



Theses and Dissertations

---

2020-04-06

## Portuguese and Chinese ESL Reading Behaviors Compared: An Eye-Tracking Study

Logan Kyle Blackwell  
*Brigham Young University*

Follow this and additional works at: <https://scholarsarchive.byu.edu/etd>



Part of the [Arts and Humanities Commons](#)

---

### BYU ScholarsArchive Citation

Blackwell, Logan Kyle, "Portuguese and Chinese ESL Reading Behaviors Compared: An Eye-Tracking Study" (2020). *Theses and Dissertations*. 8435.

<https://scholarsarchive.byu.edu/etd/8435>

This Thesis is brought to you for free and open access by BYU ScholarsArchive. It has been accepted for inclusion in Theses and Dissertations by an authorized administrator of BYU ScholarsArchive. For more information, please contact [ellen\\_amatangelo@byu.edu](mailto:ellen_amatangelo@byu.edu).

Portuguese and Chinese ESL Reading Behaviors Compared: An Eye-Tracking Study

Logan Kyle Blackwell

A thesis submitted to the faculty of  
Brigham Young University  
in partial fulfillment of the requirements for the degree of  
Master of Arts

Troy L. Cox, Chair  
Grant T. Eckstein  
Robert J. Reynolds

Department of Linguistics  
Brigham Young University

Copyright © 2020 Logan Kyle Blackwell

All Rights Reserved

## ABSTRACT

### Portuguese and Chinese ESL Reading Behaviors Compared: An Eye-Tracking Study

Logan Kyle Blackwell  
Department of Linguistics, BYU  
Master of Arts

While reading behaviors have been studied extensively in L1 reading studies through the use of eye-tracking and L2 reading has been measured through inherently indirect means, there is a relative lack of research done on early and late reading measures of ESL readers. Eye-tracking technology, available to researchers only in the past few decades, has opened the field to a new means of measuring these early and late measures of reading in second language learners. This study investigates the reading behaviors of 34 native Portuguese and Chinese readers who read in both their native languages (L1) and in their second language (L2), which is English. It was found that readers processed their reading differently in response to different text difficulties and varied between the different native languages.

Keywords: ESL, reading, reading proficiency, eye-tracking, Portuguese, Chinese

## ACKNOWLEDGEMENTS

I would like to thank all those that made this paper a reality. I'd like to thank Dr. Troy Cox for getting me started on this project and for giving me the guidance on how to turn an academic idea into a polished paper. I also have the rest of my committee to thank for giving me advice on developing myself as an academic and professional in the field. In the same way, I am grateful for the whole ICAMRA research group for making this interdisciplinary work possible and for laying the groundwork that I have benefited from. The instrument used in this study was designed in coordination with the ICAMRA research group under Dr. Troy Cox, with help from Michael Child in designing the Portuguese section, Rachel Liu in designing the Chinese section, and Steve Luke and Grant Eckstein in setting up the layout of the design itself. Finally, I have my wife Jennifer to thank for supporting me and for believing in me through sleepless nights over the last couple of years.

## TABLE OF CONTENTS

Portuguese and Chinese ESL Reading Behaviors Compared: An Eye-Tracking Study .....	i
ABSTRACT.....	ii
ACKNOWLEDGEMENTS.....	iii
TABLE OF CONTENTS.....	iv
LIST OF TABLES.....	vi
LIST OF FIGURES .....	vii
PREFACE.....	viii
<b>Introduction</b> .....	1
<b>Review of Literature</b> .....	2
Eye-tracking.....	2
Overview of Reading in Assessment .....	5
Reading Proficiency.....	6
Portuguese.....	7
Chinese.....	8
Research Questions.....	9
<b>Methodology</b> .....	12
Participants.....	12
Texts.....	13
Materials .....	14

Equipment .....	16
Procedures .....	16
Data Measurement and Analysis .....	17
<b>Results</b> .....	18
Portuguese Measures of Reading.....	18
Chinese Measures of Reading.....	23
Portuguese and Chinese .....	27
<b>Discussion</b> .....	30
Portuguese Early and Late Measures of Reading .....	30
Chinese Early and Late Measures of Reading .....	30
Interlanguage Comparison .....	31
ESL Reading Proficiency.....	31
<b>Conclusion</b> .....	32
Summary .....	32
Limitations .....	32
Future Research .....	33
References.....	34
<i>Appendix A</i> .....	38

## LIST OF TABLES

Table 1. Participant Reading Levels in English by Native Language .....	13
Table 2. First Fixation Duration per word for Portuguese ESL readers .....	20
Table 3. Gaze Duration for Portuguese ESL readers.....	21
Table 4. Total Dwell Time for Portuguese ESL readers.....	22
Table 5. First Fixation Duration for Chinese ESL readers. ....	24
Table 6. Gaze Duration for Chinese ESL readers.....	25
Table 7. Total Dwell Time for Chinese ESL readers.....	26
Table 8. First Fixation comparison between Chinese and Portuguese ESL readers.....	27
Table 9. Gaze Duration comparison between Chinese and Portuguese ESL readers. ....	28
Table 10. Total Dwell Time comparison between Chinese and Portuguese ESL readers.....	29

## LIST OF FIGURES

Figure 1. Measures of Interest. ....	3
Figure 2. First fixation duration results for Portuguese ESL readers. ....	19
Figure 3. Gaze duration results for Portuguese ESL readers. ....	20
Figure 4. Total dwell time results for Portuguese ESL readers. ....	22
Figure 5. First fixation duration results for Chinese ESL readers. ....	23
Figure 6. Gaze duration results for Chinese ESL readers. ....	25
Figure 7. Total dwell time results for Chinese ESL readers. ....	26
Figure 8. First fixation duration comparison between Chinese and Portuguese ESL readers. ....	27
Figure 9. Gaze duration comparison between Chinese and Portuguese ESL readers. ....	28
Figure 10. Total dwell time comparison between Chinese and Portuguese ESL readers. ....	29



## PREFACE

This thesis was written with the purpose of being submitted as a manuscript to be published in a journal. The following journals were chosen as options for publication: *Language Assessment Quarterly*, *Language Testing*, *Reading in a Foreign Language*, and *Second Language Research*

## Introduction

College bound ESL (English as Second Language) students have a strong interest in learning how to read fluently. Fluent reading is a complex and multifaceted concept that can be difficult to define. Grabe (2009) defines fluent reading in terms of ten processes that include being rapid, efficient, comprehending, interactive, strategic, flexible, purposeful, evaluative, learning, and linguistic. For ESL learners, fluent reading can be as simple as the kind of reading that native readers do. For those that read in their first language (L1), it can be frustrating to not be able to read as fluently in their second language (L2). This is especially true when ESL students pursue academics where reading skills are more valued than oral skills (Johns, 1981). However, measuring reading proficiency can be a difficult and inaccurate science. Multiple choice questions may only explain a test-taker's ability to answer those questions. Even short answer questions are only a proxy of what a reader is processing. Until the last few decades, however, there hasn't been a more direct way of measuring reading proficiency.

Eye-tracking technology has revolutionized the way that language processing is studied. Eye-trackers are high resolution cameras that measure eye positions on a screen by capturing images of where a person looks from millisecond to millisecond. The Eye-Mind Hypothesis postulates that what we observe is what we process in our mind, and how long we observe a given point reflects how much processing we utilize for it (Rayner, 1998). For decades, studies have used eye-tracking models (1) to understand the natural reading process by modeling the effects of higher language processing while reading (Reichle, Warren, & McConnell, 2009), (2) to inform language test validity (Bax & Chan, 2019) and even (3) to differentiate between the native languages of ESL learners (Berzak et al., 2017). However, with all the research conducted on eye-tracking and reading, relatively few researchers have investigated ESL reading measures

using eye-tracking. Reading measures used in other studies to investigate ESL reading behaviors were examined.

## Review of Literature

### Eye-tracking

There are hundreds of eye-tracking measures that can be used to glean information from movements in the eyes, but for this study, those pertinent to reading will be discussed. As a person reads, their eye fixates briefly on letters, words, phrases, etc., revealing both what they are comprehending and how much processing is being done. The areas being research are called areas of interest (AOI) and can be as small as a single letter but could also consist of multiple sentences or areas of screen. Each AOI that is fixated on is referred to as a *fixation*, while the movement between AOIs is referred to as *saccades*. A saccade can go forward or backward which is called a *regression*. Eye-trackers measure the number and amount of times an AOI is fixated upon as well as the path the eyes take among other measures such as pupil size.

In Figure 1, we see a hypothetical example of the familiar phrase, “The quick brown helicopter jumps over the lazy dog.” In this example, the AOIs are the words in the sentence, though another research might define the AOIs as the whole sentence or even single letters. The numbers 1 to 9 indicate each of the fixations. The first fixation is the word “the”, and the second is the word “brown.” The word “quick” was skipped, thus indicating a saccade. The fixations 3, 4, and 6 are all on the same AOI— “helicopter.” This is unsurprising as that word replaced the expected “fox” that is normally used. The sum of the time spent in fixations on 3 and 4 is the gaze duration. Fixation 5 is the word “jumps,” but then there is a regression back to the helicopter AOI. Thus, the total reading time for the helicopter AOI is the sum of the time spent with fixations 3, 4 and 6.

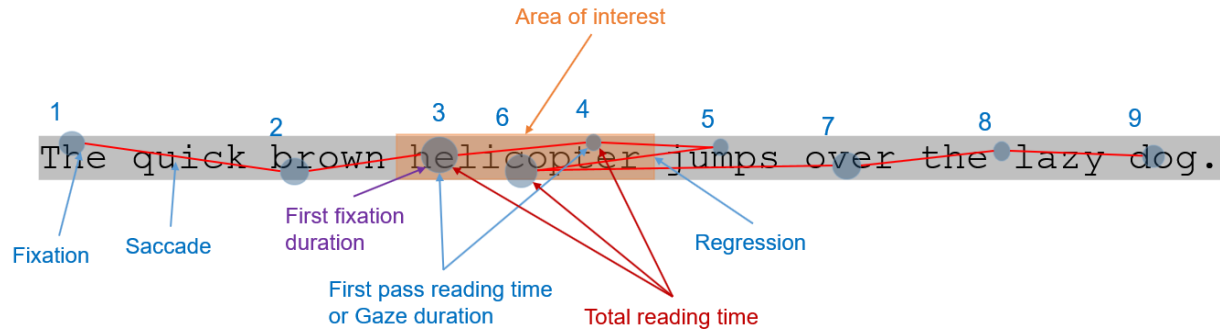


Figure 1. Measures of Interest.

Although there are hundreds of measures, including first fixation duration, regressions in and out of an area of interest, saccade length, number of fixations, the skip rate, etc., most reading research divides the list into early and late measures of reading comprehension (Clifton, Staub, & Rayner, 2007; Rayner 1998, 2009). Early measures include letter and word recognition and are the first signs that a word is recognized. They predict length, frequency, and predictability of the fixated word (Kliegl, Nuthmann, & Engbert, 2006). The most commonly used early measures are *first fixation duration* (the time that is spent in the area of interest for the first fixation, such as fixation 3 in Figure 1.), *gaze duration* (the sum of all the fixations on a word the first time that it is read, such as fixations 3 + 4 in Figure 1.) Late measures look more at whether a word is fully comprehended (Rayner et al., 1989; Roberts & Siyanova-Chanturia, 2013) of which *total reading time* (the sum of all the fixations in an area of interest, such as fixations 3 + 4 + 6 in Figure 1.) is the most common (Rayner, 1998; 2009).

In the past 15 years, eye-tracking research has provided further understanding of reading patterns with these measures. Rayner (1998) found that only 10-15% of saccades are regressions for native English readers. In first language (L1) reading, Ashby, Rayner, and Clifton (2005) found that less skilled readers were differentiated from fluent readers by shorter saccades and longer fixations. Similar conclusions were reached about ESL readers by Cop et al. (2015). In

their 2015 study, they gave Dutch L1/English L2 readers as well as English monolingual readers a novel to read in English and in Dutch in four one and a half hour sessions. They found that the readers in the L2 had longer sentence reading times, more fixations per sentence, longer fixation durations, shorter saccade lengths, and were less likely to skip AOIs compared to in the L1.

In one study it was concluded that those with more L2 reading exposure in contrast to those with less L2 reading exposure had faster reading rates in the L2, but decreased L1 reading rates (Whitford & Titone, 2012, 2015). They divided participants into two groups: those that had a high degree of current exposure to L2 input relative to their L1 input, and those that were the reverse (the L2 input was relatively low compared to their L1 input). They computed the L1/L2 ratio using an assessment where participants had to make lexical decisions about animate and inanimate nouns in counterbalanced L1 and L2 blocks and found that the rate of L2 exposure relative to L1 exposure influenced general reading behaviors in the L1.

In another study, Krstić et al. (2018) found that students that did well on the PISA (the Programme for International Student Assessment) had a smaller variance in measures such as average fixation duration and saccade amplitude. Serbian students aged 14 and 15 were given four PISA reading literacy units with six multiple-choice items and were measured using an eye-tracker. The reading units varied in proficiency difficulty and were conducted in the L1 of the students. The lower scoring students had shorter saccade amplitudes, which mirror the results found in other studies where young children and those with dyslexia have smaller saccades. The findings indicate that lower scoring students can be distinguished from higher scoring students based on saccade length, if not by fixation duration, in their L1. In both studies, it can be noted that proficiency in the language played a role in measured differences in eye behavior and

movement. Proficiency studies that use eye-tracking measures have found differences among reading proficiency levels.

### **Overview of Reading in Assessment**

In the testing field, much discussion has been made about the measurement of reading proficiency of L2 learners. Reading is seen as a necessary component of English in many areas, including in academia (Johns, 1981), and while certain aspects of reading are universal, such as phonological processing, readers in different languages prioritize different strategies in how fluent reading is accomplished (Grabe, 2009). Since reading is a receptive skill, assessment has been through indirect means such as comprehension questions (e.g. multiple-choice, true-false, and cloze tasks (Paris & Stahl, 2005)), with the understanding that an indirect measure of reading can be used to infer the proficiency of the reader. These item types are practical in terms of scoring efficiency and allow for a comparison between large numbers of students, however, are only proxy measures and may not reflect the construct well.

Many have criticized the overuse of multiple-choice comprehension questions as a reliable means of testing readers (Bernhardt, 2011). Rupp, Ferne, and Choi (2006) found that test takers of the TOEFL did not use reading processes that matched reading processes outside of a testing environment. They found that “different MC questions do not merely tap but, indeed, create very particular comprehension and response processes” (p. 470). The participants approached the questions as problem-solving tasks rather than reading comprehension tasks. In both instances, however, the interaction with the tasks was being used to infer the reading process itself and threatens validity claims. Cohen and Upton (2007) also found that test takers of the TOEFL (Test of English as a Foreign Language) were completing it with the primary goal of answering questions correctly, not learning or acquiring new knowledge from the texts they read.

In both studies, participants demonstrated that without careful alignment of question and reading, assumptions about student proficiency may be flawed.

Thus, there is a potential challenge in construct validity when assessing reading through comprehension questions. Even though the criticisms of these indirect methods of assessing reading comprehension are clear, it is less clear how to assess reading comprehension without them. With the productive skills, such as speaking or writing, what is created is what is being tested, allowing for direct access to the proficiency of the language. But for the receptive skills, direct access to measure proficiency is impossible, though carefully designed proxies can reveal much of the behaviors of the reading process.

### **Reading Proficiency**

The American Council on the Teaching of Foreign Languages (ACTFL) separates language proficiency into four skill areas: Reading, Writing, Listening, and Speaking (ACTFL, 2012). Language, as understood under this framework is the ability that a person has to perform an action or function in a language. Proficiencies are divided into five main levels (Novice, Intermediate, Advanced, Superior, and Distinguished), with the first three being additionally divided into three sublevels (Low, Mid, and High). In the High sublevel, the learner is able to reach the next level up most of the time, with occasional failure, leading them to have many of the same characteristics as that of the next level up.

Proficient reading is defined “as the active, automatic, far-transfer process of using one’s internalized language and culture expectancy system to efficiently comprehend an authentic text for the purpose for which it was written.” (Clifford & Cox, 2013, p. 50) Thus, the view of the reading vis-à-vis a proficiency lens requires identifying both the author’s intent and the real-world communicative function reader should perform. But, as reading itself is a receptive skill, it can only be measured through comprehension questions that match the proficiency of the text

(ACTFL, 2012). While some have challenged the validity of the proficiency levels (Alderson, 2000; Brown & Hudson, 2002), others have shown that there are significant differences among them (Clifford & Cox, 2013). The proficiency guidelines for receptive skills for both reading and listening are based on the assumption that the author or speaker has an intended function, and that the reader or listener is expected to understand that function at different levels of precision. That is, at different proficiency levels, the reader or listener has to match the level of the written or spoken text.

One of the major claims of the ACTFL proficiency scale is that it intended to be language agnostic. The General Preface of the ACTFL Proficiency Guidelines (2012) state, “these Guidelines present the levels of proficiency as ranges, and describe what an individual can and cannot do with language at each level, regardless of where, when, or how the language was acquired” (p. 3) and can be applied to an L1 as well as an L2. Thus, well-educated, highly articulate L1 readers of English will be able to perform the same functions as well-educated, highly articulate L1 readers of Portuguese, Russian and Chinese. However, these functions are global, and may not be done at the microlevel for any single language. Berzak et al. (2017) found L2 readers of English (Japanese, Chinese, Spanish, and Portuguese) had significantly different eye-tracking patterns between the languages when looking at the first fixation, gaze duration, and total reading time with AOIs that were tagged by part of speech. They suggest that further research could indicate other means of distinguishing the languages using eye-tracking.

### **Portuguese**

Portuguese, an official language of ten countries, is a Romance language that is the tenth most common first language of English Language Learners in the United States (National Center for Education Statistics, 2019). It shares many cognates from Latin with English and many Portuguese speakers learn English as a second language.



There are also differences between Portuguese and English that might contribute to difficulties in reading behaviors for L2 learners of English. In terms of orthography, Portuguese words include the accent marks, both acute and grave, the circumflex, the tilde, and the cedilla. These special characters allow for a more transparent spelling system, as opposed to English, which through centuries of change, has developed many non-transparent spellings for words. This can lead to Portuguese learners of English and other learners from a phonemic L1 complaining about the non-phonemic nature of English.

However, eye-tracking research done by Grahl et al. (2019) suggests that while English learners of Portuguese read slower in their L2 than in their L1, there was not a “significant interaction between readers’ performance in either of the languages” (p. 42) when it came to proficiency. The readers’ performance in this case is defined as the average first fixation duration, gaze duration, and total reading time per word of the reader. This lack of difference between the proficiency levels in the L2 was attributed to the fact that both English and Portuguese are alphabetic languages. English readers of Portuguese were able to recognize words just as quickly when they were at a lower proficiency as when they were at a higher proficiency level.

### **Chinese**

Mandarin Chinese is used as an official language in five countries and is the third most common language for English Language Learners in the United States at over two percent of all ELLs (National Center for Education Statistics, 2019). Of the five countries that use it as an official language, three countries use the simplified Chinese characters as the default script for the language.

One of the most salient features of Chinese that could interfere with the process of reading in English for Mandarin Chinese readers is the logographic script used in Chinese. In

previous eye-tracking studies done on native Chinese readers, it has been found that native Chinese readers have an average saccade length of around 2.6 characters (Inhoff & Liu, 1997, 1998) compared to the average length in English of seven to eight characters (Rayner 1998). Chinese characters condense information into a much smaller space, which could potentially cause them to spend more time on a single character. However, first fixation durations do not seem to be affected by this. Yan et al. (2006) found that average first fixation duration for native English and Chinese readers were comparable.

While there have been many more studies on native Chinese readers than on native Portuguese readers, there has not been a study comparing either of them as ESL readers. Portuguese sharply contrasts with Chinese due to the writing systems of English and Portuguese being more similar to each other than they are to Chinese. Both English and Portuguese use the Latin alphabet, while Chinese readers learn the logographic characters. Additionally, Chinese learners of English do not get the vocabulary cognates that Portuguese L1 readers do. Other studies have looked at Chinese and Portuguese ESL readers using think-aloud protocols (Li & Suen, 2015), but a comparable study has not been done with eye-tracking technology. In their study, four Chinese ESL learners and three Romance-language ESL learners took the Michigan English Language Assessment Battery (MELAB) reading test while performing a think-aloud study. The researchers found that the Chinese ESL learners used test-taking strategies instead of vocabulary knowledge to answer questions, while the Romance-language ESL learners were more likely to take advantage of their vocabulary knowledge in English.

### **Research Questions**

With all of this study on eye-tracking measures, and specifically on first and second language reading, there are only a few studies that go in depth on second language reading rates in eye-tracking studies. Looking at the sentence level, the study by Cop et al. (2015)

demonstrates that there is some similarity between first language early and late reading measures and second language early and late reading measures.

In addition, in her thesis, Grahl (2019) used an eye-tracker to measure the differences found in English learners of Portuguese when reading texts of varying difficulty in both their L1 and L2. She found that while there were differences when reading at varying text difficulties, there were not many differences between the proficiency levels of the participants. When it came to the texts, lower level texts tended to be processed more quickly in the L2 than higher level texts.

Rybakova (2019) furthered Grahl's research (2019) by examining English learners of Russian who were reading texts written in the Cyrillic alphabet. Rybakova used the same methodology and instruments (translated into Russian) and found that there also was a significant difference between proficiency levels of the participants in their reading behaviors between the Intermediate and Advanced leveled readers in their L2 (Russian), but that there was not a significant difference between Advanced readers and Superior readers in their L2 of Russian. She also found that English readers of Russian were much slower in their L2 than in their L1. This difference was larger than that found by Grahl (2019) was attributed to the differences in scripts. Neither study looked at ESL readers, however.

In this study, by contrast, early and late measures of ESL reading were examined to determine differences in proficiency level and between first languages. With all the research pursued in eye-tracking and reading, there has been relatively little work conducted on ESL reading measures using eye-tracking when it comes to proficiency. In the study, the behaviors of readers of English as a second language with respect to three variables were measured using differences between the readers based on their L1 (Chinese and Portuguese), their proficiency

level (comparing more advanced readers to their less advanced counterparts), and the difficulty of the text.

These issues lead to the exact nature of this study. The purpose of this study to investigate the differences in early and late reading measures of native Portuguese and Chinese ESL readers. In order to do that, observations occurred at the word level using first fixation duration, gaze duration, and total reading time to stand in as proxies for early and late reading measures when ESL learners with different L1 backgrounds and with different proficiency levels read texts at different proficiency levels in English. In order to achieve a direct comparable replication of the studies done by Grahl (2019) and Rybakova (2019), the methodology of this study mirrored those studies as closely as possible with the research questions below. In those studies, native readers of English who were learning Portuguese and Russian respectively were measured using the same three measures that were used here (first fixation duration, gaze duration, and total reading time). However, this study also intended to probe further into comparisons between languages as well. Hence, the following research questions:

1. What are the differences between early and late measures of reading of native Portuguese ESL readers in their native languages and English?
  - a. How are those differences moderated by their ESL reading proficiency and the passage proficiency level?
2. What are the differences between early and late measures of reading of native Chinese ESL readers in their native languages and English?
  - a. How are those differences moderated by their ESL reading proficiency and the passage proficiency level?

3. What differences exist between native Portuguese and native Chinese readers with early and late measures of reading both in their L1 and in English?

## **Methodology**

### **Participants**

The participants in this study were 34 students from a large western American university. There were 11 males and 23 females, aged between 18 and 60 years old. Participants were native speakers of Portuguese (n=18) and Chinese (n=16). Of the Portuguese speakers 1 was from Angola and 16 were from Brazil. Of the Chinese participants, one was from Malaysia, five were from Taiwan, and 11 were from Mainland China. They were recruited from English classes at an intensive English program and from student organizations at the university. Participants English reading proficiency ranged from Novice-High to Advanced-Mid, based on the proficiency guidelines from the American Council of the Teaching of Foreign Languages (ACTFL).

In the ACTFL Proficiency Guidelines a “High” sublevel is reserved for a person who can meet the major level above it most of the time, but not all the time. A High sublevel reader will often look similar to a low reader of the next level up. Both can read at the higher major level most of the time. The difference is that the High sub-level reader of the lower level will break down occasionally. Clifford and Cox (2013) found that there was a significant difference in scores on a test based on the ACTFL Guidelines between the Intermediate level and the Advanced level. Therefore, the participants were split into two groups, those at an Intermediate High level up to those at Superior level, and those from Novice Mid up to Intermediate Mid. The former group was labeled the Upper group and the latter group was labeled the Lower group. This both allowed for analysis of the most similar participants and in addition allowed for comparisons between similarly sized groups of participants to be made. If the division had been

made at another level, the two groups would have been lop-sided either toward the Upper group or the Lower group.

Finally, all final participants either had normal or corrected vision and were able to read a computer screen. Due to visual impairments that prevented use of the eye-tracker, five participants (not included in the data) were excluded from this study. For a breakdown of the groups according to their Grouping, English Reading Proficiency level, and native language, refer to Table 1.

*Table 1. Participant Reading Levels in English by Native Language*

	Lower Proficiency Group		Upper Proficiency Group		
	Novice	Intermediate Low and Mid	Intermediate High	Advanced	Superior
Portuguese	4	7	5	2	0
Chinese	6	3	5	1	1

## **Texts**

For the languages used, Portuguese and Chinese were chosen due to the difference in the writing systems. Portuguese uses the same alphabetic Latin script that English does with only a couple of minor differences in the number of letters. Chinese was chosen as a logographic system that has as little phonetic transcription as possible.

The other two aspects studied involve proficiency testing. For this experiment, a reading proficiency test based on the ACTFL Proficiency Guidelines that had been validated (Clifford & Cox, 2013) was used to determine both the proficiency level of the participant and the difficulty of the text. The ACTFL Proficiency Guidelines outline five major levels of proficiency: Novice, Intermediate, Advanced, Superior, and Distinguished, with sub-levels of Low, Mid, and High for the lowest three levels (ACTFL, 2012).

In this study, both the difficulty of the text and the proficiency of the readers were measured using this scale. The texts were separated into Intermediate, Advanced, and Superior,

as those are the levels that generally concern non-native readers of English (Boldt et al., 1992). However, for the readers' proficiency the approach described above was employed, splitting the participants into Upper and Lower Proficiency groups.

### **Materials**

The materials for this study cover the three sections of the experiment. There were 3 instruments used in this study: the Eye-Tracking and Language Background Survey, the English Reading Proficiency Test, and the Eye-Tracking Experiment Software.

#### *Eye-Tracking and Language Background Survey*

In the Eye-Tracking and Language Background Survey, participants filled out an online survey created with Qualtrics, which removed participants who would not be able to participate due to visual issues reading a computer screen, asked for demographic information, and provided a Language Background Questionnaire (Appendix A). Further information of the removal of participants can be found in the Discussion Section. In addition, in order to collect this data, an electronic consent form was provided for participants.

#### *English Reading Proficiency Test*

Participants took an adaptive English Reading Proficiency Test validated by Clifford and Cox (2013) for the purpose of determining their reading proficiency. The test, based on the ACTFL Guidelines (2012), was written completely in English, was computerized, and took around half an hour to complete. The test had 30 questions, divided into three levels. Participants read passages on the left side of the screen with five multiple choice options on the right side.

#### *Eye-Tracking Experiment Software*

Participants also utilized an eye-tracker to complete the experiment. In that part of the experiment, participants read 34 reading passages with half in the L1 (Portuguese or Chinese) and half in the L2 (English). These reading passages were also developed in conjunction with

Clifford and Cox (2013). The passages were divided into three groups according to ACTFL levels: Intermediate, Advanced, and Superior. Each passage was verified by an expert to be at the level intended. The Intermediate passages totaled at 12 passages per language with a word count range of 50 to 70 words, with the Portuguese and Chinese being translations of English texts and the English being different but equivalent passages. The Advanced passages numbered four and averaged 150 to 180 words each, with the same division of English and non-English passages. Each language also had one Superior level text, ranging between 300 and 400 words. The Superior reading passage was unique to each language as one defining characteristic of Superior level reading is cultural understanding that varies from language to language.

Participants were calibrated to the machine using a 9-point calibration. The experiment had a true randomization for each participant between the L1 of the participant and their L2 (English). They were then presented two practice readings, which were not measured. After the practice readings, the participants were shown the passages. In addition to being randomized by language as a group, the passages that the participants read were also randomized within the language, with the exception that the Superior text was prevented from appearing first, in order to minimize reader fatigue. In the case of a reading being longer than one screen, participants were not permitted to return to a previous screen.

In order to focus the participant on reading, a comprehension question was administered on the screen following the reading but was not used in the analysis. Participants were also not permitted to return to the reading after seeing the question. Each question, written in the same language as the reading they had just completed, had five potential answers, with the last one being “I don’t know” and the first four being split into three distractors and one correct answer. This was added to the study in order to ensure that readers had a purpose to their reading, while



avoiding some potential issues common to multiple-choice comprehension questions, such as participants searching for key words in the question. Participants could not return to the reading after seeing the question, so they didn't know exactly what they were looking for. In addition, the questions were also leveled according to ACTFL standards so that they matched the difficulty of the text.

### **Equipment**

Participants completed the study individually using one of two identical eyetracking machines in separate locations. Both labs were equipped to house identical SR Research Eye-link 1000 Plus eye-trackers with a 35mm lens, using forehead and chin rests. The chin rests were set at 63 inches from a 21-inch computer monitor with a resolution of 1920 x 1080 pixels. Fixations were measured at 1000Hz. The fixations are accurate to the millisecond and within 5 degrees of visual angle.

### **Procedures**

This study was divided into three sessions taken by participants individually: the Eye-Tracking and Language Background Survey, the English Reading Proficiency Test, and the Eye-Tracking Session.

Participants took the Eye-Tracking and Language Background Survey as the first step of the study, after which they were contacted by email to continue with the study by taking the English Reading Proficiency Test. Participants took the English Reading Proficiency Test in a proctored testing lab, after which participants were contacted by email to schedule the Eye-Tracking Session. Once there, participants were seated in front of the eye-tracking machine. Participants completed a written consent form and received oral and written instructions on how to complete the experiment. After the experiment was finished, participants received a deposit

into a university account worth \$25 as well as the certificate for the English Reading Proficiency Test.

### **Data Measurement and Analysis**

For the purpose of this study, the differences between early and late measures of reading were examined according to an eye-tracker. The early measures were operationalized as first fixation duration (early measure), gaze duration (early measure), and late measure was operationalized as total time (late measure).

Data collected by this study were cleaned according to standard practices of eye-tracking research. The data were reformatted both spatially and temporally to account for systematic errors in data collection. Temporal cleaning removed any fixations smaller than 80 milliseconds or larger than 1600 milliseconds resulting in loss of 6% of fixations. The upper limit was raised from 800 milliseconds in order to account for the longer fixations of the Chinese participants. However, all data reformatting occurred equally on every participant.

In order to address the research questions of this study, a Linear Mixed Effect Model (LME) was computed using R 3.6.0. (R Core Team, 2019), using the lme4 package and the lmerTest packages (Bates, Mächler, Bolker, & Walker, 2016) to get *p* values. LME modelling is a type of linear regression modelling that considers fixed and random effects. Fixed effects in this study were language, proficiency of the participant, and text difficulty. Random effects were words and participants.

In an LME model, interactions would obscure large-scale effects if they were not significant. In order to account for this, the model was run first with all interactions. In instances where none of the three-way interactions were significant, they were eliminated from the model and the model was re-run, looking at two-way interactions. If none of the two-way interactions were significant, it was re-run only looking at the highest level.

In this study, two dependent variables were considered, early and late measures. They were operationalized into three measures of reading: (1) first fixation duration (early measure), (2) gaze duration (early measure), and (3) total time (late measure). For research questions one and two, the independent variables of participants' native language (L1), whether they were reading in English or their L1, and participant English reading proficiency were used. For their reading proficiency, the participants were separated into two groups, Upper and Lower. Participants that reached a proficiency score of Intermediate High or higher, or who showed evidence of having that level by being admitted into an English-speaking university, were placed into the Upper group. Participants who had scores of Intermediate Mid or lower were placed into the Lower group. An ANOVA performed on the Upper and Lower groups found the Portuguese Upper and Lower groups and the Chinese Upper and Lower groups were not significantly different ( $F = .303, p = .586$ ). For research question three, the groups were split up by L1 and looked at separately. The independent variables used were text difficulty, proficiency level, and text language.

## **Results**

It was found that there were significant results in Portuguese and Chinese relating to text difficulty as well as comparisons that could be made between the two languages in terms of first fixation duration, gaze duration, and total time. The results of this study will be presented by first looking at Portuguese and text difficulty, then looking at Chinese and text difficulty, and then looking at inter-language comparisons.

### **Portuguese Measures of Reading**

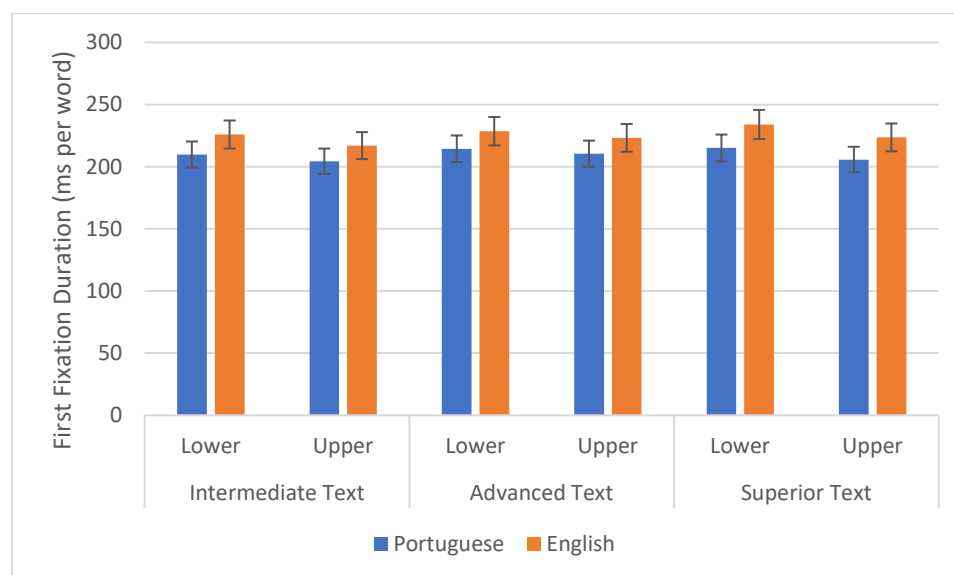
Portuguese ESL readers varied between text difficulties in terms of their early and late measures. The Advanced English text as read by Upper Proficiency readers served as the

baseline for the comparisons. This permitted comparison between higher and lower levels of the study.

### *Early Reading Measures*

#### First Fixation Duration

For the first fixation duration, Portuguese readers of both Upper and Lower proficiencies read in Portuguese significantly more quickly than they did in English. In addition, the AOIs in the Intermediate texts had a significantly lower first fixation duration than the AOIs in the Advanced texts or the AOIs in the Superior texts. This would indicate that the words were decoded more quickly at the Intermediate difficulty than at higher difficulties. However, there was no difference between the first fixation in the AOIs at the Advanced and Superior text levels. In addition, there was no difference between Upper and Lower proficiency groups. Refer to Figure 2 and Table 2 for more information.



*Figure 2. First fixation duration results for Portuguese ESL readers.*

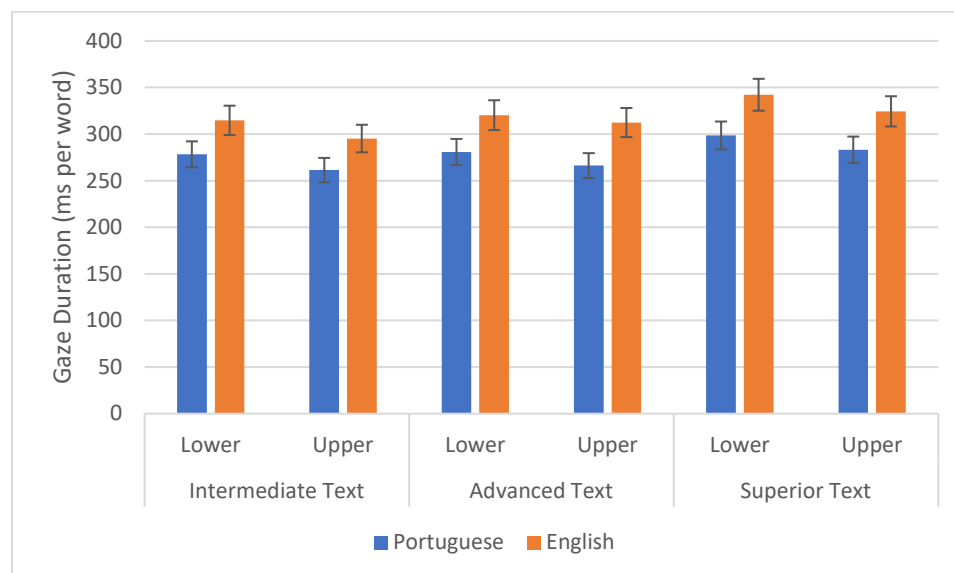
*Table 2. First Fixation Duration per word for Portuguese ESL readers*

	<i>b</i>	<i>SE</i>	<i>df</i>	<i>t</i> -value	<i>p</i> -value
Upper Prof. reading English Adv. Text (Baseline)	5.329	0.041	17.063	129.427	<.001*
Upper Prof. reading Portuguese Adv. Text	-0.054	0.009	17.031	-6.009	<.001*
Upper Prof. reading English Int. Text	-0.022	0.005	17.173	-4.714	<.001*
Upper Prof. reading English Sup. Text	0.004	0.008	16.861	0.459	0.652
Lower Prof. reading English Adv. Text	0.018	0.051	16.101	0.357	0.726

Note. \* = <.05

### Gaze Duration

For the second early reading measure gaze duration, the Portuguese readers continued the pattern of reading significantly more quickly in their native language than they did in English. Intermediate difficulty gaze duration was significantly faster than Advanced difficulty and Superior gaze duration was significantly slower than Advanced. Proficiency level was not significant. Refer to Figure 3 and Table 3 for more information.



*Figure 3. Gaze duration results for Portuguese ESL readers.*

*Table 3. Gaze Duration for Portuguese ESL readers.*

	<i>b</i>	<i>SE</i>	<i>df</i>	<i>t</i> -value	<i>p</i> -value
Upper Prof. reading English Adv. Text (Baseline)	5.57	0.05	22.91	119.14	<.001*
Upper Prof. reading Portuguese Adv. Text	-0.11	0.02	17.03	-6.20	<.001*
Upper Prof. reading English Int. Text	-0.02	0.01	18.12	-2.85	0.01*
Upper Prof. reading English Sup. Text	0.04	0.01	16.90	3.81	0.001*
Lower Prof. reading English Adv. Text	0.03	0.05	18.91	0.67	0.51

Note. \* = <.05

### *Late Reading Measure*

#### Total Time

For average reading time per word, the Portuguese readers were significantly faster on average in their native language than they were in English. They also were significantly slower at reading Intermediate texts and Superior Texts in English than Advanced texts, while in Portuguese, they read the Intermediate and Superior texts more quickly than Advanced Texts. More notably, there was a difference between the Upper and Lower groups. While reading the Superior texts, the Upper group spent more time on average integrating the word into the context than the Lower proficiency group did. They put more time into integrating the words into the larger context while reading the Intermediate text than they did the Advanced texts.

Portuguese readers were able to decode words and letters equally well at broad Upper and Lower proficiencies and they decoded them faster in Portuguese than they did in English. Decoding happened more quickly for texts at the Intermediate level than at Advanced or Superior levels of text difficulty. However, when it came to reading the whole word in context of the sentence, there tended to be a much more progressive increase in the difficulty between Intermediate, Advanced, and Superior. It was only at the point where the Portuguese readers were trying to integrate the words into the larger context that they began to interact less with the

Superior texts. They spent less time integrating the words in the Intermediate texts and Superior texts. Refer to Figure 4 and Table 4 for more information.

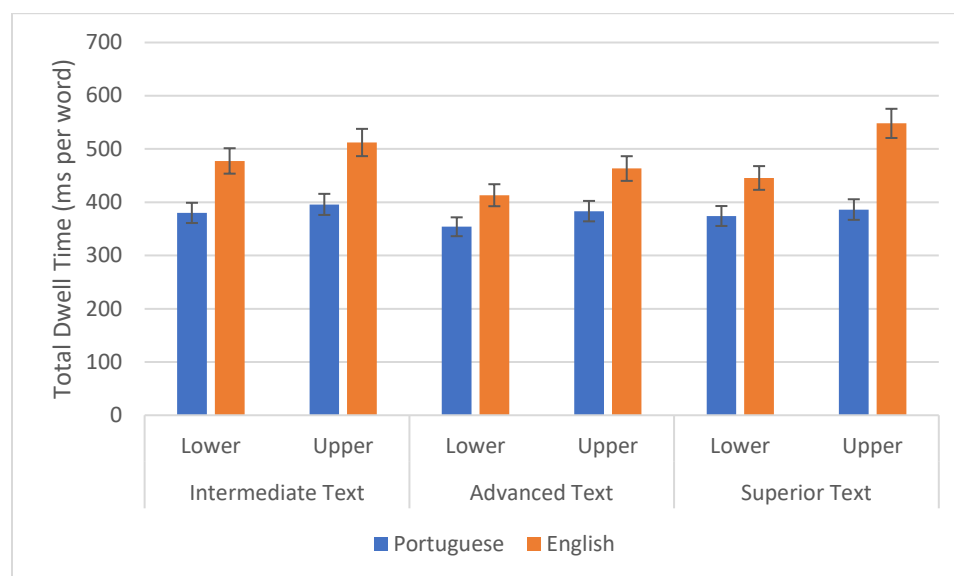


Figure 4. Total dwell time results for Portuguese ESL readers.

Table 4. Total Dwell Time for Portuguese ESL readers.

	<i>b</i>	<i>SE</i>	<i>df</i>	<i>t</i> -value	<i>p</i> -value
Upper Prof. reading English Adv. Text (Baseline)	5.898	0.052	16.180	112.866	<.001*
Upper Prof. reading Portuguese Adv. Text	-0.168	0.033	18.250	-5.044	<.001*
Upper Prof. reading English Int. Text	0.077	0.030	19.680	2.579	0.018*
Upper Prof. reading English Sup. Text	0.115	0.035	20.250	3.283	0.004*
Lower Prof. reading English Adv. Text	-0.076	0.067	16.180	-1.132	0.274
Upper Prof. reading Portuguese Int. Text	-0.046	0.019	55100.000	-2.445	0.014*
Upper Prof. reading Portuguese Sup. Text	-0.105	0.026	53790.000	-4.069	<.001*
Lower Prof. reading Portuguese Adv. Text	0.034	0.042	18.210	0.802	0.433
Lower Prof. reading English Int. Text	0.029	0.038	19.660	0.766	0.453
Lower Prof. reading English Sup. Text	-0.070	0.045	20.180	-1.559	0.135
Lower Prof. reading Portuguese Int. Text	-0.013	0.024	55100.000	-0.524	0.600
Lower Prof. reading Portuguese Sup. Text	0.098	0.032	54610.000	3.029	0.002*

Note. \* = <.05

## Chinese Measures of Reading

Similar to the Portuguese readers, the Chinese readers had the Upper proficiency group reading the Advanced English Texts used as the baseline for early and late measures of reading. This permitted comparison between the different text difficulty levels.

### *Early Measures of Reading*

#### First Fixation Duration

For first fixation duration, there was no significant difference between Chinese ESL readers reading English and Chinese. They did, however, spend significantly less time in their first fixations on the Superior texts in Chinese than they did in English, and the Lower proficiency group had a relatively smaller gap between English and Chinese texts in the Intermediate texts and the Superior texts than they did in the Advanced texts. Refer to Figure 5 and Table 5 for more information.

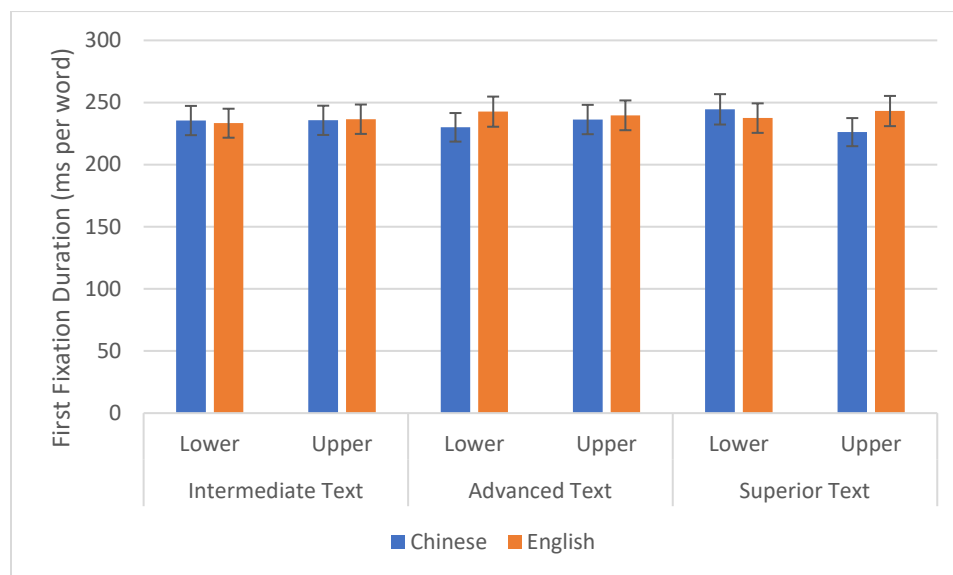


Figure 5. First fixation duration results for Chinese ESL readers.



*Table 5. First Fixation Duration for Chinese ESL readers.*

	<i>b</i>	<i>SE</i>	<i>df</i>	<i>t</i> -value	<i>p</i> -value
Upper Prof. reading English Adv. Text (Baseline)	5.388	0.039	14.120	139.024	<.001*
Upper Prof. reading Chinese Adv. Text	-0.021	0.026	15.760	-0.825	0.421
Upper Prof. reading English Int. Text	-0.007	0.012	22.620	-0.555	0.584
Upper Prof. reading English Sup. Text	0.021	0.013	28.150	1.635	0.113
Lower Prof. reading English Adv. Text	-0.007	0.052	14.080	-0.143	0.889
Upper Prof. reading Chinese Int. Text	0.002	0.014	42550.000	0.125	0.900
Upper Prof. reading Chinese Sup. Text	-0.061	0.019	42340.000	-3.293	<.001*
Lower Prof. reading Chinese Adv. Text	-0.028	0.035	15.680	-0.800	0.435
Lower Prof. reading English Int. Text	-0.015	0.016	21.710	-0.946	0.354
Lower Prof. reading English Sup. Text	-0.034	0.017	26.410	-2.004	0.055
Lower Prof. reading Chinese Int. Text	0.044	0.019	42490.000	2.341	0.019*
Lower Prof. reading Chinese Sup. Text	0.116	0.024	41500.000	4.778	<.001*

Note. \* = <.05

#### Gaze Duration

In gaze duration, Chinese ESL readers showed significantly more time spent on English texts than their L1 texts. In addition, English Intermediate texts were read significantly faster on the first pass overall compared to the Advanced texts, though Superior texts were given the same amount of time as the Advanced texts. However, when looking at Chinese texts, the Superior texts were read significantly faster in the gaze duration than the Advanced texts, while the Chinese Intermediate texts were not significantly different from the Advanced text gaze durations. Refer to Figure 6 and Table 6 for more information.

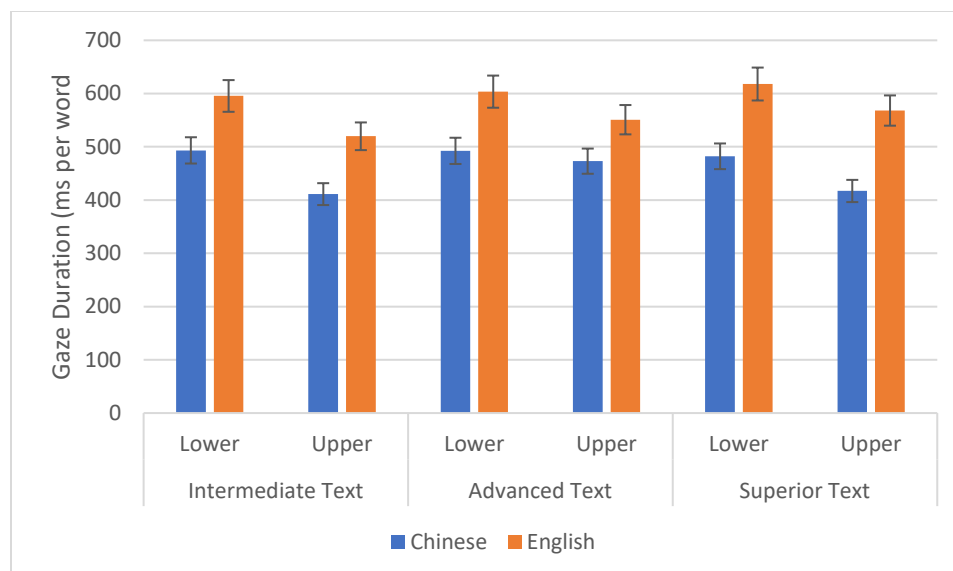


Figure 6. Gaze duration results for Chinese ESL readers.

Table 6. Gaze Duration for Chinese ESL readers.

	<i>b</i>	<i>SE</i>	<i>df</i>	<i>t</i> -value	<i>p</i> -value
Upper Prof. reading English Adv. Text (Baseline)	5.740	0.073	14.320	78.952	<.001*
Upper Prof. reading Chinese Adv. Text	-0.162	0.043	16.210	-3.761	0.002*
Upper Prof. reading English Int. Text	-0.036	0.017	42830.000	-2.138	0.033*
Upper Prof. reading English Sup. Text	0.016	0.022	42820.000	0.711	0.477
Lower Prof. reading English Adv. Text	0.056	0.097	14.290	0.581	0.570
Upper Prof. reading Chinese Int. Text	0.049	0.027	42840.000	1.810	0.070
Upper Prof. reading Chinese Sup. Text	-0.097	0.035	42820.000	-2.761	0.006*
Lower Prof. reading Chinese Adv. Text	-0.054	0.057	16.110	-0.935	0.364
Lower Prof. reading English Int. Text	0.032	0.022	42830.000	1.452	0.147
Lower Prof. reading English Sup. Text	-0.018	0.028	42820.000	-0.619	0.536
Lower Prof. reading Chinese Int. Text	0.037	0.036	42840.000	1.033	0.302
Lower Prof. reading Chinese Sup. Text	0.128	0.046	42830.000	2.782	0.005*

Note. \* = <.05

### Late Measures of Reading

#### Total Time

For the total time, Chinese texts were read significantly faster overall than English texts, though there was no difference between proficiency groups or between text difficulties. In the

Intermediate texts, there was a significantly larger difference between L1 and L2 than for other proficiencies, both for the Upper group and the Lower group. Refer to Figure 7 and Table 7 for more information.

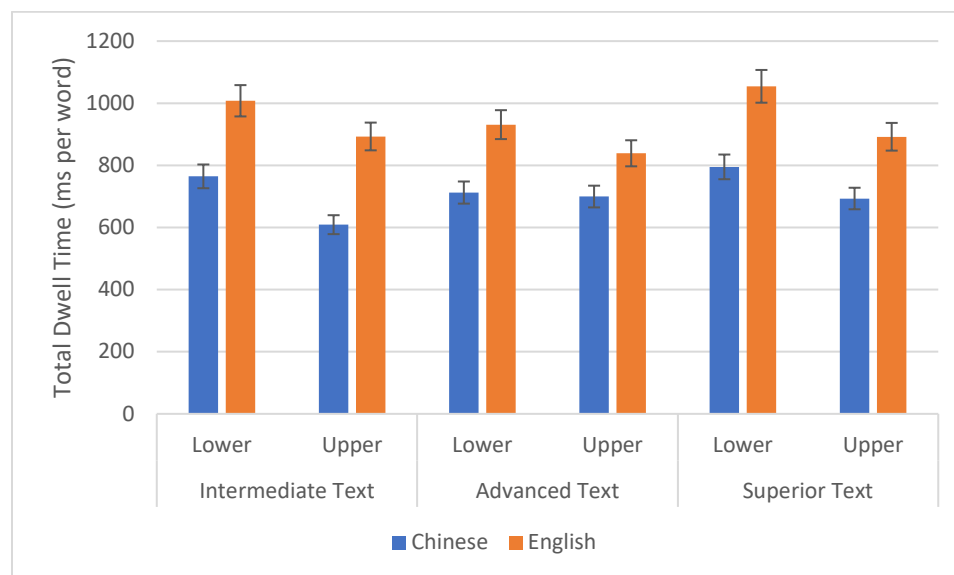


Figure 7. Total dwell time results for Chinese ESL readers.

Table 7. Total Dwell Time for Chinese ESL readers.

	<i>b</i>	<i>SE</i>	<i>df</i>	<i>t</i> -value	<i>p</i> -value
Upper Prof. reading English Adv. Text (Baseline)	6.102	0.053	14.400	114.888	<.001*
Upper Prof. reading Chinese Adv. Text	-0.218	0.050	16.620	-4.346	<.001*
Upper Prof. reading English Int. Text	0.049	0.031	20.120	1.596	0.126
Upper Prof. reading English Sup. Text	0.021	0.039	22.030	0.533	0.600
Lower Prof. reading English Adv. Text	0.053	0.071	14.320	0.754	0.463
Upper Prof. reading Chinese Int. Text	-0.066	0.032	42660.000	-2.085	0.037*
Upper Prof. reading Chinese Sup. Text	-0.070	0.041	42760.000	-1.683	0.092
Lower Prof. reading Chinese Adv. Text	-0.084	0.067	16.510	-1.262	0.225
Lower Prof. reading English Int. Text	0.034	0.041	19.510	0.844	0.409
Lower Prof. reading English Sup. Text	0.031	0.051	21.190	0.601	0.554
Lower Prof. reading Chinese Int. Text	0.130	0.042	42620.000	3.107	0.002*
Lower Prof. reading Chinese Sup. Text	0.044	0.054	42690.000	0.816	0.415

Note. \* = <.05

## Portuguese and Chinese

### Early Measures of Reading

There were also significant differences when comparing Portuguese readers to Chinese readers. When looking at first fixation duration, the only significant difference was that both groups spent more time in English than they did in their L1. There were not significant differences between Portuguese readers and Chinese readers, nor were there differences between Upper and Lower Proficiency groups. Refer to Figure 8 and Table 8 for more information.

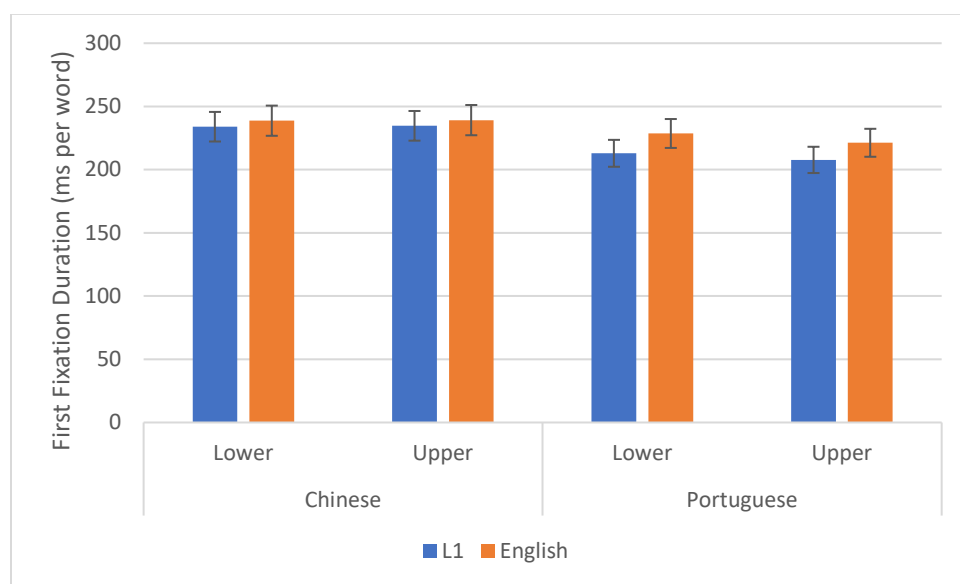


Figure 8. First fixation duration comparison between Chinese and Portuguese ESL readers.

Table 8. First Fixation comparison between Chinese and Portuguese ESL readers.

	<i>b</i>	<i>SE</i>	<i>df</i>	<i>t</i> -value	<i>p</i> -value
Upper Chinese reading English Text	5.378	0.034	31.281	160.183	<.001*
Upper Portuguese reading English Text	-0.053	0.036	31.001	-1.474	0.151
Upper Chinese reading Chinese Text	-0.042	0.009	33.088	-4.608	<.001*
Lower Chinese reading English Text	0.009	0.037	31.003	0.232	0.818

Note. \* = <.05

For gaze duration, fixations were significantly shorter in the L1 than in English, regardless of proficiency. Far more time was spent on Chinese L1 fixations than on Portuguese

L1 fixations and the gap between L1 and L2 for Chinese readers was significantly larger than the gap between L1 and L2 gaze duration for Portuguese readers. Refer to Figure 9 and Table 9 for more information.

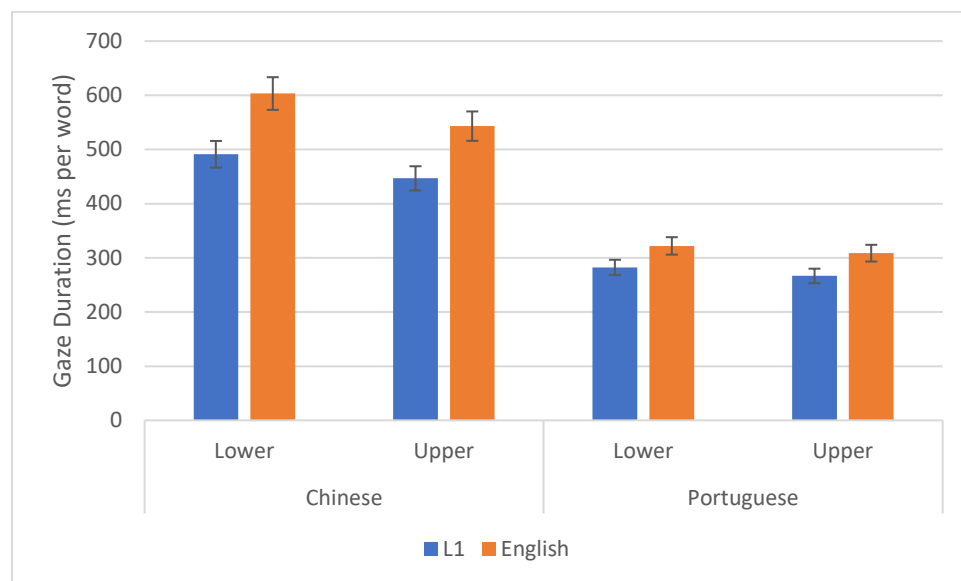


Figure 9. Gaze duration comparison between Chinese and Portuguese ESL readers.

Table 9. Gaze Duration comparison between Chinese and Portuguese ESL readers.

	<i>b</i>	<i>SE</i>	<i>df</i>	<i>t</i> -value	<i>p</i> -value
Upper Chinese reading English Text	5.738	0.053	33.552	108.216	<.001*
Upper Portuguese reading English Text	-0.185	0.059	31.270	-3.120	0.004*
Upper Chinese reading Chinese Text	-0.173	0.023	32.259	-7.614	<.001*
Lower Chinese reading English Text	0.052	0.055	31.013	0.947	0.351
Upper Portuguese reading Portuguese Text	0.066	0.031	31.806	2.120	0.042*

Note. \* = <.05

#### Late Measures of Reading

Total time was largely in line with gaze duration, with total time in the L1 being significantly shorter than in English, regardless of proficiency. As before, total time on Chinese L1 was significantly higher than time spent on Portuguese L1 fixations and the gap between L1 and L2 was larger for Chinese than for Portuguese. Refer to Figure 11 and Table 10 for more information.

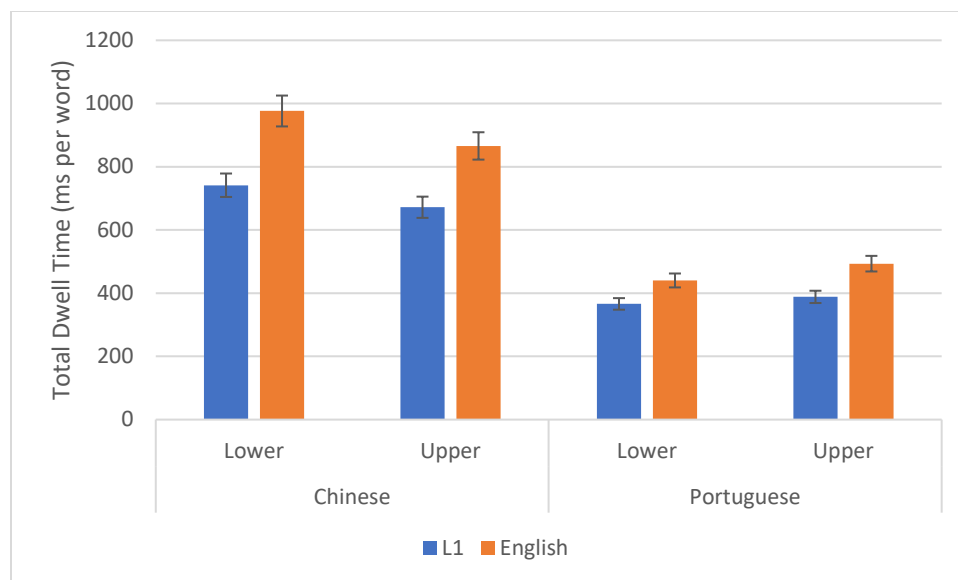


Figure 10. Total dwell time comparison between Chinese and Portuguese ESL readers.

Table 10. Total Dwell Time comparison between Chinese and Portuguese ESL readers.

	<i>b</i>	<i>SE</i>	<i>df</i>	<i>t</i> -value	<i>p</i> -value
Upper Chinese reading English Text	6.163	0.046	34.143	133.684	<.001*
Upper Portuguese reading English Text	-0.265	0.052	31.356	-5.103	<.001*
Upper Chinese reading Chinese Text	-0.269	0.026	32.732	-10.146	<.001*
Lower Chinese reading English Text	-0.003	0.047	31.023	-0.058	0.954
Upper Portuguese reading Portuguese Text	0.096	0.036	32.255	2.659	0.012*

Note. \* = <.05

While both Portuguese and Chinese ESL learners had similar word and letter recognition times, which took significantly longer in their L2 than it did in their L1, Chinese readers spent more time reading through the entire word the first time than Portuguese readers did. This was true both in the L1 and in the L2. Portuguese readers read both in their L1 and in English more quickly for the first pass than their Chinese counterparts did, with less of a gap between the L1 and L2. In addition, Portuguese ESL readers spent significantly less time integrating the word into the context of the sentence than their Chinese counterparts, both in their L1 and in their L2.

## **Discussion**

### **Portuguese Early and Late Measures of Reading**

Results for the Portuguese L1 ESL learners show that these readers processed texts in both languages in different ways at varying proficiency levels. When reading words in an intermediate text, they performed more quickly in all three measures compared to more Advanced texts. They spent significantly less time in both early and late measures of reading in the L1 than in the L2 (English). For the Superior texts they recognized the letters just as quickly as with Advanced texts, but spent more time reading the word in the first pass. In addition, by spending less time integrating words into the larger context in the Superior texts in the L1, Portuguese readers potentially demonstrated that they were more likely to give up on understanding the text at all.

### **Chinese Early and Late Measures of Reading**

The Chinese ESL readers had a much more complicated set of findings than the Portuguese ESL learners did. However, it is still clear that while there was no significant difference between L1 texts and English texts when it came to quick letter recognition, there was a much larger gap between how quickly the Chinese readers read words in Chinese and in English. Like with Portuguese, when reading words in English for the first time, intermediate texts were read more quickly than Advanced texts or Superior texts, regardless of proficiency level. Moreover, there was a larger gap between English and Chinese for the Lower proficiency group when accounting for the average of all fixations per word in the Intermediate texts relative to the others, potentially reflecting the fact that the Lower proficiency group focused more on texts that they felt they could answer, while the Upper proficiency group was more confident at those levels, and needed less time to understand the words in the text overall.

### **Interlanguage Comparison**

Overall, it was clear that ESL readers with Chinese and Portuguese L1s both read faster in their L1 than in English. In addition, Portuguese readers and Chinese readers read at different speeds in their L1. Because Taiwanese readers used the Traditional Chinese script and this experiment used the Simplified Chinese script, there was some concern that the fact that some of the participants had been born in Taiwan was influencing the result. However, an Independent Samples T-test revealed that there was no significant difference between Taiwanese and other Chinese readers (First fixation duration  $p = .704$ , Gaze duration  $p = .675$ , Total time  $p = .699$ ). Assessments of reading, such as the ACTFL, operate under the assumption that L1 readers are reading at functionally the same speed, but these results put that belief into question.

Another difference that was found between the two languages was that while Portuguese and Chinese learners had equivalent first fixation durations, the Chinese readers spent more time per word overall when reading the text, indicating that they struggled more with the comprehension of the words in context. One explanation for this could be that because Chinese has no spaces between words, more time to parse words beyond the first fixation duration could be required.

### **ESL Reading Proficiency**

While there were few major differences between Upper and Lower proficiency groups, this could be attributed to the large groupings that were used. The Upper and Lower groups each covered a large subset of participants at varying proficiency levels. While the four groups (Chinese Upper, Chinese Lower, Portuguese Upper, and Portuguese Lower) were equivalent between languages, that is, Chinese Upper was comparable to Portuguese Upper, the variance of the Chinese participants was greater in terms of reading proficiency than the variance of the Portuguese participants. In short, while this study did not find major differences between the



Upper and Lower proficiency levels, it is not necessarily certain that differences do not exist. In Grahl (2019) there was also no difference found among L2 learners, but in Rybakova (2019) Intermediate and Advanced readers were significantly different. It may be that the three metrics used as early and late measures, while useful in L1 studies, were not predictive of reading proficiency levels of L2 readers.

## **Conclusion**

### **Summary**

This study had the goal of investigating the reading behaviors of Portuguese and Chinese at various proficiency levels reading in both their native language and in English as a second language to determine what differences exist. In order to do this research as accurately as possible, eye-tracking was used to measure proxies of how they read and decoded words, how they comprehended words as they were read, and how they understood words in context. There were significant differences between Portuguese and Chinese in reading behaviors both in their L1 and in their L2. In addition, Portuguese and Chinese readers of English generally slowed the rate of reading as text difficulty increased.

### **Limitations**

The largest limitation of a study of this nature is in the number of participants. Data was collected from 34 participants overall. However, in order to measure the relevant data points, the groups were divided into several sub-groups. This made the largest difference in the proficiency levels of the sub-groups, as the participants needed to be divided into two equivalent groups. While the groups were equivalent, the large range of the groups made it difficult to state with confidence the value of the lack of findings.

## **Future Research**

This research presents a variety of potential future research options. One such option would be for other languages to be added to this study. While Portuguese and Chinese were used, other writing systems could provide valuable insight. A language with an alphabetic writing system that was less closely related to English would be especially helpful. A second application would be to use a larger population sample so that proficiency scores could be analyzed in more detail. A third option would be to investigate whether other languages vary in their L2 reading speeds. There was a difference between Chinese and Portuguese, but some of that difference was from some of the participants being less familiar with the form of written Chinese used. Looking at an L1 with fewer script complications would be helpful. Finally, the accuracy rate of the participants in the comprehension questions was not computed; they were only added to standardize the purpose for the reading. A future study should investigate how the rate of correct answers interacts with each of the independent variables of this study.

## References

- ACTFL. (2012). *ACTFL Proficiency Guidelines 2012*. White Plains, NY.
- Alderson, J. C. (2000). Assessing Reading. *Assessing Reading*.  
<https://doi.org/10.1017/cbo9780511732935>
- Ashby, J., Rayner, K., & Clifton, C. (2005). Eye movements of highly skilled and average readers: Differential effects of frequency and predictability. *Quarterly Journal of Experimental Psychology Section A: Human Experimental Psychology*, 58(6), 1065–1086.  
<https://doi.org/10.1080/02724980443000476>
- Bates, D., Maechler, M., Bolker, B., Walker, S., Christensen, R. H. B., & Singmann, H. (2016). lme4: Linear mixed-effects models using Eigen and S4. 2014. *R package version, 1*(4).
- Bax, S., & Chan, S. (2019). Using eye-tracking research to investigate language test validity and design. *System*, 83, 64–78. <https://doi.org/10.1016/j.system.2019.01.007>
- Bernhardt, E. (2011). *Understanding Advanced Second-Language Reading*. Routledge, Taylor and Francis Group.
- Berzak, Y., Nakamura, C., Flynn, S., & Katz, B. (2017). Predicting native language from gaze. *ACL 2017 - 55th Annual Meeting of the Association for Computational Linguistics, Proceedings of the Conference (Long Papers), 1*, 541–551.  
<https://doi.org/10.18653/v1/P17-1050>
- Boldt, R. F., Larsen-Freeman, D., Reed, M. S., & Courtney, R. G. (1992). *Distribution of ACTFL ratings by TOEFL score ranges* (Vol. 41). Princeton.
- Brown, J. D., & Hudson, T. (2002). *Criterion Referenced Language Testing*. Cambridge University Press.

- Clifford, R., & Cox, T. L. (2013). Empirical validation of reading proficiency guidelines. *Foreign Language Annals*, 46(1), 45–61. <https://doi.org/10.1111/flan.12033>
- Cohen, A. D., & Upton, T. A. (2007). “I want to go back to the text”: Response strategies on the reading subtest of the new TOEFL®. *Language Testing*, 24(2), 209–250. <https://doi.org/10.1177/0265532207076364>
- Cop, U., Drieghe, D., & Duyck, W. (2015). Eye movement patterns in natural reading: A comparison of monolingual and bilingual reading of a novel. *PLoS ONE*, 10(8), 1–39. <https://doi.org/10.1371/journal.pone.0134008>
- Godfroid, A., & Spino, L. A. (2015). Reconceptualizing reactivity of think-alouds and eye tracking: Absence of evidence is not evidence of absence. *Language Learning*, 65(4), 896–928. <https://doi.org/10.1111/lang.12136>
- Grahl. (2019). *L1 and L2 Reading Behaviors by Proficiency Level. An English-Portuguese eye-tracking study* (Master’s Thesis).
- Grabe, W. (2009). *Reading in a second language: Moving from theory to practice*.
- Inhoff, A. W., & Liu, W. (1997). *The perceptual span during the reading of Chinese text. Cognitive processing of Chinese and related Asian languages*. 243–266.
- Inhoff, A. W., & Liu, W. (1998). The Perceptual Span and Oculomotor Activity during the Reading of Chinese Sentences. *Journal of Experimental Psychology: Human Perception and Performance*, 24(1), 20–34. <https://doi.org/10.1037/0096-1523.24.1.20>
- Johns, A. M. (1981). Necessary English: A faculty survey. *TESOL Quarterly*, 15(1), 51–57. <https://doi.org/10.2307/3586373>
- Krstić, K., Šoškić, A., Ković, V., & Holmqvist, K. (2018). All good readers are the same, but every low-skilled reader is different: an eye-tracking study using PISA data. *European*

*Journal of Psychology of Education*, 33(3), 521–541. <https://doi.org/10.1007/s10212-018-0382-0>

Li, H., & Suen, H. K. (2015). How do Chinese ESL learners recognize English words during a reading test? A comparison with romance-language-speaking ESL learners. *International Multilingual Research Journal*, 9(2), 93–107.

<https://doi.org/10.1080/19313152.2014.995013>

Paris, S. G., & Stahl, S. A. (2005). *Children's Reading Comprehension and Assessment*. Routledge.

Prichard, C., & Atkins, A. (2018). L2 readers' global processing and selective attention: An eye tracking study. *TESOL Quarterly*, 52(2), 445–456. <https://doi.org/10.1002/tesq.423>

Rayner, K. (2009). Eye movements and attention in reading, scene perception, and visual search. In *Quarterly Journal of Experimental Psychology* (Vol. 62).

<https://doi.org/10.1080/17470210902816461>

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin*, 124(3), 372–422. <https://doi.org/10.1037/0033-2909.124.3.372>

Rayner, K., Ashby, J., Pollatsek, A., & Reichle, E. D. (2004). The effects of frequency and predictability on eye fixations in reading: Implications for the E-Z reader model. *Journal of Experimental Psychology: Human Perception and Performance*, 30(4), 720–732.

<https://doi.org/10.1037/0096-1523.30.4.720>

Rayner, K., Sereno, S. C., Morris, R. K., Rene Schmauder, A. R., & Clifton, C. (1989). Eye movements and on-line language comprehension processes. *Language and Cognitive Processes*, 4(3–4), SI21–SI49. <https://doi.org/10.1080/01690968908406362>

- Reichle, E. D., Warren, T., & McConnell, K. (2009). Using E-Z reader to model the effects of higher level language processing on eye movements during reading. *Psychonomic Bulletin and Review*, *16*(1), 1–21. <https://doi.org/10.3758/PBR.16.1.1>
- Roberts, L., & Siyanova-Chanturia, A. (2013). Using eye-tracking to investigate topics in L2 acquisition and L2 processing. *Studies in Second Language Acquisition*, *35*(2), 213–235. <https://doi.org/10.1017/S0272263112000861>
- Rupp, A. A., Ferne, T., & Choi, H. (2006). How assessing reading comprehension with multiple-choice questions shapes the construct: A cognitive processing perspective. *Language Testing*, *23*(4), 441–474. <https://doi.org/10.1191/0265532206lt337oa>
- Rybakova. (2019). *Measuring L2 (Russian) reading proficiency across various levels using eye tracking* (Master's Thesis).
- Whitford, V., & Titone, D. (2012). Second-language experience modulates first- and second-language word frequency effects: Evidence from eye movement measures of natural paragraph reading. *Psychonomic Bulletin and Review*, *19*(1), 73–80. <https://doi.org/10.3758/s13423-011-0179-5>
- Whitford, V., & Titone, D. (2015). Second-language experience modulates eye movements during first- and second-language sentence reading: Evidence from a gaze-contingent moving window paradigm. *Journal of Experimental Psychology: Learning Memory and Cognition*, *41*(4), 1118–1129. <https://doi.org/10.1037/xlm0000093>
- Yan, G., Tian, H., Bai, X., & Rayner, K. (2006). The effect of word and character frequency on the eye movements of Chinese readers. *British Journal of Psychology*, *97*(2), 259–268. <https://doi.org/10.1348/000712605X70066>

*Appendix A*

**ICAMRA: Mapping the Path to Advanced Second Language Literacy in Adults Using Eye-Tracking** You are invited to participate in a research project sponsored by the Interdisciplinary Cohort for the Assessment of Multilingual Reading Ability (ICAMRA). This study investigates characteristics that distinguish L2 readers of English and how those characteristics differ across those languages in various proficiency levels through the use of eye-tracking.

Do you agree to take the background survey to see if you are eligible to participate?

- Yes (1)
- No (2)
- 

**Eye Health**

The following questions will ask you about your eye health to see if current eye-tracking technology will work with any conditions you might have.

Please select any that apply.

- I have glaucoma (4)
- I have cataracts (5)
- I have eye implants (6)
- I have permanently dilated pupils (7)
- I need assistive technology (screen reader, magnifier, etc.) to read a computer screen (9)
- I need contacts/glasses to read a computer screen (8)
- None apply-I can read a computer screen without glasses or contacts. (10)
-

Sorry, since you have glaucoma, you are not eligible to participate in this study. We thank you for your time.

---

Sorry, since you have cataracts, you are not eligible to participate in this study. We thank you for your time.

---

Sorry, since you have eye implants, you are not eligible to participate in this study. We thank you for your time.

---

Sorry, since your pupils are permanently dilated, you are not eligible to participate in this study. We thank you for your time.

---

Sorry, since you need assistive technology to use the computer and the Web, you are not eligible to participate in this study. We thank you for your time.

---

Are your glasses/contacts \_\_\_\_\_?

- single lens for reading (1)
  - single lens for distance (2)
  - no line bifocals (4)
  - lined bifocals/trifocals, layered lens, or regression lens (3)
- 

These glasses/contacts will not work with current eye-tracking technology. Do you have a pair of single lens glasses/contacts you can wear for the study?

- Yes (4)
  - No (5)
-



Can you read a computer screen and the Web **without difficulty**?

Yes (1)

No (2)

---

Sorry, since you can't read a computer screen and the Web without difficulty with your contacts and/or eyeglasses on, you are not eligible to participate in this study. We thank you for your time.

---

Page Break



How old are you?

▼ 17 or younger (1) ... 60 or older (44)

---

Sorry, since you are not at least 18 years old, you are not eligible to participate in this study. We thank you for your time.

---

Page Break

What is your given (first) name?

---

What is your surname (last name)?

---

How many years of formal education do you have?

- Less than High School (1)
- High school graduate (2)
- Some college (3)
- 2 year degree (4)
- 4 year degree (5)
- Professional degree (6)
- Doctorate (7)



Choose one of the following:

- Male (1)
- Female (2)
- Prefer not to answer (3)

In which country were you born?

▼ Please select below... (1) ... Other (195)

Please indicate when you'd be available for the next step of the study.

	Mornings (8am-10am) (1)	Midmorning (10am-12pm) (2)	Early afternoon (12pm-2pm) (3)	Late afternoon (2pm-4pm) (4)	Evening (4pm-7pm) (5)
Monday (1)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Tuesday (2)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Wednesday (3)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Thursday (4)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Friday (5)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Saturday (6)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

How can we contact you to set up an appointment? (Select all that apply)

- Email (1) \_\_\_\_\_
- Text Message/Phone Call (2) \_\_\_\_\_
- Other (4) \_\_\_\_\_

End of Block: GeneralInfoBlock

Start of Block: LanguageListBlock

**The following questions will ask you about the languages you speak.**

What language(s) do you speak or read?

Select all the languages speak natively or non-natively and/or have studied either formally or informally to any degree.

- English (4)
- ASL (16)
- Arabic (1)
- Chinese (Mandarin) (2)
- Chinese (Cantonese) (26)
- Dutch (13)
- Finnish (18)
- French (5)
- German (6)
- Haitian Creole (21)
- Hindi (22)
- Italian (7)
- Japanese (8)
- Korean (9)
- Norwegian (19)
- Persian (25)
- Polish (23)
- Portuguese (10)
- Russian (11)
- Spanish (12)

- Swedish (17)
- Tagalog (14)
- Urdu (24)
- Vietnamese (20)
- Other:** (15) \_\_\_\_\_



Please rank the languages you know **in order of dominance (1= Most Dominant/Strongest):**  
**Drag and Drop**

Dominance
_____ English (x4)
_____ ASL (x16)
_____ Arabic (x1)
_____ Chinese (Mandarin) (x2)
_____ Chinese (Cantonese) (x26)
_____ Dutch (x13)
_____ Finnish (x18)
_____ French (x5)
_____ German (x6)
_____ Haitian Creole (x21)
_____ Hindi (x22)
_____ Italian (x7)
_____ Japanese (x8)
_____ Korean (x9)
_____ Norwegian (x19)
_____ Persian (x25)
_____ Polish (x23)

\_\_\_\_\_ Portuguese (x10)  
\_\_\_\_\_ Russian (x11)  
\_\_\_\_\_ Spanish (x12)  
\_\_\_\_\_ Swedish (x17)  
\_\_\_\_\_ Tagalog (x14)  
\_\_\_\_\_ Urdu (x24)  
\_\_\_\_\_ Vietnamese (x20)  
\_\_\_\_\_ **Other:** (x15)

---



How old were you when you started to learn the language?

English (xx4)	▼ 0 (1) ... 25 or older (26)
ASL (xx16)	▼ 0 (1) ... 25 or older (26)
Arabic (xx1)	▼ 0 (1) ... 25 or older (26)
Chinese (Mandarin) (xx2)	▼ 0 (1) ... 25 or older (26)
Chinese (Cantonese) (xx26)	▼ 0 (1) ... 25 or older (26)
Dutch (xx13)	▼ 0 (1) ... 25 or older (26)
Finnish (xx18)	▼ 0 (1) ... 25 or older (26)
French (xx5)	▼ 0 (1) ... 25 or older (26)
German (xx6)	▼ 0 (1) ... 25 or older (26)
Haitian Creole (xx21)	▼ 0 (1) ... 25 or older (26)
Hindi (xx22)	▼ 0 (1) ... 25 or older (26)
Italian (xx7)	▼ 0 (1) ... 25 or older (26)
Japanese (xx8)	▼ 0 (1) ... 25 or older (26)
Korean (xx9)	▼ 0 (1) ... 25 or older (26)
Norwegian (xx19)	▼ 0 (1) ... 25 or older (26)
Persian (xx25)	▼ 0 (1) ... 25 or older (26)
Polish (xx23)	▼ 0 (1) ... 25 or older (26)
Portuguese (xx10)	▼ 0 (1) ... 25 or older (26)
Russian (xx11)	▼ 0 (1) ... 25 or older (26)
Spanish (xx12)	▼ 0 (1) ... 25 or older (26)
Swedish (xx17)	▼ 0 (1) ... 25 or older (26)
Tagalog (xx14)	▼ 0 (1) ... 25 or older (26)
Urdu (xx24)	▼ 0 (1) ... 25 or older (26)
Vietnamese (xx20)	▼ 0 (1) ... 25 or older (26)
<b>Other:</b> (xx15)	▼ 0 (1) ... 25 or older (26)



What do you consider to be your native language(s)?

- English (1)
- ASL (2)
- Arabic (3)
- Chinese (Mandarin) (4)
- Chinese (Cantonese) (5)
- Dutch (6)
- Finnish (7)
- French (8)
- German (9)
- Haitian Creole (10)
- Hindi (11)
- Italian (12)
- Japanese (13)
- Korean (14)
- Norwegian (15)
- Persian (16)
- Polish (17)
- Portuguese (18)
- Russian (19)
- Spanish (20)
- Swedish (21)

- Tagalog (22)
- Urdu (23)
- Vietnamese (24)
- Other:** (25) \_\_\_\_\_

End of Block: LanguageListBlock

Start of Block: LanguageInfo

**The following question(s) will gather more information about how you learned  $\{\text{lm://Field/1}\}$ .**

How did you learn  $\{\text{lm://Field/1}\}$ ? Check all that apply.

- Informally (native language, from family, living abroad, friends, independent study, etc.) (4)
- School (3)
- Mission (1)
- Study Abroad (2)
- Other, please specify how and for how long: (8)
- \_\_\_\_\_

### Informal Language Learning

Please list the years and months you spent in each INFORMAL  $\{\text{lm://Field/1}\}$  language environment.

	Years

A family/home in which $\${lm://Field/1}$ is spoken. (11)	▼ 0 to 6 months (1 ... 11 or more (8))
A country where $\${lm://Field/1}$ is spoken. (2)	▼ 0 to 6 months (1 ... 11 or more (8))
A work environment in which $\${lm://Field/1}$ is spoken. (3)	▼ 0 to 6 months (1 ... 11 or more (8))
With friends who spoke $\${lm://Field/1}$ (12)	▼ 0 to 6 months (1 ... 11 or more (8))
Through the independent study of $\${lm://Field/1}$ (13)	▼ 0 to 6 months (1 ... 11 or more (8))
Other, specify: (10)	▼ 0 to 6 months (1 ... 11 or more (8))

Please provide any other comments on how you learned  $\${lm://Field/1}$  if the previous questions did not cover it:

---



---



---



---



---

### School (formal education, language classes, etc.):

Please list the years and months you spent in each  $\${lm://Field/1}$  language learning environment.

	Years	Months

Elementary school (2)	▼ 0 (1 ... 6 or more (7))	▼ 0 (1 ... 12 (13))
Middle School/Junior High (3)	▼ 0 (1 ... 6 or more (7))	▼ 0 (1 ... 12 (13))
Highschool (4)	▼ 0 (1 ... 6 or more (7))	▼ 0 (1 ... 12 (13))
Undergraduate (6)	▼ 0 (1 ... 6 or more (7))	▼ 0 (1 ... 12 (13))
Graduate (7)	▼ 0 (1 ... 6 or more (7))	▼ 0 (1 ... 12 (13))
Postgraduate (8)	▼ 0 (1 ... 6 or more (7))	▼ 0 (1 ... 12 (13))
Other, specify: (9)	▼ 0 (1 ... 6 or more (7))	▼ 0 (1 ... 12 (13))
Other, specify: (10)	▼ 0 (1 ... 6 or more (7))	▼ 0 (1 ... 12 (13))

---

Please provide comments on your study of  $\{lm://Field/1\}$  in schools (if any).

---



---



---



---



---



---

Page Break

**Mission information:**

---

Where did you serve your  $\${lm://Field/1}$ -speaking mission?

---

How many months did you speak  $\${lm://Field/1}$  on your mission?

0 2 5 7 10 12 14 17 19 22 24

Months ()	
-----------	--

---

Please provide comments on your mission in  $\${lm://Field/1}$  (if any).

---

---

---

---

---

---

Page Break

**Study abroad**

---

Which country did you go to study  $\${lm://Field/1}$ ?

---

How long did you stay in the country where you studied  $\${lm://Field/1}$ ?

Term (Spring/Summer) (1)

Semester (Fall/Winter) (2)

Other (3) \_\_\_\_\_

---

Please provide comments on your study abroad in  $\${lm://Field/1}$  (if any).

---

---

---

---

---

What is your level of speaking proficiency in  $\{\text{lm://Field/1}\}$ ?

- 10—I can interact and negotiate to resolve abstract and complex matters on a wide variety of topics in a variety of situations with NO PATTERN of linguistic errors. (10)
  - 9 (9)
  - 8—I can hold complex conversations and tell detailed stories in the past, present and future time frames accurately with both my pronunciation and grammar. (12)
  - 7 (7)
  - 6 (6)
  - 5—I can hold simple conversations through asking and answering questions related to my daily life. (5)
  - 4 (4)
  - 3 (3)
  - 2—I can produce memorized words & phrases and formulaic utterances about personal topics. (2)
  - 1—Very Little (1)
- 

What is your level of reading proficiency in  $\{\text{lm://Field/1}\}$ ?

- 10—I can follow the flow of ideas and infer meaning from unfamiliar, abstract topics in complex language in informational texts. (10)
- 9 (9)
- 8—I can identify the underlying message and MOST supporting details across major time frames in texts. (12)
- 7 (7)
- 6 (6)
- 5—I can understand the main idea and key information short straightforward texts. (5)
- 4 (4)
- 3 (3)
- 2—I can recognize letters and understand lists of words & phrases by general topic. (2)
- 1—Very Little (1)

---

Page Break

End of Block: LanguageInfo

---