half a year of Chinese instruction and all of the participants had less than two years of instruction. However, suggesting that at least some experience is necessary for a writing effect, Chang et al. (2014) did not observe a writing effect on the learning of Chinese characters with participants who only had eight weeks of Chinese instruction. In a study of first, third, and fifth grade Japanese students learning pseudo-Chinese characters and Arabic graphs, Naka (1998) found that the benefit of copying decreased with age in learning pseudo-Chinese characters, with which they had some experience. In contrast, the benefit of copying was equally prevalent across the age groups when the students were learning novel Arabic graphs. In the context of these previous studies, my novice adult learners encounter graphs that they do not know

[9] I refer to the language Hindi because it is the most well-known language that uses the Devanagari script. However, there are other languages that use the Devanagari script (e.g., Marathi, Sanskrit, Nepali languages) (Sinha & Mahabala, 1979). In fact, I use one akshara (ळ) which is used in Marathi but not in Hindi (Rathod, Dhore, & Dhore, 2013). Because the stimuli are modeled on Devanagari, but there is nothing specific to Hindi per se, the results of this study are equally applicable to all languages that use Devanagari.

(similar to Naka (1998), which showed a copying effect) and they engage in writing and copying at the very beginning of their instruction (similar to Chang et al. (2014), which showed no writing effect).

3.1.3 Benefit of testing

I also wanted to test the hypothesis that writing is more beneficial than mere copying, because of writing's greater demands on memory for the graphic form. There is limited evidence to support that hypothesis. For example, Naka (1998) argued that the need to momentarily hold the orthographic form in memory was why copying was more beneficial than tracing. Because writing requires participants to hold the orthographic form in memory to a greater extent than copying does, I predicted that writing would be more beneficial than copying. However direct comparisons of writing and copying are lacking. Some studies have tested writing (Cao et al., 2013; Chang et al., 2014; Guan et al., 2011; Ouellette & Tims, 2014; Vaughn et al., 1993; Xu et al., 2013) whereas others have tested copying (Cunningham & Stanovich, 1990; James, 2010; James & Gauthier, 2006; Longcamp et al., 2005; Naka, 1998; Vaughn et al., 1992). The present study directly compares writing and copying.

To more clearly differentiate the writing and copying conditions, the writing condition is instantiated slightly differently from previous studies (Cao et al., 2013; Chang et al., 2014; Guan et al., 2011; Ouellette & Tims, 2014; Vaughn et al., 1993; Xu et al., 2013). In those studies, the character/word was shown, removed from the sight, and then participants had to write it down. In the present study, participants are shown the two simple akshara and asked to produce the complex akshara they comprise from memory (e.g., participants shown क + य = ; participants need to write क्य). Therefore, rather than briefly holding the visual form in working memory, participants had to retrieve the visual form from long-term memory. Thus, the writing condition was similar to the testing effect: the phenomenon that retrieving information from memory strengthens the memory for that information (Roediger & Karpicke, 2006). The testing effect is a domain-general learning mechanism that has been demonstrated in many educational contexts, including spelling (Rieth et al., 1974). Observing a writing effect here could be construed as a special case of a testing effect.

3.1.4 Study overview and hypotheses

To increase experimental control, the present study utilized an artificial orthography strongly modeled on Devanagari and novice participants. Although this approach lacks the authenticity of real literacy contexts, there are two specific benefits of the artificial orthography. First, in Devanagari, there are significantly more transparent complex akshara than opaque complex akshara. Therefore, it can be difficult to test the effect of opacity because the small number of opaque graphs would reduce statistical power. In contrast, the artificial orthography can be designed such that half the complex akshara are opaque. Second, learning a new language entails learning both a new phonological system and a new orthographic system. Because I was interested in orthographic learning, I paired my simple graphs with English phonemes and my complex graphs with English consonantal clusters. Using the English phonological system eliminated the need for participants to learn a new phonological system and allowed me to isolate orthographic learning.

This study compared four learning methods. The first two methods were multiple-choice (MC) methods that do not benefit from motor encoding. In the first method, decompose complex akshara MC, the complex akshara was provided and participants had to choose the corresponding simple akshara from a set of choices. In the second method, compose complex akshara MC, the simple akshara were provided and the participants had to choose the corresponding complex akshara from a set of choices. Comparing these two methods allowed me to test if emphasis on part-whole versus whole-part relationships matters. The third and fourth methods involved copying and writing, respectively.

This study used a variety of outcome measures to isolate specific skills. Some measures (orthographic legality and writing tests) required high-quality orthographic representations whereas others required a connection between orthography and phonology (hear and choose, reading, and writing tests). Finally, this study used both transparent and opaque graphs.

I hypothesized that participants would perform better on transparent graphs than opaque graphs on tests that require a connection between orthography and phonology. Furthermore, because of the benefit of testing, participants would perform better when they learn via writing than via copying on tests that require a connection between orthography and phonology, especially for opaque graphs. Because of the benefit of motor encoding, participants would perform better when they learn via copying/writing than via MC on tests that require high quality orthographic representations. I did not predict any differences between the two MC conditions because I did not expect emphasis on part-whole versus whole-part relationships to have an effect.

3.2 METHODS

3.2.1 Participants

Participants were 47 undergraduate students at the University of Pittsburgh. Six participants did not finish all three sessions and were thus excluded. The remaining 41 participants were aged 18-

28 years (average age = 19 years) and there were 26 males. All participants were from the subject pool and received course credit for participating. Participants were also given a $10 bonus for attending all three sessions and "performing well" (the bonus was not contingent on performance, but participants were told it was to encourage them to try their best). All participants were native English speakers and had no experience with any languages that use alphasyllabic orthographies.

3.2.2 Materials

Materials included 15 simple akshara (all consonants) and 80 complex akshara that were formed using pairs of simple akshara. Half of the complex akshara were transparent (i.e., they looked like a combination of their components, e.g., क + य = क्य) and half were opaque (i.e., they did not look like a combination of their components, although for some, one of the components was visible, e.g., क + स = के; ग + य = ए)[10]. All of the akshara were real Devanagari akshara, but some were paired with different sounds so that non-native phonology would not confuse the participants. For example, ण is pronounced as /ɳ/ in Hindi. Because that phoneme is not present in English, it would have been difficult for participants to pronounce. Therefore, it was paired with the phoneme /m/. Furthermore, the mappings between some of the pairs of simple akshara and their corresponding complex akshara were invented. For example, in Hindi, ग + य = ग्य, but the participants learned ग + य = ए. This was done to ensure that half of the pairing were transparent and the other half were opaque. See Table 13 for all of the akshara used in the study.

[10] Note: These are not the mappings between simple and complex akshara in Hindi; they are the mappings used in the present experiment.

Complex akshara can represent either consonantal blends or consonantal syllabic breaks. This study focused on teaching akshara in isolation, not embedded within words. Because it is very difficult to pronounce isolated consonantal syllabic breaks, but not to pronounce isolated consonantal blends, only consonantal blends were used in this study. To ease pronunciation, most of the consonantal blends used can be found in English words.

Table 13: Simple and complex akshara used in the experiment. The simple akshara and their pronunciations are in the first two rows and first two columns. The first akshara in a consonantal blend is shown in the first column and the second akshara is shown in the first row. For example, क्य is pronounced as /kj/. The complex akshara the simple akshara comprise are shown in the grid. Some of the blends occur at the beginning of English words (e.g., /sk/ occurs at the beginning of *skip*). These blends were pronounced with a minimal schwa after the second consonant to ease pronunciation. Other blends are pronounced at the end of English words (e.g., /ps/ occurs at the end of *cups*). These blends were pronounced with a minimal schwa before the first consonant to ease pronunciation. The black complex akshara are transparent and the red complex akshara are opaque.

		य	क	प	स	ष	ण	न	म	व	ग	त	ध	भ	श	घ
		j	k	p	s	l	m	n	w	d	g	t͡θ	z	t͡ʃ	ʃ	r
य	j															
क	k	क्य			**के**	क्ष			**को**							क्घ
प	p	**आ**			प्स	**उ**			प्म							**ऊ**
स	s	स्य	**श्र**	स्प		स्ष	**सा**	**सः**	स्म							
ष	l	**र्ण**	ष्क	**षु**	**क्ष**		ष्ण	**षू**		ष्व	**ऌ**	ष्त	ष्ध	**भ्र**		
ण	m	ण्य		ण्प					**झ**			**इ**	**ई**			
न	n	न्य	**ने**						न्म	**नै**	न्ग	**त्र**	**द्ध**	न्भ		
म	w															
व	d	**ख**							**द्व**				व्ध			व्घ
ग	g	**ए**				**ऐ**			ग्म				ग्ध			ग्घ
त	t͡θ	त्य				त्ष	**त्त**	त्न	**ह**							**हृ**
ध	z	**थ**							ध्म							ध्घ
भ	t͡ʃ								भ्म							**ळ**
श	ʃ	श्य	श्क	**फ**		**ल**	श्ण	श्न	**शी**							**शि**
घ	r		**ट**	घ्प	घ्स	**ठ**	**ड्ड**	घ्न		**ड**	**ढ**	घ्त	**द**	घ्भ		

3.2.3 Learning methods

Participants learned the complex akshara using four learning methods: decompose complex akshara MC, compose complex akshara MC, copy, or write. The goal of all four learning methods was to promote intra-akshara awareness by drawing attention to the component akshara. Importantly, phonology was not included in any of the learning methods. Instead, the learning methods focused on teaching participants about the component akshara, which were already associated with their pronunciations. If participants can identify the component akshara and if they know the pronunciation of the component akshara, they can pronounce the complex akshara.

3.2.3.1 Decompose complex akshara MC

In this learning method, participants were shown a complex akshara. They had to choose the two simple akshara that comprise the complex akshara, from among four choices, in the correct order. They could not move on to the next akshara until they got the answer correct. This is similar as to what was done in the mobile game.

3.2.3.2 Compose complex akshara MC

In this learning method, participants were shown two simple akshara and had to choose the correct complex akshara from among four choices. They could not move on to the next akshara until they got the answer correct.

3.2.3.3 Copy

In this learning method, participants were shown the two simple akshara and the complex akshara they comprise. They had to write the complex akshara on a sheet of paper. There was no demonstration of how to write the akshara in this condition to more closely mirror the multiple choice conditions.

3.2.3.4 Write

In this learning method, participants were shown the two simple akshara. They had to write the corresponding complex akshara onto a sheet of paper from memory, as best they could. They were then shown a static image of the correct complex akshara so they could check their answer. If they got the answer incorrect, they were asked to write the correct answer down. There was no demonstration of how to write the akshara in this condition to more closely mirror the multiple choice conditions.

Comparing the decompose complex akshara MC and compose complex akshara MC learning methods allowed me to determine if emphasis on part-whole versus whole-part relationships matters. Comparing the copy and write learning methods allowed me to determine if testing benefits learning. Comparing the decompose complex akshara MC/compose complex akshara MC and copy/write learning methods allows me to determine if motor encoding benefits learning.

3.2.4 Tests

Participants were tested on the complex akshara using four tests: hear and choose, orthographic legality, reading, and writing. The hear and choose test was administered on Day 2 (tested on all complex akshara) and Day 3 (tested on ¼ of the complex akshara). The other three tests were administered only on Day 3 and tested ¼ of the akshara each.

3.2.4.1 Hear and choose

Participants heard a pronunciation and then saw two akshara. Participants had to quickly choose the akshara that matches the pronunciation via button press. This test requires a connection between orthography and phonology.

3.2.4.2 Orthographic legality

Participants saw two versions of the same akshara, one correct and one with an orthographic error (see Figure 3). Participants had to quickly choose the correct akshara via button press. This test requires high quality orthographic representations.

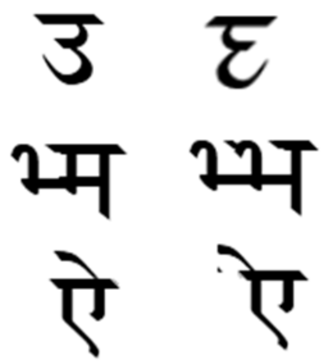

Figure 3: Examples from the orthographic legality test. Participants have to choose the orthographically legal akshara (in this case, always the one on the left). The examples illustrate some of the different types of orthographic problems present in the test (from top to bottom: mirror images, incorrect components, incorrectly positioned components).

3.2.4.3 Reading

Participants had to say the pronunciation of the complex akshara aloud. The answers were recorded to check for scoring accuracy. This test requires a connection between orthography and phonology.

3.2.4.4 Writing

Participants heard the pronunciation of an akshara and had to write the corresponding akshara on a piece of paper. This test requires a connection between orthography and phonology and high quality orthographic representations.

3.2.5 Procedure

The experiment took place over three consecutive days. One hour was allotted on the first and third days and two hours on the second day. Participants typically took most/all of the allotted time on the first two days but finished early on the third day (see Table 14 for an overview of the procedure).

Table 14: Procedure overview. The black items are associated with learning simple akshara. The blue items are associated with learning complex akshara. The red items are complex akshara tests.

Day 1	**Day 2**	**Day 3**
Learn 15 simple akshara	Simple reading test	Learn ¼ of akshara using each learning method (1x)
Simple reading test	Simple writing test	Hear and choose test (¼)
(Repeat if necessary)	Introduced to complex akshara	Orthographic legality test (¼)
Simple writing test	Learn ¼ of akshara using each learning method (2x)	Reading test (¼)
(Repeat if necessary)	Learn ¼ of akshara using each learning method (1x)	Writing test (¼)
Introduced to complex akshara	Hear and choose test (all)	

3.2.5.1 Day 1

On the first day, participants first learned the 15 simple akshara. They were shown each akshara on the screen and heard its pronunciation. They then repeated the pronunciation aloud and copied the akshara onto a sheet of paper. They went through all of the akshara twice in a random order. The first time through, the experimenter demonstrated how the akshara was written before the participant copied the akshara.

After the learning phase, participants were tested on the akshara. The first test was a reading test; they were shown the simple akshara and had to say its corresponding pronunciation aloud. The answers were recorded to check for scoring accuracy. If they got at least 80% (12/15) correct, they were then given the writing test. If not, they went through the learning phase again (but only saw each akshara once) and then re-took the reading test. This cycle continued until they passed the reading test or took the reading test three times, whichever came first. The second test was a writing test; they heard an akshara's pronunciation and had to write the akshara. If they did not get at least 80% correct, they re-did the learning phase and then took the test again. This cycle continued until they passed the writing test or took the writing test three times, whichever came first. For both tests, the correct answers were provided after the participants gave their answers.

Next, participants were introduced to the complex akshara. First, they saw a complex akshara and heard its pronunciation. Then, they saw the complex akshara's composition in equation form (i.e., simple akshara + simple akshara = complex akshara) and heard the pronunciations of all three akshara. They then repeated the pronunciation of only the complex akshara aloud. The akshara were presented in a randomized order.

3.2.5.2 Day 2

Participants began the second day by re-taking the simple reading and writing tests to refresh their memories. After that, the participants were introduced to the complex akshara one more time. Then, the participants began to learn the complex akshara. Participants learned a quarter of the akshara using each learning method (decompose complex akshara MC, compose complex akshara MC, copy, write). First, they went through all four learning methods once, with the akshara associated with the learning method presented twice. Then, they went through the four learning methods one more time, with the akshara associated with that learning method presented once.

The akshara were presented in a randomized order. Which akshara were paired with which learning method and the order of the learning methods were counterbalanced across participants. Half of the graphs associated with each learning method were opaque and half were transparent. After that, the participants completed the hear and choose test using all of the graphs.

3.2.5.3 Day 3

On the third day, the participants cycled through each learning method one more time and practiced each akshara associated with that learning method one time. Then they completed the hear and choose, orthographic legality, reading, and writing tests. Each test covered ¼ of the graphs. The graphs on each test were evenly distributed among the four learning methods and the opaque/transparent distinction. Therefore, there were 20 akshara tested on each test; 5 from each learning method. 2-3 of those were opaque and 2-3 were transparent (see Table 15 for distribution of stimuli across learning conditions and tests).

Note that there are only 2-3 items in each test from each learning condition/transparency combination. This would not be sufficient if the statistical tests were done with a typical by-participants ANOVA, which averages across all items in a given condition for each participant. However, I used linear mixed effects modeling, in which the unit of observation is the individual trial rather than an average for each participant. Therefore, linear mixed effects models yield more power than do traditional ANOVAs.

Table 15: Distribution of stimuli across tests. There were 40 transparent complex akshara (T1-40) and 40 opaque complex akshara (O1-40). The four learning methods are shown in the left column and the tests are shown in the middle and right columns. All participants saw the same akshara on the tests (e.g., /chr/ was always on the orthographic legality test). However, which learning method a given akshara was associated with was counterbalanced across participants (e.g., one participant may have learned /chr/ via copying, another via writing, etc.). The order of the learning methods was also counterbalanced across participants. The four Day 3 tests were always presented in the same order.

Decompose complex akshara MC		Day 2 Hear and Choose		Day 3 Hear and Choose	
T1	O1	T1	O1	T1	O1
T2	O2	T2	O2	T2	O2
T3	O3	T3	O3	T11	O3
T4	O4	T4	O4	T12	O11
T5	O5	T5	O5	T13	O12
T6	O6	T6	O6	T21	O21
T7	O7	T7	O7	T22	O22
T8	O8	T8	O8	T31	O23
T9	O9	T9	O9	T32	O31
T10	O10	T10	O10	T33	O32
Compose complex akshara MC		T11	O11	Day 3 Orthographic Legality	
		T12	O12		
T11	O11	T13	O13	T3	O4
T12	O12	T14	O14	T4	O5
T13	O13	T15	O15	T5	O13
T14	O14	T16	O16	T14	O14
T15	O15	T17	O17	T15	O15
T16	O16	T18	O18	T23	O24
T17	O17	T19	O19	T24	O25
T18	O18	T20	O20	T25	O33
T19	O19	T21	O21	T34	O34
T20	O20	T22	O22	T35	O35
Copy		T23	O23	Day 3 Reading	
T21	O21	T24	O24	T6	O6
T22	O22	T25	O25	T7	O7
T23	O23	T26	O26	T16	O8
T24	O24	T27	O27	T17	O16
T25	O25	T28	O28	T18	O17
T26	O26	T29	O29	T26	O26
T27	O27	T30	O30	T27	O27
T28	O28	T31	O31	T36	O28
T29	O29	T32	O32	T37	O36
T30	O30	T33	O33	T38	O37
Write		T34	O34	Day 3 Writing	
T31	O31	T35	O35	T8	O9
T32	O32	T36	O36	T9	O10
T33	O33	T37	O37	T10	O18
T34	O34	T38	O38	T19	O19
T35	O35	T39	O39	T20	O20
T36	O36	T40	O40	T28	O29
T37	O37			T29	O30
T38	O38			T30	O38
T39	O39			T39	O39
T40	O40			T40	O40

3.3 RESULTS

3.3.1 Simple akshara learning

On the first day, 14 participants did not pass the reading test after three attempts. Of the participants who did pass, they took an average of 2.52 attempts to pass. One participant did not pass the writing test after three attempts. Of the participants who did pass, they took an average of 1.56 attempts to pass.

On the second day, participants re-took the simple reading and writing tests. Any participant who got fewer than 11/15 correct on either test was excluded from analysis because the complex akshara learning methods are predicated on knowing the simple akshara well. If a participant knows the pronunciations of the simple akshara, and knows which simple akshara comprise a complex akshara, he/she can deduce the pronunciation of the complex akshara. If a participant does not know the pronunciations of the simple akshara, the methods of learning the complex akshara will not be useful. Based on this exclusion criterion, six participants were excluded. Of these six participants, four had failed to pass the reading test and one had failed to pass both the reading and writing tests on the first day. This left me with a final sample of 35 participants (age range: 18-28 years; average age = 19 years; 23 males). Of the remaining participants, average accuracy on the Day 2 reading and writing tests was high, 90.3 and 90.9% respectively.

3.3.2 Data analysis for tests of complex akshara learning

The hear and choose, orthographic legality, and reading data were analyzed using the same model, described here. The writing data were analyzed using a different model, described in the relevant section. The accuracy data were analyzed using binomial linear mixed effects models and the reaction time data were analyzed using general linear mixed effects models. The models included an interaction between learning method and transparency. Learning method was tested using three orthogonal contrasts: the comparison between composing complex akshara/decomposing complex akshara and copying/writing (motor effects), the comparison between copying and writing (testing effects), and the comparison between composing complex akshara and decomposing complex akshara (emphasis on part-whole versus whole-part relationships). Random effects that significantly contributed to the models as indicated by Aikake Information Criterion were included: a random intercept for participant, a random intercept for items, and the effect of transparency was allowed to vary by participant.

For the reaction time data, only reaction times to correct answers were considered. Any reaction times more than three standard deviations from the participants' mean were removed. Six, one, and four RTs were trimmed from the Day 2 hear and choose, Day 3 hear and choose, and Day 3 orthographic legality tasks respectively (all $< 0.75\%$).

For all analyses, all significant effects are reported. See Table 16 for mean performance on the tasks and Table 17 for a summary of the hypotheses and results.

Table 16: Means (standard deviations) on the complex akshara tests, by learning method. The results for transparent graphs are on top and for opaque graphs on bottom. Accuracy data are displayed as percents and reaction time data are displayed as time in milliseconds. Acc = Accuracy, RT = reaction time, D2 = Day 2, D3 = Day 3.

Measure	Write	Copy	Compose complex akshara MC	Decompose complex akshara MC
D2 Hear and Choose Acc	84.00 (13.33) 68.57 (15.74)	87.71 (12.62) 64.86 (18.21)	85.43 (11.97) 60.29 (17.23)	88.29 (12.48) 62.29 (16.46)
D2 Hear and Choose RT	2213.06 (558.44) 2298.44 (578.82)	2158.72 (540.64) 2393.55 (625.90)	2236.82 (526.03) 2441.39 (673.95)	2204.33 (521.87) 2408.75 (594.48)
D3 Hear and Choose Acc	98.57 (8.45) 68.10 (31.67)	98.10 (7.85) 69.05 (33.12)	93.33 (16.76) 70.95 (30.88)	95.24 (13.75) 67.14 (32.96)
D3 Hear and Choose RT	1947.48 (466.80) 2100.98 (740.72)	1935.80 (484.25) 1990.52 (596.40)	1968.78 (508.00) 2132.16 (722.65)	2097.81 (539.72) 2013.08 (714.47)
D3 Orthographic Legality Acc	85.71 (23.96) 81.90 (24.04)	85.71 (23.96) 86.67 (21.69)	85.71 (22.19) 71.90 (28.52)	76.19 (26.90) 75.71 (32.68)
D3 Orthographic Legality RT	1943.82 (604.79) 1825.89 (692.24)	1893.40 (591.91) 1861.25 (482.00)	2017.45 (516.53) 1978.75 (668.11)	1853.13 (571.33) 1862.34 (708.27)
D3 Reading Acc	75.24 (37.13) 28.57 (32.48)	65.24 (37.13) 19.05 (30.02)	74.29 (28.68) 23.33 (32.89)	69.52 (35.58) 19.05 (26.86)
D3 Writing Acc	69.52 (33.21) 10.48 (23.25)	58.57 (35.56) 11.90 (18.33)	45.71 (27.81) 2.86 (9.47)	41.90 (31.41) 3.81 (10.76)

Table 17: Summary of expected and actual results. The four hypotheses are listed at the top. For each hypothesis, the left column shows the expected results and the right column shows the actual results. In the last row, the check marks and X's indicate whether the hypothesis was supported or not supported, respectively. O = effect only seen for opaque graphs; T = effect only seen for transparent graphs; O⇔P = tasks requiring a strong connection between orthography and phonology; HQOR = tasks requiring high quality orthographic representations; Acc = Accuracy; RT = reaction time; D2 = Day 2; D3 = Day 3. **p* < .05, ~*p* < .10. Note that if the interaction was a trend but the post-hoc test was significant, this table reports a trend.

Measure	Transparent > Opaque for O⇔P		Write > Copy for O⇔P (especially for opaque graphs) because testing improves memory		Copy/Write > MC for HQOR because motor encoding benefits orthographic representations		Compose complex akshara MC (C) = Decompose complex akshara MC (D)	
D2 Hear and Choose Acc	*	*	O*			O~		D > C ~
D2 Hear and Choose RT	*	*	O*			~		
D3 Hear and Choose Acc	*	*	O*			T~		
D3 Hear and Choose RT	*		O*					
D3 Orthographic Legality Acc					*	*		C > D for T~
D3 Orthographic Legality RT					*	*		
D3 Reading Acc	*	*	O*	*				
D3 Writing Acc	*	*	O*		*	*		
	☑		☒		☑		☑	

3.3.2.1 Day 2 hear and choose

Accuracy. Participants[11] were more accurate on transparent graphs than on opaque graphs, z = 7.899, p < .001. Participants were marginally more accurate when they learned via decomposing complex akshara than via composing complex akshara, z = -1.666, p = .096. There was a marginal interaction between method (copy & write/decompose complex akshara & compose complex akshara) and transparency, z = -1.762, p = .078. There was no difference between the two learning conditions for transparent graphs, z = -0.641, p = .521, but motor encoding was associated with better performance for opaque graphs, z = 2.084, p = .037. There was also a marginal interaction between method (copy/write) and transparency, z = -1.759, p = .079, but none of the post-hocs separating by transparency were significant. Thus, the interaction may have been a spurious result.

Reaction time. Participants were faster on transparent graphs than on opaque graphs, t (61) = -3.208, p = .002. Participants were marginally faster when they learned via copying & writing than via composing complex akshara & decomposing complex akshara, t (1999.7) = -1.669, p = .095.

3.3.2.2 Day 3 hear and choose

Accuracy. Participants[12] were more accurate on transparent graphs than on opaque graphs, z = 4.892, p < .001. The comparison between composing complex akshara & decomposing complex akshara and copying & writing interacted with transparency at a trend level, z = 1.670, p

[11] Although the model did not converge, the relative gradient was less than .001.
[12] Although the model did not converge, the relative gradient was equal to .001.

= .095. Copying & writing were marginally better than the two multiple choice methods for transparent graphs, z = 1.692, p = .091, but not for opaque graphs, z = -0.083, p = .934.

Reaction time. There were no significant main effects nor interactions.

3.3.2.3 Day 3 orthographic legality

Accuracy. Participants[13] were more accurate when they learned via copying & writing than via composing complex akshara & decomposing complex akshara, z = 2.947, p = .003. The odds of answering correctly were 1.941 times higher if people learned via copying & writing than via composing complex akshara & decomposing complex akshara. There was also a trend-level interaction between the composing complex akshara/decomposing complex akshara comparison and transparency, z = 1.788, p = .074. The two learning methods were equivalent for opaque graphs, z = -0.493, p = .622, but composing complex akshara outperformed decomposing complex akshara for transparent graphs, z = 1.961, p = .050.

Reaction time. Participants were approximately 110.98 ms faster if they learned via copying & writing than via composing complex akshara & decomposing complex akshara, t (523.1) = -2.066, p = .039.

3.3.2.4 Day 3 reading accuracy

Participants[14] were more accurate on transparent graphs than on opaque graphs, z = 8.262, p < .001. Participants were also more accurate if they learned via writing than via copying, z = 2.301,

[13] Although the model did not converge, the relative gradient was less than .001.
[14] Although the model did not converge, the relative gradient was less than .001.

p = .021. The odds of answering correctly were 1.950 times higher if people learned via writing than via copying.

3.3.2.5 Day 3 writing accuracy

The writing analyses required a slightly different model. Because accuracy on opaque graphs was very low (7.4%), transparency could not be included as a fixed effect although it was retained as a random slope. Furthermore, the random slope that allowed learning method (copy/write) to vary across items explained a large portion of the variance so that was included as well[15]. Although transparency could not be included as a fixed effect, participants were more accurate on transparent graphs (54.3%) than opaque graphs (7.4%).

Participants were more accurate when they learned via copying & writing than via composing complex akshara & decomposing complex akshara, z = 5.048, $p < .001$. The odds of answering correctly were 3.096 times higher if people learned via copying & writing than via composing complex akshara & decomposing complex akshara.

3.3.3 Time on task

I did not control for time on task; I allowed for natural variation as would happen in an educational setting. However, because the start time of each task was recorded, time on task could be estimated using the intervals between consecutive tasks. The second time through each learning condition was used to estimate time on task; because participants did not need instructions the second time through, the estimates were more accurate. The average amounts of time spent on the compose

[15] Although the model did not converge, the relative gradient was less than .001.

complex akshara MC, decompose complex akshara MC, copying, and writing conditions were 4m 40s, 4m 58s, 5m 22s, and 8m 44s, respectively. Paired t-tests showed that the two MC conditions did not vary from each other, t (34) = 1.052, $p = .300$. Participants took longer to copy than to compose complex akshara MC, t (34) = 3.205, $p = .003$, but copy and decompose complex akshara MC did not vary from each other, t (34) = 0.965, $p = .341$. Writing took longer than the other three conditions, all ts > 8.996 and all *p*s < .001.

3.4 DISCUSSION

3.4.1 Transparency

I predicted that participants would perform better on transparent graphs than opaque graphs, especially on tasks that require a strong connection between orthography and phonology. This hypothesis was strongly supported; on the hear and choose task participants were more accurate (Days 2 and 3) and faster (Day 2) on transparent graphs. For the reading and writing tasks, participants were more accurate on transparent graphs. The only task that did not show an effect of transparency was the orthographic legality task, which was expected because the task does not require any phonological knowledge.

I expected that the type of learning method would matter more for opaque graphs, which are harder to learn. Although the expected interaction was present in the Day 2 hear and choose data, for the Day 3 hear and choose test, learning method mattered more for transparent graphs. I especially expected to see an interaction on the writing assessment because knowing that a complex graph is transparent and what its two components are allows easy derivation of the orthographic form. In contrast, the opaque graphs have to be individually memorized, which is very difficult. It is possible that the expected interaction was not present because half of the graphs were transparent and half were opaque. So, although it was easy to derive the orthographic form of transparent graphs, it was difficult to remember which graphs were transparent and which were opaque. Devanagari has more transparent akshara than opaque akshara. It is possible that the expected interaction between learning method and transparency would be present with a more natural stimulus set.

3.4.2 Benefit of motor encoding

I predicted that the copying /writing conditions would perform better than the MC conditions on tasks that require high-quality orthographic representations because motor encoding has been shown to increase orthographic knowledge. This effect was particularly robust. On the orthographic legality test, participants were significantly more accurate and faster if they learned via copying & writing than via MC. Similarly, on the writing test, participants were more accurate if they learned via copying & writing. Copying & writing was even beneficial on the hear and choose task, a task on which I was not expecting to see effects. The only task that did not show a benefit of copying & writing was reading. This finding was expected because reading does not require high-quality orthographic representations; it is possible to read words you cannot spell (Martin-Chang, Ouellette, & Madden, 2014).

Overall, the data support prior research (Cunningham & Stanovich, 1990; Guan et al., 2011; Longcamp et al., 2008, 2005; Naka, 1998; Ouellette, 2010; Xu et al., 2013) that suggests that motor encoding is valuable for strengthening orthographic representations. Most importantly, the data suggest that copying & writing is beneficial when one is given a phonological form and needs to produce/recognize the orthographic form (i.e., writing and hear and choose) or when one needs pure orthographic knowledge (i.e., orthographic legality) but not when one is given the orthographic form and needs to produce the phonological form (i.e., reading).

3.4.3 Benefit of testing

I predicted that the writing condition would perform better than the copying condition on tasks that require a strong connection between orthography and phonology because testing is beneficial for memory. This hypothesis was largely not supported, although the expected effect was seen on the reading test. It is surprising that writing did not outperform copying because participants spent significantly more time on the writing condition than on the copying condition.

A large body of work has shown that testing improves memory performance for a variety of materials, including orthographic information (Rieth et al., 1974). The expected effect was not present. One possible reason could be that the previous study did not control for exposure, whereas the present study did.

Furthermore, there are some exceptions to the testing effect. For example, if participants get an answer incorrect during an initial test, it can strengthen the incorrect response and lead to more incorrect responses on a final test (see Roediger & Karpicke, 2006 for review). Butler, Marsh, Goode, and Roediger (2006) found that when the initial test was easy and participants got most of the answers correct, testing benefitted learning. In contrast, when the initial test was difficult and participants got many incorrect answers, testing produced additional costs. In this case, the initial test was quite difficult and accuracy was low. Therefore, testing could have strengthened incorrect responses, thus negating the benefits of testing.

I did not expect testing to strengthen the incorrect response in this study for two reasons: 1) immediate feedback was given and 2) the task involved orthographic learning. Previous research has shown that incorrect responses are not problematic when immediate feedback is given (Butler & Roediger, 2006), as was the case in this study. Furthermore, much of the research on the negative effects of testing has used word lists and texts as the to-be-learned material. Studies that focus on orthographic learning often fail to demonstrate the same effect. For example, Ehri, Gibbs, and Underwood (1988) had experimental participants generate spellings for difficult-to-spell pseudowords (thus generating many misspellings) whereas control participants rested or performed an alternate task. Then, all the participants studied the correct spellings. Generating the misspellings did not interfere with experimental participants' ability to remember the correct spellings.

Although I did not expect poor performance on the initial test to strengthen the incorrect response in this study, post-hoc analyses suggest that this may have occurred. I did a median split of the participants based on how many questions they got correct during the writing learning condition to see whether it interacted with the copying/writing contrast. There was a marginal interaction on the writing test, in that writing accuracy during learning interacted with the copying/writing contrast, z = 1.848, p = .065. People who got many answers correct during the writing learning condition found writing marginally more beneficial than copying, z = 1.859, p = .063, whereas people who got many answers incorrect found copying slightly more beneficial than writing, z = -0.795, p = .427. Thus, this study suggests that even with immediate feedback and orthographic information as the to-be-learned material, poor performance on an immediate test can strengthen incorrect responses. Future studies should do more initial learning so that the writing condition elicits fewer incorrect responses and see if the benefits of testing are apparent in those conditions.

Testing is one particular example of a desirable difficulty; a task that makes encoding more difficult but leads to more robust learning. However, as stated before, whether a difficulty is desirable depends on properties of the subjects, materials, and criterial tasks (McDaniel & Butler, 2011). This study used novice participants who made many errors during learning. Thus, the writing learning condition may have been too difficult for them, and therefore not desirable. For more experienced participants, the difficulties induced by the writing learning condition may be desirable. The post-hoc analysis described in the preceding paragraph suggests that this is the case.

3.4.4 Time on task

I did not control for time on task; I allowed it to vary naturally as it would in an educational setting. However, I did estimate time on task so I could test if time on task was driving my results. Overall, composing complex akshara, decomposing complex akshara, and copying took approximately the same amount of time (although copying took more time than composing complex akshara, decomposing complex akshara did not significantly vary from the two other learning methods). However, writing took significantly more time than the three other learning methods. Time on task did not predict post-test outcomes. Although copying took approximately the same amount of time as the two MC conditions, it resulted in better post-test outcomes. Furthermore, although writing took significantly more time than copying, both of those learning conditions resulted in similar post-test outcomes. Therefore, the results were not an artifact of time on task. Furthermore, the time on task data suggest that copying is a particularly beneficial pedagogical tool for beginning learners; it is time-efficient and leads to comparably high levels of learning.

3.4.5 Discrepancies in the prior literature

The results from this study can help resolve some of the discrepancies in the prior literature. Specifically, the results suggest that copying is helpful for novice learners who need to perform tasks that require pure orthographic knowledge or the production of an orthographic form when given a phonological form. Writing is beneficial for experienced learners performing those same tasks.

Copying was beneficial in the present study because the participants were novice learners. The benefit was most apparent on tasks that required pure orthographic knowledge (orthographic legality) and tasks that required producing an orthographic form when given a phonological form (hear and choose, writing). Copying was also beneficial in Cunningham and Stanovich's (1990) study in which relatively novice learners (1st graders) performed a task which required producing an orthographic form given a phonological form (spelling). Similarly, copying was beneficial in Longcamp et al.'s (2005) study in which novice learners (4 years olds) performed a task which required pure orthographic knowledge (orthographic legality). Longcamp et al. (2008) also used an orthographic legality test and novice learners (English-speakers learning akshara) and found copying to be beneficial. Finally, Naka (1998) and Naka and Naoi (1995) found that copying was beneficial when Japanese-speaking adults and children learned Arabic graphs (which were completely novel to them). The task, a free recall of the graphs, required pure orthographic knowledge. Additionally, Naka (1998) demonstrated that copying was very beneficial for first graders learning pseudo-Chinese characters (with which they had some experience). Together, these studies suggest that copying is beneficial on tasks that require pure orthographic knowledge or the production of an orthographic form when given a phonological form. Copying is beneficial for novice learners, either those learning a new script or very beginning learners of their first language script (4 year olds – first graders).

The studies that have not found a benefit of copying either used advanced learners or tasks that required the production of a phonological form given an orthographic form. For example, Naka (1998) did not find a large copying benefit for Japanese-speaking children in grades 3-5 learning pseudo-Chinese characters (with which they had some experience). Similarly, Vaughn et al. (1992) did not find a copying benefit for 2nd-3rd grade children. Furthermore, Naka and Naoi (1995) did not find a copying benefit when Japanese-English adult bilinguals learned Japanese and English words, nor when English-speaking adults learned English words. Furthermore, no copying benefit was found when the Japanese-English bilinguals needed to recall the pronunciations of novel Arabic letters. In fact, on this task, reading was a more beneficial learning method than was copying.

In contrast to copying, writing appears to be beneficial for advanced learners. Ouellette (2010) found that writing was beneficial for second graders when the post-test was a spelling task. Similarly, Guan et al. (2011) found that writing was beneficial for English learners of Chinese who were enrolled in Elementary Chinese II. This benefit was apparent on tasks that required pure orthographic knowledge, lexical decision and partial cue-based character recognition. Xu et al. (2013) studied a similar population (students were enrolled in their 2nd semester of a Chinese course) and found that the writing was beneficial on a task that required pure orthographic knowledge (lexical decision) and on a task that required participants to write the character given a pinyin prompt.

Writing is not beneficial for novice participants or for tasks that require participants to produce a phonological form from an orthographic form. Chang et al. (2014) found that writing was not beneficial for English speakers enrolled in their first semester of Chinese classes (they had eight weeks of prior instruction). Thus, it seems that at least one semester of Chinese instruction is necessary before writing becomes beneficial. Furthermore, Guan et al. (2011) and Xu et al. (2013) did not find a benefit of writing on tasks that required participants to produce a phonological form from an orthographic form. In the Guan et al. (2011) study, participants were required to produce the pinyin for the learned characters. In the Xu et al. (2013) study, participants were shown a character and then saw a pinyin representation combined with a voice pronouncing the syllable. Participants had to decide whether the character matched the pronunciation. Because the character was shown first, this task primarily relied on producing a phonological form for the character.

Although the framework I proposed explains most of the discrepancies in the literature, it does not explain all of them. Notably, both Ouellette and Tims (2014) and Vaughn et al. (1993) used writing practice with relatively advanced learners (2^{nd} — 4^{th} graders learning their first language, English). The tasks involved either pure orthographic knowledge (orthographic legality) or producing an orthographic form from a phonological form (spelling). Nevertheless, a writing benefit was not found. It is unclear why this is the case. Two of the studies that demonstrated a writing benefit did so for participants learning individual Chinese characters (Guan et al., 2011; Xu et al., 2013). Ouellette (2010) found that writing was more beneficial than reading for 2^{nd} graders learning English pseudowords. In contrast, Ouellette and Tims (2014) compared writing to typing and Vaughn et al. (1993) compared writing to both tracing and typing. Thus, it is possible that writing is very beneficial for advanced learners learning visually complex isolated graphs (such as Chinese characters). However, for advanced learners learning words composed of a relatively small set of visually simple graphs (such as English letters), writing is more beneficial than reading but as effective as typing and tracing. This hypothesis may be especially true for participants who are proficient in typing (Ouellette & Tims, 2014). Thus, more research comparing writing to various other encoding methods in a variety of orthographies with participants varying in proficiency is necessary to identify the conditions in which writing is beneficial.

3.4.6 Use of artificial orthographies

This experiment was conducted with an artificial orthography and novice participants. These conditions allowed me to fully control the distribution of transparent and opaque graphs, the phonological difficulty of the graphs, and the background knowledge of participants. However, I do acknowledge some shortcomings to this approach. First, I chose to have equal numbers of transparent and opaque complex graphs to ease comparisons between the two. However, the distributional properties of the orthography may affect learning. Therefore, it is necessary to replicate this study with a more natural stimulus set.

Second, I chose to use novice participants to better control for prior knowledge and because copying has been shown to be more beneficial for novice participants. My results and previously published data (Chang et al., 2014; Cunningham & Stanovich, 1990; Guan et al., 2011; Longcamp et al., 2005; Naka, 1998; Vaughn et al., 1992; Xu et al., 2013) suggest that the benefit of copying may decrease with experience whereas the benefit of writing may increase with experience. In my study, students spent significantly less time copying than writing, but both tests led to equivalent post-test outcomes. Studies with more experienced students are needed to determine if writing is more beneficial than copying for those students.

The focus of the current study was on learning orthography-phonology connections for single akshara. Because Hindi has a highly (although not completely) transparent orthography, learning the orthography-phonology connections is a critical first step in learning to read Hindi.

Once a person can pronounce singleton akshara, it is relatively easy to pronounce whole words and to access the semantic referents of the words. Because my focus was on pronouncing isolated akshara, I did not embed the akshara in words and I did not teach semantics. However, I acknowledge that, although pronouncing single akshara is a critical precursor to reading development, word-level effects also play a role in reading. First, an understanding of word-level prosody and semantic support is important for correct pronunciation of the schwa vowel, which is not expressed in the orthography (Bhide, Gadgil, Zelinsky, & Perfetti, 2013; Nag, 2014; Pandey, 2014). Second, it is difficult to pronounce single complex akshara that cross syllable boundaries, but those complex akshara do occur in words. Studies of word reading have shown that blends are easier to learn than syllabic breaks (Nag, 2014). Therefore, it is important to do similar research with words rather than singleton akshara so that consonantal syllabic breaks can be included and the effects of word-level prosody and semantic support can be studied.

3.4.7 Educational implications

This study suggests that motor encoding is very important for early learning, especially for building pure orthographic knowledge and the ability to produce an orthographic form given a phonological form. The results of this study suggest that perhaps adding copying to the mobile game will improve orthographic learning. However, more testing is required before a strong recommendation can be made. The participants in the current study were much older than the demographic that will be using the game. Furthermore, the participants in the current study had no Hindi experience, whereas the game is targeted towards people with a couple of years of Hindi experience. Thus, it would be best to develop a version of the mobile game that includes copying and to compare it against the current version. If the version that includes copying produces superior results, as the results of this study suggest it will, then copying should be fully integrated into the game.

3.5 CONCLUSION

To conclude, this study suggested that transparent graphs are easier to learn that opaque graphs. Testing (instantiated in this study as writing) is not important for orthographic learning, although future studies need to elucidate whether this finding generalizes to other instances, or is only applicable during initial learning when accuracy is low. The fact that the benefit of copying decreases with experience but the benefit of writing increases with experience (Chang et al., 2014; Cunningham & Stanovich, 1990; Guan et al., 2011; Longcamp et al., 2005; Naka, 1998; Vaughn et al., 1992; Xu et al., 2013) suggests that more experienced students who have higher accuracy when writing find testing beneficial, but novice students who have lower accuracy when writing find copying beneficial.

Furthermore, this study has identified the specific instances in which motor encoding is helpful for learning. Specifically, I propose that motor encoding is helpful for tasks that require pure orthographic knowledge or the ability to produce an orthographic form given a phonological form. Motor encoding is less helpful when a person is given the orthographic form and asked to produce the phonological form. Copying and writing are equally beneficial during early learning, but copying is more time-efficient. Based on this research and that of others, I recommend that teachers use copying practice when teaching complex akshara to beginning learners.

4.0 EXPERIMENT 3: DISCRIMINATION OF PHONEMIC CONTRASTS

4.1 INTRODUCTION

The first experiment sought to improve phonological knowledge by including close phonological foils. However, qualitative analyses suggest that this was not sufficient to develop high quality phonological representations. Specifically, confusion between dental/retroflex and aspirated/unaspirated consonants was pervasive on the post-tests. Experiment 3 seeks to understand how pedagogical differences and individual differences predict second language (L2) phonemic perception. Specifically, it asks three questions: 1) can orthographic support aid L2 phonemic perception; 2) can manipulated utterances improve L2 phonemic perception; and 3) do individual differences predict L2 phonemic perception?

4.1.1 Hindi and Marathi phonemic contrasts

The experiment focuses on phonological representations for place of articulation and aspiration differences that do not occur in English. These phonemic differences are found in both Marathi and Hindi. Previous research has examined these contrasts in a Hindi context; the present study examines them in a Marathi context.

For place of articulation, discrimination of d̪/ɖ and t̪/ʈ is examined. Place of articulation of stop consonants refers to location at which the airflow is blocked. For example, to make the sound /p/, airflow is blocked by the lips. In both d̪ and t̪, airflow is blocked by the tongue touching the teeth, therefore they are referred to as dental stops. For both ɖ and ʈ, airflow is blocked by the tongue touching the back of the roof of the mouth and the tongue is slightly curled so that the underside touches the roof of the mouth (Cibelli, 2015; Verma & Chawla, 2003; Werker, 1989) They are called retroflex stops. The English realization of the /d/ and /t/ sounds is somewhere in between the two Hindi/Marathi realizations; the tongue touches the middle of the roof of the mouth (the alveolar ridge). The difference between the 'd' sounds and the 't' sounds is that the 'd' sounds have pre-voicing, the vocal folds begin to vibrate before the air release.

For aspiration, discrimination of k/k^h is examined. Aspiration or VOT refers to time between the release of the stop closure in the oral cavity and the onset of vocal fold vibration. In that period of time, the vocal tract is completely unobstructed and the air is released producing an aspiration noise. Whereas the presence versus the absence of aspiration in English cues a two-way voicing contrast (absence of aspiration signals voiced sounds while presence of aspiration cues voiceless stops), in Hindi and Marathi aspiration cues a three-way voicing contrast. Voiced sounds are defined by the absence of aspiration (/g/), as they are in English. Moreover, voiceless sounds with shorter aspirations (/k/) contrast with voiceless sounds with longer aspirations ($/k^h/$). The length of the aspiration in English voiceless sounds falls in between the short- and long-voiceless sounds in Hindi and Marathi.

Place of articulation contrasts in Hindi (especially of stop consonants) have been studied extensively. Because the dental/retroflex place of articulation is not contrastive in English, English-speaking adults find the contrast extremely difficult. Previous research has shown that English-speaking infants (~6 months old) can distinguish the contrast, but English speaking children of 4, 8, and 12 years of age as well as English-speaking adults cannot (Werker, Gilbert, Humphrey, & Tees, 1981; Werker & Tees, 1983).

Furthermore, previous research has shown that it is incredibly difficult to teach English-speaking adults to discriminate dental/retroflex place of articulation contrasts. Tees and Werker (1984) (see also Werker et al., 1981) had participants complete a category change discrimination paradigm with the t̪/ʈ distinction. Their participants heard a continuous string of one phoneme, followed by the other phoneme. They had to press a button when they detected the change occurring. On pre-test, only 1/15 participants showed evidence of successful discrimination. After 300 trials of training, only 6/14 participants could discriminate the phonemes. On a delayed post-test, only 3/15 participants could discriminate the phonemes. Thus, training had little influence on successful discrimination. Further studies showed that students with 1 year of a Hindi as a foreign language class also could not distinguish the phonemes, but students with 5 years of Hindi classes were able to. Werker and Tees (1984) studied the same contrast with an AX paradigm. They found that participants could discriminate the phonemes above chance (A’ score ~.8) but that training did not result in statistically significant improvement. Polka (1991) studied discrimination for place of articulation for all combinations of voicing and aspiration using an AX test. She found that discrimination for the t̪/ʈ distinction was significantly better than for the other three contrasts (~40% errors). Unlike previous research, she did find evidence for improved discrimination of the t̪/ʈ contrast as the experiment progressed. However, she did not find evidence of improvement for the other three contrasts.

As for aspiration contrasts, previous research suggests that these are relatively easy for native English speakers to discriminate even without training (Aggarwal, 2012; Guion & Pederson, 2007). This high level of performance likely stems from the fact that the contrast uses a voice onset time distinction that is similar to, although not identical to, the English g/k contrast (Guion & Pederson, 2007). More generally, aspiration contrasts are primarily distinguished by temporal acoustic dimensions (voice onset time) (S. Rosen, 1992). In contrast, although the dental and retroflex sounds do vary in terms of voice onset times (Hamann, 2003; Polka, 1991; Verma & Chawla, 2003), these temporal differences are very small. Formant information is a better cue of place of articulation than is durational information (S. Rosen, 1992). Burnham (1986) has proposed that sound contrasts that are distinguished primarily on temporal dimensions, such as the aspiration difference, are "robust" contrasts that can easily be re-learned in adulthood. In contrast, the dental/retroflex place of articulation is a "fragile" contrast that is more difficult to learn as an adult.

4.1.2 Orthographic support

Previous studies used purely auditory designs. The present study examines whether adding orthographic labels to the members of different categories can help with discrimination. Orthographic labels may help because previous research has shown that orthographic representations can shape phonological representations (Bhide et al., 2013; Castles, Holmes, Neath, & Kinoshita, 2003; Ehri & Wilce, 1980; Prakash, Rekha, Nigam, & Karanth, 1993; Seidenberg & Tanenhaus, 1979).

Many studies have demonstrated that orthographic representations can affect L2 learners' representations of L2 phonology. For example, Arabic students learning English will sometimes pronounce silent letters in words such as "walk" (Jesry, 2005). Likewise, German obstruents are unvoiced in word-final positions, but English learners of German often pronounce them as voiced because they are influenced by their spellings (e.g., "Bund") (Young-Scholten, 2002). Similarly, pinyin represents /p/ using the letter "b" and /p^h/ using the letter "p". However, English learners of Chinese often continue to pronounce these letters as they are pronounced in English (Meng, 1998). Furthermore, L2 learners of Chinese tend not to pronounce vowels that are not represented in pinyin (Bassetti, 2007; Ye, Cui, & Lin, 1997). They also omit these vowels in phoneme counting and phoneme segmentation tasks (Bassetti, 2006). A few lines of evidence suggest that these effects are due to orthography, and not due to articulatory constraints. For example, vowels are only omitted in pinyin in certain word positions. When the vowels are orthographically represented, the students have no trouble pronouncing them (Bassetti, 2007). Similarly, although Polish consonant clusters are difficult for both children and L2 learners to pronounce, they handle the difficulty differently because L2 learners are also exposed to orthographic input. Children often delete some of the consonants whereas L2 learners add sounds in between the consonants. Because L2 learners can see all of the consonants in the spelling, they make an effort to pronounce all of them (Young-Scholten, 1997; Young-Scholten, Akita, & Cross, 1999). An online vocabulary learning task in which some participants were exposed to the L2 orthography and phonology, whereas other participants only received the phonology, demonstrated that orthography can help learners develop more target-like phonological representation (Steele, 2005).

Although research has shown that L2 orthography can affect L2 phonological representations, less research has been done on how first language (L1) phonological transliterations of L2 words can affect L2 phonological representations. For example, when students learn Hindi as an L2 in the United States, they not only learn Hindi orthography but also L1 transliterations (e.g., when they learn the vocabulary word /t̪al/ (rhythm) they learn both the Hindi orthographic representation (ताल) and the English transliteration (taal) (Bhatia, 2008). However, this orthographic representation is problematic because the English "t" is alveolar, but the Hindi word uses a dental consonant. It is possible that the English transliteration will cause proactive interference and therefore support an (incorrect) alveolar phonological representation.

Transliterations are also ubiquitous in India; people are often exposed to English transliterations on signs and billboards (see Figure 4). Experiment 3 will examine how using L1 transliterations can affect L2 phonological representations. This research is of interest theoretically because it will further our knowledge about how orthography can affect phonological representations by being the first study (to the best of my knowledge) to examine how transliterations can affect phonological representations. It is also practically applicable because if the use of L1 transliterations negatively affects L2 phonological representations, it suggests that foreign language teachers should avoid using transliterations when their focus is on phonology and native-like pronunciation.

Figure 4: Examples of Hindi movie posters. These posters are written with English graphs, demonstrating the ubiquity of English transliterations in India. Furthermore, although the names of both movies begin with the letter 'T', the name of the movie on the left begins with the retroflex sound and the name of the movie on the right begins with the dental sound. The images for the posters were obtained from https://en.wikipedia.org/wiki/File:Tashan_Poster.jpg and https://en.wikipedia.org/wiki/Taal_(film)#/media/File:Taal.jpg. The use of scaled-down, low-resolution images of posters to provide commentary on the posters in a non-commercial, academic paper qualifies for fair use under copyright law of the United States.

4.1.3 Individual differences in learning phonemic contrasts

In addition to being interested in whether orthography can aid phoneme discrimination, I also wanted to know if individual differences affect L2 phonemic perception. I was specifically interested in two individual differences, rise time discrimination and English phonological awareness/decoding ability.

4.1.3.1 Rise time discrimination

Rosen (1992) suggested that phonemes are distinguished using envelope and fine structure, as well as other cues. Envelope refers to fluctuation in overall amplitude between 2 and 50 Hz whereas fine structure refers to faster fluctuations, between 600 Hz and 10 kHz. k/k^h is a voicing distinction. One cue for distinguishing voicing differences is envelope; voice onset time is a component of the envelope and /k^h/ has a longer voice onset time than does /k/. The t̪/ʈ and d̪/ɖ contrasts vary based on place of articulation. Place of articulation differences are primarily cued by fine structure. In terms of fine structure, there are spectral differences in the burst as well as formant differences in the following vowel that help distinguish dentals and retroflexes (Cibelli, 2015; Guion & Pederson, 2007; Hamann, 2003; Polka, 1991; Stevens & Blumstein, 1975; Verma & Chawla, 2003). One of the most stable differences that distinguishes dentals and retroflexes is that retroflexes have a lowered third formant during the vowel transition (Guion & Pederson, 2007; Hamann, 2003). However, there are some envelope differences between the contrasts as well; dental sounds have a longer voice onset time than do retroflex sounds (Hamann, 2003; Polka, 1991; Verma & Chawla, 2003).

Rise time discrimination is a measure of auditory perception that tests how sensitive people are to differences in duration in the time between sound onset and maximum volume. Rise time specifically refers to the onset of the amplitude envelope, and is therefore a component of the envelope structure. Differences in rise time discrimination thresholds predict phonological awareness and reading ability (Goswami, 2011; Goswami et al., 2002, 2010; Hämäläinen, Leppänen, Torppa, Müller, & Lyytinen, 2005; Muneaux, Ziegler, Truc, Thomson, & Goswami, 2004; Surányi et al., 2009). Although rise time discrimination has primarily been used to explain differences in syllable awareness, it can also explain differences in phonemic perception (Goswami, 2011). First, some phonemes vary in rise time (e.g., /ba/ is fast, /wa/ is slow). Differences in rise time discrimination could impair perception of those phonemic differences (Goswami, 2011; Goswami, Fosker, Huss, Mead, & Szucs, 2011). The contrasts used in the present study are all stop consonants, and thus do not have greatly different rise times. However, they do have different voice onset times, which is a portion of the sound envelope. If rise time discrimination measures sensitivity to differences in envelope structure generally, then rise time discrimination thresholds may predict ability to learn the sound contrasts in the present study.

The rise time discrimination task utilizes non-speech tones. It is unclear whether or not L2 phonetic learning is speech-specific. Díaz, Baus, Escera, Costa, and Sebastián-Gallés (2008) found that participants varying in L2 phonemic discrimination did not vary in mismatch negativity (MMN[16]) responses to tones of different lengths, tones of different frequencies, or differences in tone patterns. However, they did vary in MMN responses to L2 phonemic pairs. Thus, they suggest that individual differences are speech-specific, and cannot be explained by general auditory perception. In contrast, Slevc and Miyake (2006) found that tonal perception and production ability predicted L2 receptive and productive phonology. Specifically, they found that, in Japanese second language learners of English, tonal ability predicted their ability to discriminate and produce the r/l contrast, their ability to detect mispronunciations, and the severity of their foreign accent. Therefore, their study suggests that there is a relationship between the perception of non-speech

[16] The MMN is an event-related potential (ERP) component that is found during oddball paradigms. For example, if participants hear one different phoneme embedded in a string of the same phoneme, a large MMN will be found on that one different phoneme. The MMN can be used to test whether or not participants can discriminate two phonemes; an MMN response will only be seen if participants can make the discrimination (Näätänen, 2001).

and speech sounds. To summarize, previous research into whether the perception of non-speech sounds predicts L2 phonemic discrimination is mixed. It is important to note that neither of the studies described above examined rise time discrimination as a predictor, which is what the present study is measuring. Thus, it is unclear whether non-speech perception in general, and rise time discrimination in particular, will be able to predict learning of the Marathi phonemic contrasts. However, because rise time discrimination predicts L1 phonological skills, I predict that it will also predict learning of the Marathi phonemic contrasts.

4.1.3.2 Phonological awareness and decoding

Phonological awareness is awareness of the sounds in language and decoding is the ability to use grapheme-phoneme correspondences and blending to sound out novel words. I believe that English phonological skills may predict the ability to learn non-native phonemes, but I am unsure of the direction of the relationship.

Phonological awareness and decoding may positively predict the learning of non-native contrasts because there is large correlation between L1 and L2 phonological awareness (see Melby-Lervåg & Lervåg, 2011 for a meta-analysis). In contrast, there may also be a negative relationship between phonological skills and learning non-native contrasts. Strong English skills are predicated on having firm category boundaries for English phonemes; dyslexics often perceive phonemes allophonically rather than categorically (Bogliotti, Serniclaes, Messaoud-Galusi, & Sprenger-Charolles, 2008; Serniclaes, Heghe, Mousty, Carre, & Sprenger-Charolles, 2004; Serniclaes, Sprenger-Charolles, Carre, & Demonet, 2001). To learn the Hindi contrasts, participants must identify differences within one English phonemic category. Thus, people with more sensitivity to this allophonic variation (and thus lower English phonological skills) may be more successful at learning the non-native contrasts.

There is neural evidence to support both hypotheses. For example, Golestani et al. (2007) found that people who are better able to learn novel phonemic contrasts have larger white matter volumes in their left Heschl's gyrus, the location of the auditory cortex. Smaller primary Heschl's gyri are associated with reading and oral language deficits in the L1 (Leonard et al., 2002, 2011). Thus, it is possible that the same neural structures that support first language development also support the learning of novel contrasts. However, people who can quickly learn non-native contrasts are more likely to have a duplicated left Heschl's gyrus (Golestani et al., 2007). The presence of a duplicated Heschl's gyrus is associated with phonological dyslexia (Leonard et al., 2001). These results suggest that people with poor phonological skills may by better at learning non-native contrasts. Thus, the current neuroimaging data do not strongly bias us towards one hypothesis.

There is some recent behavioral evidence that provides strong evidence for the first hypothesis, that there will be a positive relationship between L1 phonological skills and the ability to learn novel sound contrasts. Gabay and Holt (2015) found that dyslexic adults were impaired at learning novel sound categories, as compared to control participants. Furthermore, both English phonological awareness and decoding positively predicted people's abilities to learn the novel sound categories. They believe that dyslexic participants' difficulty with acquiring the novel categories is associated with dyslexic individuals' impairment with procedural learning more generally. However, there a couple of important differences between their study and the present study. Most importantly, in their study, participants learned to categorize non-speech sounds. Therefore, participants did not need to learn to hear differences within an English phonemic category, which is what the present study requires. Furthermore, their participants included both dyslexic and typically-developing adults, whereas the present study only includes typically-developing adults. Finally, their study used an implicit learning paradigm, whereas the present study uses an explicit learning paradigm.

Given the evidence provided by Gabay and Holt (2015), I predict that there will a positive relationship between English phonological skills and learning of the Hindi contrasts. However, there is a small possibility that I will find a negative relationship instead.

4.1.4 Pilot study

The first iteration of this study used natural utterances by native Marathi speakers. Participants completed AX tasks in which they heard phonemes by two different speakers and had to indicate if they were saying the same phoneme (e.g., both saying /k/) or saying different phonemes (e.g., one said /k/, other said /k^h/). Although participants were able to learn the k/k^h contrast, they were not able to learn the d̪/ɖ and t̪/ʈ contrasts. To make the current iteration easier, I only had participants compare utterances within speaker. Furthermore, I artificially synthesized the t̪ and d̪ to make them more distinctive from their retroflex counterparts. Emphasizing differences has been shown to help people learn other non-native contrasts under certain conditions (Jamieson & Morosan, 1986; McCandliss et al., 2002). The present experiment tests if the use of manipulated utterances will lead to more learning than was seen in previous studies.

4.2 METHODS

4.2.1 Participants

Participants were 77 undergraduate students at the University of Pittsburgh. Three participants did not finish all four sessions and were thus excluded. Five participants were excluded due to experimenter error. The remaining 69 participants were aged 18-26 years (average age = 19 years) and there were 24 males. All participants were from the subject pool and received course credit for participating. Participants were also given a $10 bonus for coming to all four sessions on time, not rescheduling their appointments at the last minute, and remaining engaged throughout the tasks. All participants were native English speakers and had normal or corrected-to-normal vision and hearing, no diagnoses of language/reading disorders or speech impediments, no experience with a second language before the age of 12, were not proficient in a second language, and had no experience with any South Asian languages.

4.2.2 Overview

Participants learned three contrasts: t̪/ʈ, k/k^h, and d̪/ɖ. For brevity's sake, from now on the contrasts will be called the T, K, and D contrasts, respectively. They learned these contrasts in three orthography conditions: English orthography, Marathi orthography, and no- orthography. Which contrast was paired with which orthographic condition and the order of the contrasts were counterbalanced across participants. For the Marathi orthography, the corresponding Devanagari akshara were used. For the English orthography, the English letters used to transliterate Devanagari akshara in textbooks were used (Bhatia, 2008). See Table 18 for the corresponding graphs.

Table 18: The graphs in the two orthography conditions

Phoneme	Marathi graph	English graph	Phoneme	Marathi graph	English graph
t̪	त	t	ʈ	ट	T
k	क	k	k^h	ख	k^h
d̪	द	d	ɖ	ड	D

4.2.3 Procedure

The experiment took place over the course of four days. On the first day, participants completed a pre-test, learning phase, and post-test for the first contrast. On the second day, they completed a delayed post-test on the first contrast and completed a pre-test, learning phase, and post-test for the second contrast. On the third day, they completed a delayed post-test on the second contrast and completed a pre-test, learning phase, and post-test for a third contrast. On the fourth day, participants completed a delayed post-test on the third contrast and completed several individual difference measures.

The tasks included in the pre-test, learning phase, and post-test are detailed in Table 19.

Table 19: The tasks included in the pre-test, learning, and post-test phases of the experiment. Note that the same tasks were used for the immediate and delayed post-tests. Related tasks are displayed in the same color. FC ID = Forced Choice Identification

Pre-test	Learning	Post-test
Repetition	FC ID Phonemes Learning	Choose Correct Word Test
Introduce Phonemes	FC ID Words Learning	Repetition
FC ID Phonemes Test	Learning Words	FC ID Phonemes Test
FC ID Words Test	Choose Correct Word Learning	FC ID Words Test
	Learning Words	

4.2.3.1 Repetition

In this task, participants had to say the phonemes aloud to the best of their ability. First, they heard both phonemes. Then, they heard the first phoneme said three times by three different speakers and they had to say it aloud once into a recorder. Then, they heard the second phoneme said three times by three different speakers and they had to say it aloud into the recorder.

4.2.3.2 Introduce phonemes

This task was a learning activity, but it was done before the forced-choice identification phonemes pre-test because it provided information participants would need to do that task. In this task, participants heard both phonemes. Then they heard the first phoneme and, if they were in an orthography condition, saw its corresponding graph. The experimenter then taught them how to pronounce it. This process was repeated for the second phoneme.

Participants were taught how to articulate the phonemes in the following manner. For the D and T contrasts, participants were asked to produce the English alveolar realization. Then, they were asked to either move their tongue forward to touch their teeth or further back to produce the dentals and retroflexes respectively. For the K contrast, participants were asked to produce the English realization with slight aspiration while holding their hand in front of their mouth to feel the air release. They were then taught to produce "no air" or "lots of air" to produce the unaspirated and aspirated forms respectively. The experimenters worked with the participants until they produced the phonemes correctly at least once.

After learning how to articulate the phonemes, participants heard both phonemes said by six speakers in a random order. If they were in an orthography condition, they saw the corresponding graphs as well.

During this task, participants were given reference cards (see Figure 5). Participants could use these reference cards during all tasks except for the repetition pre-test.

t	Put tongue by teeth	Beginning of sound sharper
T	Put tongue in back of mouth	Beginning of sound rounder

द	Put tongue by teeth	Beginning of sound sharper
ड	Put tongue in back of mouth	Beginning of sound rounder

	Don’t release air	Sound cuts off promptly
	Release air	Sound ends breathily

Figure 5: Examples of the reference cards participants could use during the experiment. The top reference card is the T contrast in the English orthography condition, the middle reference card is the D contrast in the Marathi orthography condition, and the bottom reference card is the K contrast in the no- orthography condition.

4.2.3.3 Forced-choice identification phonemes

Forced-choice identification phonemes test. In this task, participants heard the same person pronounce both phonemes in the contrast. Participants had to indicate which phoneme was the dental (in the case of the D or T contrast) or unaspirated (in the case of the K contrast). The instructions varied by learning condition. In the English and Marathi orthography conditions, participants were shown the graph of the phoneme they should choose. In the no- orthography condition, participants heard which phoneme they should select and which phoneme they should not select. The phoneme that participants should select was always listed first on their reference cards and experimenters made sure they were aware of that.

Phonemes from six different people were used, with each pair played four times, for a total of twenty-four stimuli. For each person, two of the pairs were in one order (e.g., dental followed by retroflex) and two were in the opposite order. The stimulus pairs were presented in a random order.

Forced-choice identification phonemes learning. The learning phase was the same as the test but participants were given feedback after every response. In the no- orthography condition, they were shown correct/incorrect. In the orthography condition, they were shown correct/incorrect and some information about the graphs (e.g., *Correct! The first sound was d, the second sound was D*).

4.2.3.4 Forced-choice identification words

This task was very similar to forced-choice identification phonemes, but participants heard words that began with the phonemes of interest. Participants had to indicate which word began with the dental (in the case of the D or T contrast) or unaspirated (in the case of the K contrast) phoneme.

There were twenty stimulus pairs in total. Half of the stimulus pairs began with one phoneme (e.g., d̪) when spoken correctly in Marathi, half of the stimulus pairs began with the other phoneme (e.g., ɖ). The words were grouped by speaker to make the task easier. There was a test version (no feedback) and a learning version (with feedback; feedback varied by orthography condition in the same way that the forced-choice identification phonemes feedback varied by orthography condition). The same words were used in the test and learning versions, but they were said by different speakers.

4.2.3.5 Learning words

In this task, participants learned Marathi words that began with the phoneme of interest. They saw a picture that depicted the meaning of the word and heard the word pronounced. In the orthography conditions, they also saw the word spelled using English or Marathi graphs. The word was in black and the graph of interest was in red. Participants were instructed to repeat the word out loud to the best of their ability, while focusing on listening for, remembering, and accurately pronouncing the first sound.

There were twenty words, half of the words began with one phoneme and half the words began with the other. The participants cycled through each word three times in a random order. Each time, the word was pronounced by a different speaker.

4.2.3.6 Choose correct word

Choose correct word test. In this task, participants saw the picture that depicted the meaning of the Marathi word they had just learned. They then heard two pronunciations of the word, one correct and one incorrect in that its initial phoneme was incorrect. Participants had to indicate which pronunciation was correct. The speaker was never the same as the three speakers in the learning condition. The participants went through each of the twenty words once and the words were presented in a random order.

Choose correct word learning. The learning phase was the same as the test but participants got feedback after every response. In the orthography conditions, participants saw correct and incorrect and the graph representing the phoneme the word should begin with (e.g., *Correct! That word should begin with a d sound*). In the no- orthography condition, they saw *Correct/Incorrect! That word should begin with [audio recording of the phoneme]*.

4.2.3.7 Individual difference measures

On the fourth day, participants completed four individual difference measures: Language History Questionnaire, phonological awareness test, decoding test, and rise time discrimination test. All participants took the Language History Questionnaire. The other three individual difference measures were added halfway through data collection so that data is only available from 39 participants. The following semester, the participants who did not complete those individual difference measures were invited to return the laboratory and complete those tasks. This testing took 30 minutes and participants were compensated $10. Seven participants elected to return, so individual difference measures are available from 46 participants in total.

Language History Questionnaire. The participants' language background was assessed using the Language History Questionnaire (Tokowicz, Michael, & Kroll, 2004).

Phonological awareness test. The phonological awareness test (PHAT; Olson, Wise, Conners, & Rack, 1989; Perfetti & Hart, 2002) was used to measure phonological awareness. In this test participants are given a word (e.g., middle) and are asked to remove a phoneme to form another word. (e.g., *middle* without the /d/ is *mill*). They are then asked to add another phoneme in its place (e.g., add /s/ to make *missile*). The stimuli are purposely chosen such that orthography is minimally helpful in order to get a pure measure of phonological awareness. The test is scored based on the number of correct answers.

Decoding test. The Real Word Test (Olson et al., 1989; Perfetti & Hart, 2002) was used to measure decoding skill. In this test, participants are shown a list of pseudowords. They have to mark which words are phonologically identical to real words (e.g., *serkyouler* sounds like *circular*, but *dofter* does not sound like a real word). The score is a d' score.

Rise time discrimination test. Rise time discrimination threshold was measured using the "dinosaur game" with an adaptive staircase procedure (Huss, Verney, Fosker, Mead, & Goswami, 2011; Levitt, 1971). Participants had to judge which of three sounds had a slower rise time (began more softly). The test is scored based on the millisecond threshold participants could discriminate and a lower score indicates better performance. Listen to Supplementary Audio 1 to hear a tone with a fast rise time and Supplementary Audio 2 to hear a tone with a slow rise time.

4.2.4 Materials

Materials were audio recordings from six native speakers of Marathi (3 male). Each speaker produced each phoneme three times and the best recording from among them was chosen. Each speaker said each word once. All six audio recordings of each word were not needed, so for each word the best examples were chosen. For a few stimuli, there were not enough good recordings from the six speakers. For those stimuli, additional recordings were taken from another native speaker of Marathi (female) and a heritage speaker of Marathi (female). All of the recordings were cut and cleaned of background noise.

Piloting showed that the D and T contrasts were particularly difficult for people. To make them easier, Praat was used to manipulate the sound files. Previous research has shown that dentals have longer voice onset times than do retroflexes (Hamann, 2003; Polka, 1991; Verma & Chawla, 2003). To emphasize this difference, the voice onset time for all dentals was lengthened (see Figures 7 and 8).

Furthermore, previous research has shown that English speakers have difficulty distinguishing the voicing contrast when listening to the ‘d’ and ‘t’ sounds (Polka, 1991). One way to help participants hear the pre-voicing in ‘d’ sounds (so they do not sound like ‘t’ sounds) is to increase the duration of the pre-voicing. Because dental d̪ sounds have shorter pre-voicing than retroflex ɖ sounds (Verma & Chawla, 2003), the pre-voicing duration of the dental d̪ sounds was increased.

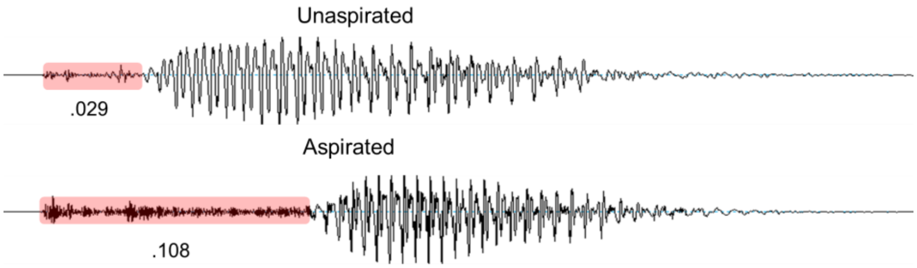

Figure 6: Example sound waveforms of K phoneme recordings from one speaker. The red box shows the voice onset time (s). The aspirated phoneme has a longer VOT than does the unaspirated phoneme. Listen to Supplementary Audio 3 and 4 to hear the /k/ and /k^h/ sounds, respectively.

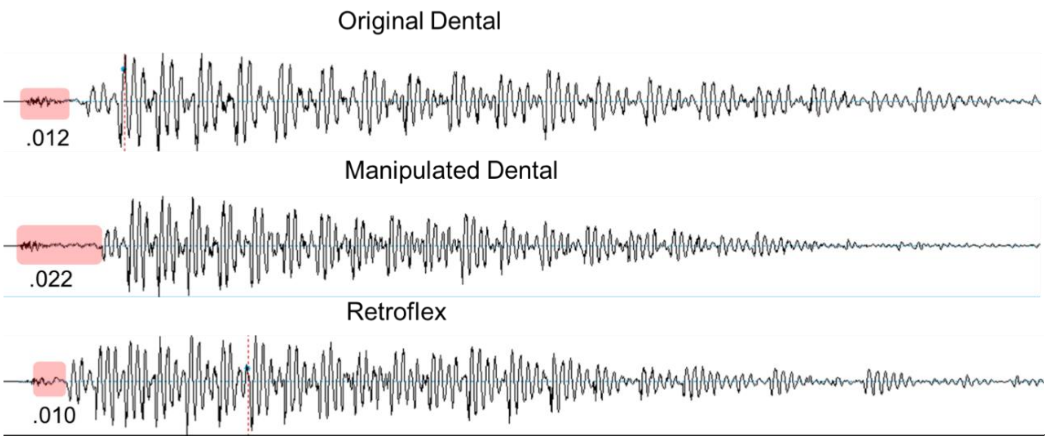

Figure 7: Example sound waveforms of T phoneme recordings from one speaker. The red box shows the voice onset time (VOT). The retroflex has a slightly shorter VOT than the original dental. The length of the VOT is increased in the manipulated dental. Listen to Supplementary Audio 5, 6, and 7 to hear the Original Dental, Manipulated Dental, and Retroflex 't' sounds, respectively.

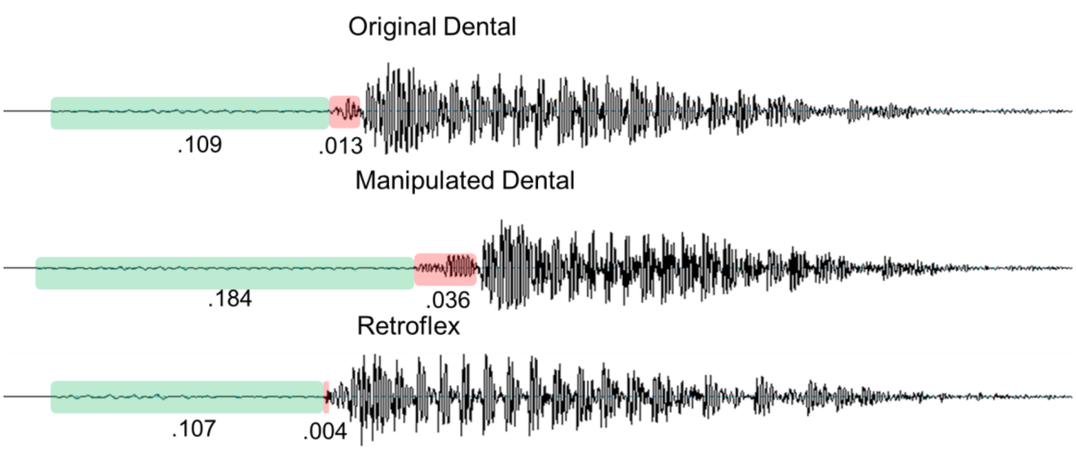

Figure 8: Example sound waveforms of D phoneme recordings from one speaker. The red box shows the voice onset time (s). The retroflex has a slightly shorter VOT than the original dental. The green box shows the pre-voicing (s). The retroflex and original dental have similar pre-voicing durations. The length of the pre-voicing duration and the VOT is increased in the manipulated dental. Listen to Supplementary Audio 8, 9, and 10 to hear the Original Dental, Manipulated Dental, and Retroflex 'd' sounds, respectively.

The duration values are in the Tables 20-22. Illustrations of what the duration values are measuring are shown in Figures 6-8. For the Learning Words task, unpaired t-tests were used to test whether the duration values significantly varied across conditions. For the all the other tasks, because the same words were used across conditions, paired t-tests were used to test whether the duration values significantly varied across conditions.

Table 20: The voice onset time (in s) of the K stimuli. The durations are displayed as mean (standard deviation).

	Unaspirated	Aspirated	t-test
Phonemes	0.033 (0.005)	0.116 (0.027)	t (5) = 6.652, *p* = .001
FC ID Word Learning	0.029 (0.010)	0.108 (0.024)	t (19) = 13.788, *p* < .001
FC ID Word Test	0.030 (0.009)	0.094 (0.039)	t (19) = 7.344, *p* < .001
Learning Words	0.023 (0.008)	0.095 (0.026)	t (28) = 24.933, *p* < .001
Choose Correct Word	0.027 (0.010)	0.099 (0.021)	t (19) = 12.214, *p* < .001

Table 21: The voice onset time (in s) of the T stimuli

	Mean (standard deviation)			t-test		
	Original Dental	Manipulated Dental	Retroflex	Original Dental Compared to Manipulated Dental	Original Dental Compared to Retroflex	Manipulated Dental Compared to Retroflex
Phonemes	0.014 (0.004)	0.023 (0.006)	0.011 (0.002)	t (5) =5.116, *p* = .004	t (5) = 1.697, *p* = .150	t (5) = 3.895, *p* = .011
FC ID Word Learning	0.018 (0.008)	0.030 (0.012)	0.012 (0.003)	t (19) = 6.935, *p* < .001	t (19) = 5.163, *p* < .001	t (19) = 8.223, *p* < .001
FC ID Word Test	0.016 (0.006)	0.031 (0.010)	0.014 (0.004)	t (19) = 10.777, *p* < .001	t (19) = 2.550, *p* = .020	t (19) = 9.566, *p* < .001
Learning Words	0.016 (0.006)	0.031 (0.010)	0.014 (0.004)	t (29) = 10.777, *p* < .001	t (28) = 2.550, *p* = .020	t (28) = 9.566, *p* < .001
Choose Correct Word	0.018 (0.006)	0.022 (0.007)	0.017 (0.006)	t (19) = 2.241, *p* = .037	t (19) = 1.159, *p* = .261	t (19) = 2.516, *p* = .021

Table 22: The voice onset time and pre-voicing duration (in s) of the D stimuli

		Mean (standard deviation)			t-test		
		Original Dental	Manipulated Dental	Retroflex	Original Dental Compared to Manipulated Dental	Original Dental Compared to Retroflex	Manipulated Dental Compared to Retroflex
Pre-voicing	Phonemes	.130 (.052)	.166 (.049)	.176 (.108)	t (5) = 4.441 *p* = .007	t (5) = 1.969 *p* = .106	t (5) = 0.328 *p* = .756
	FC ID Word Learning	.143 (.048)	.164 (.054)	.138 (.037)	t (19) = 6.728 *p* < .001	t (19) = 0.540 *p* = .596	t (19) = 2.841 *p* = .010
	FC ID Word Test	.130 (.061)	.143 (.066)	.134 (.035)	t (19) = 2.595 *p* = .018	t (19) = 0.357 *p* = .725	t (19) = 0.741 *p* = .468
	Learning Words	.110 (.038)	.141 (.044)	.115 (.031)	t (29) = 15.235 *p* < .001	t (28) = 0.471 *p* = .641	t (28) = 2.781 *p* = .010
	Choose Correct Word	.117 (.047)	.145 (.055)	.114 (.027)	t (19) = 5.859 *p* < .001	t (19) = 0.192 *p* = .850	t (19) = 2.584 *p* = .018
Voice Onset Time	Phonemes	.011 (.003)	.032 (.007)	.007 (.004)	t (5) = 6.314 *p* = .001	t (5) = 1.808 *p* = .130	t (5) = 6.322 *p* = .001
	FC ID Word Learning	.011 (.003)	.025 (.007)	.008 (.004)	t (19) = 8.896 *p* < .001	t (19) = 2.728 *p* = .013	t (19) = 11.678 *p* < .001
	FC ID Word Test	.011 (.004)	.021 (.008)	.008 (.005)	t (19) = 5.050 *p* < .001	t (19) = 2.393 *p* = .027	t (19) = 5.394 *p* < .001
	Learning Words	.013 (.005)	.028 (.011)	.008 (.004)	t (29) = 7.265 *p* < .001	t (28) = 4.510 *p* < .001	t (28) = 12.895 *p* < .001
	Choose Correct Word	.010 (.003)	.023 (.005)	.008 (.003)	t (19) = 8.122 *p* < .001	t (19) = 1.833 *p* = .083	t (19) = 10.868 *p* < .001

As you can see from Table 20, the voice onset times of the /k^h/ stimuli are reliably longer than the voice onset times of the /k/ stimuli. Table 21 shows that the original dental t̪ tokens typically have longer voice onset times than do the retroflex ʈ tokens, although the difference was not reliable in every comparison. However, the manipulated t̪ tokens do have reliably longer voice onset times than do the retroflex ʈ tokens. Table 22 shows that the original dental d̪ tokens typically have longer voice onset times than the retroflex ɖ tokens, although the difference was not always reliable. The manipulated d̪ tokens reliably have longer voice onset times than the ɖ retroflex tokens. Unlike previous research (Verma & Chawla, 2003), in these stimuli, the dental d̪ tokens and the retroflex ɖ tokens have equivalent pre-voicing durations. However, this is not problematic because the goal of increasing the pre-voicing duration was to help participants distinguish the voicing contrast.

The materials were validated by having a heritage Marathi speaker perform all the tasks. She performed the forced-choice identification tests first, followed by the force-choice identification learning activities, so that the feedback during the learning would not influence her. She did not do the forced-choice identification phonemes learning because the same items were used at learning and test. However, she did do the forced-choice identification words learning because different items were used at learning and test. Then, she went through the Learning Words once before taking the Choose Correct Word Tests. She performed well on all the measures (all scores $\geq$ 87.5%; see Table 23), suggesting that the materials were sound. Furthermore, she reported not being able to tell that some of the materials had been digitally manipulated.

Table 23: One heritage Marathi speaker's performance on the tasks. She performed well on all the tasks, demonstrating that materials were good examples of the contrasts.

Measure	Score K	Score T	Score D
FC ID Phonemes Test (out of 24)	24	21	22
FC ID Words Test (out of 20)	20	20	20
FC ID Words Learning (out of 20)	20	18	18
Choose Correct Word (out of 20)	20	18	19

4.3 RESULTS

4.3.1 Language History Questionnaire

Because language background strongly shapes phonemic perception, it is important to ensure that differences in language background did not drive performance on the current task. On the first day of the experiment, participants confirmed that they were native speakers of American English, had not learned a second language before the age of 12, were not fluent in a second language, and had no experience with a South Asian language. To further probe language history, the Language History Questionnaire (Tokowicz et al., 2004) was administered.

Importantly, on the questionnaire, no participants reported any experience with any South Asian languages nor travel to South Asia.

I was most concerned about languages that participants were moderately proficient in and languages that participants heard from a young age. The reason I was concerned about languages that participants heard at a young age is because studies have shown that language exposure during the first couple of years of life can improve phonemic perception for the relevant contrasts, even when people do not hear the language later in life, never spoke the language, and cannot understand the language (Oh, Jun, Knightly, & Au, 2003; Tees & Werker, 1984). To determine which languages participants were moderately proficient in, I looked at a series of four questions that asks participants to report which languages they speak/read/write/understand fluently. To assess which languages participants had been exposed to from a young age, I looked at a series of three questions that asks participants “What languages were spoken in your home while you were a child and by whom?; Please list the language(s) your mother/father speak”.

All participants reported that the only language they could speak fluently was English. Three participants reported by able to read Spanish and one participant reported being able to read minimal French. One participant reported being able to write Spanish and another reported being able to write minimal French. In terms of being able to understand a spoken language, three understood German, eighteen understood Spanish, one understood some Russian, one understood Hebrew, and one understood minimal French. Thus, the languages with which the participants had some level of proficiency were Spanish, German, Russian, Hebrew, and French. Importantly, none of those languages use the contrasts used in this study. All five languages have a velar, unvoiced, unaspirated stop consonant (/k/), but do not have an aspirated counterpart (Bolozky, 1997; Defior & Serrano, 2017; Grigorenko, Kornilov, & Rakhlin, 2017; Landerl, 2017; Walker, 1984). All five languages have ‘d’ and ‘t’ phonemes, but they only have one place of articulation. In Hebrew they are realized in a dento-alveolar position (Bolozky, 1997), in German they are realized in an alveolar position (Landerl, 2017), in Spanish they are realized in a dental position (Defior & Serrano, 2017), in Russian they are realized in a dental/alveolar position and can vary by palatalization (Grigorenko et al., 2017), and in French they are realized in an apico-dental position (Walker, 1984). Although they are realized in slightly different positions in the five languages, what is important is that the languages do not contrast within ‘d’ and ‘t’ sounds by place of articulation. Thus, although some participants had some degree of proficiency in these five languages, this language background should not help them with the present task.

In terms of language exposure at a young age, two people reported that German was occasionally spoken in their households, one person reported that her parents spoke Korean with each other, one person reported that his parents spoke Tagalog with each other, one person reported that Polish was spoken in her household, one person reported that her grandmother occasionally used Slovakian phrases, one person reported that Spanish was spoken in household, and one person reported that her mother knows German and Latin. Thus, the languages the participants may have been exposed to at a young age are German, Korean, Tagalog, Polish, Slovakian, Spanish, and Latin. As explained above, German and Spanish do not use any of the contrasts of interest, and therefore exposure to those languages should not affect performance in the current study. In Slovakian, ‘t’ and ‘d’ are realized in an alveolar position (Hanulíková & Hamann, 2010), in Polish they are realized in a dental position and can vary by palatalization (Gussmann, 2007), and in

Tagalog and Latin they are realized in a dental position (Llamzon, 1966; McCullagh, 2011). Korean only has ‘t’, it does not have the voiced counterpart ‘d’. The ‘t’ is realized dentally (Chung, 2007). Importantly, the four languages do not contrast within ‘d’ and ‘t’ sounds by place of articulation. Tagalog, Polish, and Slovakian have a velar, unvoiced, unaspirated stop consonant (/k/), but do not have an aspirated counterpart (Gussmann, 2007; Hanulíková & Hamann, 2010; Llamzon, 1966). However, Korean and Latin have both /k/ and /k^h/ (Chung, 2007; McCullagh, 2011). Therefore, languages participants were exposed to at a young age should not help them learn the place of articulation contrast in this study. Two participants may have been able to better hear the aspiration contrast because they were exposed to Korean and Latin. However, because their language exposure should only help with one contrast, and that contrast was relatively easy for many participants, I elected to keep them in the study.

4.3.2 Descriptive statistics

The descriptive statistics for the sound contrast tasks and individual difference measures can be found in Tables 24 and 25, respectively.

Table 24: Accuracy on the sound contrast tasks. The top row shows average (standard deviation) and the bottom row shows the range in percent accuracy. Note that not all participants were included in the choose correct word column. Only the data of participants who scored over 70% on the immediate post-test of the forced-choice identification phonemes and words tests were analyzed and therefore only their scores are displayed.

Contrast	Test	FC ID Phonemes	FC ID Words	Choose Correct Word
K	Pre-test	86.5 (14.7) 41.7 — 100	79.4 (20.0) 20.0 — 100	
	Immediate Post-test	93.3 (13.3) 16.7 — 100	88.6 (12.8) 35.0 — 100	74.1 (16.8) 35.0 — 100
	Delayed Post-test	94.8 (8.7) 54.2 — 100	91.7 (9.3) 60.0 — 100	69.5 (17.6) 35.0 — 100
T	Pre-test	71.7 (21.3) 20.8 — 100	56.2 (23.7) 15.0 — 100	
	Immediate Post-test	80.1 (18.2) 33.3 — 100	68.3 (21.3) 20.0 — 100	66.5 (13.3) 40.0 — 90.0
	Delayed Post-test	79.1 (20.8) 12.5 — 100	67.5 (23.0) 20.0 — 100	62.1 (14.3) 30.0 — 90.0
D	Pre-test	58.2 (23.6) 4.2 — 100	52.2 (18.0) 15.0 — 85.0	
	Immediate Post-test	73.1 (18.4) 33.3 — 100	62.8 (16.4) 25.0 — 90.0	67.6 (16.6) 35.0 — 95.0
	Delayed Post-test	74.9 (19.5) 8.3 — 100	62.5 (17.9) 25.0 — 100	59.2 (15.4) 30.0 — 90.0

Table 25: Performance on the individual difference measures. Only the 46 participants for whom the data is available are included. The top row shows average (standard deviation) and the bottom row shows the range.

Phonological Awareness Score (out of 38)	Decoding (d')	Rise Time Discrimination (Threshold in ms)
29.522 (6.765) 12 — 38	2.410 (0.774) -0.314 — 3.726	90.398 (68.608) 26.573 — 260.829

A few trends are easily visible just by looking at the descriptive statistics. For both forced choice identification tasks, participants improved from pre-test to immediate post-test and maintained that level of performance on the delayed post-test. The K contrast was easier than the T contrast which in turn was easier than the D contrast. In general, the forced-choice identification phonemes was easier than the forced-choice identification words. For the choose correct word task, participants were able to learn the words and performed at an above-chance level. Although their performance decreased slightly between the immediate and delayed post-tests, they still remembered some of the words.

Remember that, for all tasks, chance is 50%. On the D forced-choice identification phonemes pre-test and the D and T forced-choice identification words pre-tests, participants were, on average, close to chance. On the other tasks, participants were, on average, generally above chance. It is interesting to note that on some of the forced-choice identification tasks, some participants were significantly below chance. Thus, these participants were able to discriminate the phonemes, but were classifying them incorrectly. It is unlikely that the participants were simply confused about the instructions because the forced-choice identification learning task gave them feedback on every trial. If they were confused about the instructions, they would have realized this immediately on the learning task and performed above chance on post-test. Thus, I believe that these participants were able to discriminate the phonemes but were classifying them incorrectly.

For both the sound contrast tasks and the individual difference measures, a large degree of variability is evident. It was especially promising to see significant individual differences in the rise time discrimination task because that task is often used to compare dyslexics to typically developing readers. This wide distribution makes it easier to see if the orthography manipulation or individual differences are able to predict the learning of the sound contrasts. Floor effects are not evident. There may be ceiling effects for the K contrast on the forced-choice identification post-tests that need to be accounted for during analysis.

4.3.3 Pronunciation

Three audio files are missing because they were inaudible and six audio files are missing because they were not backed up properly.

The D and T contrasts were coded by a heritage speaker of Hindi. The K contrast was coded by a heritage speaker of Marathi. The coders marked each utterance as either correct or incorrect, and if it was incorrect, they marked which phoneme was uttered. To gauge inter-rater reliability, the heritage Hindi speaker coded a random selection of 50 K recordings. Similarly, the heritage Marathi speaker coded a random selection of 50 D and 50 T recordings. The kappa values for the accuracy rating for the D, T, and K contrasts were .508, .517, and .674, respectively. According to Altman's kappa benchmark scale (Altman, 1991), these kappa values are in the moderate to good range.

The K contrast was coded as follows: when participants had minimal aspiration it was coded as /k/ (correct pronunciation of Marathi unvoiced, unaspirated phoneme), when participants had slight aspiration it was coded as /k^h/ (slight aspiration as done in English), when participants had a lot of aspiration it was coded as /k^{hh}/ (correct pronunciation of Marathi unvoiced, aspirated phoneme), and when participants had pre-voicing it was coded as /g/.

When participants were meant to produce the /k/ phoneme, they were correct 50.5% of the time. The /k^{hh}/, /k^h/, and /g/ errors were made 31.1%, 12.6%, and 5.8% of the time respectively. When participants were meant to say /k^{hh}/, they were correct 97.6% of the time. 2.4% of the time they mistakenly said /k^h/.

When participants were supposed to say /d̪/, /ɖ/, /t̪/, and /ʈ/, they were correct 33.0%, 62.6%, 53.1%, and 47.8% of the time, respectively. The error types are listed in Table 26.

Table 26: Frequency of response types (in percents) when participants were attempting to say the D and T contrasts

		Response												
		Correct	d	ɖ	d̪	t	ʈ	t̪	k	g	n	p	b	tʃ
Phoneme	d̪	33.0	20.7	25.1		1.5	4.9	13.8	0	0	0	0.5	0.5	0
	ɖ	62.6	15.8		6.9	0	9.9	3.4	0.5	0.5	0.5	0	0	0
	t̪	53.1	7.7	1.9	1.9	20.3	11.6		2.9	0.5	0	0	0	0
	ʈ	47.8	13.0	13.5	1.0	15.0		5.8	1.9	1.4	0	0	0	0.5

The pronunciation data were analyzed using binomial linear mixed effects models. Two models were used: one with all the participants and one with all the participants for whom individual difference measures were available.

For the models, all significant and marginal effects are reported.

4.3.3.1 Pronunciation (all participants)

A binomial linear mixed effects model with accuracy as the dependent measure was used to analyze the data. The predictors were phoneme contrast (D, T, K) and the interaction between orthography (no-, English, Marathi) and test (pre-test, immediate post-test, delayed post-test). For phoneme contrast, orthogonal contrasts comparing K to D & T and comparing D and T to each other were used. For orthography, orthogonal contrasts comparing English orthography to no- & Marathi and comparing no- and Marathi to each other were used. For test, orthogonal contrasts comparing pre-test to both post-tests and comparing the post-tests to each other were used. The model also included a random intercept for both subjects and items.

Because participants said each phoneme three times, there was only a small amount of data and hence not every variable of interest was able to be included in the analysis. Specifically, the order in which the phoneme was learned was not included, but exploratory analysis showed that this was not a significant predictor. Ideally, a three-way interaction between phoneme, orthography, and test would have been included, but this was not possible so only the two-way interaction was included. Finally, although the random slope that allowed phoneme to vary by subjects explained a large portion of the variance, the model did not converge when this random slope was included[17]. The other random slopes did not explain a large portion of the variance.

Participants were marginally better at the K contrast than the D & T contrasts, z = 1.877, $p = .061$. The odds of pronouncing the K contrast correctly were 7.681 times higher than pronouncing the D & T contrasts correctly. Participants improved from pre-test to post-tests, z = 4.646, $p < .001$. The odds of pronouncing the phonemes correctly were 1.972 times higher on the post-tests than the pre-test.

4.3.3.2 Pronunciation with individual differences

A very similar model as the one used for Pronunciation (all participants) was used for this analysis. Participants were restricted to those for whom individual difference measures were available. Because the scores on the Real Word Test and the PHAT were highly and significantly correlated (r = .530, $p < .001$), those measures were combined into a composite measure. Specifically, scores on both tests were converted into z-scores and then the z-scores were averaged to form a composite phonological awareness/decoding (PAD) measure. Rise time discrimination was not correlated

[17] The statement that the random slope explained a large portion of the variance is based on the model output from the non-convergent model.

with either the Real Word test or the PHAT test, r = .067, p = .654 and r = .016, p = .914, respectively. Rise time discrimination was z-scored. Interactions between test and both PAD and rise time discrimination were included in the model.

Importantly, the same pattern of effects from the model with all participants was present in the model with individual differences. Participants performed better on the K contrast than on the D & T contrasts, z = 2.268, p = .023. Participants improved from pre-test to post-test, z = 3.348, p = .001.

Importantly, both rise time discrimination and PAD score predicted pronunciation ability. People with better rise time discrimination (z = -3.405, p = .001) and better phonological awareness/decoding (z = 3.231, p = .001) performed better on the pronunciation task. A one standard deviation increase in both rise time discrimination and PAD were associated with a 1.509 and 1.565 times increase in the odds of correctly pronouncing the phonemes, respectively.

4.3.4 Forced-choice identification analyses

The forced-choice identification data were analyzed using binomial linear mixed effects models. For both the forced-choice identification phonemes and forced-choice identification words, two models were run: one with all the participants and the second with all the participants for whom individual difference measures were available.

For the models, all significant and marginal effects are reported.

4.3.4.1 Forced-choice identification phonemes

Forced-choice identification phonemes (all participants). A binomial linear mixed effects model with accuracy as the dependent measure was used to analyze the data. The predictors were order (1st, 2nd, 3rd contrast learned), phoneme contrast (D, T, K), orthography (no-, English, Marathi), and test (pre-test, immediate post-test, delayed post-test). For phoneme contrast, orthogonal contrasts comparing K to D & T and comparing D and T to each other were used. For order, the first contrast learned was used as the baseline. For orthography, orthogonal contrasts comparing English orthography to no- & Marathi orthography and comparing no- and Marathi to each other were used. For test, orthogonal contrasts comparing pre-test to both post-tests and comparing the post-tests to each other were used. Ideally, the model would have included an interaction between phoneme contrast, orthography, and test. However, it was not possible to include the three-way interaction because participants were nearly at ceiling on the K contrast at post-test. Therefore, the model included an interaction between orthography and test only. The model also included all random effects explaining a large portion of the variance: 1) random intercept for subjects, 2) random intercept for the sound pair, 3) random intercept for the sound pair in a particular order, 4) the effect of phoneme contrast was allowed to vary by subjects, and 5) the effect of test was allowed to vary by subjects[18].

The K contrast was easier than the T contrast which in turn was easier than the D contrast. The odds of getting the answer correct were 6.166 times higher in the K condition than in the D & T conditions, z = 6.547, $p < .001$. The odds of the getting the answer correct were 1.947 times higher in the T condition than in the D condition, z = 2.323, $p = .020$.

Participants significantly improved from pre-test to post-test, z = 8.970, $p < .001$. The odds of getting the answer correct were 2.276 times higher on the post-tests than on the pre-test.

Participants performed significantly better on the second and third phonemes they learned than on the first phoneme they learned, z = 2.076, $p = .038$ and z = 3.840, $p < .001$, respectively.

[18] Although the model did not converge, the relative gradient was equal to 0.001.

Forced-choice identification phonemes with individual differences. A very similar model as the one used for forced-choice identification phonemes was used to analyze these data. The participants were restricted to those for whom individual difference measures were available. The individual difference measures were transformed in the same manner as they were for pronunciation with individual differences. Interactions between test and both rise time discrimination and PAD[19] were included.

Importantly, the same pattern of effects in the model with all participants was also present in the model with individual differences. Participants performed better on the K contrast than on the D & T contrasts, z = 5.434, $p < .001$. Participants also performed better on the T contrast than on the D contrast, z = 2.583, $p = .010$. Participants improved from pre to post-tests, z = 7.974, $p < .001$. Participants performed better on the second and third contrasts they learned than on the first contrast they learned, z = 1.757, $p = .079$ and z = 2.616, $p = .009$.

There were two marginal effects in the present model that were not present in the model with all participants. First, there was a main effect of orthography, with the no- & Marathi orthography conditions outperforming the English orthography condition, z = 1.779, $p = .075$. Specifically, the English orthography condition lowered the odds of answering correctly by 1.376 times. Second, there was an interaction between test (pre/post contrast) and orthography (English/no- & Marathi contrast), z = 1.888, $p = .059$. This interaction was driven by the fact that, although there was overall improvement from pre-test to post-tests, there was more improvement in the no- & Marathi orthography conditions.

Importantly, individual differences significantly predicted performance on the task. There was a significant interaction between test (pre/post contrast) and PAD, z = 2.494, $p = .013$. There

[19] Although the model did not converge, the relative gradient was less than 0.001.

was also a marginal interaction between test (pre/post contrast) and rise time discrimination, z = -1.822, $p = .068$. Participants with better PAD and rise time discrimination improved more from pre-test to post-tests.

4.3.4.2 Forced-choice identification words

Forced-choice identification words (all participants). This model had the same fixed effects structure as the model for forced-choice identification phonemes. The model also included all random effects that explained a large portion of the variance: 1) random intercept for subjects, 2) random intercept for the speaker saying the words, 3) random intercept for word itself, 4) the effect of phoneme contrast was allowed to vary by subjects, and 5) the effect of test was allowed to vary by subjects. Because the random slope that allowed change from immediate to delayed post-test to vary by subjects explained nearly zero variance, that random effect was removed and only the random effect that allowed change from pre-test to post-tests to vary by subjects was retained.

The odds of getting the answer correct were 5.016 times higher in the K condition than in the D & T conditions, z = 6.550, $p < .001$.

Participants significantly improved from pre-test to post-tests, z = 9.787, $p < .001$. The odds of getting the answer correct were 1.912 times higher on the post-tests than on the pre-test.

Participants performed significantly better on the second and third phoneme contrasts they learned than on the first phoneme contrast they learned, z = 2.858, $p = .004$ and z = 4.854, $p < .001$, respectively.

There was a marginal interaction between orthography (English/ no- & Marathi contrast) and test (pre/post-test), z = 1.845, $p = .065$. This interaction was driven by the fact that, although there was overall improvement from pre-test to post-tests, there was more improvement in the no- & Marathi orthography conditions.

Forced-choice identification words with individual differences. The model used to analyze these data was very similar to the models used for forced-choice identification words. Participants were restricted to those for whom individual difference measures were available. The individual difference measures were transformed in the same manner as they were for pronunciation with individual differences. Interactions between test and both rise time discrimination and PAD were included.

Importantly, the same pattern of effects from the model with all participants was present in the model with individual differences. Participants performed better on the K contrast than on the D & T contrasts, z = 5.668, $p < .001$. Participants improved from pre to post-tests, z = 6.720, $p < .001$. Participants performed better on the second and third contrasts they learned than on the first contrast they learned, z = 2.344, $p = .019$ and z = 3.629, $p < .001$, respectively. There was a marginal interaction between orthography (English/no- & Marathi contrast) and test (pre/post-tests), z = 1.891, $p = .059$.

There was one effect seen in the present model that were not seen in the model with all participants. There was an interaction between test (delayed/immediate contrast) and orthography (no-/Marathi contrast), z = 1.808, $p = .071$. This interaction reflected the fact that participants in the no- orthography condition improved from immediate to delayed post-test but participants in the Marathi orthography condition did not.

Importantly, individual differences predicted performance on the task. There was a marginal main effect of rise time discrimination, z = -1.855, $p = .064$. A one standard deviation increase in rise time discrimination ability improved the odds of getting a correct answer by 1.217 times.

4.3.5 Choose correct word

For this task, I was interested in seeing how the different orthography conditions affected participants' abilities to learn the correct pronunciations for words. To perform well on the test, you must be able to hear the difference between the phonemes and remember the correct pronunciation. To focus the analysis on memory for the correct pronunciation, I restricted my analysis to participants who could successfully discriminate the phonemes. Specifically, I restricted the analysis to cases in which participants scored above a 70% on the immediate post-test for both the forced-choice identification phonemes and forced-choice identification words tests. If participants met the criterion for one phoneme but not another, I included the phoneme for which they met the criterion. This left me with D data from 25 participants, T data from 33 participants, and K data from 64 participants. I did not analyze if individual differences predicted performance on this task because, by restricting my sample to participants who could discriminate the phonemes, I also restricted the data to participants who generally had good phonological skills and rise time discrimination. Furthermore, I did not have sufficient data from participants who met the accuracy criterion and completed the individual difference measures.

The data were analyzed using a binomial linear mixed effects model with accuracy as the dependent measure. Although the data were restricted to participants who got over 70% correct, variability in the ability to distinguish phonemes may still affect performance. Therefore, performance on the immediate post-test of the forced-choice identification words task was included as a predictor. Because performance on the forced-choice identification words task absorbed the variance explained by phoneme, phoneme was removed as a predictor. Thus, the predictors were order (1st, 2nd, 3rd contrast learned), and the interaction between centered performance on the forced-choice identification words immediate post-test, orthography (no-, English, Marathi), and test (immediate post-test, delayed post-test). For order, the first contrast learned was used as the baseline. For orthography, orthogonal contrasts comparing no-orthography to English & Marathi and comparing English and Marathi to each other were used. For test, contrast coding was used to compare the immediate and delayed post-tests. The model also included random intercepts for subjects and items and the effect of test to vary by subjects[20].

Performance on the forced-choice identification words task was positively correlated with performance on the choose correct word task, z = 7.014, $p < .001$. There was also a marginal interaction between performance on the forced-choice identification words task and test, z = 1.782, $p = .075$. This interaction reflects the fact that performance on the forced-choice identification words task was more predictive of performance on the immediate test than on the delayed test. The stronger relationship on the immediate test could stem from the fact that forced-choice identification word task performance on the immediate test was used as a predictor.

Participants performed better in the two orthography conditions than in the no- orthography condition, z = 3.812, $p < .001$. The odds of a correct answer was 1.350 times lower in the no-orthography condition than in the two orthography conditions. Participants performed better on the immediate post-test than on the delayed post-test, z = 4.208, $p < .001$. The odds of a correct answer was 1.325 times higher on the immediate post-test than on the delayed post-test. Both of these main effects were qualified by an interaction between the orthography (no-/English & Marathi contrast) and test, z = 2.000, $p = .045$. This interaction was driven by the fact that

[20] Although the model did not converge, the relative gradient was less than 0.001.

participants in the no- orthography condition performed equally well on the immediate and delayed post-tests but participants in the orthography conditions decreased in performance from immediate to delayed post-test. However, the lack of a decrease in the no- orthography condition likely reflected the fact that participants in the no- orthography condition performed much worse on the immediate post-test. In fact, participants in both orthography conditions outperformed participants in the no- orthography condition at delayed post-test, but the margin was smaller than at immediate post-test.

4.4 DISCUSSION

4.4.1 Manipulated utterances

Many previous studies that have attempted to teach participants dental/retroflex contrasts failed to do so (Polka, 1991; Tees & Werker, 1984; Werker et al., 1981; Werker & Tees, 1984). In fact, Tees and Werker (1984) found that people with one year of Hindi instruction could not distinguish the phonemes! A similar result was found in the pilot study, in which participants attempted to learn non-manipulated utterances. In contrast, when I used manipulated utterances, the participants were able to learn the contrasts and maintained their performance on the delayed post-test. Thus, this result suggests that using manipulated utterances that exaggerate small durational differences may be a useful tool for teaching non-native speakers to distinguish the contrasts.

One large difference between dental and retroflex consonants is that retroflexes have a lowered third formant during the transition to the vowel (Guion & Pederson, 2007; Hamann, 2003). This lowered third formant is a consequence of three properties of retroflexes: posteriority, sublingual cavity, and retraction. Posteriority refers to the fact that the tongue touches an area in the back of the roof of the mouth, behind the alveolar region. This posterior tongue placement open up space beneath the tongue known as the sublingual cavity. Retraction is defined as backward displacement of the tongue towards the velum or pharynx. Posteriority and sublingual cavity lengthen the front cavity of the vocal tract, thus lowering the resonance frequencies, including F3. Retraction results in velarization, which lowers the F3 (Hamann, 2003). Thus, the backward curling of the tongue necessary to articulate a retroflex results in a lowered F3, one of the main acoustic cues used to identify retroflexes. English does not have retroflex stop consonants (although English occasionally uses a retroflex rhotic (Hamann, 2003)), so English speakers may find it difficult to use the F3 cue to distinguish dentals and retroflexes.

Another difference between dental and retroflex consonants is that dental consonants have a longer voice onset time (Hamann, 2003; Polka, 1991; Verma & Chawla, 2003). Differences in voice onset time is a 'stable' cue that is easy for second language learners to use. In contrast, differences in formant structure is a more 'fragile' cue that is difficult for adult second language learners to use (Burnham, 1986). Thus, I increased the voice onset time of the dentals to make it more salient to the participants. This manipulation helped the participants learn to distinguish the phonemes, both when they were isolated and when they were in word contexts. An important next step is to see if it is beneficial to begin training people on very manipulated utterances and gradually fade amount of manipulation. Previous research has shown that using an adaptive training procedure with manipulated contrasts is, under certain conditions, more beneficial than using only fixed utterances when training Japanese speakers on the English r/l distinction (McCandliss et al., 2002). It would be interesting to see if the same pattern holds for the Marathi dental/retroflex contrasts.

Because I emphasized the durational differences and not the formant differences, it is unclear if the manipulated utterances are helping participants learn the formant differences. If the manipulated utterances are not helping people learn the formant differences, training with manipulated utterances may not transfer to natural stimuli. A Hebbian account (Hebb, 1949) would predict that the manipulated utterances can help people learn the formant differences. Specifically, the Hebbian account predicts that "when a pattern of neural activity at a peripheral level elicits a cortical pattern corresponding to a percept, the strength of the connections from the peripheral neurons to the neurons constituting the percept, and among the neurons constituting the percept, is increased" (McCandliss et al., 2002, p. 91). Thus, the retroflex tokens should activate peripheral neurons corresponding to the shorter VOT and the lowered third formant during the transition to the vowel. In contrast, the dental tokens should activate peripheral neurons corresponding to the longer VOT and the non-lowered third formant. Both of these peripheral representations should activate different cortical representations. Thus, the connection between the "lowered third formant" peripheral neurons and the "retroflex" cortical neurons should be strengthened, as should the connection between the "non-lowered third formant" peripheral neurons and the "dental" cortical neurons. This change in connection strength should help people learn to distinguish dentals and retroflexes using fine structure information.

Although the Hebbian account predicts that manipulating durational information should help people learn to use fine structure information when distinguishing dentals and retroflexes, this experiment did not test that prediction. Future studies could train participants on durationally-manipulated contrasts. Then, they could digitally create phonemes that vary in terms of fine structure but not in terms of duration. They could test whether training on the manipulated contrasts transfers to the phonemes that vary only in terms of fine structure.

This study replicates previous research that demonstrates that the T contrast is easier than the D contrast (Polka, 1991). This replication is important because it demonstrates that the manipulated utterances elicited similar results as do natural utterances.

This study also replicates prior findings suggesting that the K contrast is easier than place of articulation contrasts (Guion & Pederson, 2007). However, unlike the previous studies, there were no ceiling effects at pre-test although ceiling effects were present at post-test (Aggarwal, 2012; Guion & Pederson, 2007). However, it is interesting to note that some participants continued to do poorly on this contrast even after training, so it was not universally easy. One important difference between this study and previous studies is that I trained participants on articulation so they knew that /k/ was different than /g/. I also had participants produce the phonemes and found that participants said /g/ only 5.8% of the time. Thus, it is likely that my participants would be able to discriminate /k/, k^h/, and /g/. In contrast, because previous studies did not train participants on articulation nor have them produce the sounds, it is possible that participants were perceiving the /k/-/k^h/ contrast as /g/-/k/. Thus they would not be able to discriminate the three Marathi sounds /k/, k^h/, and /g/. In fact, Guion and Pederson (2007) found that the /g/-/k/ contrast was harder for participants than the /k/-/k^h/ contrast, suggesting that it is common for English speakers to confuse /k/ and /g/ in the absence of explicit instruction.

4.4.2 Orthography

For the forced-choice identification tasks there were small effects of orthography. Participants seemed to learn more in both the no- and Marathi orthography conditions as compared to the English orthography condition. This result could be because people already associated the English graphs with their English pronunciations and thus it was difficult to learn a new mapping. More generally, this finding is an example of how proactive interference can impede learning. However, the effect sizes were small and all of the effects were marginal. Thus, more research is needed to confirm if this is a true effect.

It is possible that the manipulation was too subtle to see significant effects. For example, on the forced-choice identification learning tasks in the English condition, participants received feedback that said *Correct! The first sound was t and the second sound was T*. However, participants did not need to pay attention to that second sentence; simply the correct/incorrect provided them with all the information they needed. The manipulation would have been more robust if the participants performed an AX task and the feedback read *Correct! Both sounds are t* or *Correct! The first sound was t and the second sound was T*. In this case, the sentence would provide necessary information and participants would need to pay attention to it.

Another option would be to compare two Marathi classes, one that commonly uses English transliterations and one that solely uses Marathi graphs. In a classroom setting, students may have more motivation to learn Marathi. Furthermore, the longer exposure that a class affords may lead to more robust differences. Additionally, in the study, participants know that the study is focused on pronunciation and are explicitly taught how to pronounce the two phonemes. In a classroom situation, where the focus in on communication and not on pronunciation, larger differences may be more apparent.

The orthography manipulation more strongly predicted performance on the choose correct word task. Participants performed better in both orthography conditions than in the no-orthography condition. When participants were learning the words, participants in the no-orthography condition simply heard the correct pronunciation. If they had trouble discriminating the phonemes, they would not learn anything. In contrast, the orthography conditions provided information about which phoneme the word started with, so participants could learn even if they could not discriminate the phonemes. That is one reason participants may have performed better in the orthography conditions. However, I think that explanation is unlikely. I restricted my analysis to participants who could distinguish the phonemes well and included discrimination ability as a predictor in the regression. Therefore, I believe that the multisensory representation of the orthography combined with phonology helped participants remember which phoneme the word started with. This result is in line with the Lexical Quality Hypothesis (Perfetti & Hart, 2002), which posits that there are three components of lexical representation: orthography, phonology, and semantics. Participants in the English and Marathi orthography conditions had access to all three representations, whereas participants in the no- orthography condition did not have access to orthography. Thus, participants in the no- orthography condition had more difficulty forming lexical representations.

The results from the choose correct word task corroborate other studies demonstrating that the presence of orthographic representations during word learning improves learning of phonological representations (Phillips, 2011; Ricketts, Bishop, & Nation, 2009). However, this study was the first to test the quality of the phonological representation. Phillips (2011) demonstrated that participants were better able to learn Arabic vocabulary words if they had both English transliterations and an oral representation, as opposed to just an oral representation. In her post-test, participants had to orally produce the Arabic word that went with an English translation. All reasonably close pronunciations (in which the participant's intention was clear) were coded as correct. Ricketts et al. (2009) demonstrated that children better learned to associate non-words with pictures if they were provided with both orthographic and phonological representations of the non-words, as opposed to just phonological representations. In their post-test, students heard a phonological representation and had to choose which picture matched that representation from an array of four pictures. Thus, both of these studies demonstrated that orthographic support strengthens the connection between phonology and semantics. However, neither of the studies measured quality of the phonological representation. In contrast, the present study demonstrated that orthographic support improves the quality of the phonological representation.

Broadly speaking, the results are congruent with literature on multisensory learning that suggests that multisensory training is more effective than unisensory training, but only when the information coming from the different senses is congruent (Shams & Seitz, 2008). The presence of written representations benefitted remembering the pronunciations of the words because it provided multisensory input. The Marathi orthography supported discrimination learning better than English orthography because the pronunciations were inconsistent with English grapheme-phoneme correspondences the participants already knew. The results of this study suggest that foreign language teachers should use Marathi graphs during instruction because they help people discriminate the contrasts and remember which phonemes are in vocabulary words. However, because the evidence suggesting that English graphs impair contrast learning was weak, replications are needed before strong recommendations can be made.

4.4.3 Individual differences

Phonological awareness/decoding and rise time discrimination ability both positively predicted L2 phonemic learning. Phonological awareness and decoding were combined into one measure because they were positively correlated in this sample. Interestingly, rise time discrimination was not correlated with either measure in this sample. Research with English-speaking dyslexic and typically developing children has shown that good rise time discrimination predicts phonological awareness and decoding (Goswami, 2011; Goswami et al., 2002). Rise time discrimination deficits have also been found in compensated adult dyslexics (Corriveau, Pasquini, & Goswami, 2007; Goswami, 2011). The lack of a significant correlation in this sample could reflect the fact that this sample was composed of only typically developing adults.

4.4.3.1 Rise time discrimination

Rise time discrimination ability was positively correlated with ability to learn the phonemic contrasts. Participants with better rise time discrimination performed better on both the pronunciation and forced-choice identification words tasks overall and improved marginally more on the forced-choice identification phonemes tasks.

In this task, participants learned to discriminate stop consonants which do not greatly vary in terms of rise time. However, the contrasts do vary in voice onset time, which is one component of envelope structure. Thus, the rise time discrimination task was likely measuring sensitivity to envelope structure more generally. People who were more sensitive to differences in envelope structure were better able to learn the contrasts.

It is important to note that I manipulated the dental sounds by increasing their voice onset time and thus accentuating differences in envelope structure. Future studies should see if rise time discrimination predicts ability to discriminate natural utterances of the dental/retroflex contrast. The K contrast used natural utterances so rise time discrimination should predict ability to distinguish that contrast.

It is also important to note that I did not measure any other auditory perceptual abilities such as duration and frequency discrimination. Future studies should determine if rise time discrimination can predict the learning of non-native contrasts while controlling for other auditory perceptual abilities.

My results support and expand on the results of Slevc and Miyake (2006), who found that individual differences in L2 phonological perception and production could be predicted by differences in tonal perception and production. Both their study and the current study found that differences in the perception of non-speech sounds can predict the L2 perception and production of consonantal minimal pairs. The present study expanded on their results by using rise time discrimination as a predictor, which is very different than tonal abilities. Furthermore, the present study focused on English speakers learning Marathi, whereas their study focused on Japanese speakers learning English.

My results contrast with those of Díaz et al. (2008) who found that individual differences in auditory perception do not predict differences in L2 phonemic perception. However, they studied duration and tone, not rise time discrimination. Thus, the discrepant results could be due to the different general auditory measures used. However, their study examined vowel perception and the present study examined consonant perception. Thus, the discrepant results could also be due to differences in the specific contrasts studied. Future studies including a wide variety of auditory measures and L2 phonemic contrasts are needed.

4.4.3.2 Phonological awareness and decoding

I had two contrasting hypotheses on the relationship between phonological skills and the learning of non-native contrasts. On one hand, L1 and L2 phonological skills are highly correlated (Melby-Lervåg & Lervåg, 2011). Thus, one might predict a positive relationship between L1 phonological awareness/decoding and learning of non-native contrasts. On the other hand, allophonic perception is often associated with poor phonological skills (Bogliotti et al., 2008; Serniclaes et al., 2004, 2001). However, the ability to hear differences within a phonemic category may assist in learning non-native contrasts. Thus, one might predict a negative relationship between L1 phonological awareness/decoding and learning of non-native contrasts. It is important to note that I did not directly test phonemic versus allophonic perception. Furthermore, allophonic perception has been primarily noted in the dyslexia literature (Bogliotti et al., 2008; Serniclaes et al., 2004, 2001). It is unclear if relatively poor reading is associated with allophonic perception within the range of normal reading ability. In the future, allophonic versus phonemic perception should be directly measured and participants with dyslexia should be included.

My results support the first hypothesis; I found a positive correlation between English phonological awareness/decoding and ability to learn Marathi contrasts. Phonological skills significantly predicted both overall accuracy on the pronunciation task and improvement on the forced-choice identification phonemes task.

My results support Gabay and Holt's (2015) finding that auditory category learning is positively predicted by both L1 phonological awareness and decoding. Taken together, the results from the present study and that of Gabay and Holt (2015) suggest that L1 phonological awareness and decoding can predict the learning of both speech and non-speech categories. Furthermore, the relationship between phonological skills and learning of novel auditory categories can be found in both dyslexic and typically-developing populations. Finally, the relationship between phonological skills and learning of novel auditory categories can be found using both implicit and explicit learning paradigms. Gabay and Holt (2015) hypothesized that the relationship between L1 phonological skills and the learning of novel auditory categories reflected the fact that dyslexics have impaired procedural learning (and more specifically, impaired category learning). Because my participants were all typically-developing readers, it is unclear if a deficit in procedural learning is driving the relationship between phonological skills and novel auditory category learning.

4.5 CONCLUSION

Overall, the results of this study suggest that people can learn Marathi dental/retroflex contrasts when using manipulated utterances. More research is needed to see if training on manipulated utterances transfers to natural utterances. Furthermore, the study suggests that using Marathi graphs may be the most beneficial to foreign language learners, but more studies are needed to confirm that finding. Finally, the study suggests that both rise time discrimination and English phonological skills positively predict L2 phonemic perception. More research is needed to see if these individual difference measures are predictive of learning natural utterances, or only of learning manipulated utterances. Furthermore, the inclusion of more individual difference measures including allophonic perception, dyslexia status, category learning, and tone and length discrimination could help elucidate mechanism.

5.0 GENERAL DISCUSSION

This dissertation sought to compare instructional methods for teaching Indic languages to identify best practices. The pedagogical methods incorporated different instructional principles known to be effective in other contexts to test if these instructional principles apply to the teaching of Indic languages.

5.1 SUMMARY OF EXPERIMENT 1

Experiment 1 demonstrated that the mobile game improves akshara recognition and the spelling and reading of words that contain complex akshara. It compared two versions of the mobile game, narrow and wide spacing. The two versions of the game tested the instructional principle of spacing, and based on previous research I expected the wide spacing version to be more effective (Cepeda et al., 2006; Underwood, 1961). However, both versions resulted in comparable post-test outcomes, although the narrow spacing version was typically played faster. I was surprised that I did not see the benefit of spacing in my data. There are two main reasons why the expected spacing effect may not have been present. First, the spacing manipulation was implemented differently than previous research has. In most studies, the same or very similar problems are presented in either a spaced or distributed manner. In contrast, in the present study, the narrow spacing version consecutively presented the same akshara in two very different contexts (isolated and in a word context) and the wide spacing version spaced out those two different contexts. Second, even the narrow spacing version incorporated some distributed practice; participants had to repeat levels if they did not finish in time or did not score enough points. Furthermore, some akshara repeated in different levels. Thus, the spacing may not have been implemented in a sufficiently robust manner to see an effect.

Although there was no effect of the spacing manipulation on post-test outcomes, there was an effect of the spacing manipulation on game play. Participants in the wide spacing version generally played slower, which seems to be driven by the fact that they required more clicks to get word problems correct. This difficulty with the word problems is logical; participants who played the narrow spacing version knew that the complex akshara they had just seen would be in the word, whereas participants in the wide spacing version did not have that advantage. From the perspective of desirable difficulties (McDaniel & Butler, 2011), this should make the wide spacing version more effective. However, the narrow spacing version's consecutive ordering may have helped students learn isolated akshara and helped them understand how the akshara functioned in word contexts. The wide spacing version may have obscured this relationship, making the additional difficulty undesirable. Thus, the data suggest that the narrow spacing version is both efficient and effective.

5.2 SUMMARY OF EXPERIMENT 2

Although Experiment 1 showed that the game was beneficial, the reading and spelling gains were relatively modest. Experiment 2 compared different pedagogical methods for teaching people complex akshara. The different pedagogical methods tested two specific instructional principles, motor encoding and testing. It found that motor encoding was a particularly beneficial method for teaching people complex akshara. Specifically, motor encoding was beneficial for tasks that required pure orthographic knowledge or required people to produce an orthographic form when given a phonological form. Moreover, the data suggest that testing is not a particularly effective instructional practice in this context, which is surprising because the benefits of testing have been demonstrated in many contexts, including orthographic learning (Rieth et al., 1974; Roediger & Karpicke, 2006). Post-hoc analyses suggest that testing was not beneficial in this instance because testing strengthened incorrect responses. Participants who were highly inaccurate during the learning phase found copying slightly more beneficial than writing, whereas participants who had higher accuracy during the learning phase found writing slightly more beneficial than copying. Thus, these data suggest that testing may be more beneficial for advanced learners whose accuracy is higher.

Overall, the results suggest that, if copying were incorporated in the game, the copying would benefit spelling skills, but not reading skills. However, further experimentation with the demographic for whom the game is designed is needed.

5.3 SUMMARY OF EXPERIMENT 3

Experiment 1 demonstrated that students struggled with discriminating phonologically similar consonants. The analysis of the word-level game data showed that students often confused aspirated/unaspirated consonantal pairs. Post-test data showed that students often confused both aspirated/unaspirated and dental/retroflex consonantal pairs. Experiment 3 sought to better understand individual differences in ability to discriminate these sounds and to identify pedagogical methods for improving discrimination. Experiment 3 showed that manipulating utterances to exaggerate differences can help people learn the difficult dental/retroflex contrast.

Furthermore, it demonstrated that both rise time discrimination and phonological skills positively predict phonemic discrimination. Rise time discrimination is an auditory perceptual ability that can predict language and reading development in a first language (Goswami, 2011). Experiment 3's findings suggest that auditory perceptual abilities can predict non-native phonemic perception and that rise time discrimination is an important predictor in both L1 and L2 language development. The fact that English phonological skills positively predict non-native phonemic perception suggests that innate category learning abilities may underlie the learning of both L1 and L2 phonology. My research expands on prior research that demonstrated that L1 phonological skills predict the learning of non-speech auditory categories (Gabay & Holt, 2015) by demonstrating that L1 phonological skills also predict the learning of non-native speech categories.

The results of Experiment 3 suggest that using the Marathi orthography is most beneficial for both discriminating phonemes and remembering which phonemes are in vocabulary words. More specifically, the results suggest that English orthography impaired people's abilities to discriminate phonemes, although the results were relatively weak and need to be replicated. The English orthography likely led to interference because participants already associated the graphs with their English pronunciations and had difficulty mapping them onto a different pronunciation. Similar interference from first language grapheme-phoneme mappings has been reported in other studies as well (Meng, 1998; Young-Scholten, 2002).

Participants were able to better remember which phoneme was in a given vocabulary word when they had orthographic support. These results did not reflect differences in phonemic perception, because that was controlled for in the model. Thus, the orthographic support was benefitting memory itself. The multisensory representation, in which the correct pronunciation is represented both auditorily and visually, could have driven these effects. Thus, my finding supports prior research demonstrating that multisensory representations benefit memory (Shams & Seitz, 2008) and that orthographic support benefits word learning (Phillips, 2011; Ricketts et al., 2009). Taken together, the discrimination and memory results suggest that the Marathi orthography is the most beneficial condition.

This knowledge can be applied to the game by having the game begin with a phonological pre-training module in which children hear a manipulated utterance and have to choose the correct Hindi akshara. The amount of manipulation can gradually fade to the point at which children can successfully discriminate natural utterances. Students can then use their phonological training throughout the game.

5.4 THEORETICAL DISCUSSION

5.4.1 Multisensory encoding

Overall, my research suggests that multisensory encoding is very beneficial for learning. For example, motor encoding, which combines motor and visual encoding, benefits orthographic learning. Similarly, orthographic support, which combines auditory and visual encoding, helps people remember which phonemes are in words.

5.4.1.1 Motor encoding

My research suggests that motor encoding is particularly important for building orthographic knowledge, likely because it adds a motor trace to the visual representation. Although I demonstrated the benefit of motor encoding for learning Devanagari, similar results have been shown for learning English, Chinese, and Arabic (Cunningham & Stanovich, 1990; Guan et al., 2011; Longcamp et al., 2005; Naka, 1998; Ouellette, 2010; Xu et al., 2013). In Devanagari, which has many visually complex akshara, motor encoding should benefit graph learning in older age groups who have had several years of Hindi instruction. In English, which has relatively few, visually simple graphs, motor encoding benefits graph learning at younger ages (Longcamp et al., 2005). However, motor encoding continues to benefit whole word orthographic representations in older age groups (Cunningham & Stanovich, 1990; Ouellette, 2010).

5.4.1.2 Orthographic support

The results from the choose correct word task in Experiment 3 suggest that orthographic support helps people remember which phonemes are in words, thus improving the quality of the phonological representation. These results suggest that the presence of both visual and auditory traces benefit learning. Furthermore, the results are in line with predictions from the Lexical Quality Hypothesis (Perfetti & Hart, 2002). Specifically, the Lexical Quality Hypothesis predicts that people will have the highest quality lexical representations when they have access to orthographic, phonological, and semantic representations. The two orthography conditions provided all three of these representations, but the no- orthography condition only provided phonological and semantic representations. Without orthographic support, phonological representations suffered.

5.4.2 Reducing interference

5.4.2.1 Reducing interference from prior knowledge

My findings suggest that using L1 transliterations impairs learning of phonemic contrasts when the pronunciation of a certain graph is slightly different in the L1 than in the L2. For example, using the graph ‘k’ may be confusing for English-speakers who are learning Marathi because in English that letter is associated with a slightly aspirated velar stop whereas in Marathi that letter is associated with an unaspirated velar stop. Thus, the use of L1 transliterations may cause proactive interference from prior knowledge about a graph’s pronunciation. The need to suppress the L1 graph-phonology mappings may tax the working the memory system (V. M. Rosen & Engle, 1998). Due to the high working memory load, participants have fewer resources available for attending to the acoustic differences and learning from the feedback.

5.4.2.2 Reducing interference from incorrect responses

When considering the results from Experiments 1 and 2, one can conclude that increasing the difficulty of the learning situation and increasing time on task does not necessarily benefit learning. In Experiment 1, students played the wide spacing version slower because it was more difficult. In Experiment 2, participants required more time to learn via writing than via copying because it was more difficult. However, in Experiment 1, both versions of the game resulted in equivalent post-test outcomes. In Experiment 2, both copying and writing resulted in nearly identical learning outcomes; writing only outperformed copying on one task, reading. In both cases, the slower and more difficult version may have strengthened incorrect responses. In the game, participants required more clicks to correctly spell the word, meaning that they were more likely to choose foils or to incorrectly order the akshara. In Experiment 2, participants often wrote an incorrect complex akshara during the writing learning phase, which they then had to correct. Although the correct answer was eventually provided in both cases, this additional difficulty in arriving at the correct answer was not beneficial. This finding was surprising because prior research has shown that, although incorrect answers during an initial test can impair learning (Butler et al., 2006; Roediger & Karpicke, 2006), this typically is not problematic when immediate feedback is given (Butler & Roediger, 2006) and the to-be-learned information is orthographic in nature (Ehri et al., 1988). Nevertheless, Experiments 1 and 2 suggest that interference from prior incorrect answers can be problematic during akshara learning.

5.5 EDUCATIONAL IMPLICATIONS

5.5.1 Multisensory encoding

5.5.1.1 Motor encoding

Currently, technology such as laptops and iPads is becoming more prevalent in classrooms. Technology has many educational benefits, such as allowing for individualized learning. However, most of these technologies use typing or selecting responses on a touch screen. Thus, they do not incorporate motor encoding, which is critically important for building orthographic knowledge. Thus, educators should include traditional pencil-and-paper methods or adopt touch screen technology which allows students to write on the screen when teaching graphs and word spellings. Motor encoding may be especially important for children with poor typing skills (Ouellette & Tims, 2014).

5.5.1.2 Orthographic support

My results suggest that foreign language teachers should use orthography in the classroom because multisensory input helps people remember which phonemes are in words. Although this approach is very different from how people learn their first language (purely auditorily), it appears to benefit second language learning. However, teachers must be cognizant of which orthography they use and how it may cause inference. This point is elaborated on in the next section.

5.5.2 Reducing interference

5.5.2.1 Reducing interference from prior knowledge

Foreign language teachers need to be very cognizant of reducing interference from the first language. This is widely acknowledged in other aspects of foreign language learning, for example, teachers know to explicitly highlight false cognates when teaching vocabulary (e.g., when teaching Spanish to English speakers, it is important to explain that “éxito” means “success”, not “exit”) (Frantzen, 1998). The present study applies the basic concept of reducing L1 interference to orthographic learning. When teaching a language that uses a different script than the students’ first language, educators should use the second language script rather than transliterations because transliterations can cause interference from the first language. Although students may initially need more time to learn the new script, they will reap the benefits later when they begin learning second language phonology.

However, some language teachers may not have this option. For example, English speakers learning Spanish must use the Roman alphabet for both languages. Even when students learn Chinese, they may need to also learn pinyin, the phonetic coding system that uses the Roman alphabet and is widely used in China. In these cases, educators need to be aware that students may experience interference from their first language. They should clearly explain when certain graphs are pronounced differently in the students' first and second languages because explicitly addressing when prior knowledge may cause interference is beneficial (Dochy, Segers, & Buehl, 1999).

5.5.2.2 Reducing interference from incorrect responses

The results from Experiments 1 and 2 suggest that students experience negative interference from incorrect answers, even when feedback is provided immediately. Thus, educators should design assignments that are within students' zone of proximal development so students can get the majority of the answers correct. Then, as students gain proficiency, educators can gradually make the assignments more difficult. This design reduces the chance that students will get many problems incorrect and strengthen the incorrect response.

5.6 CONCLUSION

The goal of this dissertation was to test which instructional principles apply to teaching people the orthographic and phonological systems of Indic languages. Broadly, one can conclude that multisensory encoding is critical. For example, copying and writing, which incorporate both motor and visual encoding, benefit orthographic learning. Similarly, orthographic support, which incorporates both visual and auditory encoding, benefits phonological learning. Furthermore, it is very important to reduce interference from both inapplicable prior knowledge and prior incorrect responses. For example, it is inadvisable to use L1 transliterations, which may induce interference from prior knowledge. Similarly, it is inadvisable to create a particularly difficult learning environment during early learning, for instance by incorporating spacing or testing. Students tend to make more errors in difficult learning environments, which may strengthen the incorrect response and create interference. Students will be able to better acquire a second language if their instructors incorporate multisensory instruction and make every effort to reduce interference.

BIBLIOGRAPHY

Aggarwal, A. (2012). Acquisition of Hindi contrasts by English speakers: An optimality theoretic account. *Acta Linguistica Asiatica*, *1*(3), 9–24. http://doi.org/10.4312/ala.1.2.9-24

Altman, D. (1991). *Practical statistics for medical research*. Chapman and Hall.

Bassetti, B. (2006). Orthographic input and phonological representations in learners of Chinese as a foreign language. *Written Language and Literacy*, *9*(1), 95–114.

Bassetti, B. (2007). Effects of hanyu pinyin on pronunciation in learners of Chinese as a Foreign Language. In A. Guder, X. Jian, & Y. Wan (Eds.), *The cognition, learning and teaching of Chinese characters*. Beijing, China: Beijing Language and Culture University Press. Retrieved from http://eprints.whiterose.ac.uk/69151/

Bhatia, T. K. (2008). *Colloquial Hindi: The complete course for beginners* (2nd ed.). New York, NY: Routledge.

Bhide, A., Gadgil, S., Zelinsky, C. M., & Perfetti, C. A. (2013). Does reading in an alphasyllabary affect phonemic awareness? Inherent schwa effects in Marathi-English bilinguals. *Writing Systems Research*, *6*(1), 73–93. http://doi.org/10.1080/17586801.2013.855619

Bhide, A., & Perfetti, C. (2017). Marathi. In R. M. Joshi & C. McBride (Eds.), *Handbook of literacy in akshara orthographies*. Springer. Chapter submitted for publication.

Bogliotti, C., Serniclaes, W., Messaoud-Galusi, S., & Sprenger-Charolles, L. (2008). Discrimination of speech sounds by children with dyslexia: Comparisons with chronological age and reading level controls. *Journal of Experimental Child Psychology*, *101*(2), 137–155. http://doi.org/10.1016/j.jecp.2008.03.006

Bolozky, S. (1997). Israeli Hebrew phonology. In A. S. Kaye & P. T. Daniels (Eds.), *Phonologies of Asia and Africa* (pp. 287–311). Winoa Lake: Eisenbrauns.

Burnham, D. K. (1986). Developmental loss of speech perception: Exposure to and experience with a first language. *Applied Psycholinguistics*, *7*(3), 207–239.

Butler, A. C., Marsh, E. J., Goode, M. K., & Roediger, H. L. (2006). When additional multiple-choice lures aid versus hinder later memory. *Applied Cognitive Psychology*, *20*(7), 941–956. http://doi.org/10.1002/acp.1239

Butler, A. C., & Roediger, H. L. (2006). *Feedback neutralizes the detrimental effects of multiple-choice testing.* Poster presented at the 18th Annual Convention for the Association for Psychological Science, New York, NY.

Cao, F., Vu, M., Chan, D. H. L., Lawrence, J. M., Harris, L. N., Guan, Q., … Perfetti, C. A. (2013).

Writing affects the brain network of reading in Chinese: A functional magnetic resonance imaging study. *Human Brain Mapping, 34*(7), 1670–1684. http://doi.org/10.1002/hbm.22017

Castles, A., Holmes, V. M., Neath, J., & Kinoshita, S. (2003). How does orthographic knowledge influence performance on phonological awareness tasks? *The Quarterly Journal of Experimental Psychology*, *56*(3), 445–467. http://doi.org/10.1080/02724980244000486

Cepeda, N. J., Pashler, H., Vul, E., Wixted, J. T., & Rohrer, D. (2006). Distributed practice in verbal recall tasks: A review and quantitative synthesis. *Psychological Bulletin*, *132*(3), 354–380. http://doi.org/10.1037/0033-2909.132.3.354

Chang, L.-Y. (2015). *Visual orthographic variation and learning to read across writing systems* (Doctoral dissertation). University of Pittsburgh, PA, USA.

Chang, L.-Y., Xu, Y., Perfetti, C. A., Zhang, J., & Chen, H.-C. (2014). Supporting orthographic learning at the beginning stage of learning to read Chinese as a second language. *International Journal of Disability, Development and Education*, *61*(3), 288–305. http://doi.org/10.1080/1034912X.2014.934016

Chung, I. (2007). *Ecology of PF: A study of Korean phonology and morphology in a derivational approach* (Doctoral dissertation). University of Connecticut, CT, USA.

Cibelli, E. S. (2015). *Aspects of articulatory and perceptual learning in novel phoneme acquisition* (Doctoral dissertation). University of California Berkeley, CA, USA.

Cook, D. L. (1962). The Hawthorne effect in educational research. *The Phi Delta Kappan*, *44*(3), 116–122.

Corriveau, K., Pasquini, E., & Goswami, U. (2007). Basic auditory processing skills and Specific Language Impairment: A new look at an old hypothesis. *Journal of Speech, Language, and Hearing Research*, *50*, 647–666.

Cunningham, A. E., & Stanovich, K. E. (1990). Early spelling acquisition: Writing beats the computer. *Journal of Educational Psychology*, *82*(1), 159–162. http://doi.org/10.1037/0022-0663.82.1.159

Defior, S., & Serrano, F. (in press). Learning to read Spanish. In L. Verhoeven & C. A. Perfetti (Eds.), *Reading acquisition across languages and writing systems* (pp. 270–298). Cambridge, UK: Cambridge University Press.

Díaz, B., Baus, C., Escera, C., Costa, A., & Sebastián-Gallés, N. (2008). Brain potentials to native phoneme discrimination reveal the origin of individual differences in learning the sounds of a second language. *Proceedings of the National Academy of Sciences of the United States of America*, *105*(42), 16083–16088. http://doi.org/10.1073/pnas.0805022105

Dochy, F., Segers, M., & Buehl, M. M. (1999). The relation between assessment practices and outcomes of studies: The case of research on prior knowledge. *Review of Educational Research*, *69*(2), 145–186.

Ehri, L. C., Gibbs, A. L., & Underwood, T. L. (1988). Influence of errors on learning the spellings of English words. *Contemporary Educational Psychology*, *13*(3), 236–253. http://doi.org/10.1016/0361-476X(88)90024-0

Ehri, L. C., & Wilce, L. S. (1980). The influence of orthography on readers' conceptualization of the phonemic structure of words. *Applied Psycholinguistics*, *1*(4), 371–385. Retrieved from http://journals.cambridge.org/abstract_S0142716400009802

Fishman, E. J., Keller, L., & Atkinson, R. C. (1968). Massed versus distributed practice in computerized spelling drills. *Journal of Education & Psychology*, *59*(4), 290–296. http://doi.org/10.1037/h0020055

Frantzen, D. (1998). Intrinsic and extrinsic factors that contribute to the difficulty of learning false cognates. *Foreign Language Annals*, *31*(2), 243–254.

Gabay, Y., & Holt, L. L. (2015). Incidental learning of sound categories is impaired in developmental dyslexia. *Cortex*, *73*, 131–143. http://doi.org/10.1016/j.cortex.2015.08.008

Golestani, N., Molko, N., Dehaene, S., LeBihan, D., & Pallier, C. (2007). Brain structure predicts the learning of foreign speech sounds. *Cerebral Cortex*, *17*(3), 575–582. http://doi.org/10.1093/cercor/bhk001

Goswami, U. (2011). A temporal sampling framework for developmental dyslexia. *Trends in Cognitive Sciences*, *15*(1), 3–10.

Goswami, U., Fosker, T., Huss, M., Mead, N., & Szucs, D. (2011). Rise time and formant transition duration in the discrimination of speech sounds: The Ba-Wa distinction in developmental dyslexia. *Developmental Science*, *14*(1), 34–43.

Goswami, U., Thomson, J., Richardson, U., Stainthorp, R., Hughes, D., Rosen, S., & Scott, S. K. (2002). Amplitude envelope onsets and developmental dyslexia: A new hypothesis. *Proceedings of the National Academy of Sciences of the United States of America*, *99*(16), 10911–10916. http://doi.org/10.1073/pnas.122368599

Goswami, U., Wang, H.-L. S., Cruz, A., Fosker, T., Mead, N., & Huss, M. (2010). Language-universal sensory deficits in developmental dyslexia: English, Spanish, and Chinese. *Journal of Cognitive Neuroscience*, *23*(2), 325–37.

Gough, P. B., & Tunmer, W. E. (1986). Decoding, reading, and reading disability. *Remedial and Special Education*, *7*(1), 6–10. http://doi.org/10.1177/074193258600700104

Grigorenko, E. L., Kornilov, S. A., & Rakhlin, N. (in press). Learning to read Russian. In L. Verhoeven & C. A. Perfetti (Eds.), *Reading acquisition across languages and writing systems* (pp. 393–415). Cambridge, UK: Cambridge University Press.

Guan, C. Q., Liu, Y., Chan, D. H. L., Ye, F., & Perfetti, C. A. (2011). Writing strengthens orthography and alphabetic-coding strengthens phonology in learning to read Chinese. *Journal of Educational Psychology*, *103*(3), 509–522. http://doi.org/10.1037/a0023730

Guion, S. G., & Pederson, E. (2007). Investigating the role of attention in phonetic learning. In O.-S. Bohn & M. J. Munro (Eds.), *Language experience in second language speech learning* (pp. 57–77). Philadelphia, PA: John Benjamins Publishing Company.

Gussmann, E. (2007). *The phonology of Polish*. New York, NY, USA: Oxford University Press.

Hämäläinen, J., Leppänen, P. H. T., Torppa, M., Müller, K., & Lyytinen, H. (2005). Detection of sound rise time by adults with dyslexia. *Brain and Language, 94*(1), 32–42. http://doi.org/10.1016/j.bandl.2004.11.005

Hamann, S. (2003). *The phonetics and phonology of retroflexes* (Doctoral dissertation). Utrecht University, Netherlands.

Hanulíková, A., & Hamann, S. (2010). Slovak. *Journal of the International Phonetic Association*, *40*(3), 373–378. http://doi.org/10.1017/S0025100310000162

Hebb, D. O. (1949). *The organization of behavior*. New York: Wiley.

Huss, M., Verney, J. P., Fosker, T., Mead, N., & Goswami, U. (2011). Music, rhythm, rise time percepion and developmental dyslexia: Perception of musical meter predicts reading and phonology. *Cortex*, *47*(6), 674–689.

James, K. H. (2010). Sensori-motor experience leads to changes in visual processing in the developing brain. *Developmental Science*, *13*(2), 279–288. http://doi.org/10.1111/j.1467-7687.2009.00883.x

James, K. H., & Atwood, T. P. (2009). The role of sensorimotor learning in the perception of letter-like forms: Tracking the causes of neural specialization for letters. *Cognitive Neuropsychology*, *26*(1), 91–110. http://doi.org/10.1080/02643290802425914

James, K. H., & Engelhardt, L. (2012). The effects of handwriting experience on functional brain development in pre-literate children. *Trends in Neuroscience and Education*, *1*(1), 32–42. http://doi.org/10.1016/j.tine.2012.08.001

James, K. H., & Gauthier, I. (2006). Letter processing automatically recruits a sensory-motor brain network. *Neuropsychologia*, *44*(14), 2937–49. http://doi.org/10.1016/j.neuropsychologia.2006.06.026

Jamieson, D. G., & Morosan, D. E. (1986). Training non-native speech contrasts in adults: Acquisition of the English /ð/-/θ/ contrast by francophones. *Perception & Psychophysics*, *40*(4), 205–215. http://doi.org/10.3758/BF03211500

Jesry, M. M. (2005). Theoretically-based practical recommendations for improving EFL / ESL students’ pronunciation. *Journal of King Saud University*, *18*, 1–33.

Joshi, R. M. (2013). Literacy in Kannada, an alphasyllabic orthography. In H. Winskel & P. Padakannaya (Eds.), *South and Southeast Asian psycholinguistics* (pp. 184–191). New York, NY: Cambridge University Press.

Kachru, Y. (2006). *Hindi*. Philadelphia, PA, USA: John Benjamins Publishing Company.

Khan, I., Gupta, S. K., Rizvi, S. H. S., & Gupta, A. (1991). Statistics of printed Hindi text characters: Preliminary results. *IETE Journal of Research*, *37*(3), 268–275.

Kyle, F., Kujala, J., Richardson, U., Lyytinen, H., & Goswami, U. (2013). Assessing the effectiveness of two theoretically motivated computer-assisted reading interventions in the United Kingdom: GG Rime and GG Phoneme. *Reading Research Quarterly*, *48*(1), 61–76. http://doi.org/10.1002/rrq.038

Landerl, K. (in press). Learning to read German. In L. Verhoeven & C. A. Perfetti (Eds.), *Reading acquisition across languages and writing systems* (pp. 299–322). Cambridge, UK: Cambridge University Press.

Leonard, C. M., Eckert, M. A., Lombardino, L. J., Oakland, T., Kranzler, J., Mohr, C. M., … Freeman, A. (2001). Anatomical risk factors for phonological dyslexia. *Cerebral Cortex*, *11*, 148–157. http://doi.org/10.1093/cercor/11.2.148

Leonard, C. M., Lombardino, L. J., Walsh, K., Eckert, M. A., Mockler, J. L., Rowe, L. A., … DeBose, C. B. (2002). Anatomical risk factors that distinguish dyslexia from SLI predict reading skill in normal children. *Journal of Communication Disorders*, *35*(6), 501–531. http://doi.org/10.1016/S0021-9924(02)00120-X

Leonard, C. M., Low, P., Jonczak, E. E., Schmutz, K. M., Siegel, L. S., & Beaulieu, C. (2011). Brain anatomy, processing speed, and reading in school-age children. *Developmental Neuropsychology*, *36*(7), 828–846. http://doi.org/10.1080/87565641.2011.606398

Levitt, H. (1971). Transformed up- down methods in psychoacoustics. *Journal of the Acoustics Society of America*, *49*(2), 467–477. http://doi.org/10.1121/1.1912375

Llamzon, T. A. (1966). Tagalog phonology. *Anthropological Linguistics*, *8*(1), 30–39. Retrieved from http://www.jstor.org/stable/30029139

Longcamp, M., Boucard, C., Gilhodes, J.-C., Anton, J.-L., Roth, M., Nazarian, B., & Velay, J.-L. (2008). Learning through hand- or typewriting influences visual recognition of new graphic shapes: Behavioral and functional imaging evidence. *Journal of Cognitive Neuroscience*, *20*(5), 802–15. http://doi.org/10.1162/jocn.2008.20504

Longcamp, M., Zerbato-Poudou, M.-T., & Velay, J.-L. (2005). The influence of writing practice on letter recognition in preschool children: A comparison between handwriting and typing. *Acta Psychologica*, *119*(1), 67–79. http://doi.org/10.1016/j.actpsy.2004.10.019

Luo, W., Head, A., Schneider, W., & Wang, J. (2014). e-Chimera: Rapid in-situ mobile experimentation through integrated design, test, and deployment. Retrieved from http://www.lrdc.pitt.edu/bov/documents/Wang,Jingtao_e-Chimera.pdf

Martin-Chang, S., Ouellette, G., & Madden, M. (2014). Does poor spelling equate to slow reading? The relationship between reading, spelling, and orthographic quality. *Reading and Writing*, *27*(8), 1485–1505. http://doi.org/10.1007/s11145-014-9502-7

McCandliss, B. D., Fiez, J. A., Protopapas, A., Conway, M., & McClelland, J. L. (2002). Success and failure in teaching the [r]-[l] contrast to Japanese adults: Tests of a Hebbian model of plasticity and stabilization in spoken language perception. *Cognitive, Affective & Behavioral Neuroscience*, *2*(2), 89–108. http://doi.org/10.3758/CABN.2.2.89

McCullagh, M. (2011). The sounds of Latin: Phonology. *A Companion to the Latin Language*, 81–91. http://doi.org/10.1002/9781444343397.ch6

McDaniel, M. A., & Butler, A. C. (2011). A contextual framework for understanding when difficulties are desirable. In *Successful remembering and successful forgetting: A festschrift in honor of Robert A. Bjork* (pp. 175–198). Psychology Press New York, NY.

Melby-Lervåg, M., & Lervåg, A. (2011). Cross-linguistic transfer of oral language, decoding, phonological awareness and reading comprehension: A meta-analysis of the correlational evidence. *Journal of Research in Reading*, *34*(1), 114–135. http://doi.org/10.1111/j.1467-9817.2010.01477.x

Meng, Z. (1998). Duiwai hanyu yuyin jiaoxue zhong shiyong "hanyu pinyin fang"an'de jige wenti [Some issues related to using hanyu pinyin in the teaching of Chinese as a Foreign Language]. *Yuyin Yanjiu Yu Duiwai Hanyu Jiaoxue*, 322–329.

Muneaux, M., Ziegler, J. C., Truc, C., Thomson, J., & Goswami, U. (2004). Deficits in beat perception and dyslexia: Evidence from French. *NeuroReport*, *15*(8), 1255–1259. http://doi.org/10.1097/01.wnr.0000127459.31232.c4

Näätänen, R. (2001). The perception of speech sounds by the human brain as reflected by the mismatch negativity (MMN) and its magnetic equivalent (MMNm). *Psychophysiology*, *38*, 1–21. http://doi.org/10.1111/1469-8986.3810001

Nag, S. (2007). Early reading in Kannada: The pace of acquisition of orthographic knowledge and phonemic awareness. *Journal of Research in Reading*, *30*(1), 7–22. http://doi.org/10.1111/j.1467-9817.2006.00329.x

Nag, S. (2011). The akshara languages: What do they tell us about children's literacy learning? In R. K. Mishra & N. Srinivasan (Eds.), *Language-cognition interface: State of the art* (pp. 291–310). Germany: Lincom Publishers.

Nag, S. (2014). Akshara-phonology mappings: The common yet uncommon case of the consonant cluster. *Writing Systems Research*, *6*(1), 105–119. http://doi.org/10.1080/17586801.2013.855621

Nag, S., & Sircar, S. (2008). *Learning to read in Bengali: Report of a survey in five Kolkata primary schools*. Bangalore, India: The Promise Foundation.

Nag, S., Treiman, R., & Snowling, M. J. (2010). Learning to spell in an alphasyllabary: The case of Kannada. *Writing Systems Research*, *2*(1), 41–52. http://doi.org/10.1093/wsr/wsq001

Naka, M. (1998). Repeated writing facilitates children's memory for pseudocharacters and foreign letters. *Memory & Cognition*, *26*(4), 804–809. http://doi.org/10.3758/BF03211399

Naka, M., & Naoi, H. (1995). The effect of repeated writing on memory. *Memory & Cognition*, *23*(1), 201–212.

Oh, J. S., Jun, S.-A., Knightly, L. M., & Au, T. K. (2003). Holding on to childhood language memory. *Cognition*, *86*(3), B53–B64. http://doi.org/10.1016/S0010-0277(02)00175-0

Olson, R., Wise, B., Conners, F., & Rack, J. (1989). Specific deficits in component reading and language skills: Genetic and environmental influences. *Journal of Learning Disabilities*, *22*(6), 339–348. http://doi.org/http://dx.doi.org.proxy.library.vanderbilt.edu/10.1177/002221948902200604

Ouellette, G. (2010). Orthographic learning in learning to spell: The roles of semantics and type of practice. *Journal of Experimental Child Psychology*, *107*(1), 50–58. http://doi.org/10.1016/j.jecp.2010.04.009

Ouellette, G., & Tims, T. (2014). The write way to spell: Printing vs. typing effects on orthographic learning. *Frontiers in Psychology*, *5*, 1–11. http://doi.org/10.3389/fpsyg.2014.00117

Pandey, P. (2014). Akshara-to-sound rules for Hindi. *Writing Systems Research*, *6*(1), 54–72. http://doi.org/10.1080/17586801.2013.855622

Patel, J., Bapi, R. S., & Nag, S. (2013). *Akshara counts in child directed print: A pilot study with 101 texts.* Manuscript in preparation.

Patel, P. G. (2004). *Reading acquisition in India: Models of learning and dyslexia*. New Delhi: Sage.

Perfetti, C. A., & Hart, L. (2002). The lexical quality hypothesis. In L. Verhoeven, C. Elbro, & P. Reitsma (Eds.), *Precursors of functional literacy* (pp. 189–213). Amsterdam: John Benjamins. Retrieved from http://www.amazon.com/Precursors-Functional-Literacy-Studies-Language/dp/9027218064

Phillips, A. M. (2011). *Orthographic representations affect adult second language word learning* (Master's Thesis). University of Pittsburgh, PA, USA.

Polka, L. (1991). Cross-language speech perception in adults: Phonemic, phonetic, and acoustic contributions. *The Journal of the Acoustical Society of America*, *89*(6), 2961–2977. http://doi.org/10.1121/1.400734

Prakash, P., Rekha, D., Nigam, R., & Karanth, P. (1993). Phonological awareness, orthography and literacy. In R. J. Scholes (Ed.), *Literacy and language analysis* (pp. 55–70). Hillsdale, NJ: Lawrence Erlbaum Associates.

Rathod, P. H., Dhore, M. L., & Dhore, R. M. (2013). Hindi and Marathi to English machine transliteration using SVM. *International Journal on Natural Language Computing*, *2*(4), 55–

71.

Ricketts, J., Bishop, D. V. M., & Nation, K. (2009). Orthographic facilitation in oral vocabulary acquisition. *Quarterly Journal of Experimental Psychology*, *62*(10), 1948–1966. http://doi.org/10.1080/17470210802696104

Rieth, H., Axelrod, S., Anderson, R., Hathaway, F., Wood, K., & Fitzgerald, C. (1974). Influence of distributed practice and daily testing on weekly spelling tests. *The Journal of Educational Research*, *68*(2), 73–77. http://doi.org/10.1080/00220671.1974.10884711

Roediger, H. L., & Karpicke, J. D. (2006). The power of testing memory: Basic research and implications for educational practice. *Perspectives on Psychological Science*, *1*(3), 181–210.

Rosen, S. (1992). Temporal information in speech: Acoustic, auditory and linguistic aspects. *Philosophical Transactions of the Royal Society of London*, *336*(1278), 367–373. http://doi.org/10.1098/rstb.1992.0070

Rosen, V. M., & Engle, R. W. (1998). Working memory capacity and suppression. *Journal of Memory & Language*. http://doi.org/10.1006/jmla.1998.2590

Seidenberg, M. S., & Tanenhaus, M. K. (1979). Orthographic effects on rhyme monitoring. *Journal of Experimental Psychology: Human Learning and Memory*, *5*(6), 546–554.

Serniclaes, W., Heghe, S. Van, Mousty, P., Carre, R., & Sprenger-Charolles, L. (2004). Allophonic mode of speech perception in dyslexia. *Journal of Experimental Child Psychology*, *87*, 336–361. http://doi.org/10.1016/j.jecp.2004.02.001

Serniclaes, W., Sprenger-Charolles, L., Carre, R., & Demonet, J.-F. (2001). Perceptual discrimination of speech sounds in developmental dyslexia. *Journal of Speech, Language, and Hearing Research*, *44*, 384–399.

Shams, L., & Seitz, A. R. (2008). Benefits of multisensory learning. *Trends in Cognitive Sciences*, *12*(11), 411–417. http://doi.org/10.1016/j.tics.2008.07.006

Sinha, R. M. K., & Mahabala, H. N. (1979). Machine recognition of Devanagari script. *IEEE Transactions on Systems, Man, and Cybernetics*, *9*(8), 435–441.

Slevc, L. R., & Miyake, A. (2006). Individual differences in second- language proficiency. *Psychological Science*, *17*(8), 675–681. http://doi.org/10.1111/j.1467-9280.2006.01765.x

Steele, J. (2005). Assessing the role of orthographic versus uniquely auditory input in acquiring new L2 segments. *7èmes Rencontres Internationales Du Réseau Français de Phonologie*, 2–4.

Stevens, K. N., & Blumstein, S. E. (1975). Quantal aspects of consonant production and perception: A study of retroflex stop consonants. *Journal of Phonetics*, *3*, 215–233.

Surányi, Z., Csépe, V., Richardson, U., Thomson, J. M., Honbolygó, F., & Goswami, U. (2009).

Sensitivity to rhythmic parameters in dyslexic children: A comparison of Hungarian and English. *Reading and Writing*, *22*(1), 41–56. http://doi.org/10.1007/s11145-007-9102-x

Tees, R. C., & Werker, J. F. (1984). Perceptual flexibility: Maintenance or recovery of the ability to discriminate non-native speech sounds. *Canadian Journal of Psychology*, *38*(4), 579–590. http://doi.org/10.1037/h0080868

Tiwari, S. (2011). A preliminary investigation of akshara knowledge in the Malayalam alphasyllabary: Extension of Nag's (2007) study. *Writing Systems Research*, *3*(2), 145–151. http://doi.org/10.1093/wsr/wsr013

Tokowicz, N., Michael, E., & Kroll, J. (2004). The roles of study-abroad experience and working-memory capacity in the types of errors made during translation. *Bilingualism: Language and Cognition*, *7*, 255–272.

Underwood, B. J. (1961). Ten years of massed practice on distributed practice. *Psychological Review*, *68*(4), 229–247. http://doi.org/10.1037/h0047516

Vaid, J., & Gupta, A. (2002). Exploring word recognition in a semi-alphabetic script: The case of Devanagari. *Brain and Language*, *81*, 679–690. http://doi.org/10.1006/brln.2001.2556

Vasanta, D. (2004). Processing phonological information in a semi-syllabic script: Developmental data from Telugu. *Reading and Writing: An Interdisciplinary Journal*, *17*, 59–78. http://doi.org/10.1023/B:READ.0000013830.55257.3a

Vaughn, S., Schumm, J. S., & Gordon, J. (1992). Early spelling acquisition: Does writing really beat the computer? *Learning Disability Quarterly*, *15*(3), 223–228.

Vaughn, S., Schumm, J. S., & Gordon, J. (1993). Which motoric condition is most effective for teaching spelling to students with and without learning disabilities? *Journal of Learning Disabilities*, *26*(3), 191–198.

Verma, R., & Chawla, P. (2003). Comparative analysis of Hindi retroflex and dental CV syllables and their synthesis. In *Workshop on Spoken Language Processing*. Mumbai, India: ISCA. Retrieved from http://speech.tifr.res.in/chief/publ/03eurospeech.pdf

Walker, D. C. (1984). *The pronunciation of Canadian French*. Ottawa, Canada: Ottawa Press.

Wechsler, D. (1999). Wechsler abbreviated scale of intelligence. Psychological Corporation.

Werker, J. F. (1989). Becoming a native listener. *American Scientist*, *77*(1), 54–59. http://doi.org/10.2307/27855552

Werker, J. F., Gilbert, J. H. V, Humphrey, K., & Tees, R. C. (1981). Developmental aspects of cross-language speech perception. *Child Development*, *52*(1), 349–355.

Werker, J. F., & Tees, R. C. (1983). Developmental changes across childhood in the perception of non-native speech sounds. *Canadian Journal of Psychology*, *37*(2), 278–286.

http://doi.org/10.1037/h0080725

Werker, J. F., & Tees, R. C. (1984). Phonemic and phonetic factors in adult cross-language speech perception. *The Journal of the Acoustical Society of America*, *75*(6), 1866–1878. http://doi.org/10.1121/1.390988

Wollscheid, S., Sjaastad, J., & Tømte, C. (2016). The impact of digital devices vs. Pen(cil) and paper on primary school students' writing skills - A research review. *Computers and Education*, *95*, 19–35. http://doi.org/10.1016/j.compedu.2015.12.001

Woodcock, R. W., McGrew, K. S., & Mather, N. (2001). Woodcock-Johnson tests of achievement. Itasca, IL: Riverside Publishing.

Xu, Y., Chang, L.-Y., Zhang, J., & Perfetti, C. A. (2013). Reading, writing, and animation in character learning in Chinese as a foreign language. *Foreign Language Annals*, *46*(3), 423–444. http://doi.org/10.1111/flan.12040

Ye, J., Cui, L., & Lin, X. (1997). Waiguo xuesheng hanyu yuyin xuexi duice.

Young-Scholten, M. (1997). Second-language syllable simplification: Deviant development or deviant input. In *New Sounds 97: Proceedings of the Third Symposium on the acquisition of Second-Language Speech, J. Leather and A. James (Eds.)* (pp. 351-360).

Young-Scholten, M. (2002). Orthographic input in L2 phonological development. In P. Burmeister, T. Piske, & A. Rohde (Eds.), *An integrated view of language development: Papers in honor of Henning Wode* (pp. 263–279). Trier: Wissenschaftlicher Verlag Trier.

Young-Scholten, M., Akita, M., & Cross, N. (1999). Focus on form in phonology: Orthographic exposure as a promoter of epenthesis. *Pragmatics and Pedagogy*, *2*, 227–233.