

EXPLORING THE INTERFACE OF EXPLICIT AND IMPLICIT SECOND-LANGUAGE  
KNOWLEDGE: A LONGITUDINAL PERSPECTIVE

By

MinHye Kim

A DISSERTATION

Submitted to  
Michigan State University  
in partial fulfillment of the requirements  
for the degree of

Second Language Studies – Doctor of Philosophy

2020

## ABSTRACT

### EXPLORING THE INTERFACE OF EXPLICIT AND IMPLICIT SECOND-LANGUAGE KNOWLEDGE: A LONGITUDINAL PERSPECTIVE

By

MinHye Kim

International students make up 5.5% of the student body in the US, but little is known about their language development and use. The goals of this study were twofold. First, I aimed to systematically track the amount and types of language use reported by international students immersed in the target language environment. By doing so, I hoped to provide answers to questions such as *what types of language skills receive the most (least) attention during daily life as an international student? How much individual variation exists in English use among international students in the US context?* Second, I examined the longitudinal associations between two types of second language (L2) knowledge (i.e., explicit knowledge and implicit knowledge) and their association to activities types that invite different types of processing (i.e., language-focused and meaning-focused). The exploration of the knowledge-knowledge associations and the processing-knowledge associations will inform our understanding of the interface question, which concerns how awareness of linguistic form may impact L2 learning (e.g., DeKeyser, 2007; N. Ellis, 2002, 2003, 2005; Hulstijn, 2002; Krashen, 1985; Paradis, 2009).

One hundred and twenty-two L2 English learners completed five linguistic tests that measured their explicit and implicit knowledge of L2 English at two-timepoints (T1: January–February 2019, T2: April–May 2019). The untimed written grammaticality judgment test (GJT) and metalinguistic knowledge test served as measures of explicit L2 English knowledge. The timed written GJT, oral production, and elicited imitation were administered as implicit L2

English knowledge measures. To track language engagement, participants completed self-reported language exposure logs on five days over the course of one semester.

Using a combination of confirmatory factor analysis and path analysis, I observed that explicit and implicit knowledge are in a reciprocal relationship that affected each other bi-directionally; that is, explicit knowledge at Time 1 was causally related to implicit knowledge at Time 2; and reversely, implicit knowledge at Time 1 played a facilitative role in the development of explicit knowledge at Time 2. Neither of the activity types predicted knowledge development. In addition, data on authentic language usage showed that international students are more engaged (quantitatively, in terms of hours per day spent) with L2 English than other languages. To be precise, they spend 2.2x more time using English than other languages. I also observed qualitative differences in English engagement: While students spent a comparable amount of time speaking, listening, and reading in English, they spent significantly less time writing in English. Lastly, at the individual-level, students showed wide-ranging variability in the amount and types of language engagement they reported.

The findings of this dissertation suggest that first, language acquisition is a developmental process composed of a dynamic interaction between explicit and implicit knowledge and their synergetic relationship; and second, similar affordances to engage in the L2 do not produce comparable amounts of actual L2 engagement for different individuals. These observations may reinforce that the explicit-implicit interface question, and language acquisition more generally, may be better understood when studied over time in a naturalistic context, as language acquisition in its essence is shaped by one's experience with the language in interaction with the contextual affordances in the environment.

Copyright by  
MINHYE KIM  
2020

## ACKNOWLEDGEMENTS

I am deeply indebted to many people for their guidance and help in various stages of my PhD journey. This journey was truly a wonderful one with so many eye-opening moments and inspiring interactions that have influenced me and this dissertation in the most positive way. The following is my humble attempt to address my gratitude.

First, this dissertation would not have been possible if not with the financial support of the following grants: the National Science Foundation (NSF) Doctoral Dissertation Improvement Grant, the International Research Foundation for English Language Education (TIRF) Doctoral Dissertation Grant, a National Federation of Modern Language Teachers Associations (NFMLTA) Dissertation Support Grant, and several internal funds from MSU: a Dissertation Completion Fellowship (DCF) from the College of Arts and Letters and from the Second Language Studies program. A special thank you goes to Drs. Bill Hart-Davidson and Shawn Loewen for finding ways to financially support different parts of this large-scale longitudinal study. The completion of this dissertation would have been very difficult without their help.

I am grateful to my advisor Dr. Aline Godfroid. Words cannot express the amount of support and guidance I have received from her. Among other things, Aline has introduced me to the world of psycholinguistics and made me a better researcher (and a big fan of the Faculty Diversity program!). I have many great memories of our time together in Michigan and I can only hope to provide similar type of experiences to my students. I also wish to express my gratitude to Dr. Shawn Loewen for our productive conversations about research and academia. Shawn has continuously supported my interdisciplinary endeavors, conducting collaborative projects using novel methodological approaches. He believed in me and supported me, both

academically and emotionally, in the best way possible. I am lucky to have Aline and Shawn as my mentors and my go-to people for insights and guidance.

I hope to thank the other members of my committee: Drs. Paula Winke, Sue Gass, Nick Ellis and Yuichi Suzuki. I undoubtedly had a dream team of SLA experts. I am especially appreciative of Paula's guidance. My interactions with Paula have always been so pleasant, informative, and very sincere. She showed me what a great mentor is like and as I promised to Paula, I will pay the support she has shown me forward to my future students. I also extend my thanks to Drs. Gass, N. Ellis, and Suzuki for agreeing to be on board as committee members of this project. Their feedback and comments were insightful, sharp, and constructive which fine-tuned and improved the quality of this study to a great degree.

Many others have helped me along the way. Endless thanks go to all my participants, who remained with me and this study for one entire year! Our interaction vividly reminded of why I pursue the work I do: to help adult learners acquire languages. I also thank SLS students and faculty for the opportunity to experience a supportive and collaborative working environment. Special thanks go to Dustin Crowther and Dan Isbell for being good friends and supportive colleagues.

Last and certainly not least, I am forever indebted to my family for their unconditional support and love. My family was always available to help me persevere challenging moments. Thanks especially to my lovely mother and father, for all the opportunities they have provided which shaped me into the person I am today. I love you both very much. Also my heartfelt gratitude and sincere appreciation go to my beloved uncle, Dr. Kyung-Sung Kim—my lifetime role model and a true educator. I couldn't have done this without you. I dedicate this wonderful achievement to my family.

## TABLE OF CONTENTS

|  |    |
|--|----|
| LIST OF TABLES.....  | ix |
| LIST OF FIGURES .....  | xi |
| CHAPTER 1 : REVIEW OF THE LITERATURE .....                                 | 2  |
| 1.1    Interface of L2 knowledge and processing: Theoretical stances ..... | 2  |
| 1.1.1    Synthesis .....   | 5  |
| 1.2    Interface of L2 knowledge and processing: Empirical evidence .....  | 5  |
| 1.2.1    Influence of L2 Instruction on L2 processing .....                | 6  |
| 1.2.2    Influence of L2 instruction on L2 knowledge.....                  | 7  |
| 1.2.3    Influence of L2 knowledge on L2 knowledge .....                   | 8  |
| 1.2.4    Synthesis .....   | 9  |
| 1.3    Measures of L2 Explicit and Implicit knowledge.....                 | 10 |
| 1.4    Language use in a naturalistic setting .....                        | 13 |
| 1.5    Measuring L2 usage in a naturalistic setting.....                   | 15 |
| CHAPTER 2 : CURRENT STUDY .....  | 17 |
| 2.1    Goals of this study .....   | 17 |
| 2.2    Participants.....   | 20 |
| 2.3    Materials .....   | 23 |
| 2.3.1    Target structures.....  | 23 |
| 2.3.2    Instruments.....  | 24 |
| 2.3.3    Background questionnaire .....                                    | 25 |
| 2.3.4    Motivation questionnaire .....                                    | 25 |
| 2.3.5    Oral production (OP) .....  | 26 |
| 2.3.6    Elicited Imitation (EI).....                                      | 30 |
| 2.3.7    Grammaticality Judgment Tests (GJTs) .....                        | 32 |
| 2.3.8    Metalinguistic Knowledge Test (MKT).....                          | 33 |
| 2.3.9    Language Exposure Log .....                                       | 34 |
| 2.4    Procedure .....   | 35 |
| 2.4.1    Knowledge Measures.....   | 35 |
| 2.4.2    Language Exposure Measures .....                                  | 37 |
| CHAPTER 3 : LANGUAGE USE .....   | 39 |
| 3.1    Research Questions .....  | 39 |
| 3.2    Data Preparation Details .....                                      | 40 |
| 3.3    Results.....  | 41 |
| 3.3.1    Types and Amount of Language Engagement.....                      | 41 |
| 3.3.2    Language Engagement Over Time .....                               | 45 |
| 3.3.3    Variation among learners.....                                     | 49 |

|  |     |
|--|-----|
| CHAPTER 4 : LINGUISTIC KNOWLEDGE AND LANGUAGE USE.....                   | 52  |
| 4.1 Research Questions.....  | 52  |
| 4.2 Analysis.....  | 52  |
| 4.3 Data Preparation.....  | 55  |
| 4.3.1 Missing data.....  | 55  |
| 4.3.2 Item reliability.....  | 58  |
| 4.4 Results.....   | 60  |
| 4.4.1 Descriptive Statistics for Language Tests.....                     | 61  |
| 4.4.2 Summary of Descriptive Results .....                               | 71  |
| 4.4.3 Correlations.....  | 72  |
| 4.4.4 Factor Scores.....   | 75  |
| 4.4.5 Path Analysis .....  | 83  |
| 4.4.6 Summary of Results.....  | 92  |
| CHAPTER 5 : DISCUSSION.....  | 93  |
| 5.1 Quantitative and Qualitative Differences in Language Engagement..... | 94  |
| 5.2 Explicit and Implicit Knowledge and Activity Types .....             | 96  |
| 5.2.1 Explicit-Implicit Interface.....                                   | 96  |
| 5.2.2 Implicit-Explicit Interface.....                                   | 99  |
| 5.2.3 Activity Types.....  | 100 |
| 5.3 Implications.....  | 102 |
| 5.4 Limitations and Future Directions .....                              | 104 |
| 5.5 Conclusions.....   | 108 |
| APPENDICES.....  | 109 |
| APPENDIX A. Background Questionnaire (T1) .....                          | 110 |
| APPENDIX B. Motivation Questionnaire.....                                | 112 |
| APPENDIX C. Background Questionnaire (T2) .....                          | 113 |
| APPENDIX D. Stimuli.....   | 117 |
| APPENDIX E. Language Exposure Log.....                                   | 125 |
| APPENDIX F. Instructions on the Web-Based Testing Program.....           | 127 |
| APPENDIX G. Recruitment Flyer.....                                       | 135 |
| APPENDIX H. Oral Production/EI – Coding Guidelines.....                  | 136 |
| APPENDIX I. Metalinguistic Knowledge Test: Scoring Guide.....            | 137 |
| REFERENCES .....   | 139 |



## LIST OF TABLES

|  |    |
|--|----|
| Table 2.1 Proficiency Breakdown.....   | 21 |
| Table 2.2 Background Information of The L2 Speakers .....                                | 22 |
| Table 2.3 Six Sentence Structures and Examples.....                                      | 24 |
| Table 2.4 Summary of Measures .....  | 26 |
| Table 2.5 Timeline and Test Sequence .....   | 37 |
| Table 3.1 Time Per Day (21 Hours Max) on Specific Activities.....                        | 42 |
| Table 3.2 Time Per Day (21 Hours Max) Spent on Language Focus.....                       | 44 |
| Table 4.1 Missing Data of 149 Participants in T1 And T2 .....                            | 56 |
| Table 4.2 Coefficient Alpha and Coefficient Omega .....                                  | 60 |
| Table 4.3 Descriptive Information of Five Linguistic Tests at T1 and T2 .....            | 61 |
| Table 4.4 Descriptive Statistics for the Elicited Imitation at T1 and T2 .....           | 63 |
| Table 4.5 Descriptive Statistics for the Oral Production at T1 and T2.....               | 65 |
| Table 4.6 Descriptive Statistics for the Timed Written GJT at T1 and T2.....             | 67 |
| Table 4.7 Descriptive Statistics for the Untimed Written GJT at T1 and T2 .....          | 69 |
| Table 4.8 Descriptive Statistics for Metalinguistic Knowledge Test at T1 and<br>T2 ..... | 71 |
| Table 4.9 Summary of the Descriptive Results .....                                       | 71 |
| Table 4.10 Correlational Matrix for the Five Tests at T1 and Activity Types .....        | 73 |
| Table 4.11 Correlational Matrix for the Five Tests at T2 and Activity Types .....        | 73 |

|  |     |
|--|-----|
| Table 4.12 CFA Model Fit Indices .....   | 78  |
| Table 4.13 Two-Factor Model Parameter Estimates for T1 and T2 .....                          | 80  |
| Table 4.14 Factor Scores of Explicit and Implicit Knowledge at T1 and T2 .....               | 81  |
| Table 4.15 Correlational Matrix for the Knowledge and Activity Types.....                    | 82  |
| Table 4.16 Model Fit Indices for the Interface and Non-interface Models .....                | 85  |
| Table 4.17 Model Fit Indices of the Reciprocal Interface and Reverse<br>Interface Model..... | 88  |
| Table 4.18 Model Parameter Estimates for the Reciprocal Interface Model.....                 | 89  |
| Table 5.1 Oral Production/EI – Coding Guidelines.....  | 136 |

## LIST OF FIGURES

|  |    |
|--|----|
| Figure 2.1. The Interface model (left) and the Non-interface model (right) used to test three main research hypotheses. Note. H = Hypothesis; T1 = Time 1; T2 = Time 2.....  | 20 |
| Figure 2.2. Screen shot of a part of the web-based instructions in the oral production task.....   | 27 |
| Figure 2.3. Screen shot of a part of the web-based instructions in the oral production task.....   | 28 |
| Figure 2.4. Screen shot of the oral production task during retelling with picture prompts and a progress indicator .....   | 29 |
| Figure 2.5. Screen shot of a part of the web-based instructions in the elicited imitation task.....  | 30 |
| Figure 3.1. Types and amount of language engagement during a 24-hour day (N = 154).....  | 42 |
| Figure 3.2. Mean frequency counts of language engagement data across time for reading, listening, speaking, and writing in English; using other languages; and no language use. Boxes demarcate time frames with relatively high amounts of reported engagement..... | 47 |
| Figure 3.3. Mean frequency counts of language engagement data across time for reading, listening, speaking, and writing in English. Boxes demarcate timeframes with relatively high amount of reported language engagement.....                                      | 48 |
| Figure 3.4. English engagement patterns for each hour from 6 a.m. to 3 a.m. by individuals.....  | 51 |
| Figure 4.1. Path diagram for a two-wave, two-variable path model. T1 = time 1; T2 = time 2; a & b = autoregressive paths; c & d = cross-lagged paths .....   | 53 |
| Figure 4.2. (left panel) Proportion of missing values across two time points; (right panel) Combinations of missing data patterns. Numbers on the right indicate participants who fall in a given category. Red-filled squares represent missing values .....        | 58 |
| Figure 4.3. Spaghetti plot of Elicited Imitation at T1 and T2.....   | 62 |

|   |     |
|---|-----|
| Figure 4.4. Spaghetti plot of Oral Production at T1 and T2 .....  | 64  |
| Figure 4.5. Spaghetti plot of Timed Written GJT at T1 and T2 .....  | 66  |
| Figure 4.6. Spaghetti plot of Untimed Written GJT at T1 and T2.....   | 68  |
| Figure 4.7. Spaghetti plot of Metalinguistic Knowledge Test (Explanation) at<br>T1 and T2.....  | 70  |
| Figure 4.8. Relationships among five linguistic scores at T1 and activity types.....  | 74  |
| Figure 4.9. Relationships among five linguistic scores at T2 and activity types.....  | 74  |
| Figure 4.10. Two-factor model at T1 .....   | 76  |
| Figure 4.11. Two-factor model at T2.....  | 77  |
| Figure 4.12. (top-left) The Interface Model; (top-right) The Non-interface<br>Model; (bottom-left) The Reciprocal Interface Model; (bottom-right) The<br>Reverse Interface Model. The top two are a <i>priori</i> path models constructed<br>before the analysis. The bottom two are a <i>posteriori</i> path models. Paths a and<br>b represent the explicit-implicit interface and the implicit-explicit interface,<br>respectively ..... | 84  |
| Figure 4.13. (left) The Interface Model <i>without</i> the Implicit-Explicit interface<br>path; (right) The Reciprocal Interface Model <i>with</i> the Implicit-Explicit<br>interface path.....   | 90  |
| Figure 5.1. Screenshot 1.....   | 127 |
| Figure 5.2. Screenshot 2.....   | 128 |
| Figure 5.3. Screenshot 3.....   | 129 |
| Figure 5.4. Screenshot 4.....   | 130 |
| Figure 5.5. Screenshot 5.....   | 130 |
| Figure 5.6. Screenshot 6.....   | 131 |
| Figure 5.7. Screenshot 7.....   | 132 |

|                                    |     |
|------------------------------------|-----|
| Figure 5.8. Screenshot 8.....      | 134 |
| Figure 5.9. Recruitment Flyer..... | 135 |

## INTRODUCTION

A core issue in second language acquisition (SLA) is how second language (L2) knowledge types (i.e., explicit and implicit) and L2 processing (i.e., form-focused and meaning-focused) relate. Do form-focused processing and explicit knowledge facilitate the development of implicit knowledge? This question is known as the *interface* issue and has been the object of extensive theorization (e.g., DeKeyser, 2007; N. Ellis, 2005; R. Ellis, 2009; Hulstijn, 2002; Krashen, 1985; McLaughlin, 1987; Paradis, 2009).

There are good reasons as to why understanding the relationship between L2 knowledge and processing is important. Empirical and theoretical progress in this area will shed light on how languages can best be learned and taught. It will also advance researchers' understanding of the cognitive basis of L2 acquisition and to what extent this basis differs from, or is similar to, first language (L1) acquisition. For instance, to what degree can L2 learners acquire implicit knowledge of a L2 through meaning-focused processing thought to guide L1 acquisition? Does form-focused processing facilitate the development of L2 implicit knowledge? What types of L2 teaching method promote learning processes that are conducive to implicit knowledge? These are just some of the questions that can benefit from the current research.

Language acquisition is highly context dependent and is shaped by one's experience with the language. As such, it is difficult to understand the explicit-implicit interface, and language acquisition in general, without considering the types and amount of linguistic input that the learners experience. The secondary goal of this study, therefore, was to track the amount and types of language use of international students in a US higher education context and thereby test the common belief that the relationship between what a context offers and what learners gain from the learning situation is complex and diverse.

## CHAPTER 1: REVIEW OF THE LITERATURE

### 1.1 Interface of L2 knowledge and processing: Theoretical stances

Two types of knowledge are generally believed to shape language users' performance. *Implicit knowledge* is unconscious knowledge of linguistic regularities that exists outside of one's awareness. *Explicit knowledge* is conscious-verbalizable linguistic knowledge of which language users are aware that they know it. These types of knowledge may be end-products of different processing activities (Leow, 2015); for instance, learners who focus more on the linguistic aspects of a message (i.e., *form-focused processing*) or its meaning (i.e., *meaning-focused processing*) may build different knowledge bases of the language. Sparked by the learning and acquisition distinction (Krashen, 1981, 1985), three hypotheses have been put forth on the relationship between different types of knowledge (i.e., explicit and implicit) and processes (i.e., form-focused and meaning-focused). Together, the non-interface, the weak interface, and the strong interface position constitute the interface hypothesis.

Proponents of *the non-interface position* maintain a dual system view of explicit and implicit knowledge. They assume there is a qualitative distinction between two types of knowledge representations. The argument is that explicit and implicit knowledge are developed through independent routes in a cognitive system and thus one type of knowledge does not become the other (Hulstijn, 2002, 2007; Krashen, 1981, 1985; Paradis, 1994, 2004, 2009). The non-interface point of view is most strongly associated with Krashen (1981, 1985) who distinguished learned (explicit) knowledge from acquired (implicit) knowledge. He claimed that learned knowledge and explicit instruction are helpful for monitoring and fine-tuning accuracy but have a limited role in acquisition; as such, learners with learned (explicit) knowledge may fail to use such knowledge in daily use, for which they mainly draw upon acquired (implicit)

knowledge (Krashen, 1985). In Krashen's view, when the desire is to develop L2 acquired knowledge, the types of processing and the amount of input matter. In particular, learners require ample amounts of comprehensible input, for which the focus is on understanding the meaning not on grammar (Krashen, 1981). This idea reinforces the strong link between two types of processing (i.e., form-focused or meaning-focused) and the respective knowledge representations (i.e., explicit and implicit knowledge) in the interface debate.

The non-interface of explicit and implicit knowledge is also upheld by Paradis (1994, 2004, 2009) and Hulstijn (2002, 2007, 2015). From a neurobiological perspective, explicit and implicit knowledge involve different types of representation and are substantiated in different parts of the brain; and in this sense, the two independent routes to knowledge accumulation do not interact, let alone transform one type of knowledge into another. Importantly, both scholars (Hulstijn, 2002, 2007, 2015; Paradis, 1994, 2004, 2009), and also to a limited extent Krashen, recognize the influence of explicit knowledge on implicit knowledge. They, however, maintain that this influence is indirect; that is, explicit knowledge does not interface or make contact with implicit competence but guides learners to practice the constructions that contain grammatical regularities. In this way, the repeated use of instances, driven by explicit knowledge, indirectly influences the establishment of implicit knowledge (Hulstijn, 2015; Paradis, 2009). As Hulstijn (2015) summarized, "a non-interface position in the neurophysiological sense is by no means at variance with the practice-makes-perfect maxim" (p. 36).

An emphasis on processing types in relation to knowledge types is also evident in the *strong interface position*. Building on Anderson's Skill Acquisition Theory (SAT, 1982, 1993), the strong interface position holds that explicit L2 knowledge, as other cognitive skills, becomes proceduralized, and eventually automatized with deliberate practice. In this view, the initial



establishment of explicit knowledge is essential, and form-focused processing promotes the automatization of L2 explicit knowledge (DeKeyser, 2007). In essence, the strong interface proponents claim a causal, or even a synergetic, relationship between explicit and implicit knowledge. That is, not only does explicit knowledge influence implicit knowledge, but also certain linguistic rules are more conducive to one type of processing than the other (DeKeyser, 2015, also see Williams, 2009). From this perspective, then, SAT does not reject a role for implicit learning or its importance; rather, the focus of the theory is on L2 learning that is more practical to teaching or manipulation: “SAT... focuses on how explicit learning (which is often the only realistic possibility for specific learning problems because of time constraints or logistic issues) can, ..., lead to knowledge that is functionally equivalent to implicit knowledge.” (2015, p. 247). In accordance with the non-interface camp, DeKeyser notes, explicit knowledge does not transform into implicit knowledge; rather a repeated use of one memory system enables “a gradual establishment of another memory system” (DeKeyser, 2017, p. 19). As such, more explicit knowledge does not imply less implicit knowledge nor does explicit knowledge magically convert into implicit knowledge (e.g., DeKeyser, 2009, 2014, 2017).

Between these two views are the proponents of *the weak-interface position* (e.g., N. Ellis, 2002, 2003, 2005, 2015; R. Ellis, 1990, 2008). N. Ellis’ stance for the interface question is well documented in the Associative-Cognitive CREED (N. Ellis, 2006): Language acquisition is a continuation of frequency-driven, statistical tallying of various patterns in the input. The more exposure to the patterns (higher frequency), the stronger the connections between constructions. Importantly, this adaptive fine-tuning occurs subconsciously without awareness, but explicit knowledge plays a role when implicit learning is hindered by a lack salience of the linguistic forms or prior knowledge blocking unfamiliar cues. As such, form-focused learning/instruction

enables the initial registration of L2 grammar, and creation of this conscious channel paves the way for effective meaning-focused processing, which may itself lead to the development of L2 implicit knowledge. Along with other scholars, N. Ellis maintained that explicit and implicit knowledge develop independently, and one does not convert into another (N. Ellis & Larsen-Freeman, 2006).

### **1.1.1 *Synthesis***

The review of these positions on the explicit-implicit interface provides important lessons to understand the development of L2 knowledge. Theorists across three interface views concur with two tenets of the interface issue. First, the two types of knowledge do not transform into another but co-operate in parallel fashion. As noted, the term *interface* is an “unfortunate appellation” (N. Ellis & Larsen-Freeman, 2006, p. 569) that leads to a common misunderstanding of the influence of explicit knowledge on implicit knowledge. Second, most researchers recognize the influence of explicit knowledge/processing on implicit knowledge development: the degree of its impact and underlying cognitive mechanisms may vary by theoretical views, but explicit practice generates implicit learning opportunities (Hulstijn, 2002; Krashen & Terrell, 1983; Paradis, 2009; DeKeyser, 2017) and explicit registration of linguistic forms allows implicit fine-tuning (N. Ellis, 2002, 2005).

## **1.2 Interface of L2 knowledge and processing: Empirical evidence**

In spite of the theoretical and practical significance of the interface debate, it can be challenging to identify directly relevant empirical work (DeKeyser, 2017). This confusion may result in part from conceptual ambiguity (e.g., conceptual overlap across supposedly dichotomous terms) and in part from methodological shortcomings (e.g., longitudinal nature of

the interface question; validity of L2 knowledge measures). Nevertheless, three main themes emerge from empirical studies pertaining to the interface hypothesis: (i) how L2 instruction types affect online L2 processing; (ii) how L2 instruction types influence L2 knowledge development; and (iii) how L2 explicit knowledge affects the development of other L2 knowledge types. These three themes combined may shed light on the trajectory of the interface issue (instruction → processes → knowledge).

### ***1.2.1 Influence of L2 Instruction on L2 processing***

With the advancement of new technologies, a number of researchers (e.g., Andringa & Curcic, 2015; Cintrón-Valentín & Ellis, 2016; Curcic, Andringa, & Kuiken, 2019; Hopp, 2013, 2016) have utilized eye tracking to closely examine learners' online processing patterns and understand how these patterns relate to instruction types. Evidence supporting the interface hypothesis comes from studies that show learners' online processing patterns alter as a function of explicit focus-on-form instruction. A case in point are the findings from Cintrón-Valentín and Ellis (2016). The authors examined how different types of form-focused instruction affected learners' attentional foci during L2 processing. Native English speakers received one of three types of form-focused instruction on verb inflections in Latin. They were then asked to interpret and produce the verbs' temporal reference (past, present, future encoded on the verb and/or adverb). The authors hypothesized the instructional intervention would induce a focus on the verb cues and away from the adverb cues, which also marked temporal reference. Given that native English speakers tend to attend more to adverbial cues, a change in attentional processing along with evidence of learning could be regarded as evidence for an interface (e.g., N. Ellis, 2005). The results of the eye-movement data revealed that all three treatment groups attended

more to the verb cues than the uninstructed control group, who gradually focused less on the verb cues over the course of the study.

The findings of Andringa and Curcic (2015), on the other hand, did not lend support to a knowledge interface. Using a visual world eye-tracking paradigm, the authors examined whether learners instructed in the target grammar (i.e., differential object marking) exploited the gained knowledge during an online processing task that required them to use the grammar rule predictively. Contrary to the authors' expectation, instruction did not play a role. Participants who were given the rule of differential object marking could not utilize the knowledge during online processing. Their eye-movement behavior aligned with the group that had not received explicit training on the rules. While this finding could support a non-interface view, it is equally possible that the grammar training was too brief for learners to form metalinguistic knowledge before the online processing task. Indeed, in their follow-up study, Curcic, Andringa, and Kuiken (2019) found that learners who reported to have gained awareness—in particular, awareness of a rule that determiners predict corresponding nouns patterns—showed higher levels of predictive processing than those without awareness of the determiner-noun patterns. A comparison of the two studies suggests that explicit instruction successfully needs to produce a well-entrenched explicit knowledge of the target structure in order for predictive processing to occur (Curcic, Andringa, & Kuiken, 2019; Hopp, 2013, 2016).

### ***1.2.2 Influence of L2 instruction on L2 knowledge***

Research in instructed SLA focuses on different types of instruction (i.e., form-focused and meaning-focused) and the effects thereof on L2 development. Instructed SLA research can inform the interface hypothesis, albeit indirectly, because the differential effects of instruction are often attributed to differences in cognitive processes (observed or assumed). Seen in this

light, findings from instructed SLA can illuminate the role of L2 processing types (i.e., form-focused and meaning-focused processing) in L2 knowledge development.

Three meta-analyses have synthesized findings of 92 studies on different instructional types (e.g., Goo, Granena, Yilmaz, & Novella, 2015; Norris & Ortega, 2000; Spada & Tomita, 2010). Spada and Tomita's (2010) meta-analysis, for instance, synthesized 30 empirical studies, 10 of that were also in Norris and Ortega's meta-analysis, that used tasks that induced learners to focus on form (e.g., controlled constructed tasks) or meaning (e.g., free constructed tasks). The authors found large effect sizes for form-focused instruction on free-response measures. Similarly, Goo et al. (2015) reported a large effect size for the effects of explicit instruction on learners' free production measures ( $g = 1.443$ )—a finding consistent with those of Norris and Ortega (2000). These findings may collectively point to the possibility of an explicit-implicit interface. Specifically, processing strategies induced by form-focused instruction may lead to L2 implicit knowledge gains. Yet, for this conclusion to be valid, the free-response measures ought to be valid measures of L2 implicit knowledge. Given the synthetic nature of meta-analysis, this question remains open but it highlights some of the difficulties in testing the wide scope of the interface hypothesis empirically.

### ***1.2.3 Influence of L2 knowledge on L2 knowledge***

Suzuki and DeKeyser's (2017) research is most closely related to the present dissertation. Using a battery of six linguistic knowledge tests and three aptitude measures, these authors examined the explicit-implicit knowledge interface, specifically whether automatized explicit knowledge contributes to the acquisition of implicit knowledge. To do so, they ran two structural equation models on their L2 participants' test scores and aptitude scores, termed *interface* model and *noninterface* model. The models were identical, except for a path that extended from

automatized explicit knowledge to implicit knowledge in the interface model. The two models did not differ significantly in terms of model fit; however, on a descriptive level, the interface model fit somewhat better (numerically higher or lower fit indices in the hypothesized direction). Moreover, in the interface model, the path from automatized explicit knowledge to implicit knowledge was positive and significant. Therefore, the results of this study provided suggestive evidence for an interface, which would need to be corroborated statistically in a future study. Such a study would ideally be longitudinal in nature, rather than cross-sectional, to demonstrate a true causal relationship between the different knowledge types in L2 development.

#### **1.2.4 *Synthesis***

The studies reviewed in previous sections represent important first steps towards testing the interface hypothesis directly. At the same time, these studies also point to some future research directions, which I take into account in the present study:

(1) Longitudinal designs are critical. While Suzuki and DeKeyser's (2017) results alluded to a knowledge interface, their cross-sectional design did not capture L2 learners' learning trajectory, but rather their long-term attainment. To account for the developmental aspects of L2 acquisition, which are inherent to the interface hypothesis (DeKeyser, 2017), longitudinal research will be essential.

(2) No studies have directly observed the developmental changes in knowledge types and their interface in a naturalistic L2 setting. Unlike laboratory experiments with artificial or extinct languages (Andringa & Curcic, 2015; Cintrón-Valentín & Ellis, 2016; Curcic, Andringa, & Kuiken, 2019), immersion in a naturalistic setting represents an authentic L2 learning context in which students use the L2 as a part of their daily lives. Researchers are thus able to observe the

development and/or the interface between two linguistic systems, as independent L2 learners naturally immerse themselves in the host country and are exposed to different types of L2 input.

(3) Lastly, reliable and valid measures of explicit and implicit knowledge are key to addressing any questions related to the different types of knowledge. In Suzuki and DeKeyser (2017), the three implicit knowledge measures (visual-world eye tracking, a word-monitoring task, and self-paced reading) showed a weak convergence, with factor loadings varying considerably in strength ( $.17 < r < .69$ ). This highlights methodological concerns in the measurement of implicit and explicit knowledge; a question that has generated a substantial amount of validation research in recent years.

### **1.3 Measures of L2 Explicit and Implicit knowledge**

The theoretical importance of the explicit-implicit dichotomy gave rise to an era of test validation research, starting with R. Ellis's landmark study (R. Ellis, 2005), which aimed at finding valid measures of implicit and explicit knowledge. In his study, R. Ellis put forth a set of descriptors for L2 implicit and explicit knowledge, based on which he proposed a battery of five linguistic knowledge measures. Results from a principal component analysis indicated oral production (OP), elicited imitation (EI), and a timed written grammaticality judgment test (TGJT) loaded onto one factor, which he termed implicit knowledge, whereas the untimed written grammaticality judgment test (UGJT) and metalinguistic knowledge test (MKT) loaded onto a different factor, which he labeled explicit knowledge (also see R. Ellis & Loewen, 2007).

In response to R. Ellis's (2005) psychometric study, a series of validation studies was carried out to ascertain the validity of explicit and implicit knowledge measures. For over a decade now, researchers have (1) administered a battery of tests in different contexts (i.e., Bowles, 2011; Gutiérrez, 2013; Suzuki & DeKeyser, 2015; Zhang, 2015), (2) utilized online

measures to measure implicit knowledge (i.e., Godfroid, Loewen, Jung, Park, Gass, & Ellis, 2015; Suzuki & DeKeyser, 2015, 2017; Vafae, Suzuki, & Kachinske, 2016), and (3) manipulated task features in grammaticality judgment tests (i.e., Godfroid et al., 2015; Gutiérrez, 2013; Kim & Nam, 2016; Spada, Shiu, & Tomita, 2015; Vafae et al., 2016). Thus far, the results from factor analyses were relatively consistent for certain tests: OP and online measures (i.e., WMT and SPR) likely measure implicit knowledge; the UGJT and MKT likely measure explicit knowledge. The construct validity of EI and TGJT, however, remains disputed.

Contrary to R. Ellis's initial proposal, Suzuki and DeKeyser (2015, 2017) argued that implicit knowledge measures ought to focus on meaning (as opposed to form) during online (real-time) language processing. As such, EI and TGJT may measure automatized explicit knowledge because they are time-pressured (rather than real-time) measures and, in the case of TGJT, the focus is on form (not meaning). In an effort to resolve these contradictory findings, Godfroid and Kim (under review) brought nine previously used explicit and implicit knowledge measures together in an empirical synthesis of extant validation research. The nine tests included: the word monitoring test (WMT), self-paced reading (SPR), elicited imitation (EI), oral production (OP), timed/untimed grammaticality judgment tests (GJTs) in the aural and written modes, and the metalinguistic knowledge test (MKT).

With the data from 151 non-native English speakers, we performed a series of confirmatory factor analyses, extending both R. Ellis' (2005) and Suzuki and DeKeyser's (2017) models with additional tasks previous validation studies have used. We found that both a three-factor model (Suzuki & DeKeyser, 2015, 2017) and a two-factor model (R. Ellis, 2005) provided a good fit for our data and that the two models did not differ significantly,  $X_{\text{dif}} = 1.44$ ,  $df = 1$ ,  $p = 0.23$ . This suggests that both two-factor and three-factor models may account equally well for L2



users' linguistic knowledge. Of test validation research, this study is the first to examine and find a full-fledged three-factor model of explicit, automatized explicit, and implicit knowledge measures.

In a further attempt to validate the underlying constructs of linguistic knowledge, Godfroid and Kim (under review) examined the predictive validity of explicit and implicit learning aptitudes for knowledge types. We regressed a battery of four implicit learning aptitude tests (auditory statistical learning, visual statistical learning, serial reaction time, and Tower of London) along with one explicit learning test (MLAT V) onto the nine linguistic measures mentioned above. Structural equation modeling revealed that only the alternating serial reaction time task significantly predicted performance on timed language tests (R. Ellis, 2005; R. Ellis & Loewen, 2007), but not on the reaction-time measures (Suzuki & DeKeyser, 2015; Vafaei, Suzuki, & Kachinske, 2017). These results lend support to the view that timed language tests are better measures of implicit knowledge (R. Ellis, 2005; R. Ellis & Loewen, 2007), and that reaction time measures may not be superior for measuring L2 implicit knowledge (Suzuki & DeKeyser, 2015; Vafaei, Suzuki, & Kachinske, 2017).

At the same time, it is worth pointing out that, in the measurement models (the confirmatory factor analyses), the factor loadings of the reaction time measures (WMT and SPR) and timed/untimed aural GJT were relatively low ( $.08 < r < .48$ ). This suggests that the association between latent constructs (i.e., explicit, auto-explicit, implicit knowledge) and the tests (i.e., reaction-time tasks for implicit knowledge and the timed/untimed aural GJT for auto-explicit/explicit knowledge) is relatively weak. Given the weak convergence of four tasks (WMT, SPR, TAGJT, and UnAGJT) on their corresponding latent constructs, I decided to remove them and use the remaining five most robust tasks to measure implicit and explicit

knowledge. As such, measures of explicit knowledge used in this project include the untimed written GJT and MKT and implicit knowledge measures include two oral production measures (i.e., OP and EI) and the timed written GJT.

#### **1.4 Language use in a naturalistic setting**

Linguistic repertoires are shaped by one's experience with the language. As such, it is difficult to understand the explicit-implicit interface, or language acquisition in general, without considering the types and amount of linguistic input that the learners experience. A learner's L2 experience, in turn, is closely related to language learning contexts and its affordances.

Many educators subscribe to the belief that L2 learners immersed in a naturalistic setting—where L2 is abundant in quantity and diverse in quality (Ranta & Meckelborg, 2013)—are more likely to show linguistic gains than L2 learners studying in a foreign context. This anecdotal observation has a theoretical tenet. A naturalistic learning environment provides learners with rich and authentic L2 input that keeps the focus on meaning (e.g., Krashen, 1985); creates ample opportunities to interact with interlocutors to produce output (e.g., Swain, 1985); to negotiate for meaning (e.g., Gass & Mackey, 2007), and to stimulate learners to notice the gap (e.g., Gass, 1997; Long, 1996). From a usage-based perspective, an immersion abroad exposes learners to high numbers of linguistic forms or items; and various agents (individuals engaged in communications) and configurations (groups, networks, and culture) all shape linguistic repertoires (e.g., N. Ellis, 2007; N. Ellis & Larsen-Freeman, 2006). SLA researchers, therefore, have many reasons to assume that immersion into authentic L2 learning environment creates an optimal condition for L2 development.

Given the potential benefits of immersed L2 learning, researchers have explored the effects of study abroad on L2 gains. Study abroad (SA) is broadly defined as “an academic

experience that allows students to complete part of their degree program through educational activities outside their country” (Sanz & Morales-Front, 2018, p. 1). Going one step beyond study abroad, immersion abroad (as used in this study) refers to an academic experience that allows students to complete their entire degree at a foreign university. Empirically, the effectiveness of SA on L2 development is found to be inconclusive (for a detailed review, see Isabelli-García, Bown, Plews, & Dewey, 2018; for a recent edited volume, see Sanz & Morales-Front, 2018). The mixed findings from previous studies, alongside methodological pitfalls of oversimplifying SA as a single construct, point to the importance of considering SA as multi-dimensional; that is, understanding that individuals’ L2 learning experience varies in a number of features, such as L2 exposure, social contact, and cultural experiences, and that language acquisition is shaped by a dynamic interaction of these factors (de Bot, Lowie, & Verspoor, 2007).

In recognition of the context-dependent nature of L2 acquisition, in the current project, I track the amount of L2 engagement of international students—those who pursue degree programs outside of their home country—both at a group-level and individual-level. The findings from the individual-level data, in particular, may shed light on whether similar contextual affordances to engage in the L2 produce comparable amounts of actual L2 engagement for different individuals. Furthermore, different types of L2 engagement, that is receptive (reading and listening) or productive (speaking and writing), will also be reported. This is in response to claims that production-based and receptive-based practice generate different knowledge basis and different skill sets. In particular, production-based and receptive-based practice/instruction benefit production skills and receptive skills, respectively (DeKeyser, 2007; Lightbown, 2008). As such, if the goal of L2 learning is to converse meaningfully, learners ought

to produce the L2. Data on linguistic skill usage will provide a first step towards exploring the association between linguistic skill types and L2 development.

### **1.5 Measuring L2 usage in a naturalistic setting**

A precursor to addressing questions on language use is a reliable measurement to assess individuals' immersion experience. One widely used measure of language contact/use is Freed, Dewey, Segalowitz and Halter's Language Contact Profile (LCP, 2004). The questionnaire consists of two parts: a pretest version used to estimate L2 usage at the beginning of SA and a posttest version provided at the end of a SA project. The frequency scale prompts participants to retrospect on an entire SA period to estimate "days per week" and "hours per day" on various tasks such as listening to tv, writing academic papers, reading novels etc. Using the LCP, a number of studies have documented the actual L2 contact and use in a SA context: some have reported (1) the paucity of L2 use, during SA, with comparable amount of L1 and L2 usage (e.g., Dewey, Belnap, & Hillstrom, 2013; Freed, Segalowitz, & Dewey, 2004) and (2) the amounts of L2 engagement in a SA context was significantly less than that in the intensive immersion program at home (e.g., Freed et al., 2004). On the contrary, Ranta and Meckelborg (2013) documented the superior use of L2, compared to L1, with Chinese L2 graduate students studying in Canada. A number of distinct features could have contributed to the inconsistent findings. First, the learner populations are different. Participants in Dewey et al. (2013) and Freed et al. (2004) were college exchange students who temporarily enrolled in classes for a semester or two, whereas participants in Ranta and Meckelborg (2013) were composed of degree-seeking graduate students who engaged in the L2 quantitatively more than the exchange students. Also, Dewey et al. (2013) and Freed et al. (2004) used the LCP in a retrospective manner, that is after a period of study abroad, while Ranta and Meckelborg (2013) employed a questionnaire similar to

LCP, whereby activity types in L1 and L2 were collected, but the survey was administered on a daily-basis for a 24 hour day divided into 15-minute segments. As such, the questionnaire in Ranta and Meckelborg (2013) was more fine-grained than the ones used in Dewey et al. (2013) and Freed et al. (2004). Furthermore, the results of L2 usage collected retrospectively may have obscured variations of L2 engagement because learners had to generalize their language contact over the entire period of a SA experience (McManus, Mitchell, & Tracy-Ventura, 2014). In recognition of such difficulties, I administer a Language Exposure Log (LEL), which is a modified version of the log of Ranta and Meckelborg (2013), with each daily log composed of 21 hours, from 6 a.m. to 3 a.m., divided into one-hour blocks. This way, learners need not rely on memory of an entire semester but can complete the log on an hourly basis.

## CHAPTER 2: CURRENT STUDY

### 2.1 Goals of this study

The goals of this study were threefold. My first aim was to examine the amount and types of language engagement (i.e., *language focus* and *language skills*) of international students in US higher education. A secondary goal was to explore the extent to which explicit knowledge influences the development of implicit knowledge. Lastly, I examined the association between the knowledge types and activity types. Results reported in this dissertation are based on two time-point data from a one-year longitudinal project implemented in three waves. Key terms are defined as follow:

- *Explicit knowledge* refers to conscious-verbalizable knowledge of language rules and *implicit knowledge* represents tacit-unconscious knowledge of linguistic regularities.
- *Language engagement (or language use)* is used as an umbrella term that encompasses learners' use of the L2 (English), other languages (native or third languages), and no language use.
- *Activity focus* refers to the types of focus learners adopted when engaging in activities. For instance, *language-focus* is defined as learners' engagement with an L2 through a focus on the grammatical or lexical forms of the target language; for instance, the purpose of performing the activities is to learn linguistic aspects of the target language or forms. This would include learning grammar in a language course or individually, consulting a dictionary to learn the meaning or forms of words/phrases/idioms, reviewing papers to correct the language, or analyzing language to understand the reading etc. *Meaning-focus* is defined as learners' engagement with a language for

communicative or meaning-making purposes. Examples include, watching television for pleasure, surfing the Internet to gain information, or reading or writing emails to convey messages. The average frequency of each category (i.e., language-focused activity and meaning-focused activity) served as the outcomes for analysis.

- *Language skills* references learners' use of the four language skills including reading & listening (receptive) and speaking & writing (productive).

In what follows, I present the four research questions along with the hypotheses that will guide the study. The path diagram in Figure 2.1 illustrates which paths in the statistical model will enable me to test each hypothesis.

The first and second research questions explored the qualitative and quantitative nature of language engagement.

RQ1. How much *L2 engagement* do the international students have?

- Hypothesis 1: Relative to other-language use (e.g., native or third languages), it is anticipated that international students will have more engagement with English (L2).

RQ1.1 What types of activities (*language-focus or meaning-focus*) do the international students adopt the most and the least when engaging in English?

- Hypothesis 1.1: Compared to language-focused activities, international students are expected to use more meaning-focused activities, but there will be considerable individual variations.

RQ1.2 What types of *language skills* do the international students engage in the most?

- Hypothesis 1.2: A wide-ranging variation is expected, but on average, I predict

that international students will engage more receptively with the L2.

The second research question focuses on the relationship between knowledge types.

RQ2. To what extent does explicit knowledge influence the development of implicit knowledge?

- Hypothesis 2: Explicit knowledge positively influences the development of implicit knowledge.

The third research question connects research questions 1.1 and 2; that is, it examines the effects of activity focus (i.e., language-focused and meaning-focused) on knowledge type development.

RQ3. To what extent do the types of activities (*language-focus or meaning-focus*) contribute to the development of different types of knowledge?

In general, I anticipate that both types of activity will contribute to both types of knowledge development. However, language-focused activities are expected to have a stronger predictive effect on explicit and implicit knowledge compared to the effects of meaning-focused activities on both knowledge types (e.g., Goo et al., 2015).

- Hypothesis 3: The amount of language-focused activity positively influences the acquisition of explicit knowledge.
- Hypothesis 4: The amount of language-focused activity positively influences the acquisition of implicit knowledge.
- Hypothesis 5: The amount of meaning-focused activity positively influences the acquisition of explicit knowledge.
- Hypothesis 6: The amount of meaning-focused activity positively influences the acquisition of implicit knowledge.



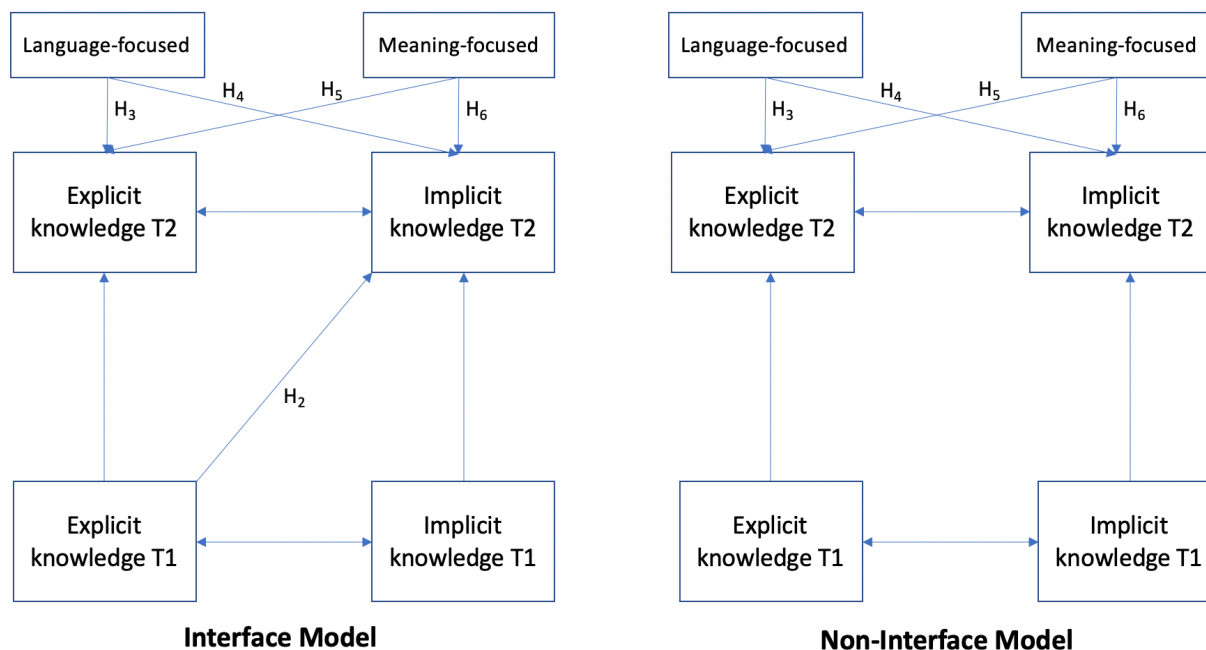


Figure 2.1. The Interface model (left) and the Non-interface model (right) used to test three main research hypotheses. Note. H = Hypothesis; T1 = Time 1; T2 = Time 2

## 2.2 Participants

One hundred and twenty two English L2 speakers studying at a large American university in the Midwest took part in two testing stages of the study.<sup>1</sup> Overall, there was 25.5 percent of participant attrition: a total of 149 participants participated in T1 and 122 participants remained in T2. All participants met the following two criteria: (1) They received a minimal score of 60 in the iBT TOEFL and (2) were physically available to visit the lab in the three testing periods (T1: January–February, T2: April–May, T3: November–December of 2019). The minimal TOEFL score was set to 60 to mirror the provisional admission for English at the institution where participants were recruited. One of the important goals of this study was to

<sup>1</sup> Two participants were excluded from all analyses as they struggled to compete the tasks.

examine learners' L2 developmental processes. It was thus important to be inclusive of L2 learners of a wide range of proficiency levels, including those in the beginning level. Other standardized tests such as ILETS, DIALANG, and TOEIC were also accepted. In such cases, the scores were converted to TOEFL scores using the reference by Educational Testing Services (<https://www.ets.org/toefl/institutions/scores/compare/>). As seen in Table 2.1, most participants had a TOEFL score above 79, and thus were considered intermediate to advanced English users. The group's average TOEFL score, including the converted scores, was 93.03 (SD = 13.11) at T1 and 93.61 (SD = 12.72) at T2.

Table 2.1 Proficiency Breakdown

|                                       | TOEFL<br>60-78 | TOEFL<br>79-90 | TOEFL<br>91-100 | TOEFL<br>101+ | Average (SD)  |
|---------------------------------------|----------------|----------------|-----------------|---------------|---------------|
| Time 1 ( <i>N</i> = 149) <sup>a</sup> | 18             | 40             | 48              | 43            | 93.03 (13.11) |
| Time 2 ( <i>N</i> = 122) <sup>a</sup> | 14             | 32             | 38              | 38            | 93.61 (12.80) |

Note. <sup>a</sup>There were four cases where graduate students had not taken standardized English tests.

In such cases, I logged the minimal required TOEFL scores required for acceptance into their programs.

All participants were recruited through one of the following recruitment routes: 1) Flyers on campus and off campus buildings (see Appendix G for a copy of a recruitment flyer), 2) the Registrar Data Request System (<https://reg.msu.edu/Forms/DataRequest/DataRequest.aspx>) that assists in email distribution to eligible participants affiliated with MSU, 3) the SONA system at the College of Communication and Science (<https://msucas.sona-systems.com/Default.aspx?ReturnUrl=%2f>) which offers a recruitment outlet to both MSU and non-MSU affiliated prospective participants, and 4) social media (Facebook, WeChat, allMSU).

Participants received 80 dollars in total: 60 dollars upon the completion of the three experiments and an additional 20 dollars upon submitting 10 language exposure logs.

The 122 participant sample who took part at T1 and T2 consisted of learners who possessed a final or current educational degree at the bachelor's level ( $n = 33$ ) and the master's or doctoral levels ( $n = 81$ ).<sup>2</sup> A majority of students reported to have learned English mainly in a language-oriented instructional setting ( $n = 68$ ) or a mixture of both types of instructional methods ( $n = 42$ ). Only two participants reported to have received meaning-oriented instruction.<sup>3</sup> The demographic information of participants enrolled at each time point is included in Table 2.2.

Table 2.2 Background Information of The L2 Speakers

| Time   | Variables                    | Mean (SD)     | Min-Max |
|--------|------------------------------|---------------|---------|
| Time 1 | Age                          | 26.56 (6.36)  | 18-49   |
|        | Length of Residence (months) | 34.78 (33.11) | 1-216   |
|        | Age of Arrival <sup>a</sup>  | 22.92 (6.05)  | 3-40    |
|        | Age of instruction           | 8.16 (3.88)   | 2-30    |
| Time 2 | Age                          | 27.02 (6.27)  | 18-49   |
|        | Length of Residence (months) | 36.19 (34.83) | 1-216   |
|        | Age of Arrival <sup>a</sup>  | 23.32 (6.07)  | 3-40    |
|        | Age of instruction           | 8.38 (4.02)   | 3-30    |

Note. <sup>a</sup>Two participants reported to have arrived in English-speaking countries before the age of 14. They had spent 1/5 to 1/3 of their lives (P1: 6.5 years; P2: 4.2 years) in English-speaking countries. I decided not to remove these participants from final analyses as they all had received

<sup>2</sup> Eight participants preferred not to specify.

<sup>3</sup> Ten participants preferred not to specify.

formal education outside of the English-speaking countries in both language- and meaning-focused instruction and thus likely to possess both explicit and implicit L2 English knowledge.

## **2.3 Materials**

### **2.3.1 *Target structures***

The target structures include six grammatical features: (1) Third person singular -s, (2) mass/count nouns, (3) comparatives, (4) embedded questions, (5) *be* passive, and (6) verb complement. Three syntactic (4-6) and morphological (1-3) structures are used to measure a range of English grammar knowledge (see Table 2.3 for examples). These structures were selected to represent early and late acquired grammatical features (e.g., R. Ellis, 2009; Pienemann, 1989) and thus were appropriate to measure participants' general English proficiency. These target structures are identical to those in Godfroid and Kim (under review) and Godfroid, Kim et al. (in preparation). A subset of these structures have also been used in previous studies such as R. Ellis (2005) and Vafaei et al (2017).

Table 2.3 Six Sentence Structures and Examples

| Structure                | Description   | Example  |
|--------------------------|---|--|
| Third person singular -s | Subject-verb agreement for a singular subject in the present tense  | * <i>The old woman <u>enjoy</u> reading many different famous novels</i>   |
| Mass(/count) noun        | Mass nouns in singular form; Mass nouns should not be marked with the -s morpheme that marks plural on countable nouns.                     | * <i>The boy had <u>rices</u> in his dinner bowl.</i>                      |
| Passive                  | Be-passive construction; Focus is on presence of <i>be</i> and form of verb following <i>be</i> (i.e., past participle).                    | * <i>The flowers were <u>pick</u> last winter for the festival.</i>        |
| Embedded question        | In an embedded clause following a <i>wh</i> -word, word order is subject-verb-object.   | * <i>He wanted to know why <u>had</u> <u>he</u> studied for the exam.</i>  |
| Comparative adjective    | Comparative adjectives are marked with the suffix <i>-er</i> (1-2 syllable adjectives) or preceded by <i>more</i> (>2 syllable adjectives). | * <i>It is <u>more harder</u> to learn Japanese than to learn Spanish.</i> |
| To-verb complement       | The target verbs (need, have, want, ask) require <i>to</i> -infinitive verb complements.  | * <i>Jim is told his parents <u>want</u> <u>buying</u> a new house.</i>    |

### 2.3.2 Instruments

I administered two questionnaires (i.e., a background and motivation questionnaire), five linguistic knowledge measures, and a Language Exposure Log (LEL). Regarding the linguistic measures, untimed written grammaticality judgment test (GJT) and metalinguistic knowledge test (MKT) were used to measure learners' explicit L2 English knowledge. The timed written GJT, oral production (OP), and elicited imitation (EI) served as implicit L2 English knowledge measures (e.g., R. Ellis, 2005; R. Ellis & Loewen, 2007; Godfroid & Kim, under review; Godfroid, Kim et al., in preparation). The types of language engagement (i.e., *activity focus* and

*language skills*) were measured with a LEL. Below, I introduce the main characteristics of the different measures. For a succinct summary, see Table 2.4.

### **2.3.3 *Background questionnaire***

Participants completed different versions of the background questionnaires at T1 and T2. The questionnaire given at T1 was developed to collect general and language-related information of the participants. There were 14 questions in total, with items related to biographical information (7 questions), language background (5 questions), and language learning (2 questions).

The questionnaires administered at T2 were developed to measure language exposure experiences on courses and weekend routines during Spring (T2) semester. For instance, questions asked about the number of English language courses enrolled (e.g., classes dedicated to the teaching of English skills in listening, speaking, reading, or writing) and the amount of language engagement during the weekends (i.e., time spent speaking, listening, writing, and reading in English). All questionnaires are included in Appendices A and C.

### **2.3.4 *Motivation questionnaire***

Along with the background questionnaire, two sets of motivation questionnaires were administered. The first set was a grit questionnaire (Duckworth, Peterson, Matthews, & Kelly, 2007) and the second set was a language mindset questionnaire (Dweck, 1999). An example of a grit scale questionnaire includes, “*I often set a goal but later choose to pursue a different one.*” or “*I am diligent. I never give up.*”; and a question on language mindset includes “*You can*

*always improve your language learning intelligence.*”.<sup>4</sup> The questionnaires are included in Appendix B.

Table 2.4 Summary of Measures

| Measures                    | Test                          | # Items (Total)      | # Items | Grammaticality | Calculation                                    | Dependent variable |
|-----------------------------|-------------------------------|----------------------|---------|----------------|--|--------------------|
| Implicit knowledge measures | Timed written GJT             | 40                   | 24      | 12 G, 12 UG    | Accuracy                                       | Factor score       |
|                             | Elicited Imitation            | 32                   | 24      | 12 G, 12 UG    | Accuracy: correct usage in obligatory contexts |                    |
|                             | Oral production               | 22 (a story context) | 22      | 22 G           | Accuracy: correct usage in obligatory contexts |                    |
| Explicit knowledge measures | Untimed written GJT           | 40                   | 24      | 12 G, 12 UG    | Accuracy                                       | Factor score       |
|                             | Metalinguistic knowledge test | 12                   | 12      | 12 UG          | Accuracy: error explanation                    |                    |
| Form-focused activity       | Self-reported language log    | 20                   | 20      | N/R            | Average hours                                  | Z-scaled frequency |
| Meaning-focused activity    | Self-reported language log    | 20                   | 20      | N/R            | Average hours                                  | Z-scaled frequency |
| L2 use                      | Self-reported language log    | 20                   | 20      | N/R            | Average hours                                  | Average hours      |

Note. G = grammatical; UG = ungrammatical

### 2.3.5 Oral production (OP)

In the web-programmed oral production task, participants read a picture-cued short story, seeded with the target structures (Godfroid & Kim, under review; Godfroid, Kim et al., in preparation). The story consisted of 18 sentences (250 words) and 10 pictures on Mr. Lee’s life. Participants read the story (with picture prompts) twice with unlimited time and importantly,

<sup>4</sup> These items were included to examine the association between the amount of L2 engagement and linguistic development. The results on the motivation questionnaire will not be included in this dissertation.

were asked not to take notes but rely on their memory. See Figure 2.2 for the web-based instructions.

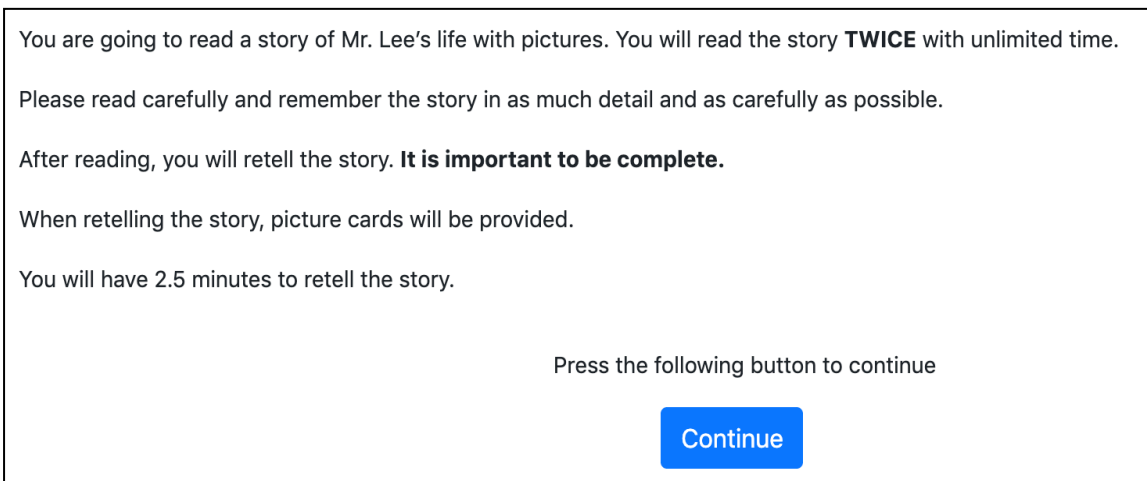


Figure 2.2. Screen shot of a part of the web-based instructions in the oral production task

A minimum of one sentence (10 words) to a maximum of four sentences (55 words) accompanied each picture. The pictures were carefully drawn to emphasize the main content so as to facilitate memory retrieval. When finished, the participants were asked to retell, in 2.5 minutes, the picture-cued story in as much detail as possible. They were informed that they would not be able to go back to the previous picture once they moved on (see Figure 2.3). During the retelling, the picture remained on the screen for participants to freely proceed to the next picture at their own pace with progress indicator (e.g., 1 out of 10 pictures) presented under each picture prompt (see Figure 2.4).



There will be ten pictures. For each picture, there will be corresponding information. Again, you have unlimited time to read the story, but you **cannot go back to a previous page once you move on.**



Press the following button to continue

Continue

Figure 2.3. Screen shot of a part of the web-based instructions in the oral production task


Mr. Lee's daily schedule

|                     |                               |
|---------------------|-------------------------------|
| 6:00 am             | Get up                        |
| 7:00 am             | Break fast (tea and bread)    |
| 8:00 am to 6:00 pm  | Work as a supermarket manager |
| 9:00 pm to 12:00 am | Work as a bartender           |

1/10

Next

Click the Record button below to start retelling the story.



Recording...

Time left: 02:26

Figure 2.4. Screen shot of the oral production task during retelling with picture prompts and a progress indicator

At the two testing points (T1 and T2), the same story prompt was used because the time interval between the two time points (T1-T2: 3-4 months) was sufficiently large to make it unlikely that participants would be able to recall the exact wordings of each sentence.

**Scoring.** Target morphosyntactic features were coded on two features: the number of times (a) a target feature is required and (b) the number of times it was correctly applied were tallied. The number correct is divided by the number required to arrive at the overall accuracy score (See Appendix D for the story prompt and Appendix H for Coding Guidelines).

### 2.3.6 *Elicited Imitation (EI)*

In the web-programmed elicited imitation task, participants were instructed to listen to a series of sentences, judge the plausibility of the sentence, and repeat each sentence in correct English after a beep sound (see Figure 2.5 for the web-based instructions).

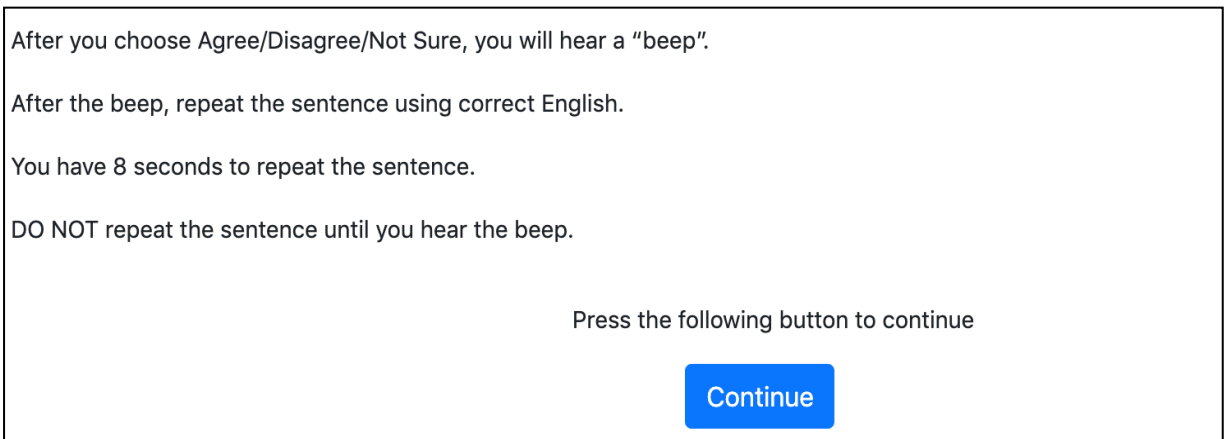


Figure 2.5. Screen shot of a part of the web-based instructions in the elicited imitation task

The task consisted of 32 sentences, 24 of which were target sentences (Godfroid & Kim, under review; Godfroid, Kim et al., in preparation). Each sentence was between 6 to 13 words. Eight practice sentences, 4 grammatical and 4 ungrammatical, and model responses to two practice sentences were played prior to the test in order to clarify the task instructions. Many participants, for instance, would alter the content of an implausible sentence to make it plausible; consequently, target structures would be omitted in the repetition. The model responses clarified such confusion. At no point did the model responses or instructions include explicit instructions or feedback on the linguistic features. Furthermore, participants were not explicitly informed that 32 sentences included ungrammatical statements.

Two counterbalanced lists of stimuli were created for the target sentences. In List 1, half of the sentences were grammatical, and half were ungrammatical and in List 2, the grammaticality was reversed from List 1. List 1 was given at T1 and List 2 was administered at T2.

At the beginning of each experimental trial, participants saw a fixation cross ('+') in the center of the screen on which to fixate their eyes. This lasted for 500 ms. The sentence then was played with a speaker icon appearing on the screen. After each sentence, a secondary task followed asking whether the participant agreed, disagreed, or was unsure about the content of the statement. They had 4 seconds to make judgments on the semantic plausibility and respond by clicking on the corresponding icon ("Agree", "Disagree", "Unsure"). A fixation cross followed for 500 ms. This was included to alert participants with the transition from the plausibility judgment task to the production task. After 500 ms, a microphone icon appeared with a beep sound and a text "Please repeat now." and "Your voice is now being recorded.". The texts were located above and under the microphone icon, respectively, and as with the oral production task, a progress marker was included (e.g., 1/32) to maintain participants' motivation. Participants had 8 seconds to repeat the sentence.

**Scoring.** As with the scoring of the oral production task, correct use of the target forms in obligatory contexts was used to calculate an overall accuracy score. See Appendix D for the experimental stimuli and Appendix H for Coding Guidelines.

### **2.3.7 Grammaticality Judgment Tests (GJTs)**

In the lab-based computerized written GJTs (untimed/timed written GJT), participants were instructed to read the sentence, either under time pressure (timed written GJT) or without time pressure (untimed written GJT), and judge its grammaticality. Both tests consisted of 40 items (24 target sentences; four for each structure; half grammatical, half ungrammatical; Godfroid & Kim, under review; Godfroid, Kim et al., in preparation). A part of the GJT items was from Vafaei, Suzuki, and Kachinske (2016), which was modified for sentence length and lexical choices to ensure comparable processing loads across sentences and to avoid lexical repetition across nine tasks used in Godfroid and Kim et al. (in preparation).

In the timed written GJT, participants were urged to make judgments as soon as possible. The time limit for each item was set based on the length of audio stimuli in the aural GJT. In particular, in Godfroid and Kim et al. (in preparation), we computed the average audio length of sentences with the same sentence length and added 50% of its median. As a result, the time limit imposed for a seven-word sentence was 4.12 seconds and that for a 14-word sentence was 5.7 seconds. Previous studies (e.g., Bowles, 2011; R. Ellis, 2005; Kim & Nam, 2016) have set the time limit for each item based on native speakers' mean response time plus 20%. Imposing a time limit based on native speakers' response rate sets an assumption that native and non-native speakers perceive linguistic ease and difficulty in a similar manner, and difficulty impacts L1 and L2 processing similarly (i.e., a linear increase in L2 processing time). Given that this is not always the case (e.g., third person singular -s is acquired late in the developmental sequence for non-native speakers but detected most saliently for native speakers of English), we used the sentence length (or number of words) instead.

The procedure was exactly the same for the untimed written GJT, but without a time limit. In so doing, the test design allowed participants to use analytic processing or explicit knowledge. Sentences appeared one at a time in their entirety (font: Helvetica; size: 44) on the computer screen. For each time point, different sets of sentences were used. See Appendix D for the experimental stimuli.

***Scoring.*** Correct responses were awarded one point.

### **2.3.8 Metalinguistic Knowledge Test (MKT)**

Participants were given twelve sentences with grammatical violations, two for each structure. The participants' task is 1) to identify the error, 2) correct the error, and 3) explain why it is ungrammatical. For the last component (explanation of ungrammaticality), participants were told to be as complete and specific as possible in their answers. Prior to the test, two practice questions were provided along with a good and a bad response to illustrate the nature of the test. Participants were also allowed to use a dictionary for translation purposes; for instance, metalinguistic terms of English grammar such as "articles" or "third person singular" in their native language but not in English. In such cases, participants were instructed to use the dictionary for English translation. For each time point, different sets of sentences were used.

***Scoring.*** Correct responses are awarded one point. For the final analysis, only the responses on the explanation component, which requires the most explicit declarative knowledge of language, were used. See Appendix D for the sentences and Appendix I for Coding Guidelines.

### 2.3.9 *Language Exposure Log*

The self-report language log was used to elicit qualitative and quantitative information on one's engagement with languages. Developed by Ranta and Meckelborg (2013), the log was designed to record daily language use in a fine-grained manner. In particular, each daily log was composed of 24 hours, from 6 a.m. to 3 a.m., divided into one-hour blocks. For every hour, learners had to provide three pieces of information. The first piece was on the general category of language usage; that is, whether they used English (Speaking, Writing, Reading, or Listening in English), native or other languages, or no languages. Overall, they had six options to choose from: Speaking, Writing, Reading, or Listening in English; using other languages; and no language use.

The second piece of information elicited specific activities of the chosen general category. There were 4-13 activity items, which differed for each option. For instance, in the Writing-in-English category, the activity items included: *writing emails*, *writing academic papers*, *messaging friends*, *creating presentation slides*, *personal writing/journaling (such as diary)*, and *others*. If the listed options did not apply, they were instructed to use the open-ended "others" option (for a full list of activity items, see Appendix E).

The last category elicited whether the chosen activity was carried out focusing on language and/or meaning. The participants were instructed to choose language-focused when the purpose of performing the activities was to learn linguistic aspects of the target language or forms. This would include learning language aspects in a language course or individually, consulting a dictionary to learn the meaning or forms of words/phrases/idioms, reviewing papers to correct the language, or analyzing language to understand the reading etc. On the other hand, meaning-focused referred to when the focus of language use was on meaning. Examples include,

watching television for pleasure, surfing the Internet to gain information, or reading or writing emails to convey messages. The average frequency of each category (i.e., language-focused activity and meaning-focused activity) served as the outcomes for analysis.

## **2.4 Procedure**

With the financial support of the Second Language Studies (SLS) program, the background and motivation questionnaires and the two oral production tasks (i.e., OP and EI) were programmed on the web for participants to complete them in the convenience of their home. No commercial software thus far supports oral recording functions on the web for remote data collection. With the help of a computational linguist, Dr. Xiaobin Chen, we programmed two oral production tasks online, and created a manual for SLA researchers who hope to incorporate customized test items to oral production and elicited imitation tasks (see Appendix F for instructions on the use of the web-based testing program). These tasks were programmed on *Java with Google Web Toolkit* (see Procedure section for details on the administrative process). The two GJTs were programmed on SuperLab 5.0, and the metalinguistic knowledge test was programmed on *Qualtrics*.

### **2.4.1 Knowledge Measures**

Table 2.5 summarizes the test procedure for the different linguistic knowledge measures at T1 (between January and February) and T2 (between April and May). As mentioned, participants completed the two oral production tasks (i.e., OP and EI) prior to coming to the lab. The GJTs and MKT were administered in the lab to prevent participants from relying on external resources when completing the tests on the web (e.g., referring to grammar books or browsing



the Internet for answers). These tasks (the two GJTs and MKT) were administered, maximum 4 participants at a time, in the Second Language Acquisition lab at MSU.

Ten days prior to the lab visits, participants received an email that contained a reminder of visit schedules (i.e., date, time, and location) and a link to a video recording with a step-by-step tutorial on how to complete the web tasks. This video instruction was recorded with a New Screen Recording function in QuickTime Player (version 10.5). Seven days prior to the lab visits, participants received an email containing a unique web link with a personalized code that directed them to an interface with the web-based versions of tasks. In the interface, participants were given general instructions, such as finding a quiet room for the next 30 minutes and using Chrome or Firefox to complete the tasks. Following the instructions were a consent form, a background questionnaire, and the two oral production tasks. A reminder of the lab visit was sent three days prior to one's lab visit. All reminders were sent using an automated mail merge function.

As for the test sequence, measures that draw attention to meaning (i.e., OP and EI) were administered prior to the measures that direct learners' attention to form (i.e., GJTs and MKT). This sequencing was meant to minimize the likelihood of participants' becoming aware of the target structures in the implicit knowledge measures. See Table 2.5 for the timeline and test sequence.

Table 2.5 Timeline and Test Sequence

| Setting | Time 1<br>(January-February, 2019) |      | Interim                      | Time 2<br>(April-May, 2019) |      |
|---------|------------------------------------|------|------------------------------|-----------------------------|------|
|         | Knowledge measures                 | Min. | Language activity measures   | Knowledge measures          | Min. |
| Web     | Background questionnaire           | 15   | <i>Language Exposure Log</i> | Background questionnaire    | 10   |
|         | Oral production                    | 10   |                              | Oral production             | 10   |
|         | Elicited Imitation                 | 15   |                              | Elicited Imitation          | 15   |
| Lab     | Timed & Untimed Written            |      |                              | Timed & Untimed Written     |      |
|         | GJT                                | 15   |                              | GJT                         | 15   |
|         | MKT                                | 15   |                              | MKT                         | 15   |

#### 2.4.2 Language Exposure Measures

With regard to the Language Exposure Log (LEL), participants recorded language usage activities on five days between the two test points: during spring semester (March–April).

A great effort was made to gather comprehensive information about learners’ language use at different days of the week and in different months. For instance, the five LELs collected between each time points represented different weekdays (e.g., Monday [log 5], Tuesday [log 4], Wednesday [log 3], Thursday [log 2], and Friday [log 1]). Also, the five LELs were distributed across the two months. For instance, between T1 and T2, three LELs were given in March and two in April.

During the lab visits, participants were trained on how to complete the LEL and how to differentiate meaning-focused from language-focused activities. A day before the LEL fill outs, participants received a reminder that included 1) a link to a recorded step-by-step video instruction, 2) a link to the questionnaire, and 3) a reminder to record their activity for every hourly segment and that they could log in and out as many times as they wanted. During the day

of fill outs, a reminder was sent twice, at 12 p.m. and 6 p.m. The questionnaire was developed via *Qualtrics*, which functioned in both web and mobile interfaces.

## CHAPTER 3: LANGUAGE USE

The aim of this chapter was to provide a detailed picture of language exposure in a naturalistic setting. By doing so, I present data on the amount of L2 engagement that international students gain in a setting with extensive opportunities for L2 usage. What types of language skills receive the most (least) attention during daily life as an international student? How much individual variation exists in English use among international students? In this section, I hope to provide empirical data to answer these questions.

For reader convenience, I define the terms again that will be used in this chapter.

- *Language engagement (or language use)* is used as an umbrella term that encompasses learners' engagement with the L2 (English), other languages (native or third languages), and no language use.
- *Activity focus* refers to the types of focus learners adopted when engaging in activities. For instance, language-focus is defined as learners' engagement with an L2 through a focus on the grammatical or lexical forms of the target language; Meaning-focus is defined as learners' engagement with a language for communicative or meaning-making purposes. Throughout the dissertation, I will use language- and meaning-focused *activity* types.
- *Language skills* (or language activity) references learners' use of the four language skills including reading & listening (receptive) and speaking & writing (productive).

### 3.1 Research Questions

The research questions addressed in this chapter are the following:

RQ1. How much language engagement do international students have?

RQ1.1 What types of activities (*language-focus or meaning-focus*) do the international students adopt the most and the least when engaging in English?

RQ1.2 What types of *language skills* (i.e., receptive [reading & listening] or productive [speaking & writing]) do international students practice the most?

### 3.2 Data Preparation Details

As mentioned in Chapter 2, the five Language Exposure Logs (LELs) were collected via *Qualtrics* which extracts data in .csv files. Using the data program *R* version 1.2.1335, I automatized the data cleaning and analysis processes in three steps. Step one was data cleaning: I read in five wide-formatted logs with 395 columns each, checked missing data/duplicates, and transposed the logs to long-formats by *time* (6 a.m. to 3 a.m.) and *skills* (reading, speaking, writing and listening in English; using other languages; and no-language use). In step two, I computed descriptive information using the *mutate* function to calculate time allocated to each skill. For instance, if three activities were performed at 12 p.m. (e.g., speaking, reading, and listening in English), I would assign 20 minutes to each skill, whereas 60 minutes would be allocated to one skill if only one skill were used at 12 p.m. Lastly, line graphs were generated with this descriptive information. Both the logs and *R* scripts will be made available through IRIS (<https://www.irisdatabase.org>) following publication of the project. Also, the raw data, meta-data, and read-me document will be deposited in Dataverse, an open source web application, with a two-year embargo period for data files uploaded.

### 3.3 Results

#### 3.3.1 *Types and Amount of Language Engagement*

The following results provide descriptive information on the amount and the types of participants' language engagement. I start by reporting the overall time that the international students used different types of language skills (i.e., reading, listening, speaking and writing in English; using other languages; and no language use) and what their language foci were (i.e., language-focused and meaning-focused). These results will provide a general picture of how much learners engage in different activities and languages on a daily basis (RQs 1, 1.1 and 1.2). For the analysis, I used data from 154 participants.

As seen in Table 3.1, out of a 21-hour day (from 6 a.m. to 3 a.m.), the international students spent the most time using English ( $M = 9.06$ ,  $SD = 3.10$ ). This was followed by 8.14 hours of not engaged in languages ( $SD = 3.21$ ) and 3.79 hours using other languages ( $SD = 2.92$ ). After assigning the three hours that were not included in the LEL (from 3 a.m. to 6 a.m.) to No Languages (assuming that all participants are asleep between 3 a.m. and 6 a.m.), we can see that, over 24 hours, an average participant day consisted of 36% English use (9.06 hours), 15% other language use (3.76 hours), and 45% no language use (11.14 hours). Figure 3.1 displays the details.

Table 3.1 Time Per Day (21 Hours Max) on Specific Activities

|                 | Mean | SD   | Min-Max     | CI<br>lower-upper<br>bound |
|-----------------|------|------|-------------|----------------------------|
| English         | 9.06 | 3.10 | 0.60–18.50  | 8.57–9.56                  |
| Speaking        | 2.58 | 1.76 | 0.00–10.25  | 2.29–2.86                  |
| Reading         | 2.45 | 1.62 | 0.00–8.00   | 2.19–2.71                  |
| Listening       | 2.72 | 1.74 | 0.00–9.33   | 2.44–3.00                  |
| Writing         | 1.32 | 1.21 | 0.00–6.40   | 1.12–1.52                  |
| Other languages | 3.79 | 2.92 | 0.00–16.00  | 3.32–4.26                  |
| No languages    | 8.14 | 3.21 | 0.25- 17.50 | 7.62–8.66                  |

Note. SD, standard deviation; CI, confidence interval

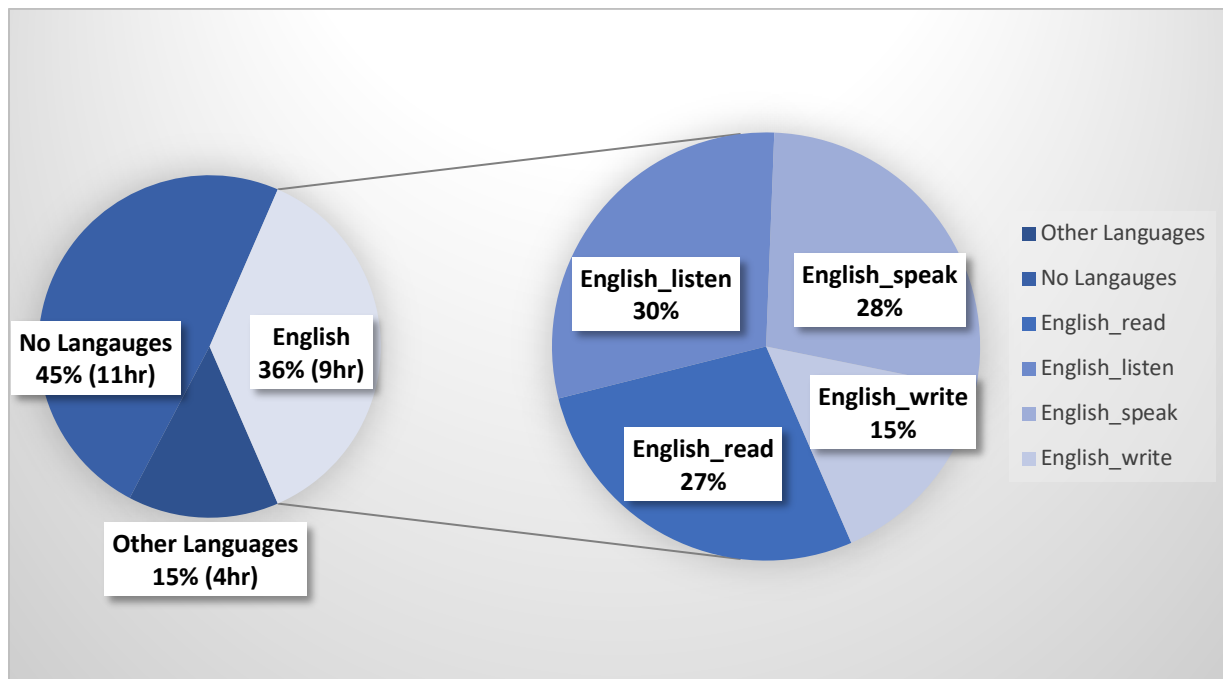


Figure 3.1. Types and amount of language engagement during a 24-hour day (N = 154)

Note that the 11.14 hours of No languages includes the time in which learners are asleep. I was curious about how international students allocate their time while awake. To this end, I excluded the recommended 7 hours of sleep from No Languages, which left an estimated 4.14 hours per day when learners were awake and not engaging in verbal communication. I conclude that, of the time international students are awake, international students spent about half the time using English (9.06 hours), which is roughly twice the amount of English use than use of Other Languages (3.79 hours) and No Languages (4.14 hours).

It is important to highlight that the time spent using English and other languages were statistically different. To be precise, international students spent 2.2x more time using English than other languages. As seen in Table 3.1, the confidence intervals (CIs) in English (8.57–9.56) and the Other (3.32–4.26) category do not overlap. This finding lends some support to a widely held assumption that immersion contexts provide learners with extensive opportunities for engagement with the target language. Even so, evidence for this statement could be strengthened by adding a control group of EFL learners at a university in a non-English-language environment.

Lastly, a wide range of values were observed for each activity type. This is particularly evident for English and No Languages with language use ranges between 0.60–18.50 and 0.25–17.50 hours/day, respectively. I will revisit this point of non-trivial individual variability in the *Variation among learners* section.

In regard to RQ 1.2 (what *language skills* the international students engage in the most), I observed qualitative differences in English engagement. At a descriptive level, engagement with listening ( $M = 2.72$  hours) and writing ( $M = 1.32$  hours) showed the highest and lowest engagement rate, respectively. A comparable amount of time was spent across reading ( $M = 2.45$



hours), listening ( $M = 2.72$  hours) and speaking ( $M = 2.58$  hours) in English, as the CIs for the average time spent on each of these skills overlapped; however, the amount of engagement with writing ( $M = 1.32$  hours) was statistically lower than the amount spent listening, reading, and speaking in the L2. As such, while international students spent a comparable amount of time speaking, listening, and reading in English, they spent significantly less time writing in English.

To address RQ 1.1 (what *activities* the international students use the most when processing English), I computed descriptive information on activity types. The international students enrolled in an English-medium university spent significantly more time engaging in English for meaning-focused purposes than for form (see Table 3.2). This is rather an expected outcome, as students presumably rely on meaning-focused processes when using language as a tool (e.g., to achieve tasks, such as ordering food or listening to a chemistry lecture). Similar patterns were observed with other-language engagements (that is, their native or L3 languages), where learners also reported processing non-English languages for meaning and not form.

Table 3.2 Time Per Day (21 Hours Max) Spent on Language Focus

|                      |                 | Mean | SD   | Min-Max   | CI<br>lower-<br>upper<br>bound |
|----------------------|-----------------|------|------|-----------|--------------------------------|
| Focus on<br>Language | English         | 2.43 | 1.88 | 0.00–8.5  | 2.13–2.73                      |
|                      | Other languages | 0.85 | 1.78 | 0.00–13.5 | 0.57–1.13                      |
| Focus on<br>Meaning  | English         | 6.76 | 3.01 | 0.33–14.8 | 6.28–7.24                      |
|                      | Other languages | 3.56 | 2.90 | 0.00–14.1 | 3.10–4.02                      |

Note. “No Language use” category does not include *language focus* options; SD, standard deviation; CI, confidence interval

In summary, a group-averaged daily engagement data showed that international students who enrolled at an English-medium university in the US were more engaged (quantitatively, in terms of hours per day spent) with L2 English than other languages. To be precise, they spent 2.2x more time using English than other languages. I also observed qualitative differences in English engagement. While students spent their time relatively evenly between speaking, listening, and reading in English, they spent significantly less time writing in English. Lastly, the international students spent significantly more time using English for meaning-focused activity than language-focused activity, consistent with the behavior they reported for other languages.

### ***3.3.2 Language Engagement Over Time***

Next, I break down the amount and types of language engagement by time. This was to collect empirical evidence of when, over a 21-hour day, the learners used their L2 or other languages, and how much. I then explore variation among individuals in the amount and types of L2 skills they used. To this end, I created line graphs of the language exposure logs of all participants separately.

As a means to explore different types of language engagement by time, I plotted the five daily language logs for each activity type separately. Figure 3.2 illustrates the frequency of six types of language engagement by time of day (i.e., reading, listening, speaking and writing in English; using other languages; and no language use). An immediately noticeable feature is the converging patterns of the five daily logs. As mentioned in Chapter 2, each language log fell on a different weekdays (log 1 [Friday], log 2 [Thursday], log 3 [Wednesday], log 4 [Tuesday] and log 5 [Monday]) across two months (logs 1, 2, and 3 in March; logs 4 and 5 in April). Seen in

this light, the convergence patterns observed across the five logs speak to the reliability of the data, specifically that the day the LELs were collected did not substantially affect the overall language engagement patterns.

Next, I focused on the different time frames in each graph to see what times of day learners were most commonly doing different activities. For ease of interpretation, I enlarged the visuals for the four English skills in Figure 3.3.

The line graphs for each activity type highlight several interesting findings. Starting with the four English skills, in a 21-hour day, the highest engagement rates in all four skills were during the typical office hours (9 a.m. to 5 p.m.) and to some extent from 8 p.m. to 11 p.m. The latter is especially evident in English reading and writing, and to some degree in listening as well. These findings mirror the daily routines of undergraduate and graduate students quite nicely: they typically attend and teach courses, engage in academia-related meetings, or interact in English with peers during the traditional working hours. Additionally, students continue to engage in some activities after working hours that typically require reading, writing, and listening skills to accomplish self-directed and less interactive tasks for school. A relatively high level of engagement in these three skills during the evening time may speak to the nature of graduate/undergraduate student life and student duties that extend beyond the commitment of typical working hours.

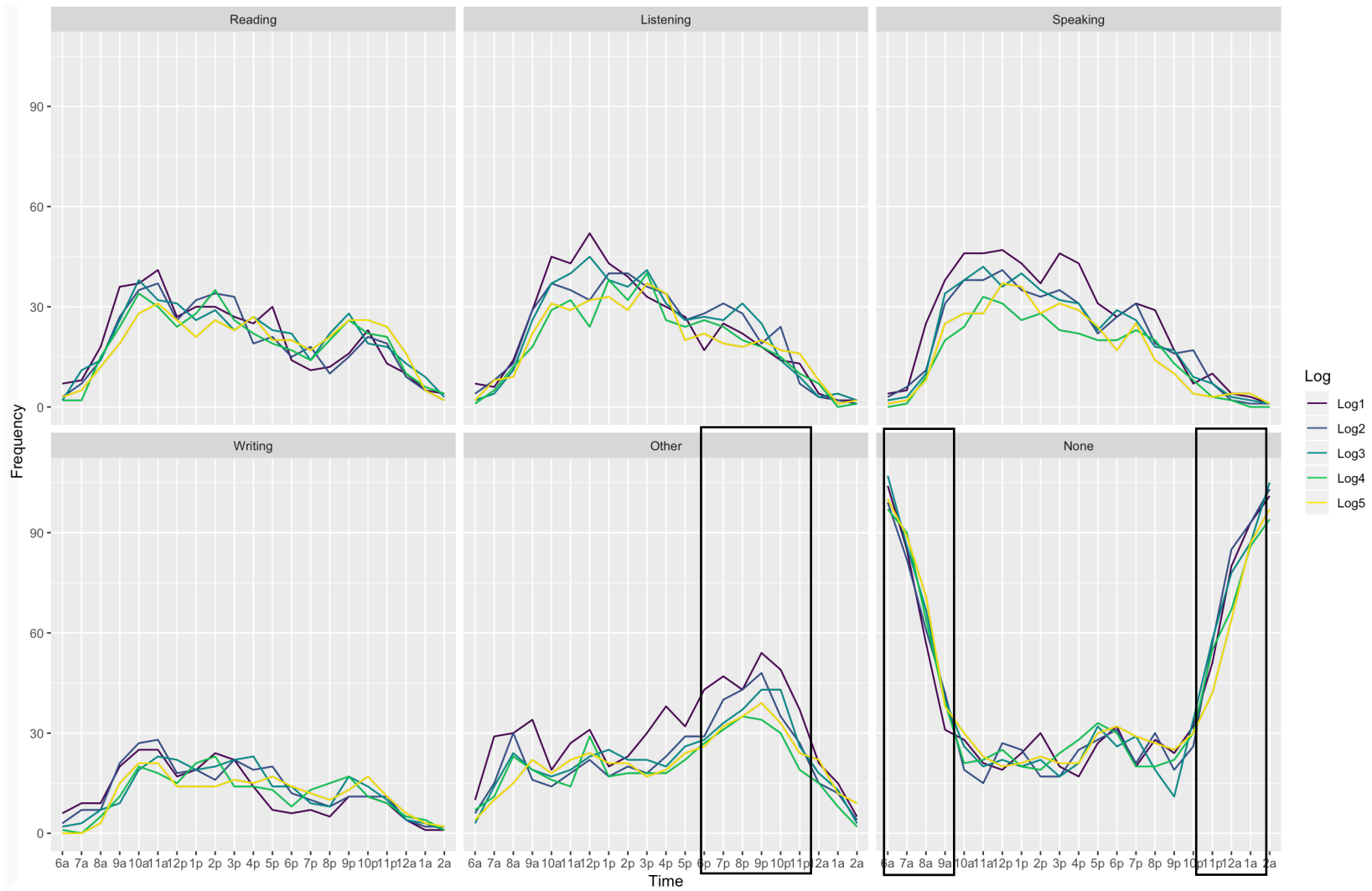


Figure 3.2. Mean frequency counts of language engagement data across time for reading, listening, speaking, and writing in English; using other languages; and no language use. Boxes demarcate time frames with relatively high amounts of reported engagement

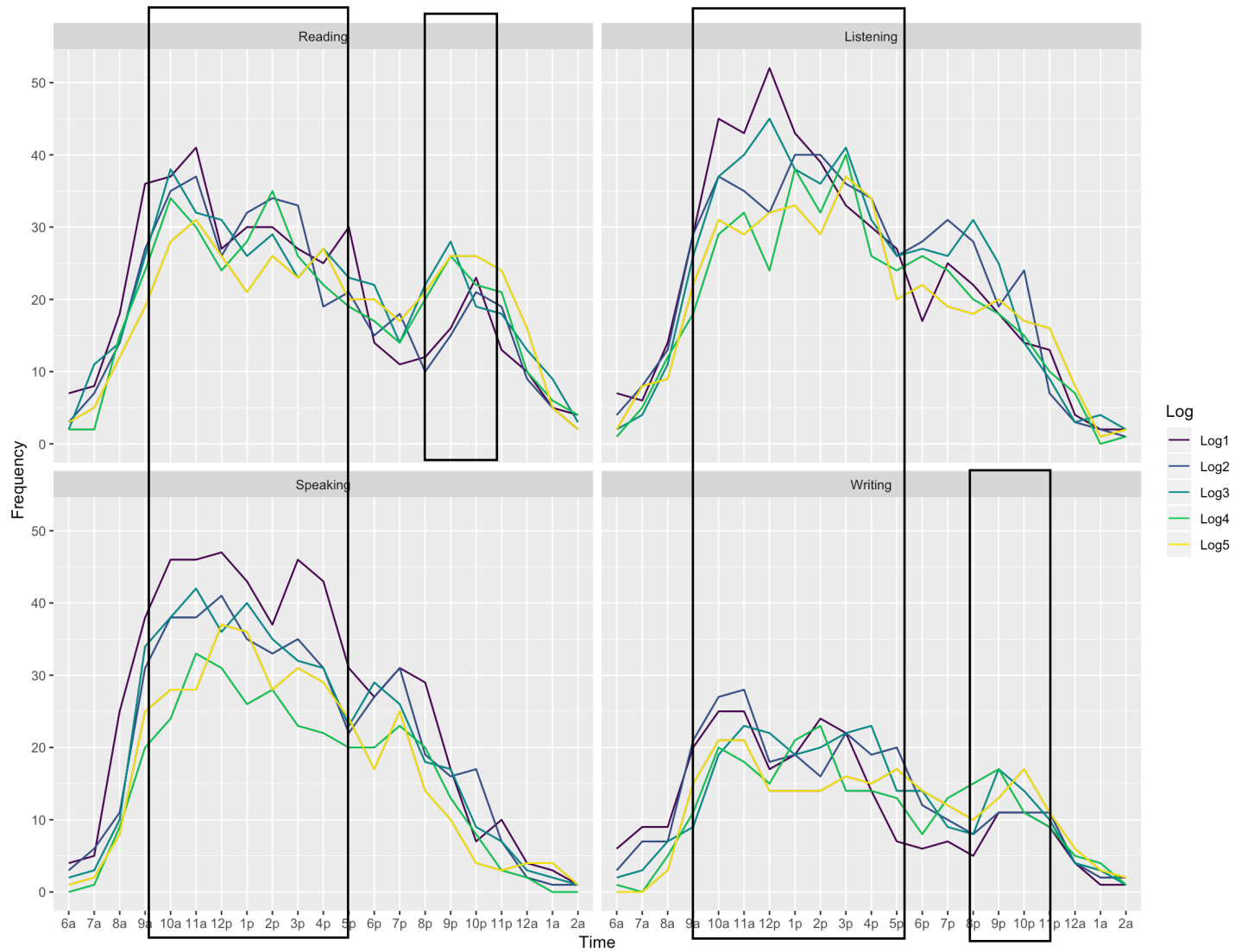


Figure 3.3. Mean frequency counts of language engagement data across time for reading, listening, speaking, and writing in English.

Boxes demarcate timeframes with relatively high amount of reported language engagement

On the other hand, there is also a steep increase in the amount of time spent using other languages in the evening. The Other panel displayed in Figure 3.2 shows that the time window with the highest amount of Other Language use (including L1 or L3) spans roughly from 6 p.m. to 11 p.m. This is plausibly, though speculatively, the time the international students spend after work communicating with family or roommates at home or abroad in a mutual language other than English. Lastly, there is a steep decline in the early morning hours in the time spent without using languages mirrored by an equally steep increase in the late evenings. This pattern is collectively seen in all five logs. Recall that activities categorized as No-Language use include sleeping, eating, doing nothing, exercising, chores and daily tasks (e.g., cooking, doing laundry, and cleaning the house, office, or yard/packing bag/organizing materials). As evident from the None panel in Figure 3.2, international students show a high level of no language engagement between 6 a.m. and 9 a.m. and from 11 p.m. to 3 a.m. In particular, the amount of no-language engagement in these two time frames shows a sharp drop-off as the time approaches the start of working hours (9 a.m.) and a steep peak after 11 p.m., which is near bedtime. Hence, the high levels of no-language engagement in the late evenings and early mornings may reflect the time in which learners are preparing to be or are asleep, respectively.

### ***3.3.3 Variation among learners***

I will now take a closer look at the extent to which individual learners vary in the quantity of language engagement. As previously noted, I observed a wide range of variability in the amount and types of language engagement (see Table 3.1 on p. 41). The amount of English and No language use ranged from 0.60 to 18.50 and from 0.25 to 17.50 hours/day, respectively. This indicates that, in a 21-hour day, some learners reported to have used English 36 minutes and others up to 18 hours and 30 minutes; or did not engage in languages anywhere from 15 minutes

and up to 17 hours and 30 minutes a day. These ranges are non-trivial. To examine the finer details of the amount of English engagement, I plotted each individual's English engagement patterns to visualize learner variability. Figure 3.4 plots 110 participants' English engagement patterns from 6 a.m. to 3 a.m.<sup>5</sup> What is perhaps most apparent and interesting in this plot is the substantial variability across individuals in the overall amount of English exposure. For instance, a drastic difference of these two features can be observed in participant 47 and 109's engagement: The total amount of English use is 36 minutes and 16.12 hours for participant 47 and 109, respectively.

The data can also be viewed in three groups: (1) Participants who are (almost) completely immersed in English (e.g., participants number 2, 10, 19, 22, 34, 41, etc), (2) those who have high engagement levels during working hours (e.g., most participants fall in this category), and lastly, (3) those who mostly use other languages (e.g., participants number 47, 110, etc).

In summary, this plot provides data on individual differences on the amount of engagement with the target language even in a naturalistic context where most learners are given relatively equal amount of affordances to engage with and be exposed to the target language. From the results, it is clear that immersion in a naturalistic setting does not necessarily result in comparable amounts of actual L2 use for all learners. I will revisit the implications of this finding in Chapter 5 Discussion.

---

<sup>5</sup> Participants with all five language logs are included in the plot.

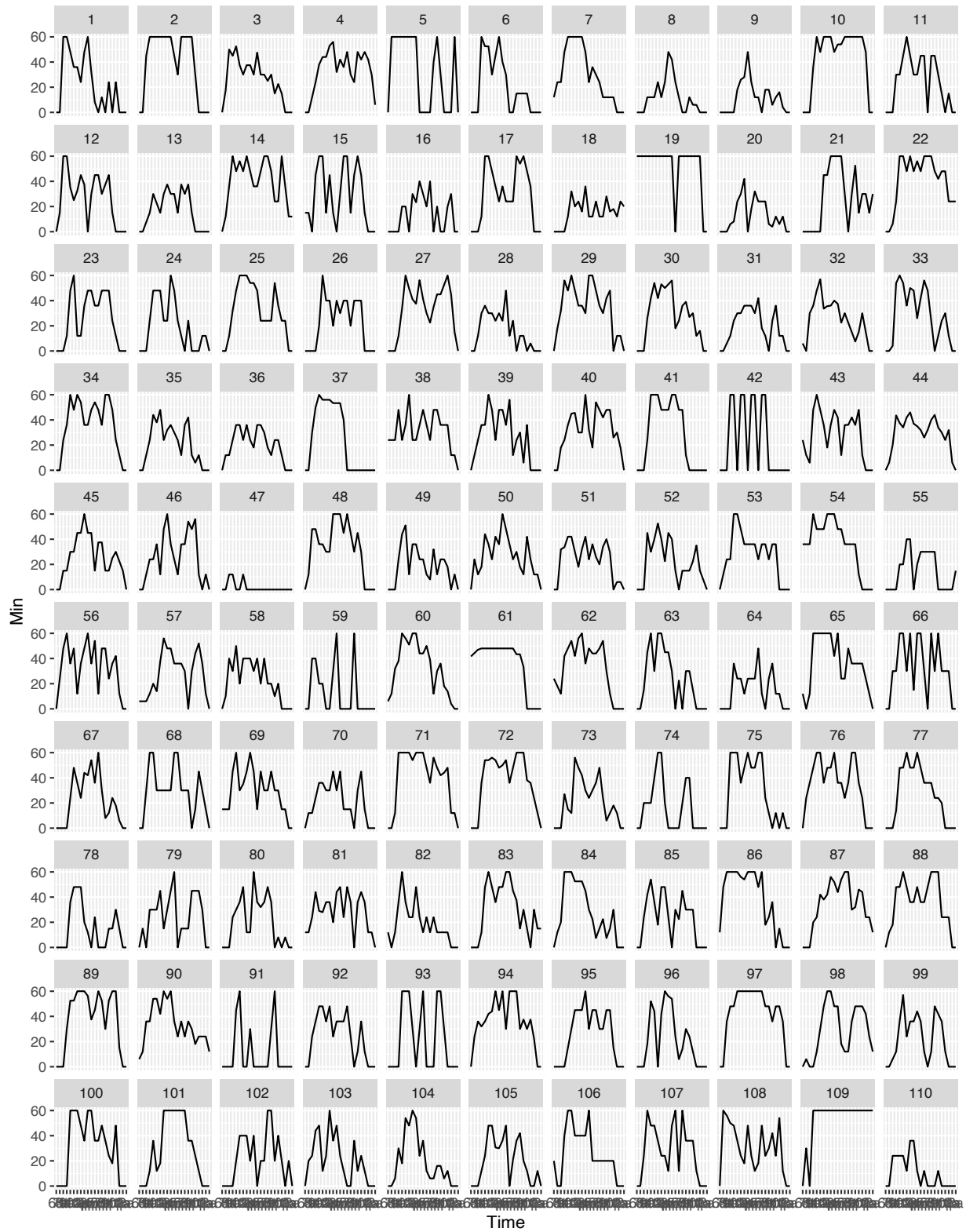


Figure 3.4. English engagement patterns for each hour from 6 a.m. to 3 a.m. by individuals



## CHAPTER 4: LINGUISTIC KNOWLEDGE AND LANGUAGE USE

The aim of this chapter was twofold. First, I examined the association between different types of English knowledge of international students in US higher education. In doing so, I address the interface question of whether explicit knowledge influences the development of implicit knowledge. The second goal was to understand how different types of activities relate to different types of knowledge; in particular, does language-focused and/or meaning-focused activity facilitate the development of explicit and/or implicit knowledge?

### 4.1 Research Questions

The two research questions addressed by the results in this chapter are the following:

RQ2. To what extent does explicit knowledge influence the development of implicit knowledge?

RQ3. To what extent do the types of *activity* (i.e., language-focused and meaning-focused) contribute to the development of different types of knowledge?

The following subsections provide information on the linguistic data on explicit and implicit English grammar knowledge at Time 1 (T1) and Time 2 (T2). I first present data preparation processes, including handling of missing data, item reliability analyses, and detailed summary statistics of all linguistic measures for the two time points. Then, I report a series of path analyses to address the two research questions.

### 4.2 Analysis

I ran two sets of cross-lagged path models (CLPM) to examine the causal influences between knowledge types over time. The term *cross* refers to paths that cross over from one variable to another (e.g., paths c and d in Figure 4.1) and *lagged* indicates the temporary

separation between the constructs. The *cross-lagged effects* then refer to the predictive effects of explicit knowledge at T1, for instance, on the implicit knowledge development at T2 (path d). It is important that the cross-lagged effect accounts for the residual variance at T2; that is, the amount of variance that is left unexplained after accounting for the effects of implicit knowledge at T1. As such, I added two autoregressive paths (e.g., paths a and b) to the cross-lagged model to control for level of the variable being predicted. Adding these two autoregressive paths is important as it offsets claims that the effects of cross-lag are simply due to the fact that explicit knowledge and implicit knowledge at T1 are highly correlated.

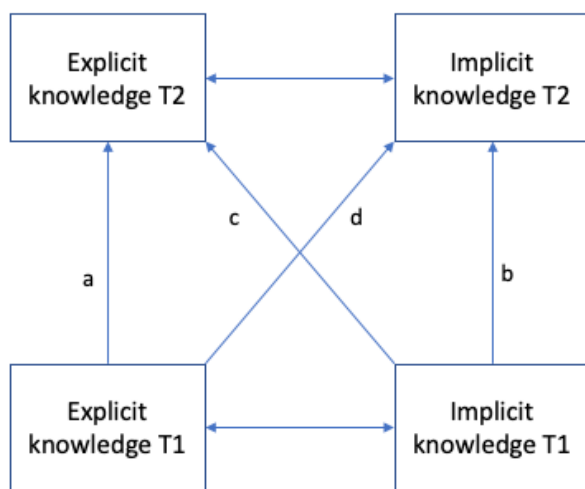


Figure 4.1. Path diagram for a two-wave, two-variable path model. T1 = time 1; T2 = time 2; a & b = autoregressive paths; c & d = cross-lagged paths

The path models utilized factor scores obtained from two confirmatory factor analyses (CFAs). Factor scores are composite variables with which researchers control for measurement errors by only extracting construct relevant scores from measurement instruments using CFA. For this reason, I modeled two CFAs (CFA at T1 and T2) where the explicit knowledge

construct was composed of two linguistic tests (i.e., UGJT, MKT) and implicit knowledge was modeled with three instruments (i.e., EI, OP, TGJT). The extracted factor scores were used as observed variables in the cross-lagged path models.

I evaluated the CFA and CLPM models based on two major aspects: (1) global goodness of fit and (2) presence or absence of localized strain. The overall goodness of fit indices provide a (global) summary of the acceptability of the model; that is, whether the model has been properly specified. Model fit indices considered were the  $\chi^2$  statistic (and the corresponding degree of freedom and  $p$  value), root mean square error of approximation (RMSEA, corrects for model complexity taking sample size into account) and its 90% confidence interval, the standardized root mean square residual (SRMR), and comparative fit index (CFI, compares the fitted model to a base model with no parameter restrictions). I followed Hu and Bentler's (1999) and Kline's (2016) guidelines for fit interpretation (i.e., RMSEA lower bound confidence interval value  $\leq 0.06$ , which yields a nonsignificant  $p$  value, SRMR values  $\leq 0.08$ , and CFI values  $\geq .95$ ).

If these model fit indices are poor (suggesting that the model significantly departs from the data), researchers pursue subsequent evaluations to diagnose/identify local areas of misspecification. I used two statistics to identify localized areas of ill fit: residuals and the modification index (MI). The standardized residuals provide specific information about the differences between model-implied and sample covariances. Generally, larger differences suggest overestimation or underestimation of the difference between model-implied and sample covariances. Since standardized residuals are  $z$ -scores, indices above the value of  $|1.96|$  are flagged, as 1.96 corresponds to a statistically significant  $z$ -score at  $p < .05$  (Brown, 2015). The MIs reflect how well the model fit ( $\chi^2$ ) would improve if a new path is added to the model.

Indices of 3.84 or greater suggest that the addition of such a path will statistically improve the overall fit of the model (a value of 3.84 indexes  $\chi^2$  at  $p < .05$ ; Brown, 2015).

The Interface and Non-interface models were compared using the above criteria. A better statistical fit of the Interface model, together with a positive, significant path from explicit knowledge (T1) to implicit knowledge (T2) would lend support to the interface position. Furthermore, the paths extending from language-focused and meaning-focused activity, respectively, will speak to the importance of amount and type of language activity in the development of linguistic knowledge. All CFA and CLPM analyses were carried out in *R* version 1.2.1335 using the *lavaan* package.

### **4.3 Data Preparation**

#### **4.3.1 Missing data**

As in most longitudinal studies, there were multiple missing values. In particular, there were missing test scores within and across sessions. For instance, participants completed the web-based tasks but did not come to the lab session to complete the remaining tests (missing test data within session) and there were participant dropouts (missing data across sessions). As seen in Table 4.1, the missingness of individual measures of 149 participants, at a test score level, ranged from 0.67% to 24.83%, with an average of 14.83%, which is not uncommon in longitudinal studies (see Schoonen, van Gelderen, Stoel, Hulstijn & de Glopper, 2011). The highest percentage of missing data came from oral productions at T1 and T2, with 14 (out of 26) and 6 (out of 37) missing data due to technical errors (i.e., poor quality or corrupted file). Regarding missing data across time, the overall attrition rate was 25.5% between T1 and T2 ( $n = 149$  at T1;  $n = 122$  at T2).

Table 4.1 Missing Data of 149 Participants in T1 And T2

| Measures | Time | Missing (n) | Missing (%) |
|----------|------|-------------|-------------|
| EI       | 1    | 10          | 0.67%       |
| OP       | 1    | 26          | 17.45%      |
| TGJT     | 1    | 8           | 5.37%       |
| UGJT     | 1    | 8           | 5.37%       |
| MKT      | 1    | 8           | 5.37%       |
| EI       | 2    | 34          | 22.82%      |
| OP       | 2    | 37          | 24.83%      |
| TGJT     | 2    | 31          | 20.81%      |
| UGJT     | 2    | 31          | 20.81%      |
| MKT      | 2    | 28          | 18.79%      |

Note. EI = elicited imitation; OP = oral production; TGJT = timed written grammaticality judgment task; UGJT = untimed written grammaticality judgment task; MKT = metalinguistic knowledge test

In handling missing values at the test levels, I used a model-based approach (i.e., full-information maximum likelihood estimation) that produces parameter estimates of models in the presence of missing data. This meant all 149 participants' data from T1 could be included without needing to remove the 28 dropouts. An important assumption made in this approach is that the data are missing at random (MAR) or completely at random (MCAR). While MCAR is not likely to be met with longitudinal data since missingness is typically correlated with

proficiency, MAR can be met when the missing data can be predicted from observed data (Little & Rubin, 1987). In my data, missing patterns are partially traceable by the types of missing tests. The left panel of Figure 4.2 visualizes the proportion of missing values across two time points and the right panel provides combinations of missing data patterns. The red-filled squares represent missing values and numbers on the right indicate participants who fall in a given category. As visualized in the right panel of Figure 4.2, seven participants without TGJT, UGJT, and MKT scores (the three lab-based tests) at T1 dropped out from the experiment; 14 participants without OP scores at T1 decided to discontinue the study. In addition to this, all variables in this data set are sufficiently correlated, which is an essential element for drawing reliable estimates of missing scores. With this evidence in mind, I discerned the data to be MAR and carried out a full-information maximum likelihood estimation in evaluating different models.

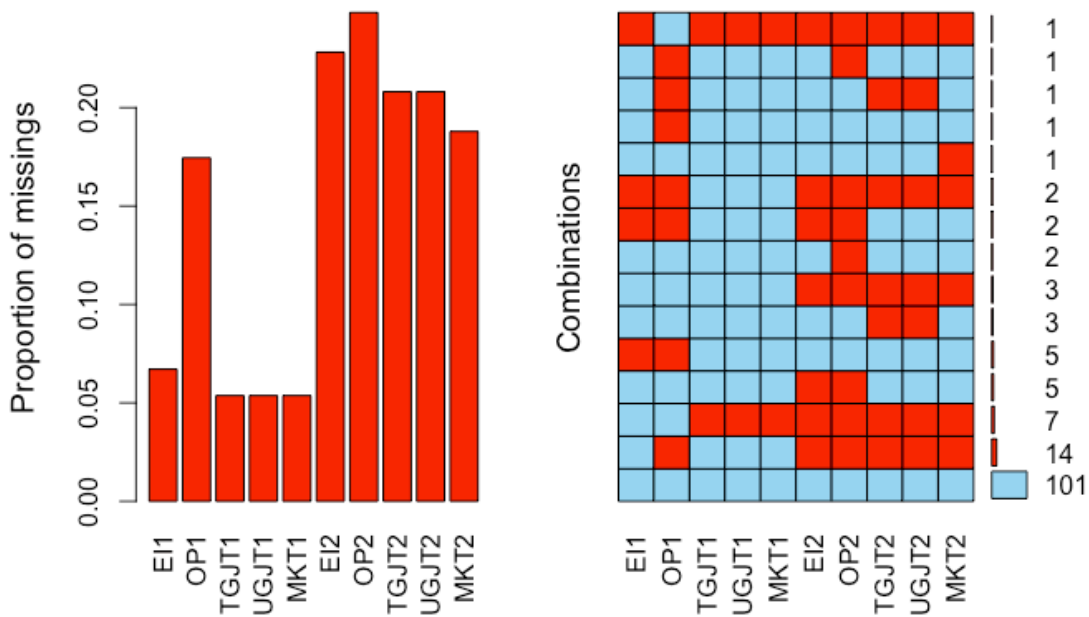


Figure 4.2. (left panel) Proportion of missing values across two time points; (right panel) Combinations of missing data patterns. Numbers on the right indicate participants who fall in a given category. Red-filled squares represent missing values

#### 4.3.2 Item reliability

For item reliability, I first inspected the item discrimination index (item-total correlation) using the *alpha()* function in psych package and removed items with negative or low values. This resulted in a number of item removals (see Table 4.2) with coefficient alphas of each test ranging within the acceptable range .62 to .79. The largest number of item removals were from the GJTs. In particular, nine items were removed from UGJT at T1, seven items from TGJT at T2, and six items from UGJT at T2. All removed items (except one item from UGJT at T2) were grammatical sentences. This finding—low correlation between learner performance on grammatical and ungrammatical items—provides some empirical support to previous findings that reported grammaticality impacts L2 learners' judgment results (e.g., Guitérrez, 2013; Vafae

et al., 2016) and adheres to a common convention to only include ungrammatical items in the UGJT (i.e., R. Ellis, 2005; R. Ellis & Loewen, 2007) or ungrammatical items in both the TGJT and UGJT (i.e., Guitérrez, 2013; Vafaei et al., 2016).

All alpha and omega values after initial item removal seemed satisfactory and thus I did not remove any further items beyond this point. For oral production and metalinguistic knowledge test, interrater reliability of two raters was also computed. Table 4.2 displays coefficient alpha and omega before and after item removal and interrater reliability values.<sup>6</sup>

---

<sup>6</sup> In addition to the coefficient alpha, I computed coefficient omega (McDonald, 1999) for each linguistic measure with items as observed variables. Cronbach's alpha can be too conservative as it imposes factor loadings for each item to be equal (tau-equivalence). Omega does not impose this assumption and thus is appropriate to use when loadings vary (congeneric). This may be the case for this study as the six target structures that comprised the total score of each test varied in the level of complexity (early and late acquired) and linguistic components (syntactic and morphological).



Table 4.2 Coefficient Alpha and Coefficient Omega

|        | Measures | Reliability with all items |             | Reliability with removed items |                          |
|--------|----------|----------------------------|-------------|--------------------------------|--------------------------|
|        |          | <i>k</i>                   | alpha/omega | <i>k</i>                       | alpha/omega              |
| Time1  | EI       | 24                         | 0.67/0.84   | 23                             | .67/.84                  |
|        | TGJT     | 24                         | 0.64/0.81   | 21                             | .68/.83                  |
|        | UGJT     | 24                         | 0.50/0.85   | 15                             | .68/.85                  |
|        | MKT      | 12                         | 0.79/0.93   | 12                             | .79/.93/.89 <sup>a</sup> |
|        | OP       | -                          | -           | -                              | .93 <sup>a</sup>         |
| Time 2 | EI       | 24                         | 0.65/0.84   | 19                             | .69/.87                  |
|        | TGJT     | 24                         | 0.49/0.8    | 17                             | .62/.71                  |
|        | UGJT     | 24                         | 0.58/0.75   | 18                             | .64/.84                  |
|        | MKT      | 12                         | 0.76/0.91   | 10                             | .77/.93/.92 <sup>a</sup> |
|        | OP       | -                          | -           | -                              | .91 <sup>a</sup>         |

Note. OP = oral production; EI = elicited imitation; TGJT = timed written grammaticality judgment task; UGJT = untimed written grammaticality judgment task; MKT = metalinguistic knowledge test. <sup>a</sup> Pearson *r* inter-rater reliability.

#### 4.4 Results

In this section I provide the overall descriptive results of all linguistic measures at T1 and T2. This is followed by a descriptive report of the linguistic measures by linguistic structure, with visualization of performance changes from T1 to T2 separately by tests. I then report a series of CFA and Path analysis results.

#### 4.4.1 Descriptive Statistics for Language Tests

Table 4.3 summarizes descriptive statistics of the five linguistic variables at T1 and T2. Numerically, a general improvement in average scores is observed between two time points for all tests. The univariate skewness and kurtosis values for many measures at T2 exceed the acceptable range of  $\pm 1$ , with untimed GJT and oral production being extreme. Different versions of multivariate normality tests (Mardia, Henze-Zirkler's, and Doornik-Hansen's MVN tests) collectively flagged the data to be disproportionate (all  $p < .001$ ). To accommodate assumption violations, I used robust maximum likelihood (MLR) estimator method for the CFA.

Table 4.3 Descriptive Information of Five Linguistic Tests at T1 and T2

| Tests | Time | n   | Mean | SD   | 95% CI     | Min  | Max  | Skewness | Kurtosis |
|-------|------|-----|------|------|------------|------|------|----------|----------|
| MKT   | T1   | 141 | 0.36 | 0.25 | [.32, .40] | 0.00 | 1.00 | 0.62     | -0.33    |
| MKT   | T2   | 121 | 0.60 | 0.28 | [.55, .65] | 0.00 | 1.00 | -0.19    | -1.15    |
| UGJT  | T1   | 141 | 0.65 | 0.18 | [.62, .68] | 0.20 | 1.00 | -0.24    | -0.62    |
| UGJT  | T2   | 118 | 0.80 | 0.14 | [.78, .83] | 0.11 | 1.00 | -1.43    | 3.71     |
| TGJT  | T1   | 141 | 0.56 | 0.17 | [.54, .59] | 0.19 | 1.00 | 0.13     | -0.42    |
| TGJT  | T2   | 118 | 0.59 | 0.18 | [.56, .62] | 0.24 | 1.00 | 0.10     | -0.76    |
| EI    | T1   | 139 | 0.64 | 0.15 | [.62, .67] | 0.17 | 0.96 | -0.33    | -0.01    |
| EI    | T2   | 115 | 0.69 | 0.17 | [.66, .72] | 0.21 | 1.00 | -0.37    | -0.40    |
| OP    | T1   | 123 | 0.89 | 0.13 | [.87, .91] | 0.33 | 1.00 | -1.93    | 4.74     |
| OP    | T2   | 112 | 0.93 | 0.07 | [.91, .94] | 0.50 | 1.00 | -2.39    | 10.17    |

Note. OP = oral production; EI = elicited imitation; TGJT = timed written grammaticality

judgment task; UGJT = untimed written grammaticality judgment task; MKT = metalinguistic knowledge test; T1 = Time1; T2 = Time2

#### 4.4.1.1 Elicited Imitation

Figure 4.3 plots individuals' performance changes in the elicited imitation task at T1 and T2. As might be expected, bi-directional changes are observed, with some participants showing (steep) improvements from T1 to T2 while others performed poorly or maintained comparable scores at T2. While the median (bolded line inside boxplots) seems comparable at both time points, participants performed significantly better, as a group, at T2,  $t = -3.6784$ ,  $df = 109$ ,  $p < .001$ ,  $d = 0.311$ . Table 4.4 displays descriptive information of elicited imitation by structures.

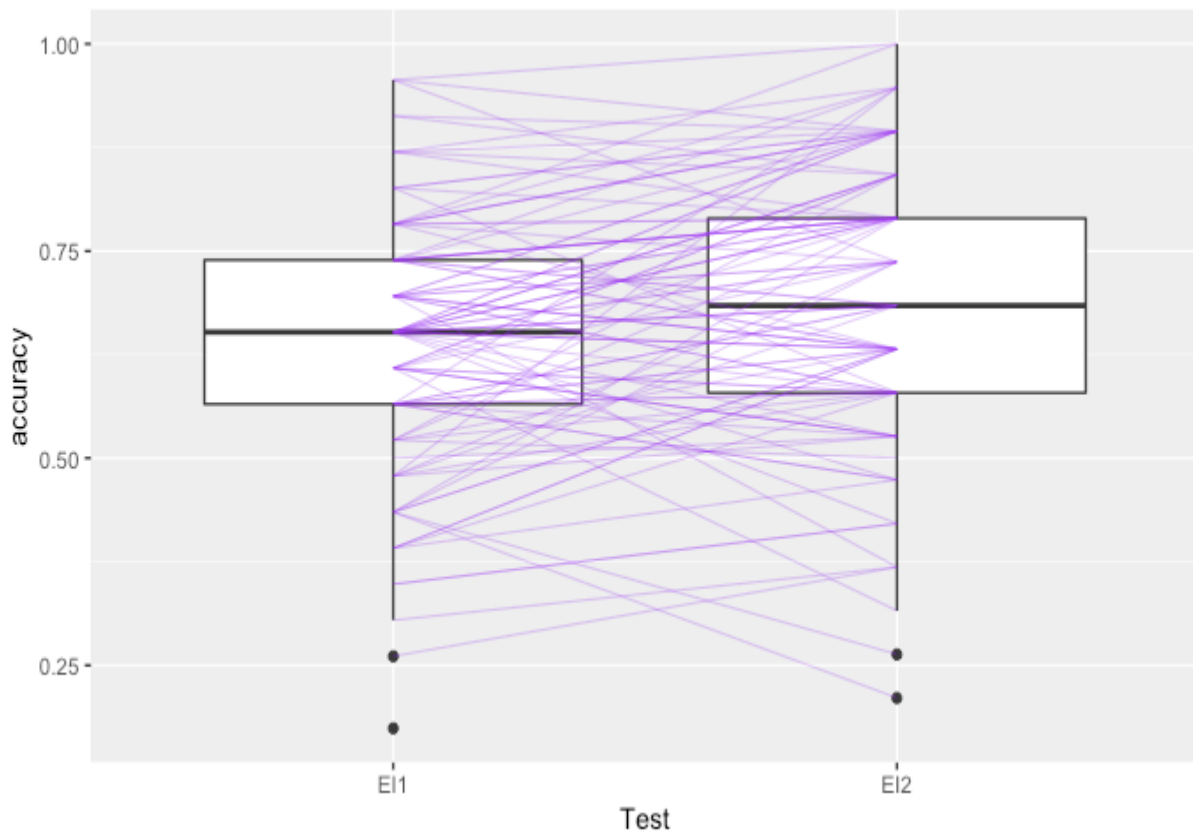


Figure 4.3. Spaghetti plot of Elicited Imitation at T1 and T2

Table 4.4 Descriptive Statistics for the Elicited Imitation at T1 and T2

| Time   | Structure    | Gramm | Mean  | SD    |
|--------|--------------|-------|-------|-------|
| Time 1 | Third Person | G     | 0.658 | 0.475 |
|        |              | UG    | 0.524 | 0.5   |
|        | Mass/Count   | G     | 0.694 | 0.462 |
|        | Nouns        | UG    | 0.507 | 0.501 |
|        | Be Passive   | G     | 0.77  | 0.422 |
|        |              | UG    | 0.482 | 0.501 |
|        | Embedded     | G     | 0.978 | 0.146 |
|        | Questions    | UG    | 0.691 | 0.463 |
|        | Comparatives | G     | 0.835 | 0.372 |
|        |              | UG    | 0.478 | 0.5   |
|        | Verb         | G     | 0.87  | 0.337 |
|        | Complements  | UG    | 0.367 | 0.483 |
| Time 2 | Third Person | G     | 0.774 | 0.419 |
|        |              | UG    | 0.548 | 0.5   |
|        | Mass/Count   | G     | 0.713 | 0.453 |
|        | Nouns        | UG    | 0.557 | 0.498 |
|        | Be Passive   | G     | 0.778 | 0.416 |
|        |              | UG    | 0.557 | 0.498 |
|        | Embedded     | G     | 0.965 | 0.184 |
|        | Questions    | UG    | 0.548 | 0.5   |
|        | Comparatives | G     | 0.856 | 0.352 |
|        |              | UG    | 0.5   | 0.501 |
|        | Verb         | G     | 0.887 | 0.318 |
|        | Complements  | UG    | 0.661 | 0.475 |

Note. G = grammatical; UG = ungrammatical

#### 4.4.1.2 Oral Production

Figure 4.4 plots individuals' performance changes in oral production at two time points. As reflected in the wide ranges of skewness and kurtosis reported in Table 4.3, scores are disproportionate in both time points with upper and lower whiskers of the boxplot stretched out unevenly alongside different sizes of boxes. Most participants scored above 60% at T1 and 80% at T2 with group means near ceiling ( $M = 89\%$  and  $M = 93\%$ , respectively). A Wilcoxon signed rank test shows that, at the group level, participants showed a significant improvement at T2 ( $V = 1437.5, p > .001$ ) with an effect size of  $d = 0.383$ . Table 4.5 presents descriptive information of oral production by structure.

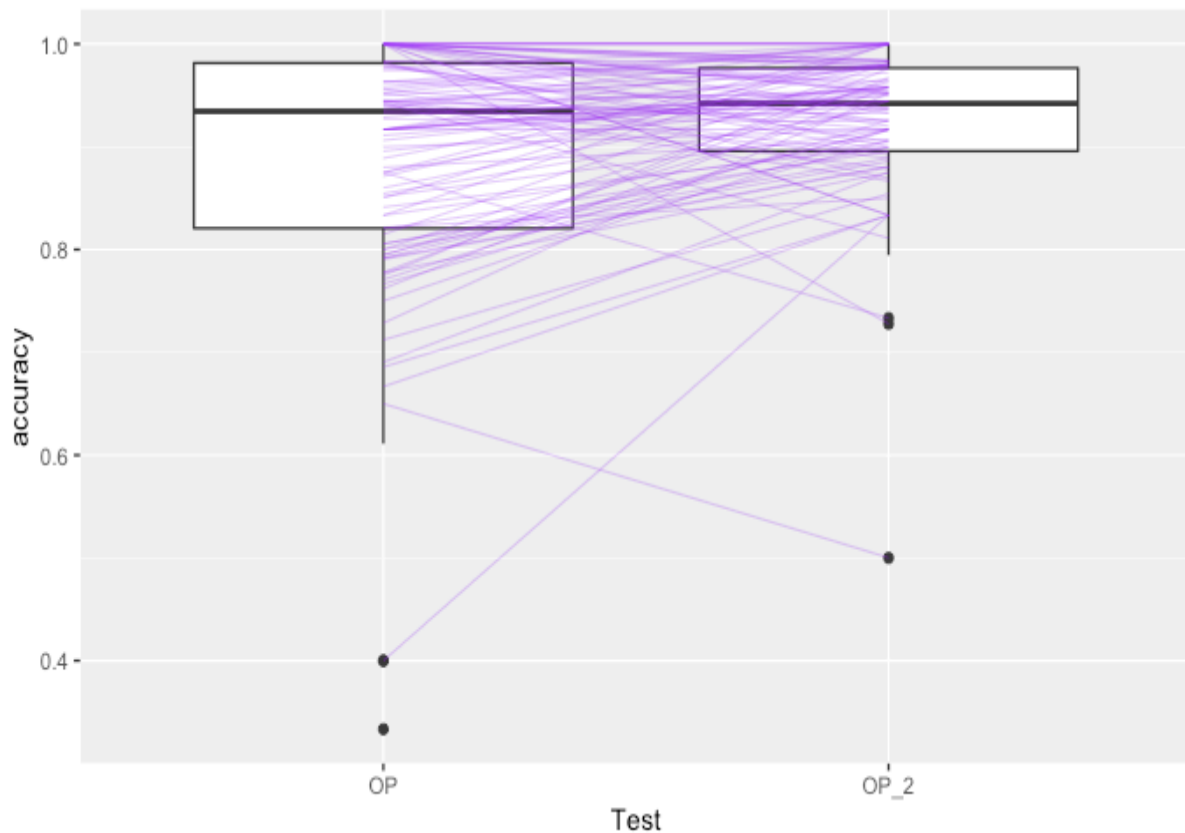


Figure 4.4. Spaghetti plot of Oral Production at T1 and T2

Table 4.5 Descriptive Statistics for the Oral Production at T1 and T2

| Time   | Structure             | Mean | SD    |
|--------|-----------------------|------|-------|
| Time 1 | Third Person          | 0.71 | 0.29  |
|        | Mass/Count<br>Nouns   | 0.98 | 0.11  |
|        | Be Passive            | 0.94 | 0.21  |
|        | Embedded<br>Questions | 0.90 | 0.26  |
|        | Comparatives          | 0.89 | 0.27  |
|        | Verb<br>Complements   | 0.99 | 0.06  |
| Time 2 | Third Person          | 0.67 | 0.27  |
|        | Mass/Count<br>Nouns   | 0.99 | 0.05  |
|        | Be Passive            | 1.00 | 0.00  |
|        | Embedded<br>Questions | 0.97 | 0.14  |
|        | Comparatives          | 0.97 | 0.13  |
|        | Verb<br>Complements   | 0.99 | 0.007 |

#### 4.4.1.3 Timed Written GJT

#### 4.4.1.3 Timed Written GJT

The boxplots on timed written GJT scores (Figure 4.5) picture a normal distribution of data with the sizes of the two boxplots comparable and whiskers stretched out evenly in both time points. A paired sample  $t$ -test returned a nonsignificant result,  $t = -1.54$ ,  $df = 117$ ,  $p = 0.127$ ,  $d = 0.17$ , suggesting there were no discernable performance differences between T1 and T2.

Table 4.6 displays the mean scores by structure.

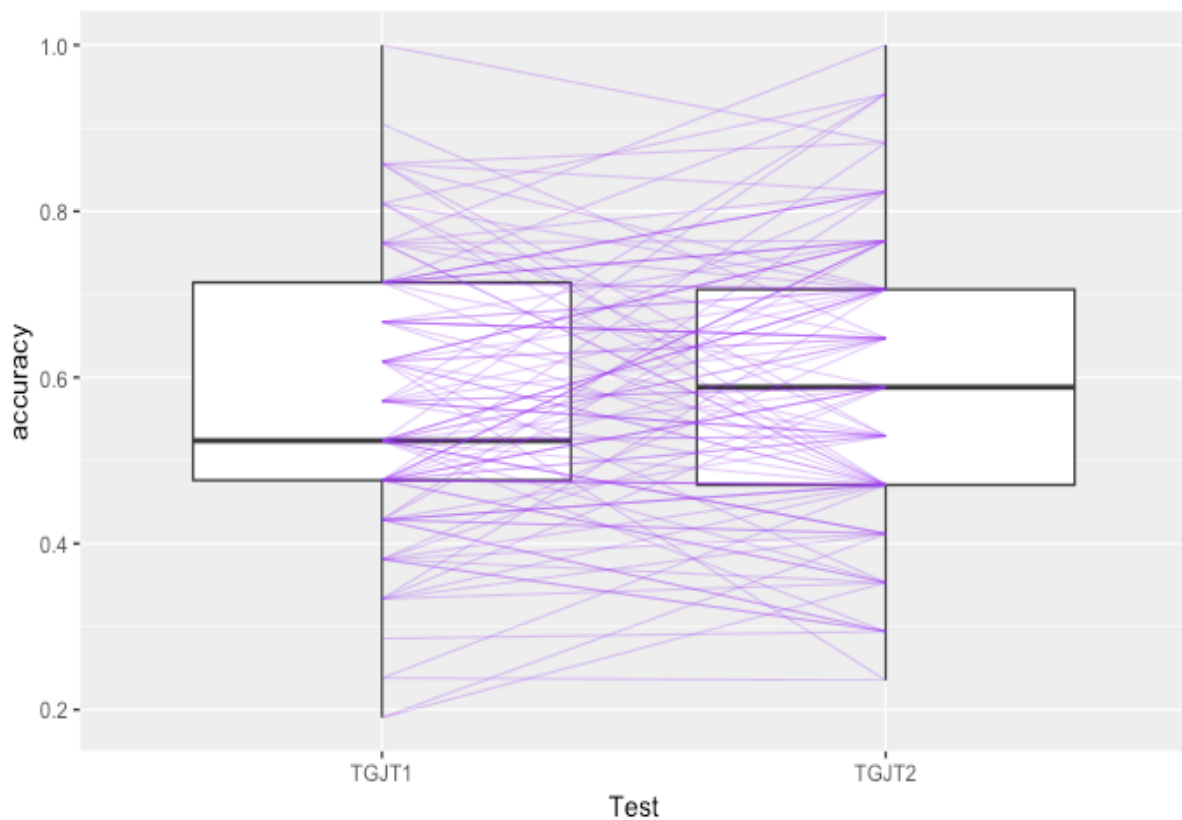


Figure 4.5. Spaghetti plot of Timed Written GJT at T1 and T2

Table 4.6 Descriptive Statistics for the Timed Written GJT at T1 and T2

| Time   | Structure             | Gramm | Mean  | SD    |
|--------|-----------------------|-------|-------|-------|
| Time 1 | Third Person          | G     | 0.727 | 0.446 |
|        |                       | UG    | 0.511 | 0.501 |
|        | Mass/Count<br>Nouns   | G     | 0.759 | 0.429 |
|        |                       | UG    | 0.337 | 0.473 |
|        | Be Passive            | G     | 0.759 | 0.429 |
|        |                       | UG    | 0.55  | 0.498 |
|        | Embedded<br>Questions | G     | 0.684 | 0.466 |
|        |                       | UG    | 0.383 | 0.487 |
|        | Comparatives          | G     | 0.745 | 0.437 |
|        |                       | UG    | 0.535 | 0.5   |
|        | Verb<br>Complements   | G     | 0.539 | 0.5   |
|        |                       | UG    | 0.426 | 0.495 |
| Time 2 | Third Person          | UG    | 0.432 | 0.496 |
|        |                       | G     | 0.568 | 0.497 |
|        | Mass/Count<br>Nouns   | UG    | 0.394 | 0.49  |
|        |                       | G     | 0.856 | 0.352 |
|        | Be Passive            | UG    | 0.589 | 0.493 |
|        |                       | G     | 0.525 | 0.5   |
|        | Embedded<br>Questions | UG    | 0.525 | 0.5   |
|        |                       | G     | 0.788 | 0.409 |
|        | Comparatives          | UG    | 0.627 | 0.485 |
|        |                       | G     | 0.525 | 0.5   |
|        | Verb<br>Complements   | UG    | 0.525 | 0.5   |

Note. G = grammatical; UG = ungrammatical

#### 4.4.1.4 Untimed Written GJT

Figure 4.6 plots the untimed written GJT scores at T1 and T2. As with other tasks, bidirectional changes are observed, with many participants showing some degree of improvement, while others showing deterioration at T2. A non-parametric t-test suggests a significant improvement at T2,  $V = 682.5$ ,  $p < .001$ ,  $d = 0.93$ . Table 4.7 presents descriptive statistics of the UGJT scores by target forms.



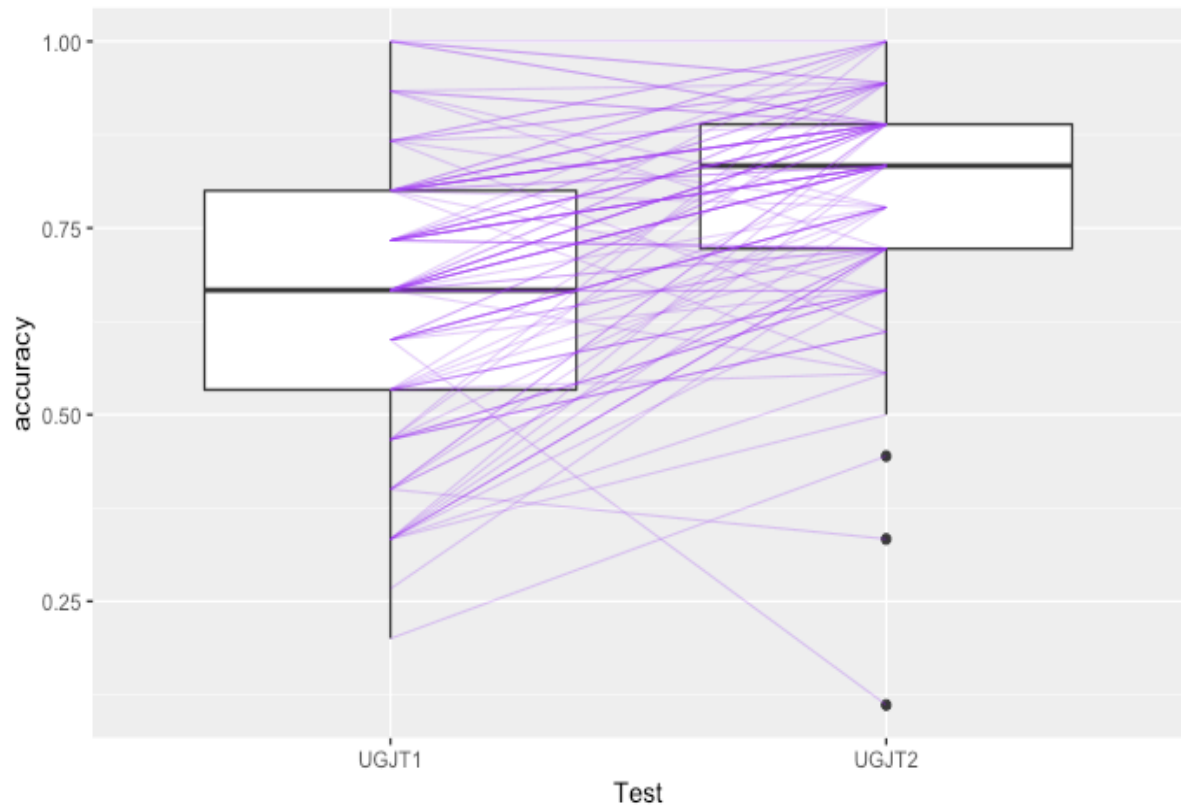


Figure 4.6. Spaghetti plot of Untimed Written GJT at T1 and T2

Table 4.7 Descriptive Statistics for the Untimed Written GJT at T1 and T2

| Time   | Structure        | Gramm | Mean  | SD    |
|--------|------------------|-------|-------|-------|
| Time 1 | Third Person     | UG    | 0.77  | 0.422 |
|        | Mass/Count Nouns | G     | 0.468 | 0.501 |
|        |                  | UG    | 0.34  | 0.475 |
|        | Be Passive       | G     | 0.936 | 0.245 |
|        |                  | UG    | 0.691 | 0.463 |
|        | Embedded         | G     | 0.879 | 0.327 |
|        | Questions        | UG    | 0.461 | 0.499 |
|        | Comparatives     | UG    | 0.773 | 0.42  |
| T2     | Verb Complements | UG    | 0.723 | 0.448 |
|        | Third Person     | G     | 0.856 | 0.352 |
|        |                  | UG    | 0.686 | 0.465 |
|        | Mass/Count Nouns | G     | 0.898 | 0.304 |
|        |                  | UG    | 0.627 | 0.486 |
|        | Be Passive       | G     | 0.915 | 0.279 |
|        |                  | UG    | 0.843 | 0.364 |
|        | Embedded         | G     | 0.932 | 0.252 |
|        | Questions        | UG    | 0.712 | 0.454 |
|        | Comparatives     | UG    | 0.797 | 0.403 |
|        | Verb Complements | G     | 0.958 | 0.202 |
|        |                  | UG    | 0.708 | 0.456 |

Note. G = grammatical; UG = ungrammatical

#### 4.4.1.5 Metalinguistic Knowledge Test

Lastly, Figures 4.7 visualizes the distribution of test scores on MKT. A noticeable improvement is observed for MKT scores at T2, which was confirmed statistically with large effect size,  $t = -12.127$ ,  $df = 120$ ,  $p < .001$ ,  $d = 0.90$ . Table 4.8 presents detailed descriptive information.

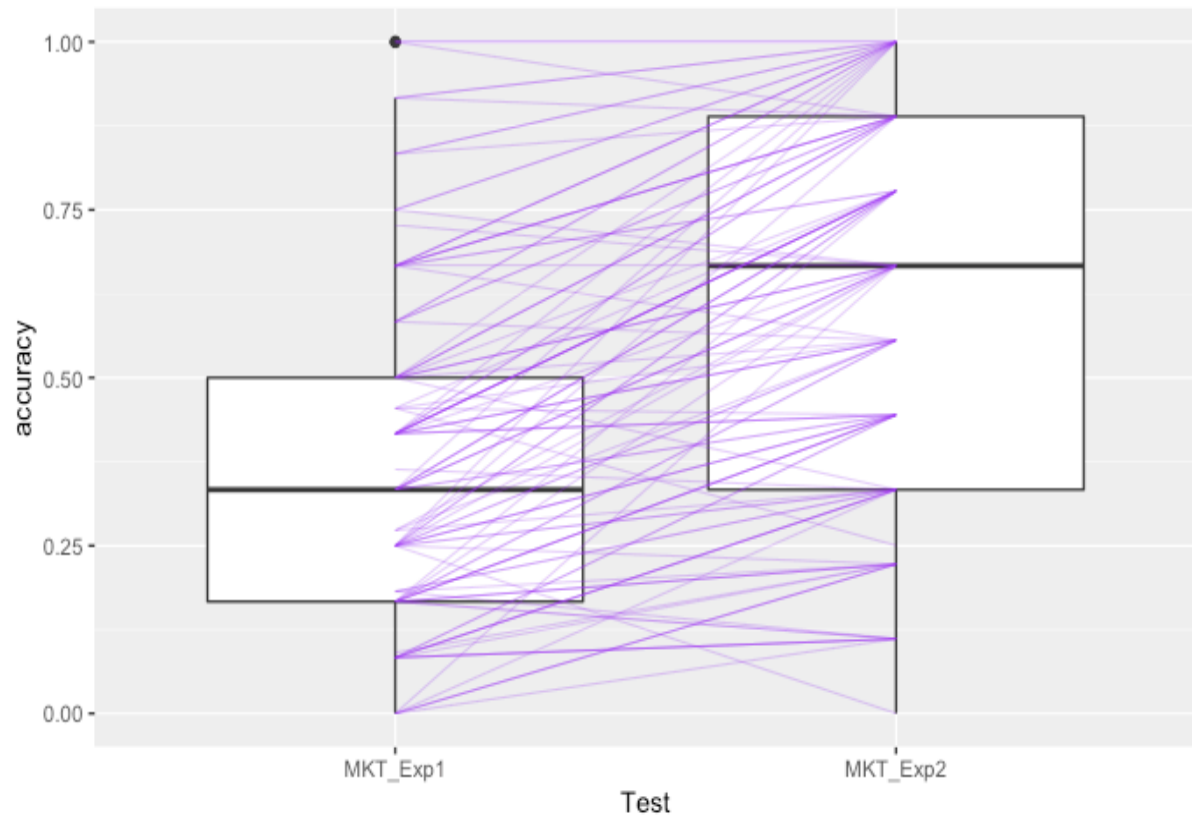


Figure 4.7. Spaghetti plot of Metalinguistic Knowledge Test (Explanation) at T1 and T2

Table 4.8 Descriptive Statistics for Metalinguistic Knowledge Test at T1 and T2

| Time   | Structure          | Mean  | SD    |
|--------|--------------------|-------|-------|
| Time 1 | Comparatives       | 0.217 | 0.413 |
|        | Mass/Count Nouns   | 0.273 | 0.446 |
|        | Embedded Questions | 0.312 | 0.464 |
|        | Verb Complements   | 0.273 | 0.446 |
|        | Be Passive         | 0.442 | 0.498 |
|        | Third Person       | 0.662 | 0.474 |
| Time 2 | Comparatives       | 0.694 | 0.463 |
|        | Embedded Questions | 0.527 | 0.5   |
|        | Verb Complements   | 0.529 | 0.5   |
|        | Be Passive         | 0.541 | 0.499 |
|        | Third Person       | 0.764 | 0.425 |

Note. All items are ungrammatical

#### 4.4.2 Summary of Descriptive Results

Table 4.9 summarizes the descriptive results of the five linguistic tests. Participants showed a significant improvement in all tests but the timed written GJT.

Table 4.9 Summary of the Descriptive Results

| Constructs         | Measures                      | Findings   | Effect size <sup>a</sup>     |
|--------------------|-------------------------------|------------|------------------------------|
| Implicit Knowledge | Elicited Imitation            | Improved   | $d = 0.31$ ; Small to Medium |
|                    | Oral Production               | Improved   | $d = 0.38$ ; Medium          |
|                    | Time WGJT                     | Comparable | $d = 0.17$ ; Small           |
| Explicit Knowledge | Untimed WGJT                  | Improved   | $d = 0.93$ ; Very large      |
|                    | Metalinguistic Knowledge Test | Improved   | $d = 0.90$ ; Very large      |

Note. <sup>a</sup> Effect size interpretation is based on Plonsky and Oswald (2014)

### 4.4.3 Correlations

Before performing two-factor CFAs, I first examined the associations between tests by computing spearman correlation coefficient ( $r_s$ ) among the five language test and two activity types separately for T1 and T2.

Table 4.10 includes the interrelationships between activity and knowledge types at T1. The values range from  $r_s = .099$  (EI & MKT) to  $r_s = .469$  (UGJT & TGJT). Pairs of tests, besides MKT and the two verbal production tasks (EI and OP), showed moderate but significant relationships (all  $r_s$  below .01). At T2, some qualitative changes (compared to T1) were observed. As shown in Table 4.11, while a relatively weak relationship remained between EI and MKT ( $r_s = .149$ ), the strong correlations between OP and the two GJTs at T1 attenuated to  $r_s = .209$  (TGJT) and  $r_s = .205$  (UGJT) at T2. At the same time, the interrelationships between activity and knowledge types are virtually nonexistent, both at T1 ( $-.128 < r_s < .122$ ) and T2 ( $-.143 < r_s < .070$ ), suggesting that knowledge types and engagement with different types of activity were essentially unrelated.

Figure 4.8 contains Spearman correlation coefficients (upper diagonal) of the five linguistic measures and activity types, scatterplots for variable pairs (lower diagonal), and density plots for each variable (diagonal) at T1. Figure 4.9 visualizes intercorrelations of the five linguistic measures and activity types at T2.

Table 4.10 Correlational Matrix for the Five Tests at T1 and Activity Types

|          | EI     | OP     | TGJT   | UGJT   | MKT  | Meaning | Language |
|----------|--------|--------|--------|--------|------|---------|----------|
| EI       | -      |        |        |        |      |         |          |
| OP       | .452** | -      |        |        |      |         |          |
| TGJT     | .421** | .415** | -      |        |      |         |          |
| UGJT     | .380** | .327** | .469** | -      |      |         |          |
| MKT      | .099   | .122   | .218** | .390** | -    |         |          |
| Meaning  | .062   | .109   | .034   | .090   | .083 | -       |          |
| Language | .048   | -.128  | .000   | -.109  | .010 | .462**  | -        |

Note. \*\* Correlation is significant at the 0.01 level (2-tailed). \* Correlation is significant at the 0.05 level (2-tailed)

Table 4.11 Correlational Matrix for the Five Tests at T2 and Activity Types

|          | EI     | OP     | TGJT   | UGJT   | MKT   | Meaning | Language |
|----------|--------|--------|--------|--------|-------|---------|----------|
| EI       | -      |        |        |        |       |         |          |
| OP       | .378** | -      |        |        |       |         |          |
| TGJT     | .429** | .209*  | -      |        |       |         |          |
| UGJT     | .340** | .205*  | .469** | -      |       |         |          |
| MKT      | .149   | .254** | .339** | .455** | -     |         |          |
| Meaning  | .070   | -.006  | -.069  | -.052  | -.036 | -       |          |
| Language | -.023  | .023   | -.143  | -.049  | -.032 | .462**  | -        |

Note. \*\* Correlation is significant at the 0.01 level (2-tailed). \* Correlation is significant at the 0.05 level (2-tailed)

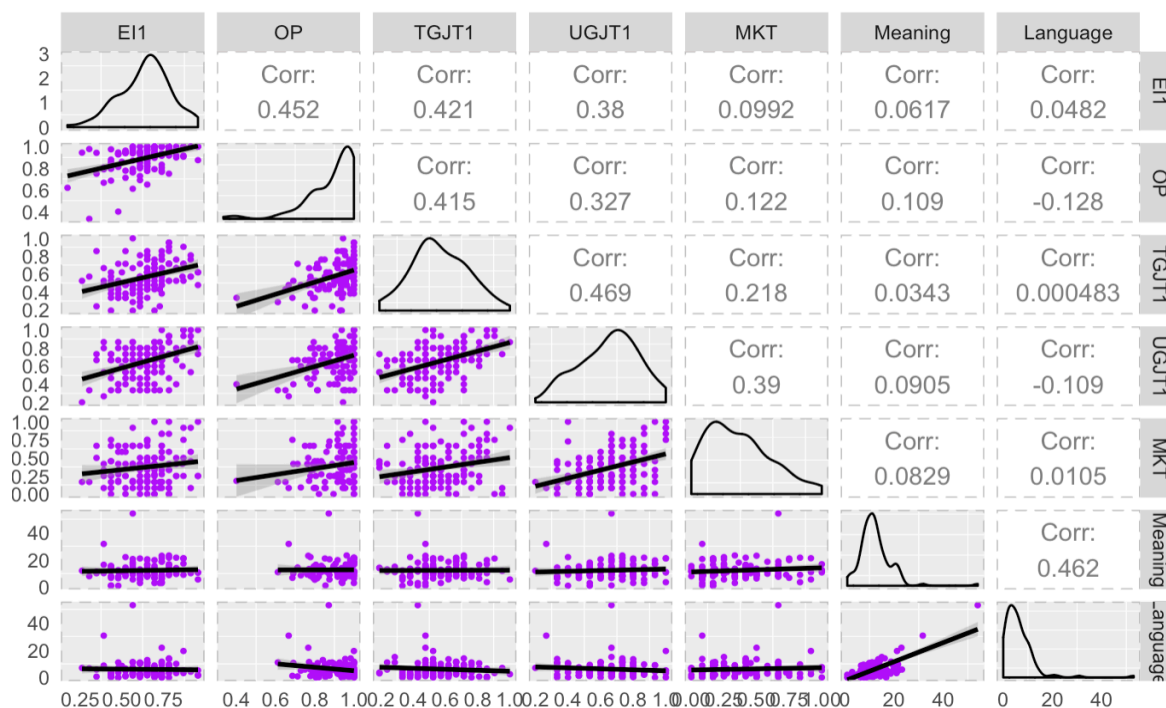


Figure 4.8. Relationships among five linguistic scores at T1 and activity types

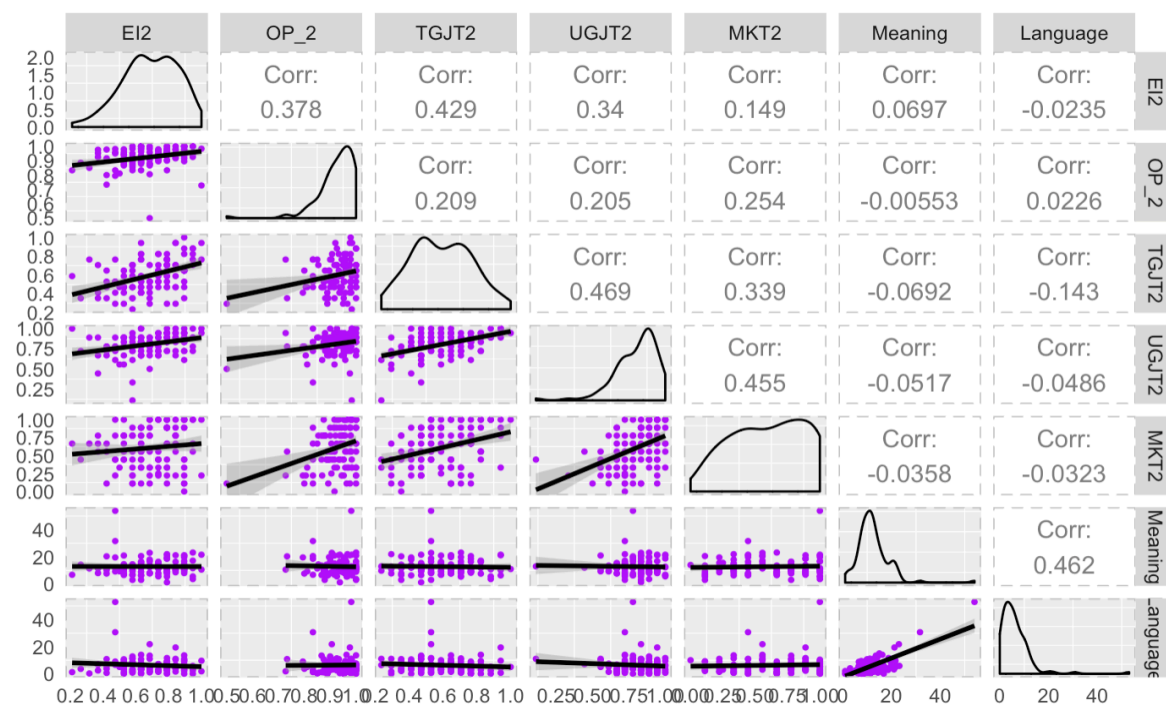


Figure 4.9. Relationships among five linguistic scores at T2 and activity types

#### **4.4.4 Factor Scores**

To examine the psychometric associations between the linguistic measures (and to extract factor scores from the results), I performed a two-factor CFA model on the linguistic tests, separately for T1 and T2. These CFA models were specified based on theory and previous empirical findings from test validation studies (e.g., R. Ellis, 2005; Godfroid & Kim, under review; Godfroid, Kim et al., in preparation).

The two-factor model was specified with two correlated latent variables, implicit and explicit L2 morphosyntactic knowledge (see Figure 4.10 for T1 and Figure 4.11 for T2). The implicit factor was represented by EI, OP, TGJT. The explicit factor was represented by UGJT and MKT. EI and UGJT served as the reference indicators, respectively, for the implicit and explicit constructs.



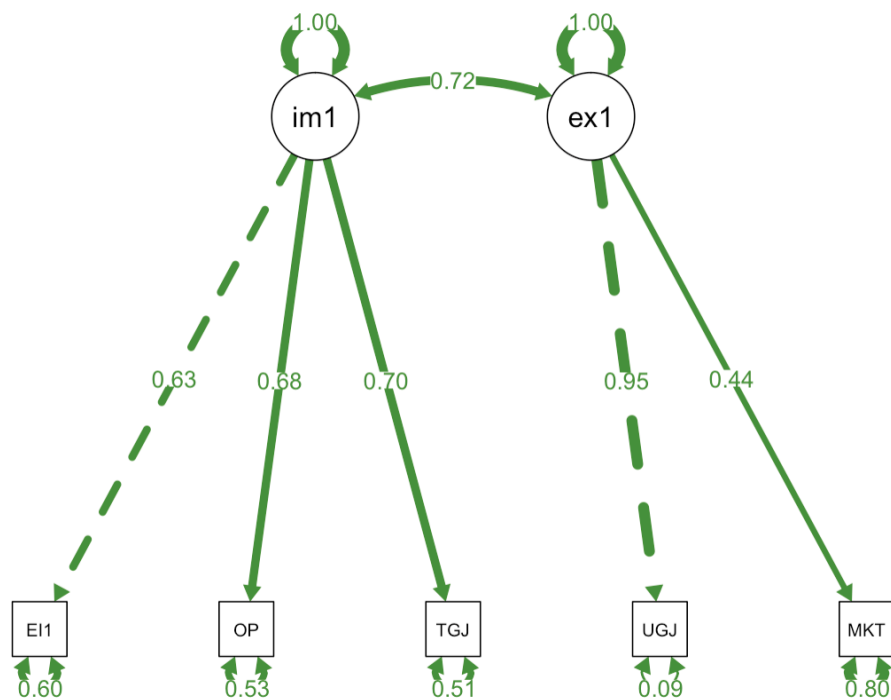


Figure 4.10. Two-factor model at T1

Note. *im1* = implicit knowledge at T1; *ex1* = explicit knowledge at T1; *E1* = elicited imitation at T1; *OP* = oral production; *TGJ* = timed written grammaticality judgment task; *UGJ* = untimed written grammaticality judgment task; *MKT* = metalinguistic knowledge test

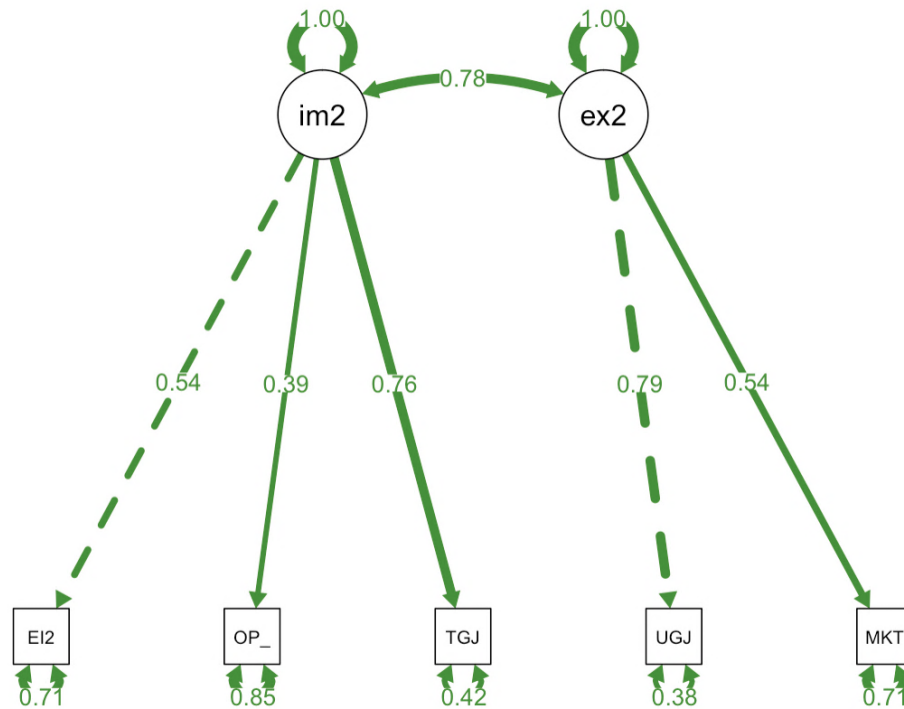


Figure 4.11. Two-factor model at T2

Note. Imp2 = implicit knowledge at T2; ex2 = explicit knowledge at T2; EI2 = elicited imitation at T2; OP\_ = oral production; TGJ = timed written grammaticality judgment task; UGJ = untimed written grammaticality judgment task; MKT = metalinguistic knowledge test

The global goodness of fit indices are summarized in Table 4.12. Overall, both models fit the data adequately, with all indices within an acceptable range. To diagnose any sources of model misspecification, I inspected the modification indices and standardized residuals. No modification indices were larger than 3.84 (largest = 1.73) and no standardized residual for any of the indicators was greater than  $|1.96|$  (largest = 0.001), suggesting an absence of localized areas of ill fit in my model specification.

Table 4.12 CFA Model Fit Indices

|                             | T1    | T2    |
|-----------------------------|-------|-------|
| Parameters (n)              | 16    | 16    |
| $\chi^2$                    | 3.234 | 7.298 |
| $\chi^2 p (> 0.05)$         | 0.519 | 0.121 |
| df                          | 4     | 4     |
| CFI ( $\geq .95$ )          | 1.000 | 0.957 |
| SRMR ( $\leq 0.08$ )        | 0.021 | 0.047 |
| RMSEA                       | 0.000 | 0.082 |
| RMSEA lower ( $\leq 0.05$ ) | 0.000 | 0.000 |
| RMSEA upper                 | 0.122 | 0.162 |

Note.  $\chi^2$  = chi-square;  $\chi^2 p$  = chi-square test p-value; df = degree of freedom; CFI =

comparative fit index; SRMR = standardized root mean square; RMSEA = root mean square error of association

As a last step to model evaluation, I explored the parameter estimates (factor loadings). In particular, I inspected the direction (positive or negative loadings), magnitude, and significance of each parameter estimate. Parameter estimates for the two-factor model are detailed in Table 4.13. All indicators show positive directions of estimates (no negative loadings); also, standardized factor loadings were all above .30, which is a threshold that is commonly operationalized as “salient” (Brown, 2015). Lastly, all factor loadings were statistically significant, indicating the construct is explaining each observed variable in a meaningful way.

At the same time, a relatively high correlation was observed between the explicit and implicit latent variables in both time points (T1:  $r = 0.719$ , T2:  $r = 0.784$ ). Although factor correlations above .80 or .85 are typically flagged for showing poor discrimination validity

(Brown, 2015, p. 116), I ran a one factor model to examine if the data is better represented as reflecting a single underlying construct. Goodness of fit and modification indices both suggest two-factor models to be superior than a one-factor at T1 (significant  $\chi^2$   $p$  value = .020; low CFI = .932; one modification index exceeding 3.84) and T2 (borderline significant  $\chi^2$   $p$  value = .063; low CFI = .928; two modification indices over 3.84). I thus conclude that the two-factor model in both time points fit the data best.

Table 4.13 Two-Factor Model Parameter Estimates for T1 and T2

|    | Parameter           | Estimate | SE    | <i>p</i> | Standardized Est. |
|----|---------------------|----------|-------|----------|-------------------|
| T1 | Implicit →          |          |       |          |                   |
|    | EI                  | 1.00     |       |          | 0.632             |
|    | OP                  | 0.884    | 0.185 | < .001   | 0.682             |
|    | TGJT                | 1.293    | 0.278 | < .001   | 0.697             |
|    | Explicit →          |          |       |          |                   |
|    | UGJT                | 1.00     |       |          | 0.951             |
|    | MKT                 | 0.623    | 0.214 | .004     | 0.444             |
|    | Correlations        |          |       |          |                   |
|    | Implicit ↔ Explicit | 0.012    | 0.003 | < .001   | 0.719             |
| T2 | Implicit →          |          |       |          |                   |
|    | EI                  | 1.00     |       |          | 0.542             |
|    | OP                  | 0.304    | 0.144 | .034     | 0.392             |
|    | TGJT                | 1.455    | 0.329 | < .001   | 0.760             |
|    | Explicit →          |          |       |          |                   |
|    | UGJT                | 1.00     |       |          | 0.789             |
|    | MKT                 | 1.341    | 0.343 | < .001   | 0.539             |
|    | Correlations        |          |       |          |                   |
|    | Implicit ↔ Explicit | 0.008    | 0.002 | < .001   | 0.784             |

Note. OP = oral production; EI = elicited imitation; TGJT = timed written grammaticality

judgment task; UGJT = untimed written grammaticality judgment task; MKT = metalinguistic

knowledge test

With an acceptable measurement solution established, I then extracted factor scores to use them as proxies for latent variables of explicit and implicit knowledge at T1 and T2. Factor scores were extracted using the *lavPredict ()* function in *R* with a regression method. For the activity measures, I used average frequency of language-focused and meaning-focused activity engagement across the five LEL logs. The activity type values were converted to z-scores to scale them to the same unit as the factor scores, which are also z-scores.

Table 4.14 displays descriptive results of the four factor scores for implicit and explicit knowledge and z-scores for the two activity measures (language-focused and meaning-focused). The mean scores of the factor scores are zero as a regression method extracts factor scores in z score form. With multivariate normality assumptions violated ( $p < .001$ ), I used the MLR estimation approach to evaluate the associations between indicators in path analyses.

Table 4.14 Factor Scores of Explicit and Implicit Knowledge at T1 and T2

|             | Mean  | SD    | Min    | Max   | Skewness | Kurtosis |
|-------------|-------|-------|--------|-------|----------|----------|
| Implicit T1 | 0.000 | 0.084 | -0.303 | 0.173 | -0.582   | 0.767    |
| Explicit T1 | 0.000 | 0.167 | -0.402 | 0.338 | -0.264   | -0.515   |
| Implicit T2 | 0.000 | 0.078 | -0.208 | 0.178 | -0.187   | -0.170   |
| Explicit T2 | 0.000 | 0.097 | -0.367 | 0.193 | -0.748   | 0.965    |
| Meaning     | 0.000 | 1.000 | -1.742 | 6.435 | 2.375    | 12.526   |
| Language    | 0.000 | 1.000 | -0.970 | 7.387 | 3.911    | 23.736   |

Note. Implicit = implicit knowledge; Explicit = explicit knowledge; T1 = time 1; T2 = time 2;

Meaning = meaning-focused activity; Language = language-focused activity

Lastly, I examined the interrelationships between the knowledge and activity types. As seen in Table 4.15, the values range from  $r_s = -.094$  (Language & Explicit T1) to  $r_s = .896$  (Explicit T2 & Implicit T2). In essence, pairs of knowledge types between the same time points

shows the highest correlation while the relationship within activity types and between knowledge types were essentially zero.

Table 4.15 Correlational Matrix for the Knowledge and Activity Types

|             | Implicit T1 | Explicit T1 | Implicit T2 | Explicit T2 | Meaning | Language |
|-------------|-------------|-------------|-------------|-------------|---------|----------|
| Implicit T1 | -           |             |             |             |         |          |
| Explicit T1 | .829**      | -           |             |             |         |          |
| Implicit T2 | .695**      | .686**      | -           |             |         |          |
| Explicit T2 | .610**      | .664**      | .896**      | -           |         |          |
| Meaning     | .105        | .102        | -.037       | -.053       | -       |          |
| Language    | -.067       | -.094       | -.092       | -.070       | .462**  | -        |

Note. \*\* Correlation is significant at the 0.01 level (2-tailed); Implicit = implicit knowledge;

Explicit = explicit knowledge; T1 = time 1; T2 = time 2; Meaning = meaning-focused activity;

Language = language-focused activity

#### 4.4.5 Path Analysis

In this section, I report a series of path analyses to address the interface question. In an attempt to share the entire model selection process, I divided the results into two sections. In Part 1, I report the findings of two competing models that were constructed *prior* to data analysis (and thus align with the models presented in the Methods section). The two competing models were,

- I. Non-interface Model: No cross-paths
- II. Interface Model: One cross-path from Explicit T1 to Implicit T2

In Part 2, I report two additional models that *later* proved to be better fitting models. The two addition models were,

- III. Reverse Interface Model: One cross-path from Implicit T1 to Explicit T2
- IV. Reciprocal Interface Model: Two cross-paths from Implicit T1 to Explicit T2 and from Explicit T1 to Implicit T2

For a path diagram of each model, see Figure 4.12.

In Part 1, two pieces of evidence will count as support for the interface position: (1) A better statistical fit of the Interface model (that includes the interface path) than the Non-Interface model (that does not include the interface path), together with (2) a positive, significant interface path from explicit knowledge (T1) to implicit knowledge (T2).

Similarly, in Part 2, evidence that lends support to the interface question includes, (1) a better statistical fit of the Reciprocal Interface Model (that includes the interface path) compared to Reverse Interface Model (that does not include the interface path), together with (2) a positive, significant path from explicit knowledge (T1) to implicit knowledge (T2).



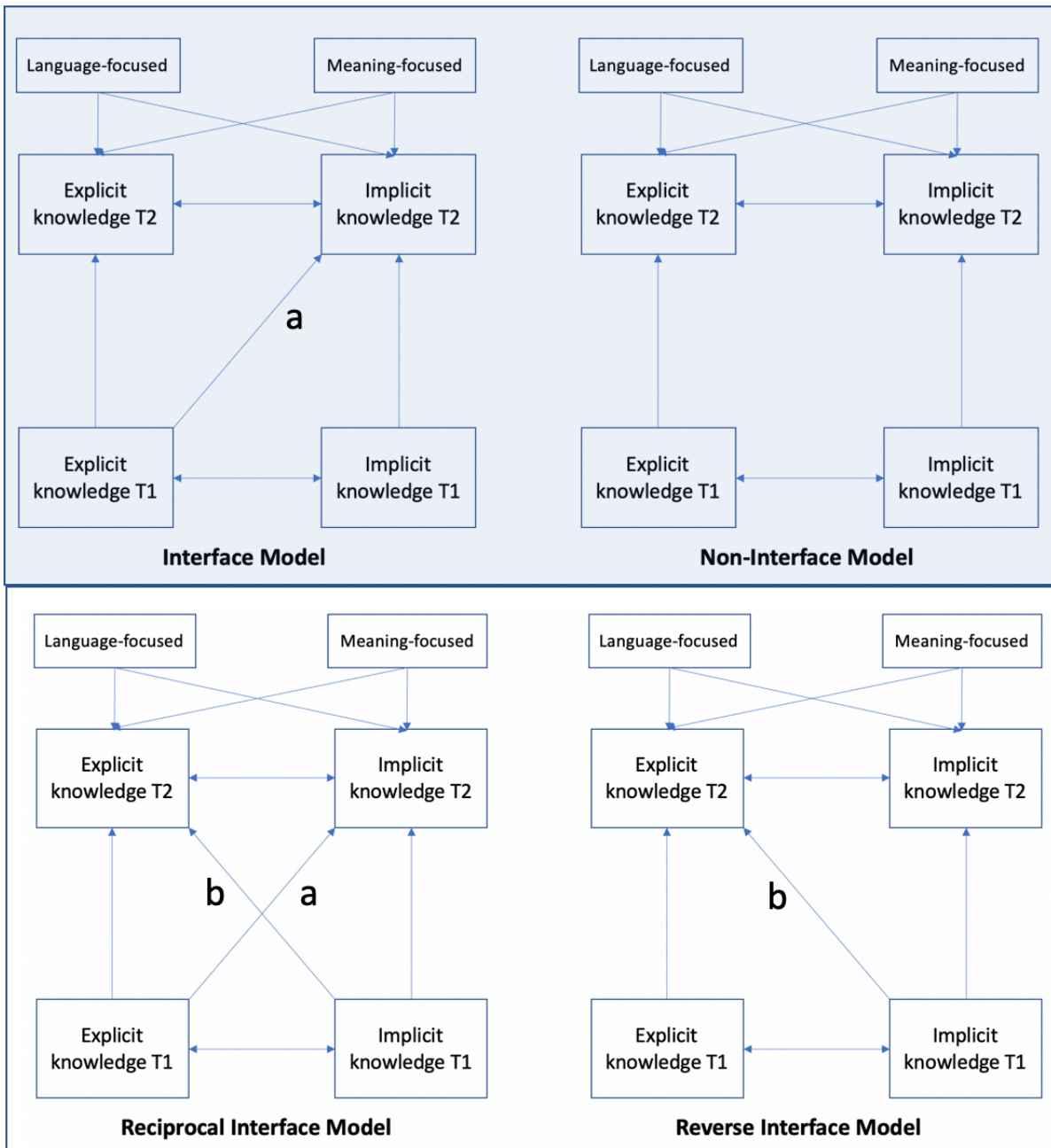


Figure 4.12. (top-left) The Interface Model; (top-right) The Non-interface Model; (bottom-left) The Reciprocal Interface Model; (bottom-right) The Reverse Interface Model. The top two are a *priori* path models constructed before the analysis. The bottom two are a *posteriori* path models. Paths a and b represent the explicit-implicit interface and the implicit-explicit interface, respectively

#### 4.4.5.1 Part 1: Interface vs. Non-interface

Two competing *priori* path models were constructed: The Interface model (with a cross-lag from Explicit T1 to Implicit T2) and the Non-interface model (no cross-lags). In both models, I regressed two predictive activity variables—meaning-focused and language-focused—onto both types of knowledge at T2. Table 4.16 summarizes the model fit indices of both models.

Table 4.16 Model Fit Indices for the Interface and Non-interface Models

|                             | Interface | Non-interface |
|-----------------------------|-----------|---------------|
| Parameters (n)              | 22        | 21            |
| $\chi^2$                    | 20.420    | 69.093        |
| $\chi^2 p (> 0.05)$         | 0.001     | 0.000         |
| df                          | 5         | 6             |
| CFI ( $\geq .95$ )          | 0.978     | 0.903         |
| SRMR ( $\leq 0.08$ )        | 0.062     | 0.170         |
| RMSEA                       | 0.131     | 0.251         |
| RMSEA lower ( $\leq 0.05$ ) | 0.075     | 0.200         |
| RMSEA upper                 | 0.192     | 0.306         |

Note.  $\chi^2$  = chi-square;  $\chi^2 p$  = chi-square test p-value; df = degree of freedom; CFI = comparative fit index; SRMR = standardized root mean square; RMSEA = root mean square error of association

As can be seen, all fit indices in the Non-interface model were poor. The Interface model, on the other hand, revealed a mixed picture: While the values of CFI and SRMR were acceptable, the lower bound of RMESA was slightly higher than the recommended cutoff point of 0.05. On top of this, the chi-square test was significant, suggesting that the model significantly departed from the data. Typically, RMSEAs with low *df* (and sample size) generate artificially

large values of the RMSEA and falsely indicate a poor-fitting model (Kenny, 2015). This is because the computational formula of the RMSEA is highly dependent on  $df$  and  $n$  size:  $\sqrt{(\chi^2 - df) / df(N-1)}$  (for this reason, Kenny, Kaniskan, and McCoach (2014) argued to not compute the RMSEA with low  $df$  models). Similarly, the significant chi-square shown in the Interface model may be a combined results of a low  $df$ , smaller samples, and multivariate non-normality (Hayduk, Cummings, Boadu, Pazderka-Robinson, & Boulianne, 2007), all of which contribute to high Type 1 error rate (i.e., concluding the model significantly departs from the data when that is not really the case) (Kenny, 2015).

A scaled chi-square difference test suggested that the Interface model was a significantly better fitting model than the Non-interface model,  $\chi^2_{dif} = 37.339$ ,  $df = 1$ ,  $p = 0.001$ . Along with the significant interface path from Explicit T1 to Implicit T2 ( $p < .001$ ), the results suggest that explicit knowledge at Time 1 influences the development of subsequent implicit knowledge at Time 2.

#### 4.4.5.2 Part 2: Reciprocal Interface vs Reverse Interface

While the above results lend support to the explicit-implicit interface, additional evidence of a significant and a *good fitting* interface model would further evidence the explicit-implicit interface. As such, I explored ways in which the model can be improved by inspecting the modification indices (MI). Of several respecification suggestions, one modification was theoretically justifiable (i.e., a path from Implicit T1 to Explicit T2; that is, implicit knowledge influencing explicit rule discovery; e.g., Bialystok, 1994, 2001; Cleeremans, 2007). No standardized residuals exceeded .10. And thus, a structural path from Implicit at T1 to Explicit at T2 was newly added to the model for model fit improvement.

Again, with the new models, two pieces of evidence count as support to the interface position: A better statistical fit of the Reciprocal Interface model (that includes a cross-lag from Explicit T1 to Implicit T2) compared to Reverse Interface model (without the interface cross-lag); also, a positive, significant path from explicit knowledge (T1) to implicit knowledge (T2) is needed.

**Explicit-Implicit Interface.** As seen in Table 4.17, the addition of a new regression path improved the model fit of the Reciprocal Interface model yielding good fit indices within acceptable range. However, the Reverse Interface model still remained as a poor fitting model. A scaled chi-square difference test (Satorra & Bentler, 2010) suggested that the Reciprocal Interface model was significantly better than the Reverse Interface model,  $\chi^2_{\text{dif}} = 8.65$ ,  $df = 1$ ,  $p = 0.003$ .

Table 4.18 presents the parameter estimates of the Reciprocal Interface model. As expected, the results indicate that all autoregressive paths were significant. In particular, the strongest predictor of current implicit knowledge is prior implicit knowledge (Std. Est.= 0.483) and the strongest predictor of current explicit knowledge is prior explicit knowledge (Std. Est.= 0.385). This suggests that participants showed steady improvement in their explicit knowledge and implicit knowledge development and individual differences in explicit knowledge and implicit knowledge were stable over the 3-4 month lag between occasions of measurement. Critically, the hypothesized interface path from Explicit T1 to Implicit T2 (the cross-lag path) was also significant. This reflects that explicit knowledge led to the increase of implicit knowledge even after controlling for previous standings of implicit knowledge. This path yielded a standardized coefficient estimate of 0.33, a predictive magnitude that is as strong as the autoregressive impact of Explicit T1 on Explicit T2 (the confidence intervals (CIs) cross over

across pairs). This finding—the interface effect revealing a comparable magnitude of an autoregressive effect, which typically has the strongest predictive magnitude—is suggestive of a strong influence of explicit knowledge on implicit knowledge development.

Table 4.17 Model Fit Indices of the Reciprocal Interface and Reverse Interface Model

|                             | Reciprocal Interface Model | Reverse Interface Model |
|-----------------------------|----------------------------|-------------------------|
| Parameters (n)              | 23                         | 22                      |
| $\chi^2$                    | 9.356                      | 20.258                  |
| $\chi^2 p$ ( $> 0.05$ )     | 0.053                      | 0.001                   |
| df                          | 4                          | 5                       |
| CFI ( $\geq .95$ )          | 0.992                      | 0.976                   |
| SRMR ( $\leq 0.08$ )        | 0.054                      | 0.057                   |
| RMSEA                       | 0.087                      | 0.138                   |
| RMSEA lower ( $\leq 0.05$ ) | 0.000                      | 0.079                   |
| RMSEA upper                 | 0.161                      | 0.203                   |

Note.  $\chi^2$  = chi-square;  $\chi^2 p$  = chi-square test p-value; df = degree of freedom; CFI =

comparative fit index; SRMR = standardized root mean square; RMSEA = root mean square error of association

Table 4.18 Model Parameter Estimates for the Reciprocal Interface Model

| Path                      | Estimate<br>[CI lower, upper] | SE    | <i>p</i> | Standardized<br>Est. |
|---------------------------|-------------------------------|-------|----------|----------------------|
| Implicit T1 →             |                               |       |          |                      |
| Implicit T2               | 0.480<br>[0.295, 0.665]       | 0.094 | 0.000    | 0.483                |
| Explicit T2               | 0.445<br>[0.156, 0.734]       | 0.147 | 0.003    | 0.366                |
| Explicit T1 →             |                               |       |          |                      |
| Explicit T2               | 0.234<br>[0.109, 0.358]       | 0.063 | 0.000    | 0.385                |
| Implicit T2               | 0.163<br>[0.071, 0.255]       | 0.047 | 0.001    | 0.329                |
| Meaning →                 |                               |       |          |                      |
| Explicit T2               | -0.012<br>[-0.028, 0.004]     | 0.008 | 0.132    | -0.120               |
| Implicit T2               | -0.010<br>[-0.023, 0.003]     | 0.007 | 0.149    | -0.118               |
| Language →                |                               |       |          |                      |
| Explicit T2               | 0.011<br>[-0.004, 0.027]      | 0.008 | 0.146    | 0.112                |
| Implicit T2               | 0.008<br>[-0.004, 0.020]      | 0.006 | 0.186    | 0.097                |
| Covariances/Correlations  |                               |       |          |                      |
| Implicit T1 ↔ Explicit T1 | 0.012<br>[0.009, 0.015]       | 0.001 | 0.000    | 0.832                |
| Implicit T2 ↔ Explicit T2 | 0.003<br>[0.002, 0.004]       | 0.001 | 0.000    | 0.828                |
| Meaning ↔ Language        | 0.710<br>[-0.067, 1.488]      | 0.397 | 0.073    | 0.716                |

**Reverse Interface.** As important as the structural path from Explicit T1 to Implicit T2 is the new regression path of Implicit T1 to Explicit T2 that was added to the model. With this path included, the Interface model has improved to an acceptable fit. Then, a new question emerges: Does Implicit T1 influence the development of Explicit T2? In other words, is learners' explicit knowledge predicted by their previous amount of implicit knowledge?

To test this hypothesis, I compared the Reciprocal Interface model (that *includes* the Implicit-Explicit interface path) to the Interface model (that *does not include* the Implicit-Explicit path). The path diagram in Figure 4.13 illustrates two competing models with and without the Implicit-Explicit cross-lag path (Path a). Two pieces of evidence will count as support for the Implicit-Explicit interface: (1) A better statistical fit of the Reciprocal Interface model than the Interface model, together with (2) a positive, significant path from implicit knowledge (T1) to explicit knowledge (T2).

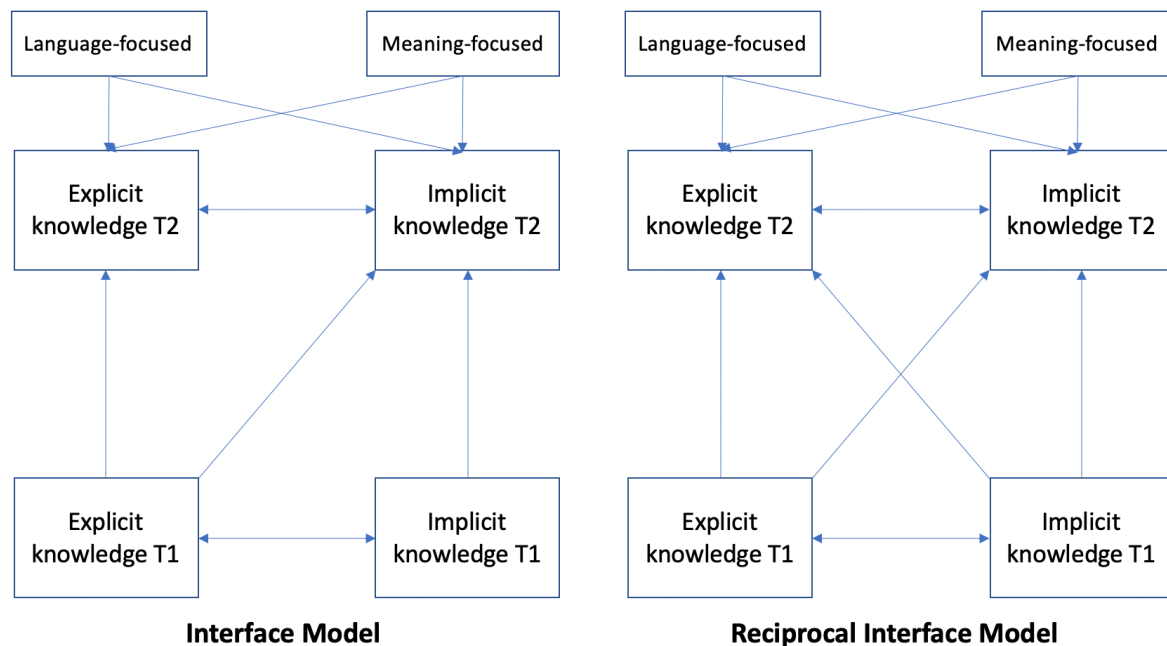


Figure 4.13. (left) The Interface Model *without* the Implicit-Explicit interface path; (right) The Reciprocal Interface Model *with* the Implicit-Explicit interface path

A scaled chi-square difference test suggested that the Reciprocal Interface model was a better fitting model than the Interface model,  $\chi_{\text{dif}}^2 = 9.356$ ,  $df = 1$ ,  $p = 0.000$ . The parameter estimates presented in Table 4.17 also yield a significant Implicit-Explicit interface path. These findings jointly suggest that learners' explicit knowledge at Time 2 was highly influenced by their previous standings in implicit knowledge at Time 1. Importantly, the predictive magnitude of Implicit T1 on Explicit T2 was comparable in magnitude to the impact of Explicit T1 on Implicit T2. As such, I conclude that both knowledge types are in a reciprocal relationship impacting each other bi-directionally with a comparable predictive value.

Lastly, none of the activity type measures predicted the development of implicit or explicit knowledge, suggesting their weak attribution to knowledge development. Intriguingly, though, the directional paths of meaning-focusing activity toward both knowledge types were negative (for explicit,  $-.120$ ; for implicit,  $-.118$ ), while positive directions were observed from language-focused activity to explicit knowledge ( $.112$ ) and implicit knowledge ( $.097$ ).



#### 4.4.6 *Summary of Results*

- Finding 1: the strongest predictor of current explicit knowledge is prior explicit knowledge; the strongest predictor of current implicit knowledge is prior implicit knowledge.
  - Evidence: The significance of the two autoregressive paths from explicit knowledge at T1 to T2 and implicit knowledge at T1 to T2.
- Finding 2: positive impact of previous explicit knowledge on the development of implicit knowledge
  - Evidence: A better statistical fit of the Reciprocal Interface Model (that includes both cross paths) compared to Reverse Interface Model (that does not include the interface path), together with a positive, significant path from explicit knowledge (T1) to implicit knowledge (T2).
- Finding 3: explicit knowledge was highly influenced by the previous levels of implicit knowledge
  - Evidence: A better statistical fit of the Reciprocal Interface model (that includes both cross paths) and the original Interface model (that does not include the Implicit-Explicit interface path), together with a positive, significant path from implicit knowledge (T1) to explicit knowledge (T2).
- Finding 4: no predictive impact of activity types on knowledge development.
  - Evidence: Both paths, language-focused and meaning-focused activity, regressed non-significantly to both knowledge types.

## CHAPTER 5: DISCUSSION

The goal of this study was to systematically track the amount and types of language use of international students immersed in the target language environment. By doing so, I aimed to provide answers to the following questions: How much time do international students spend using English as opposed to other languages? What types of language skills receive the most (least) attention during daily life abroad? How much individual variation exists in English use among the international student sample? A secondary goal was to compare the longitudinal associations between two types of knowledge (i.e., explicit knowledge and implicit knowledge) and examine how they relate to different types of activity (i.e., language-focused and meaning-focused).

Data on authentic language usage showed that international students are more engaged (quantitatively, in terms of hours per day spent) with L2 English than other languages. To be precise, they spend 2.2x more time using English than other languages. I also observed qualitative differences in English engagement. While students spent a comparable amount of time speaking, listening, and reading in English, they spent significantly less time writing in English. Lastly, wide-ranging variability was observed in the amount and types of language engagement among the international student sample.

In a two-timepoint longitudinal experiment, I demonstrated that there was a facilitative relationship between explicit and implicit L2 morphosyntactic knowledge. The best fitting model of the associations of knowledge and activity types suggested that both types of knowledge are in a reciprocal relationship affecting each other bi-directionally and functioning as causes and consequences of each other. None of the activity types predicted knowledge development.

## 5.1 Quantitative and Qualitative Differences in Language Engagement

One important contribution of this study is that I systematically tracked international students' L2 engagement patterns to test a common belief that studying abroad provides many opportunities to use the target language. The result of the group-averaged daily-engagement data showed that international students who enrolled at an English-medium university in the United States are more engaged (quantitatively, in terms of hours per day spent) with L2 English than other languages. This finding, along with the linguistic development that will be discussed in section 5.2, highlights two key messages: First, it confirms that, at a group-level, international students engage in the L2 significantly more than their native and additional languages in an immersion context; second, the results speak to the benefits of degree-achieving study abroad experiences of international students.

At the same time, the results also caution against generalizing L2 usage patterns in different learning contexts and settings. For instance, a number of studies have documented the paucity of L2 use in a study abroad (SA) context, reporting comparable amounts of L1 and L2 usage among exchange students instead (Dewey et al., 2013; Freed et al., 2004). Freed et al. (2004) also reported that the amount of L2 engagement in a SA context was significantly less than that in an intensive immersion program at home. As mentioned, the L2 engagement patterns can be affected by participant profiles. Participants in Dewey et al. (2013) and Freed et al. (2004) were college exchange students who temporarily enrolled in classes for a semester or two, whereas participants in the current research consisted of degree-seeking undergraduate and graduate students. The graduate students, who comprised most of this study sample, may have a workload at school, including their assistantship duties, that may differ from that of college exchange students. This superior use of L2, compared to L1, was also reported in Ranta and

Meckelborg (2013) with Chinese L2 graduate students studying in Canada. I would therefore be cautious of generalizing international students' engagement patterns across different settings, because language learning context alongside participant profile lead to different L2 engagement patterns.

A second important finding from the LEL logs was that individuals differed widely in the amount of L2 use. This was evident in the large standard deviations for L2 usage and the divergent patterns of individuals' L2 engagement across time. When examining the individual-level L2 engagement patterns, we can easily see that the overall group-level data are deceptive and there are a great deal of individual variations. The findings from the individual-level data, therefore, indicate that similar contextual affordances to engage in the L2 do not produce comparable amounts of actual L2 engagement for different individuals. These observations, in turn, may reinforce that language acquisition can be better understood when details of the language learning context are considered, as "the relationship between what a context offers and the nature of what an individual brings to the learning situation is both crucial and complex" (Segalowitz & Freed, 2004, p.196).

Last but not least, most L2 engagement occurred during worktime between 9 a.m. to 6 p.m. The other languages, on the other hand, were largely used in the evenings. This finding echoes that of McManus, Mitchell, and Tracy-Ventura (2014) whereby the L2 was reported to be used mostly at work and school, and engagement with the L1, with family and friends, was mostly sustained virtually; that is, via internet-based communication. It is also interesting to note that in the present study, the highest levels of L2 speaking occurred during office hours. From a practical and administrative point of view, these findings collectively confirm the importance of a structured work time that brings international students into regular contact with the L2. Seen in

this light, language program directors and classroom teachers can effectively organize students' daytime use by arranging socialization events and employment opportunities that foster abundant L2 interaction opportunities during the day while assigning tasks that involve L2 use of reading or writing that can be completed individually at home, which can in turn maximize L2 use in the evening.

## **5.2 Explicit and Implicit Knowledge and Activity Types**

### **5.2.1 *Explicit-Implicit Interface***

A review of the literature on the interface of explicit and implicit knowledge suggested that most SLA scholars, despite their varied theoretical stances, concur with the idea that explicit knowledge has a facilitative effect on implicit knowledge development. Based on this claim, a central aim of this dissertation was to examine empirically to what extent explicit L2 knowledge plays a causal role in the development of implicit L2 knowledge (RQ2). Results of a two-wave cross-lag path model confirmed that the Reciprocal Interface model with two cross-lag paths ( $\text{Imp} \rightarrow \text{Exp}$  &  $\text{Exp} \rightarrow \text{Imp}$ ) fit significantly better than the Reverse Interface model *without* the interface path ( $\text{Exp} \rightarrow \text{Imp}$ ). Given that the interface path in the Reciprocal Interface model was positive and significant, I demonstrated a facilitative effect of explicit knowledge on implicit knowledge development.

The current study makes a unique contribution to SLA research by providing one of the first empirical evidence for the explicit-implicit interface (1) using a natural language, (2) used in a naturalistic context, (3) longitudinally. These three combined components are important as acquisition of implicit L2 knowledge, and language acquisition in general, is a developmental process that is mediated by L2 exposure that is qualitatively and quantitatively different from

lab-based experiments (Paradis, 2009). From this perspective, it is important to note that the current finding extends those of lab-based intervention studies using artificial or extinct languages (Cintrón-Valentín & Ellis, 2016; Curcic, Andringa, & Kuiken, 2019) and a cross-sectional study that reported suggestive evidence of the impact of automatized explicit knowledge on implicit knowledge (Suzuki & DeKeyser, 2017).

While the exact mechanism underlying the interface is yet part of the black box, one theoretical account could be that conscious registration of linguistic patterns created a conscious channel for implicit tallying of these linguistic rules (N. Ellis, 2005; 2015). Alternatively, exposure to patterns, driven by conscious knowledge of those patterns, could have allowed for more practice and exposure to these regularities which in turn facilitated the development of their implicit representation (DeKeyser, 2009; Hulstijn, 2002, 2007, 2015; Paradis, 1994, 2004, 2009).

In relation to previous work, the present findings suggest that the predictive impact of explicit knowledge on implicit knowledge is not constrained to a more automatized version of explicit knowledge. The explicit knowledge measures used in this study—metalinguistic knowledge test (MKT) and untimed GJT—are the most controlled and most analytic measures of conscious knowledge that either directly ask for metalinguistic explanation or invite participants to employ controlled and conscious processing during problem solving. With conscious knowledge measured with these tests, the results suggest that less- or non-automatized explicit knowledge may impact the development of implicit morphosyntactic knowledge. This finding underscores the instructional values of a wide range of instructional techniques that vary in the continuum of explicitness on intuitive and spontaneous knowledge of linguistic rules (for different types of instructional activities, see Loewen, 2020).

On a descriptive level, two findings should be noted. First, in oral production (OP), most participants scored above 60% accuracy at T1 and 80% accuracy at T2 with group means near ceiling ( $M = 89\%$  and  $M = 93\%$ , respectively). Markedly high scores observed in OP are somewhat on a par with those reported in previous studies: Godfroid, Kim et al. (in preparation) reported 89 % accuracy with 151 intermediate-to-advanced L2 English speakers ( $M = 96$ , iBT TOEFL). With beginning-to-intermediate L2 learners ( $M = 6.25$ , IELTS, which converts to 60-78 iBT TOEFL), the OP mean score was 72% (R. Ellis, 2005).

Conversely, a fairly low mean score was observed in participants' MKT scores. This was quite evident in their performance at T1 where they scored 36% correctly. This accuracy rate is 17% lower than the scores reported in R. Ellis (2005) with the beginning-to-intermediate L2 learners (i.e., 53% accuracy). While the low MKT scores (alongside high OP performance) could be interpreted as participants being native-like in L2, it is important to note that the MKT used in this current study was the strictest version of explicit knowledge measures to tap into learners' "explicit declarative facts" about the language (Elder, 2009, p. 114). In particular, the scores were based on learners' ability to *provide* explanations to ungrammatical items. This design of MKT (the provision of rule explanation) is different from previous validation studies that employed a design of selecting correct explanations out of four options combined with identifying named grammatical parts in a sentence (i.e., R. Ellis, 2005; R. Ellis & Loewen, 2007) or the use of summed scores of providing metalinguistic explanations and identifying ungrammatical parts (i.e., Vafaei et al., 2017). As such, fairly low MKT scores observed in the current findings may be due to the distinct feature of the MKT design.

### 5.2.2 *Implicit-Explicit Interface*

Another notable finding of this study is that, in addition to the explicit-implicit interface, implicit knowledge was found to have a positive impact on the development of explicit L2 knowledge. In fact, the standardized coefficient of the two cross-lag paths in the Reciprocal Interface model demonstrated that the predictive impact of implicit knowledge on explicit knowledge was as comparable in magnitude to the impact of explicit knowledge on implicit knowledge (see Table 4.17). I find these results to be fascinating. While the interface debate in SLA mainly concerns the facilitative impact of explicit knowledge on implicit knowledge, the observed patterns of result (i.e., the reciprocal relationship between explicit and implicit L2 knowledge) suggests that awareness not only facilitate implicit learning/knowledge but it can be a product of implicit learning/knowledge.

The notion of “insight” is a term from the problem-solving literature referring to a sudden recognition of solutions (Mayer, 1995, p.3). In the context of SLA, this would translate to emergence of rule awareness from implicitly accrued knowledge (e.g., Bialystok, 1994, 2001; Cleeremans, 2007). Evidence of rule discovery has been empirically testified in relation to memory consolidation where awareness arises after a period of sleep (e.g., Batterink, Oudietter, Reber, & Paller, 2014; Fischer, Drosopoulos, Tsen, & Born, 2006; Wagner, Gais, Haider, Verleger, & Born, 2013; Wilhelm, Rose, Imhof, Rasch, Büchel, & Born, 2013). As such, the current findings suggest that claims from cognitive science about non-linguistic pattern discovery can also be extrapolated to L2 research context with natural language acquisition.

Two prevailing views may explain the underlying mechanisms of rule discovery: a single-system view and a multiple-system view. From a single-system perspective, no qualitative differences are assumed between the two knowledge types; instead, development of awareness is



contingent upon the quality/stability of implicit learning/knowledge. In other words, implicit knowledge gradually transitions to explicit knowledge when the quality of implicit representation of rules are stable (e.g., Cleeremans & Jiménez, 2002, Goujon, Didierjean, & Poulet, 2014; Mathews, et al., 1989). On the contrary, the proponents of a multiple-system view assume that the two knowledge types are qualitatively distinct; as such, implicit knowledge cannot become conscious by increased stability and explicit processing do not have direct access to implicit knowledge (e.g., Esser & Haider, 2017; Haider & Frensch, 2005; Rüniger & Frensch, 2008). Instead, the emergence of rule knowledge comes from a form of behavioral changes (e.g., an explicit hypothesis testing triggered by unexpected events during implicit learning). In turn, these behavioral changes might lead to rule discovery. This study was not a test of the underpinnings of consciousness, but an investigation into this topic would be a valuable avenue for future research in understanding and theorizing the role of awareness in L2 acquisition (for a similar attempt, see Williams, 2018).

### **5.2.3 Activity Types**

Lastly, in this study, activity types made little contribution to the acquisition of explicit and implicit knowledge. The standardized coefficients of both meaning-focused and language-focused activity did not predict explicit and implicit knowledge development in a significant manner (see Table 4.17). Intriguingly, though, the directional paths of meaning-focusing activity to both knowledge types were negative, while those from language-focused activity to both explicit and implicit knowledge were positive. Although the interpretation of these nonsignificant paths requires caution, positive predictive power of language-focused activity, compared to meaning-focused activity, on the acquisition of both explicit and implicit knowledge is in line with the superior effectiveness of explicit instructions reported in meta-

analytic reviews. For instance, Goo et al. (2015) reported medium effect size differences on learner performances on free production ( $g = 0.454$ ) and constrained production ( $g = 0.584$ ) with the explicit instruction being superior—a finding consistently reported in Norris and Ortega (2000) and Spada and Tomita (2010). Perhaps L2 speakers, including the more advanced learners, may benefit more from focused attention to linguistic forms at least for some morphosyntactic structures and under certain contexts.

While the above interpretation may partially explain the stronger effects of language-focused activity, it still does not explain why more meaning-focused activity induces less acquisition of L2 knowledge types. It may be the case that the relationship between meaning-oriented input and language development is linear until certain point, but beyond this threshold, the effects of language exposure may diminish. As such, an interesting replication would include extending this study with beginning or intermediate L2 speakers and explore whether the same pattern of results manifest. Another reason for the null association may relate to the measurement; in particular, the “meaning-focused activity” may be too coarse-grained and perhaps certain types of meaning-focused activity, for instance meaningful engagement during L2 comprehension or L2 production, may be more associated to language gains. On top of this, narrowing the time-segments to 30 minutes or even 15 minutes (as in Ranta & Meckelborg, 2013), instead of one hour, may provide researchers with a more comprehensive report on learners’ L2 engagement patterns. In essence, to better understand the association between processing-knowledge, a more fine-grained approach to analyzing different types of meaning-focused activity combined with learner characteristics may be needed.

### 5.3 Implications

The findings of this dissertation have methodological, pedagogical, and educational implications. Methodologically, notable technological advances were introduced. First, the web-programmed oral production tasks (oral production and elicited imitation) served as an efficient tool for remote data collection. Technical errors were quite minimal. For instance, in oral production, the task that had the highest missingness, 4.96 percent (at T2) to 9.40 percent (at T1) of data were missing due to poor quality of recordings or missing files. These numbers are relatively trivial considering the overall volume of data that can be collected online. With no commercial software that supports remote oral recording functions, the current project shows that this program, along with the manual that allows for customization (see Appendix F), can be a viable alternative to in-person data collection. With this program, data collection will become increasingly flexible even during global events such as the current COVID-19 health crisis.

Second, I sought to find a method that measures language usage by increasing the level of detail beyond what an offline questionnaire can provide and in a format that is also practical. This led to the development of a self-recorded language exposure log that collected L2 usage data on an hourly basis in real time. Through this device, I demonstrated that individuals vary substantially in how they engage with L2 in an immersion context. Similarly, I believe a host of research questions can be addressed using this device. These may include language engagement patterns of L2 speakers in EFL and ESL contexts, the quality and quantity of L2 use by cultural background, or a conceptual question such as whether there is a linear relationship between input and language development or whether there is a threshold beyond which the effects of language exposure diminish?

Second, the findings of this study have important pedagogical and educational implications. This study focused on language development and use of international students in US tertiary education, who made up 5.5% of the student body in 2019 (Institute of International Education, 2019). Improvements in linguistic skills reported in this study speak directly to the educational benefits of studying abroad and can inform classroom instructional approaches, and L2 learners' language learning techniques. In particular, the synergistic effects of explicit and implicit knowledge underscore the importance of developing both types of knowledge, since it entails that they function as a catalyst for a stronger representation of linguistic knowledge. In other words, development of explicit knowledge will impact implicit knowledge gains and implicit knowledge gains will also facilitate explicit knowledge development. As such, the results should not be construed as evidence in favor of one knowledge type over the other. Instead, instructors are recommended to provide a wide range of instructional techniques that vary in the involvement of explicit processing of rules (e.g., from a more implicit task such as consciousness-raising tasks to a more explicit task such as PPP), to cater to classroom learners of mixed explicit/implicit aptitude profiles. This way, the development of knowledge types, which can be explicit and/or implicit depending on one's aptitude profile, can impact the improvement of other type of knowledge.

Third, data on authentic usage will be of benefit to university administrators and student advisors, who can use this information to enhance and enrich the international student experience. For instance, given the large variation of L2 usage reported in this study, efforts need to be made to advocate for a greater diversity in events that bring together L1 and L2 English speakers. Potential new initiatives include cultural nights, field trips, mentorship arrangements, and volunteer work at community organizations such as the Refugee Centers or the public

libraries. The idea is to expand domestic and international students' networks and foster meaningful exchanges through regular social events. At a classroom level, data of L2 usage can function as a diagnostic for L2 classroom teachers who hope to understand international students' L2 and L1 engagement patterns outside of the classroom. As reported by Ranta and Meckelborg (2013), the pattern of conversational interaction or (un)willingness to communicate may differ by cultural background (e.g., Chinese students spend less time engaging in L2 with native speakers of English). From this perspective, data on the types of L2 engagement may provide useful information to both teachers and students to balance the use of L2 skills in and outside of the classroom.

#### **5.4 Limitations and Future Directions**

To my knowledge, this study is one of the first in the language sciences to investigate the explicit-implicit interface question longitudinally. Despite its significance, the current study had limitations that would need to be considered when interpreting the results.

First, from a statistical standpoint, the analyses of this study were based on factor scores, as opposed to latent constructs of explicit and implicit knowledge that allow for unexplained error variances. I placed an earnest effort to employ a longitudinal structural equation modelling (LSEM) with the current data set. However, the complex nature of LSEM, combined with the fairly small sample size of the current study for this type of analysis, necessitated that I impose unjustifiable and non-theory-driven constraints to the data (e.g., fixing covariances of latent variables to zero or error covariance of the same instruments across time to zero). This prevented me from achieving confident and reliable results. As such, I elected a simpler model (i.e., path model with factor scores) that is as rigorous as LSEM analysis (i.e., accounts for measurement errors that are inherent in the different linguistic measures) and thus has significant

advantages over traditional analysis of variance (i.e., ANOVA or regression). However, as mentioned earlier, this approach imposes unwarranted or, at least, untested assumptions of factorial invariance which is the notion that instruments measure the same constructs over time. Hence, the longitudinal effects should be interpreted with caution.

While a combination of factors could have contributed to the difficulty of running LSEM (e.g., power, highly correlated parameter estimates), future researchers can build on to this project by strengthening the reliability of each test and construct. In current study, the individual test reliability was high, but the composite reliability of each construct could be improved. The composite reliability for a latent construct refers to whether the observed measures consistently represent the same construct. Typically, composite reliability (omega) higher than .70 is recommended and reliabilities between .60 and .70 are acceptable. In the current result, most reliabilities fell within the acceptable range (from .61 to .70), but the explicit knowledge construct at T2 was low (reliability omega = .52). This may be due to the limited number of observed indicators in the explicit knowledge construct (i.e., untimed GJT and MKT). These values can be improved by increasing the number of items for each test (and thus for each linguistic structure) and the number of indicators for each construct. These two factors, however, go hand-in-hand with practicality and funding, and highlight the need for more collaborative, sponsored research.

Researchers can save on the cost of data-collection by carefully designing a planned missing data longitudinal study. For instance, one can randomly assign participants to have missing items (e.g., multiform designs; Graham, Hofer, & MacKinnon, 1996) or missing measurement occasions (e.g., wave missing designs; Little & Rhemtulla, 2012). These are powerful techniques that can increase power (by allowing researchers to collect more data) and

validity (by reducing fatigue and burden on participants) when used appropriately. As the current study is one of the first longitudinal examinations of the explicit-implicit interface, more research is needed to test the reproducibility and generalizability of the current findings in different learning contexts and to specific subsamples. To do so, SLA researchers must be better trained in the design and analysis of rigorous longitudinal analysis/data-collection techniques to efficiently utilize limited resources.

Second, I cannot exclude practice and retest effects in my results. This may be true for the oral production task that utilized the same story prompt across two time points and to some extent elicited imitation that differed only in the grammaticality of items across the two sets of sentences at T1 and T2. Presumably, the interval between two time points (3-4 month) was sufficiently distanced enough to make it difficult for participants to recall the exact wordings of each sentence. Also, I did not observe a stark improvement of oral production performance (T1 = 89% and T2 = 93%) and elicited imitation (T1 = 64% and T2 = 69%) which would have been the case if practice effects were notable. Nevertheless, the familiarity of the content could have positively influenced participants' test engagement and their performance of oral production and elicited imitation at T2. Generating more stories with picture prompts for oral production was not possible with the current budget, but future researchers could vary the oral production prompts at each time point ensuring comparable difficulties/complexities of the stories. Regarding elicited imitation, researchers may also employ different sets of items after carefully controlling for comparability.

Third, the findings of the current study are based on participants who had a wide range of L2 proficiency level and length of residence in L2 speaking countries. In particular, 18 participants were considered beginning to intermediate L2 users (TOEFL score between 60-78),

80 participants were intermediate to advanced (TOEFL score between 79-100), and 43 participants were considered advanced. Also these participants varied widely in how long they lived in an English speaking country (i.e., from 1 month to almost 10 years). Being inclusive of a wide range of L2 English speakers enabled me to explore the developmental changes of a relatively general population of L2 learners, but the current results may not hold when replicated with different subsamples (i.e., beginning to intermediate L2 speakers with limited experience abroad). Future studies would be served by replicating the present study design with a more tightly controlled population to examine the generalizability of the current results. Another interesting avenue for future research is to compare the *rate of change* (the slope and not the acquisition points) of explicit and implicit knowledge acquired in second and foreign language learning contexts. Ideally, researchers would build a growth curve model with at least 4 time points to explore the process of acquisition (linear or curve) to understand *how* learners arrive at the acquired knowledge. At the end of the day, a highly relevant and interesting question for practitioners would be what trajectory learners follow to reach a certain linguistic product or knowledge, rather than the linguistic product itself.

Last but not least, while the LEL measure captured L2 usage details in a relatively fine-grained manner, it still suffers from subjectivity. A logical next step is to triangulate self-reported data with an objective measure of L2 usage. An excellent device is the Electronically Activated Recorder (EAR, Mehl, Pennebaker, Crow, Dabbs, & Price, 2001) developed by Matthias Mehl and colleagues. Through an iEAR app on Android (available with the 6.0.1 version), EAR captures minimal personal information adequate for reliable coding (i.e., sampling 30 seconds of every 12 minutes), but not beyond. Certified with the NIH Certificate of Confidentiality, EAR



reduces ethical concerns and labor intensity as well (e.g., transcribing and coding 24 hours of conversation).

## **5.5 Conclusions**

In a longitudinal study, I demonstrated a significant reciprocal association between explicit and implicit knowledge alongside considerable individual variation in English engagement patterns. These findings have two implications. First, language acquisition is a developmental process composed of a dynamic interaction between explicit and implicit knowledge and their synergetic relationship; and second, similar affordances to engage in the L2 do not produce comparable amounts of actual L2 engagement for different individuals. These observations may reinforce that the explicit-implicit interface question, and language acquisition more generally, can be better understood when studied over time in a naturalistic context, as language acquisition in its essence is shaped by one's experience with the language in interaction with the contextual affordances in the environment.

## APPENDICES

## APPENDIX A. Background Questionnaire (T1)

### A. Personal information

1. Full name: \_\_\_\_\_
2. Age: \_\_\_\_\_
3. Phone number: \_\_\_\_\_
4. Email address: \_\_\_\_\_
5. Gender:      Female ☐      Male ☐      Prefer not to specify ☐
6. Your latest TOEFL score: \_\_\_\_\_
  - When did you obtain the above TOEFL scores? : \_\_\_\_\_ (e.g., 2018 summer)
7. Final (or current) education (e.g., undergraduate, graduate): \_\_\_\_\_

### B. Language Background

1. What is your native (or first) language?  
: \_\_\_\_\_
2. How long (in MONTHS, NOT YEARS) have you lived in English speaking countries (e.g., USA, UK, Australia, Canada)?  
: \_\_\_\_\_ MONTHS
3. How old were you when you moved to the above English speaking countries? (e.g., USA: 21 years old)  
: \_\_\_\_\_ years old
4. How old were you when you first learned English?  
: \_\_\_\_\_ years old
5. At present, how many hours a day do you use  
English: \_\_\_\_\_ hours  
Another language (please specify): \_\_\_\_\_ language: \_\_\_\_\_ hours  
Another language (please specify): \_\_\_\_\_ language: \_\_\_\_\_ hours

### C. Language Learning Experience

1. How many years have you studied English at school?  
: \_\_\_\_\_ years

2. What was the instruction of English classes that you received at school like? (Circle the best answer)
- A. Mainly grammar-oriented instruction (i.e. a lot of time was spent studying grammar)
  - B. Mainly communication-oriented instruction (i.e. most of the time was spent communicating in English)
  - C. A mixture of grammar- and communication-oriented instruction

## APPENDIX B. Motivation Questionnaire

Here are a number of statements that may or may not apply to you. There are no right or wrong answers, so please answer honestly, considering how you compare to most people.

|  | Not at<br>all like<br>me | Not<br>much<br>like<br>me | Some-<br>what<br>like me | Mostly<br>like me | Very<br>much<br>like<br>me |
|--|--------------------------|---------------------------|--------------------------|-------------------|----------------------------|
| 1. New ideas and projects sometimes distract me from previous ones.                                | 1                        | 2                         | 3                        | 4                 | 5                          |
| 2. Setbacks (e.g., events that delay your progress) don't discourage me. I don't give up easily.   | 1                        | 2                         | 3                        | 4                 | 5                          |
| 3. I often set a goal but later choose to pursue a different one.                                  | 1                        | 2                         | 3                        | 4                 | 5                          |
| 4. I am a hard worker.   | 1                        | 2                         | 3                        | 4                 | 5                          |
| 5. I have difficulty maintaining my focus on projects that take more than a few months to complete | 1                        | 2                         | 3                        | 4                 | 5                          |
| 6. I finish whatever I begin.  | 1                        | 2                         | 3                        | 4                 | 5                          |
| 7. My interests change from year to year.  | 1                        | 2                         | 3                        | 4                 | 5                          |
| 8. I am diligent. I never give up.   | 1                        | 2                         | 3                        | 4                 | 5                          |
| 9. I have been obsessed with a certain idea or project for a short time but later lost interest.   | 1                        | 2                         | 3                        | 4                 | 5                          |
| 10. I have overcome setbacks to conquer an important challenge.                                    | 1                        | 2                         | 3                        | 4                 | 5                          |

## APPENDIX C. Background Questionnaire (T2)

The information that you provide below will help us to better understand your language experiences.

Your honest and detailed responses will be greatly appreciated. – Kathy

1. What language(s) do you spend at home? \_\_\_\_\_ (e.g., Korean and English)

2. Have you taken any classes this Spring semester?

a. Yes: \_\_\_\_\_

b. No: \_\_\_\_\_

c. Others: \_\_\_\_\_

2.1. If yes, please fill in the blanks on the courses you took this Spring semester:

| Course name                          | Course number  | Class schedule (date and time)                        | Please estimate the time of using the skills in percentage (%)  |
|--------------------------------------|----------------|---|---|
| (e.g., Methods of Language Teaching) | (e.g., LLT307) | (e.g., Tuesday: 10:20-11:30<br>Thursday: 11:00-13:00) | Listening in English: _____%<br>Speaking in English: _____%<br>Reading in English: _____%<br>Writing in English: _____% |
|                                      |                |   | Listening in English: _____%<br>Speaking in English: _____%<br>Reading in English: _____%<br>Writing in English: _____% |
|                                      |                |   | Listening in English: _____%<br>Speaking in English: _____%<br>Reading in English: _____%                               |

|  |  |  |   |
|--|--|--|---|
|  |  |  | Writing in English: _____%  |
|  |  |  | Listening in English: _____%<br>Speaking in English: _____%<br>Reading in English: _____%<br>Writing in English: _____% |
|  |  |  | Listening in English: _____%<br>Speaking in English: _____%<br>Reading in English: _____%<br>Writing in English: _____% |

3. How many English courses (e.g., classes that teach you how to improve English skills in listening, speaking, reading, or writing) have you taken this semester? \_\_\_\_\_ classes (e.g., 2)

4. Please estimate the time of using the following skills **in the weekends this semester:**

<Saturdays>

1. How many hours do you SPEAK ENGLISH (e.g., speaking to your native English friends) on Saturdays?

0–1hrs    1–2hrs    2–3hrs    3–4hrs    4–5hrs    more than 5hrs

1.1. When SPEAKING in English, what is typically your major focus of processing?

- a. Focused on meaning (e.g., focused on the message not grammar or forms)
- b. Focused on form (e.g., any occasion when the purpose is to learn/understand language features, such as grammar or words of English)
- c. Focused on both meaning and form

2. How many hours do you LISTEN TO ENGLISH (e.g., watching TV, taking online lectures) on Saturdays?

0–1hrs    1–2hrs    2–3hrs    3–4hrs    4–5hrs    more than 5hrs

2.1. When LISTENING in English, what is typically your major focus of processing?

- a. Focused on meaning (e.g., focused on the message not grammar or forms)
- b. Focused on form (e.g., any occasion when the purpose is to learn/understand language features, such as grammar or words of English)
- c. Focused on both meaning and form

3. How many hours do you WRITE ENGLISH (e.g., writing papers, emails, texts, posting comments on social media)?

0–1hrs    1–2hrs    2–3hrs    3–4hrs    4–5hrs    more than 5hrs

3.1. When WRITING in English, what is typically your major focus of processing?

- a. Focused on meaning (e.g., focused on the message not grammar or forms)
- b. Focused on form (e.g., any occasion when the purpose is to learn/understand language features, such as grammar or words of English)
- c. Focused on both meaning and form

4. How many hours do you READ ENGLISH (e.g., reading articles, surfing the Internet, reading novels or magazines)?

0–1hrs    1–2hrs    2–3hrs    3–4hrs    4–5hrs    more than 5hrs

4.1. When READING in English, what is typically your major focus of processing?

- a. Focused on meaning (e.g., focused on the message not grammar or forms)
- b. Focused on form (e.g., any occasion when the purpose is to learn/understand language features, such as grammar or words of English)
- c. Focused on both meaning and form



5. How many hours do you use your NATIVE LANGUAGE (e.g. speaking, reading, writing, or reading in your native language)?

0–1hrs    1–2hrs    2–3hrs    3–4hrs    4–5hrs    more than 5hrs

6. How many hours do you use OTHER LANGUAGES (e.g. speaking, reading, writing, or reading in other languages)?

0–1hrs    1–2hrs    2–3hrs    3–4hrs    4–5hrs    more than 5hrs

This is the end of the background questionnaire. Thank you very much for your input.

## APPENDIX D. Stimuli

Note. Item numbers starting with F stand for fillers.

| Task Time | Item# | Sentence   |
|-----------|-------|--|
| EI_T1     | 01    | Everyone love to read comic books as a child                                     |
| EI_T1     | 02    | A good teacher make learning a joy for students.                                 |
| EI_T1     | 03    | Technology plays an important role in language learning nowadays.                |
| EI_T1     | 04    | Regular exercise helps people maintain a normal weight.                          |
| EI_T1     | 05    | Americans usually like to have breads for breakfast.                             |
| EI_T1     | 06    | Young people often seek advices from their parents about finding jobs.           |
| EI_T1     | 07    | Sometimes dogs knock over the trash when they are left alone.                    |
| EI_T1     | 08    | Women like to buy jewellery, necklaces, and rings when they get married.         |
| EI_T1     | 09    | Some seats on planes are reserved for mothers with infants.                      |
| EI_T1     | 10    | Children should not be allowed to stay out late with their friends.              |
| EI_T1     | 11    | Wedding guests should be dress in a suit and tie.                                |
| EI_T1     | 12    | Abraham Lincoln is consider one of the greatest presidents of the United States. |
| EI_T1     | 13    | Carl believes he needs preparing an extra cake for tomorrow.                     |
| EI_T1     | 14    | Jo tells them they have studying hard for each test.                             |
| EI_T1     | 15    | Before dinner she often asks to play with her friends.                           |
| EI_T1     | 16    | Every winter he wants to move to Florida from Michigan.                          |
| EI_T1     | 17    | People are not sure when will scientists find a cure for cancer.                 |
| EI_T1     | 18    | Girls always want to know how do celebrities stay fit.                           |
| EI_T1     | 19    | Kids like to ask their parents why the dinosaurs died out.                       |
| EI_T1     | 20    | Many people are curious about how The Pyramids were built.                       |
| EI_T1     | 21    | It is more harder to learn Japanese than to learn English.                       |
| EI_T1     | 22    | European people tend to be more taller than Asian people.                        |
| EI_T1     | 23    | New Zealand is greener and more beautiful than other countries.                  |
| EI_T1     | 24    | Luxury brands are cheaper in China than in the United States.                    |
| EI_T1     | F5    | People who live in Beijing do not need to worry about traffic jams.              |
| EI_T1     | F6    | It is acceptable for teachers to physically punish students.                     |
| EI_T1     | F7    | Gambling is a good way to earn money.  |
| EI_T1     | F8    | In the 1980s many Chinese people could afford luxury cars.                       |
| EI_T1     | F13   | The Chinese people were the first to land on the moon, aren't they?              |
| EI_T1     | F14   | Spending two hours at the gym is a waste of time, doesn't it?                    |
| EI_T1     | F15   | A good student must to do everything the teacher says.                           |
| EI_T1     | F16   | The software that Taylor Swift invented it changed the world.                    |

|        |     |  |
|--------|-----|--|
| EI_T2  | 01  | Everyone loves to read comic books as a child                                      |
| EI_T2  | 02  | A good teacher makes learning a joy for students.                                  |
| EI_T2  | 03  | Technology play an important role in language learning nowadays.                   |
| EI_T2  | 04  | Regular exercise help people maintain a normal weight.                             |
| EI_T2  | 05  | Americans usually like to have bread for breakfast.                                |
| EI_T2  | 06  | Young people often seek advice from their parents about finding jobs.              |
| EI_T2  | 07  | Sometimes dogs knock over the trashes when they are left alone.                    |
| EI_T2  | 08  | Women like to buy jewelleryes, necklaces, and rings when they get married.         |
| EI_T2  | 09  | Some seats on planes are reserve for mothers with infants.                         |
| EI_T2  | 10  | Children should not be allowed to stay out late with their friends.                |
| EI_T2  | 11  | Wedding guests should be dressed in a suit and tie.                                |
| EI_T2  | 12  | Abraham Lincoln is considered one of the greatest presidents of the United States. |
| EI_T2  | 13  | Carl believes he needs to prepare an extra cake for tomorrow.                      |
| EI_T2  | 14  | Jo tells them they have to study hard for each test.                               |
| EI_T2  | 15  | Before dinner she often asks playing with her friends.                             |
| EI_T2  | 16  | Every winter he wants moving to Florida from Michigan.                             |
| EI_T2  | 17  | People are not sure when scientists will find a cure for cancer.                   |
| EI_T2  | 18  | GIrls always want to know how celebrities stay fit.                                |
| EI_T2  | 19  | Kids like to ask their parents why did the dinosaurs died out.                     |
| EI_T2  | 20  | Many people are curious about how were The Pyramids built.                         |
| EI_T2  | 21  | It is harder to learn Japanese than to learn English.                              |
| EI_T2  | 22  | European people tend to be taller than Asian people.                               |
| EI_T2  | 23  | New Zealand is more greener and more beautiful than other countries.               |
| EI_T2  | 24  | Luxury brands are more cheaper in China than in the United States.                 |
| EI_T2  | F5  | People who live in Beijing do not need to worry about traffic jams.                |
| EI_T2  | F6  | It is acceptable for teachers to physically punish students.                       |
| EI_T2  | F7  | Gambling is a good way to earn money.  |
| EI_T2  | F8  | In the 1980s many Chinese people could afford luxury cars.                         |
| EI_T2  | F13 | The Chinese people were the first to land on the moon, aren't they?                |
| EI_T2  | F14 | Spending two hours at the gym is a waste of time, doesn't it?                      |
| EI_T2  | F15 | A good student must to do everything the teacher says.                             |
| EI_T2  | F16 | The software that Taylor Swift invented it changed the world.                      |
| MKT_T1 | 1   | William lives in Ann Arbor but work in East Lansing                                |
| MKT_T1 | 2   | Martin's presentation was post on Facebook by his classmates                       |
| MKT_T1 | 3   | I asked Alan when is he going to play basketball.                                  |
| MKT_T1 | 4   | Diane wants to buy new furnitures and find the cat another home                    |

|         |    |  |
|---------|----|--|
| MKT_T1  | 5  | Lucy feels she needs asking for help with learning English.                  |
| MKT_T1  | 6  | People think he is more nicer and more intelligent than Peter                |
| MKT_T1  | 7  | We were question by the immigration officer at the airport.                  |
| MKT_T1  | 8  | Everybody know that teenagers like to play computer games.                   |
| MKT_T1  | 9  | She wondered why did her boyfriend come later for dinner.                    |
| MKT_T1  | 10 | The temperature is more higher now in winter than it was ten years ago.      |
| MKT_T1  | 11 | On the weekend she asks playing video games for an hour.                     |
| MKT_T1  | 12 | Adam will get help on the homeworks, so he is not worried.                   |
| MKT_T2  | 1  | For the first time, everyone around the world were able to see a black hole. |
| MKT_T2  | 2  | The chef at the restaurant asked me what was I looking for.                  |
| MKT_T2  | 3  | Chinese buildings are often more taller due to a larger population.          |
| MKT_T2  | 4  | The food critic asks trying the specialty on the menu.                       |
| MKT_T2  | 5  | Many students in Korean classes are learning grammars to understand K-pop.   |
| MKT_T2  | 6  | A flight attendant was praise over her kindness towards a tired traveler.    |
| MKT_T2  | 7  | Before bedtime Dustin always wants having sweets.                            |
| MKT_T2  | 8  | Wild dogs are much more happier living with other dogs than alone.           |
| MKT_T2  | 9  | Regular coffee is freshly brew every day at Starbucks.                       |
| MKT_T2  | 10 | The international student didn't know what do marshmallows taste like.       |
| MKT_T2  | 11 | Our fine city continue to attract many tourists every year.                  |
| MKT_T2  | 12 | An increasing number of people in Florida are exposed to fine dusts.         |
| TGJT_T1 | 1  | Their music teacher live close to their house.                               |
| TGJT_T1 | 2  | Our neighbor drink a cold glass of water before going to bed.                |
| TGJT_T1 | 3  | My mother's friend hires new students every summer.                          |
| TGJT_T1 | 4  | The girl next door tries to catch colorful butterflies every June.           |
| TGJT_T1 | 5  | There are enough coffees to drink for each guest.                            |
| TGJT_T1 | 6  | This store sells a lot of spinaches and carrots to local restaurants.        |
| TGJT_T1 | 7  | He should eat enough lettuce and exercise to lose weight.                    |
| TGJT_T1 | 8  | John loves to eat cheese with his morning cereal.                            |
| TGJT_T1 | 9  | The refrigerator was fill with juice packs from Whole Foods.                 |
| TGJT_T1 | 10 | The car was wash at a gas station for free yesterday.                        |
| TGJT_T1 | 11 | The due date for his final project has been postponed to next week.          |
| TGJT_T1 | 12 | Regular coffee is freshly brewed every day at the cafe.                      |
| TGJT_T1 | 13 | The food critic asks trying the specialty items on the menu.                 |
| TGJT_T1 | 14 | After class the girls want shopping for Halloween costumes.                  |

|         |     |   |
|---------|-----|---|
| TGJT_T1 | 15  | Before bedtime Kelly always wants to have sweets.                                 |
| TGJT_T1 | 16  | She feels strange when her guest asks to see the old pictures.                    |
| TGJT_T1 | 17  | Everyone wonders when will he arrive at the conference.                           |
| TGJT_T1 | 18  | The clerk at my favorite store asked me where was I going.                        |
| TGJT_T1 | 19  | He wondered why his parents preferred diet Coke to regular Coke.                  |
| TGJT_T1 | 20  | The daughter asked her parents when she could have a dog.                         |
| TGJT_T1 | 21  | The room temperature is more warmer today because of the sunlight.                |
| TGJT_T1 | 22  | She insisted that her puppy is more prettier than her friend's.                   |
| TGJT_T1 | 23  | I think bagels are sometimes harder than baguettes.                               |
| TGJT_T1 | 24  | It is easier to start a relationship than to maintain it.                         |
| TGJT_T1 | F13 | I don't know how to thank you for your help.                                      |
| TGJT_T1 | F14 | I looked for Mary and Samantha at the bus station.                                |
| TGJT_T1 | F15 | Joe realized that the train was late while he was waiting there.                  |
| TGJT_T1 | F16 | They went through an intensive training for three hours.                          |
| TGJT_T1 | F21 | Because of the test, the student aren't talking in class right now.               |
| TGJT_T1 | F22 | She and her husband gave presents to kids who they came to visit.                 |
| TGJT_T1 | F23 | The researcher is famous to reporting interesting findings.                       |
| TGJT_T1 | F24 | Mary went to Italy in her late teens, and she like it there.                      |
| TGJT_T1 | F33 | She was happy because she finished her assignment ahead of time.                  |
| TGJT_T1 | F34 | He is looking forward to going home and seeing his parents.                       |
| TGJT_T1 | F35 | I enjoy riding a bike during summer because the weather is so nice.               |
| TGJT_T1 | F36 | I would like to have a cup of tea because I feel cold.                            |
| TGJT_T1 | F37 | My professor told me that I should participate in class discussion more actively. |
| TGJT_T1 | F38 | I wrote an email to my friend to congratulate her on her graduation.              |
| TGJT_T1 | F39 | I was looking for my glasses for five hours yesterday.                            |
| TGJT_T1 | F40 | My friend is getting a job in London.   |
| TGJT_T2 | 1   | Their small town continue to attract some foreign travelers.                      |
| TGJT_T2 | 2   | The grandmother send text messages to her grandchildren after lunch.              |
| TGJT_T2 | 3   | The schoolboy shows his new toy to his friend.                                    |
| TGJT_T2 | 4   | The woman in a hat bakes warm cookies in her kitchen.                             |
| TGJT_T2 | 5   | The father orders a lot of soups for his family.                                  |
| TGJT_T2 | 6   | The village was full of smokes from the forest fire.                              |
| TGJT_T2 | 7   | They will use a lot of butter to make her favorite dessert.                       |
| TGJT_T2 | 8   | Her son put some broccoli in my daughter's bowl.                                  |
| TGJT_T2 | 9   | Her apartment was clean by her husband for today's baby shower.                   |
| TGJT_T2 | 10  | The plane was delay because of a huge snow storm.                                 |
| TGJT_T2 | 11  | She was invited to a housewarming party yesterday.                                |

|         |     |  |
|---------|-----|--|
| TGJT_T2 | 12  | The door of my first apartment was painted green.                      |
| TGJT_T2 | 13  | Lizz knows she needs passing the exam to get the certificate.          |
| TGJT_T2 | 14  | Many people think women have changing their last names after marriage. |
| TGJT_T2 | 15  | Many children know they have to wash their hands before eating.        |
| TGJT_T2 | 16  | Jim is told that his parents want to buy a new house.                  |
| TGJT_T2 | 17  | I don't know when did she decide to leave us.                          |
| TGJT_T2 | 18  | The freshman did not know where was the library.                       |
| TGJT_T2 | 19  | I didn't know what a taco tastes like.                                 |
| TGJT_T2 | 20  | My grandmother did not remember where she bought her TV.               |
| TGJT_T2 | 21  | We are more louder than those on the lower level.                      |
| TGJT_T2 | 22  | My house in China is more bigger than my house in the USA.             |
| TGJT_T2 | 23  | I think I am luckier than the others.                                  |
| TGJT_T2 | 24  | People drive faster in rural areas than in the city.                   |
| TGJT_T2 | F1  | Next week my cousin is coming from Mumbai.                             |
| TGJT_T2 | F10 | Through thick and thin I will stand by you.                            |
| TGJT_T2 | F11 | People think that you are always complaining about something.          |
| TGJT_T2 | F12 | I have come to the end of my patience with you.                        |
| TGJT_T2 | F17 | Tom met the man who he comes from Japan.                               |
| TGJT_T2 | F18 | At the end of the year, he will going to travel a lot.                 |
| TGJT_T2 | F19 | Famous restaurants is located in New York City.                        |
| TGJT_T2 | F2  | My friend Surya is getting married next month in India.                |
| TGJT_T2 | F20 | He lost one of the book that I borrowed from my teacher.               |
| TGJT_T2 | F3  | I am excited because I am leaving for New York tomorrow night.         |
| TGJT_T2 | F4  | He came to my office to ask for money and help.                        |
| TGJT_T2 | F5  | Normally, I have tea in the morning and coffee in the afternoon.       |
| TGJT_T2 | F6  | They have to commute to work for five hours every day.                 |
| TGJT_T2 | F7  | I go through the newspaper headlines every day.                        |
| TGJT_T2 | F8  | Everybody makes remarkable mistakes sometimes as you know.             |
| TGJT_T2 | F9  | I think there is always a solution to any problem.                     |
| UGJT_T1 | 1   | The Spanish university professor offer candy to his young children.    |
| UGJT_T1 | 2   | The woman on TV tell us about the weather every morning.               |
| UGJT_T1 | 3   | The girl in a coat makes treats for her cat.                           |
| UGJT_T1 | 4   | Their daughter watches movies on DVD every night.                      |
| UGJT_T1 | 5   | There are enough corns planted in my mother's garden.                  |
| UGJT_T1 | 6   | His brother never buys a lot of bacons from that grocery store.        |
| UGJT_T1 | 7   | The teenagers saw them make some popcorn in the small pot.             |
| UGJT_T1 | 8   | They sold us a lot of tea from his father's herbal farm.               |
| UGJT_T1 | 9   | The construction nearby my school will be finish next year.            |

|         |     |  |
|---------|-----|--|
| UGJT_T1 | 10  | My new wallet was chew up by my puppy yesterday.                                   |
| UGJT_T1 | 11  | This red car was purchased from a Chinese car dealer.                              |
| UGJT_T1 | 12  | My essay for American history class was graded as the best.                        |
| UGJT_T1 | 13  | At the meeting she asks speaking with the president in private.                    |
| UGJT_T1 | 14  | Many teachers believe students need learning critical thinking skills.             |
| UGJT_T1 | 15  | Her close friend often asks to join the reading group on Fridays.                  |
| UGJT_T1 | 16  | Next Sunday Kimberly has to walk her dog to the park.                              |
| UGJT_T1 | 17  | The Amazon customer survey asked how would I rate my previous shopping experience. |
| UGJT_T1 | 18  | The woman asked me where did I buy my shoes.                                       |
| UGJT_T1 | 19  | The professor explained to him why he could not pass the course.                   |
| UGJT_T1 | 20  | A girl called her friend to ask what she should bring to the camp.                 |
| UGJT_T1 | 21  | The southern desert climate is more drier during the summer months.                |
| UGJT_T1 | 22  | The baker's cookies turned out more darker the second time he made them.           |
| UGJT_T1 | 23  | He is happier than last week because he has few assignments.                       |
| UGJT_T1 | 24  | My bag gets heavier as I shop at Macy's.   |
| UGJT_T1 | F29 | Based on the program, the president will be visit our company.                     |
| UGJT_T1 | F30 | People should be report stolen bikes to the police.                                |
| UGJT_T1 | F31 | They always have arguments with the neighbors who they live next door.             |
| UGJT_T1 | F32 | People can winning large amounts of money in gambling.                             |
| UGJT_T1 | F41 | Your phone has been ringing since you left the room.                               |
| UGJT_T1 | F42 | Paul is planning to go on a picnic this weekend.                                   |
| UGJT_T1 | F43 | Kathy is looking forward to throwing a surprise party for her friends.             |
| UGJT_T1 | F44 | Jeff goes out on a date with his wife every Tuesday.                               |
| UGJT_T1 | F57 | I am used to having dinner by myself.  |
| UGJT_T1 | F58 | She asked me if I could recall any good memories from my school years.             |
| UGJT_T1 | F59 | My research interests have been developing as I learn about different topics.      |
| UGJT_T1 | F60 | It is not surprising that she broke her new phone.                                 |
| UGJT_T1 | F61 | Last spring semester was the toughest but the most rewarding for me.               |
| UGJT_T1 | F62 | She recommended that I recycle the plastic bags to save the environment.           |
| UGJT_T1 | F63 | The best teacher in my life shaped my view of the world.                           |
| UGJT_T1 | F64 | The mom scolded her kids because they screamed in the restaurant.                  |

|         |     |   |
|---------|-----|---|
| UGJT_T2 | 1   | That woman in red enjoy watching movies during the weekend.               |
| UGJT_T2 | 2   | Our math teacher buy math books for us every semester.                    |
| UGJT_T2 | 3   | Her grandmother spends all her free time with her friends.                |
| UGJT_T2 | 4   | The boy in blue sells many green apples at the market.                    |
| UGJT_T2 | 5   | There are a lot of sands in the back of our garden.                       |
| UGJT_T2 | 6   | That shopping mall sells a lot of jewelries during their annual sale.     |
| UGJT_T2 | 7   | The mother put some wood outside by the campfire.                         |
| UGJT_T2 | 8   | Her friend cooks a lot of rice on weekends and holidays.                  |
| UGJT_T2 | 9   | My bedroom door is close when I leave home.                               |
| UGJT_T2 | 10  | A new safety video is create by staff memebbers every year.               |
| UGJT_T2 | 11  | The book is summarized in the first chapter of the edited volume.         |
| UGJT_T2 | 12  | Flowers in the front yard are watered by Dad every day.                   |
| UGJT_T2 | 13  | Each weekend he needs sleeping after exercising so hard.                  |
| UGJT_T2 | 14  | Paul is curious why he has working on multiple projects.                  |
| UGJT_T2 | 15  | He knows kittens need to eat wet and dry food.                            |
| UGJT_T2 | 16  | She understands why Peter wants to postpone the meeting.                  |
| UGJT_T2 | 17  | He wanted to explain to her why was he late for the class.                |
| UGJT_T2 | 18  | The student did not remember when did the first class start.              |
| UGJT_T2 | 19  | The barista asked me what I would like to order.                          |
| UGJT_T2 | 20  | She asked her boyfriend why he had not answered her call last night.      |
| UGJT_T2 | 21  | My younger brother is much more taller than last year.                    |
| UGJT_T2 | 22  | The red balloon floated more higher than the blue one.                    |
| UGJT_T2 | 23  | The client asked him to announce the results sooner than the deadline.    |
| UGJT_T2 | 24  | The floor lamp is brighter because I replaced the light bulb.             |
| UGJT_T2 | F25 | She told her staff that smoking not allowed in the office.                |
| UGJT_T2 | F26 | I have talked to the teacher who she works in the nearby school.          |
| UGJT_T2 | F27 | Not everyone cannot successfully learn music in a year.                   |
| UGJT_T2 | F28 | The population of the world increases a lot last year.                    |
| UGJT_T2 | F45 | My classmates have been working hard during this summer vacation.         |
| UGJT_T2 | F46 | Wendy got her ears pierced because she has been longing for some changes. |
| UGJT_T2 | F47 | Min is afraid of seeing dogs running around her.                          |
| UGJT_T2 | F48 | My girlfiend is picking up Korean as she watches many Korean movies.      |
| UGJT_T2 | F49 | My hometown is famous for its beautiful scenery.                          |
| UGJT_T2 | F50 | I usually chat with my friends in China when I feel homesick.             |



|         |     |   |
|---------|-----|---|
| UGJT_T2 | F51 | Many friends of mine are planning to work abroad after they graduate.   |
| UGJT_T2 | F52 | There is a lot of construction during summer in Michigan.   |
| UGJT_T2 | F53 | I have listed my goals for this academic year during the summer vacation.   |
| UGJT_T2 | F54 | Based on the weather forecast, I packed an umbrella for tomorrow.   |
| UGJT_T2 | F55 | The university announced that it will be a smoke-free campus from next year.  |
| UGJT_T2 | F56 | The heated discussion on a political topic seemed to never end.   |
| OP      |     | <p>Every morning, Mr. Lee gets up at 6:00 am. He has some tea and three slices of bread for breakfast. He starts his work at 8 am and finishes at 6 pm to make as much money as he can. From 9 pm to 12 am he works as a bartender in a local bar. He never takes a day off or goes on vacations. His friends often say that if Mr. Lee was missing from work, something would be terribly wrong. Actually, Mr. Lee's manager job is good enough for him. His friends are all curious about why Mr. Lee needs to earn so much money. They think that Mr. Lee wants to have a richer and more luxurious life in the future.</p> <p>Yesterday, Mr. Lee told them what he has been doing with the money he earned. During college he was selected among 100 students as a volunteer to teach English in China. When he was in China, Mr. Lee met a little girl. The little girl was really talented, but her family was too poor to support her schooling. The little girl told Mr. Lee that she would be so happy if she was able to finish her high school. Mr. Lee really wanted to help this little girl and send her to college so he worked harder and even more hours than before.</p> <p>Yesterday the little girl told Mr. Lee that she was accepted by one of the top universities in China. Mr. Lee's friends were so surprised and so proud of him. For Mr. Lee there was no greater achievement in life than helping the little girl, even though he had less savings than his friends.</p> |

## APPENDIX E. Language Exposure Log

| Types of general language usage | Specific <b>activities</b>  |
|---------------------------------|---|
| Speaking in English             | Academic-related conversation/discussion (e.g., on specific topic w/ academic advisors, friends, classmates etc)<br>Casual conversation/discussion (e.g., on general/personal topics)<br>Teaching<br>Giving presentations<br>Others   |
| Writing in English              | Writing emails<br>Writing academic papers<br>Messaging with friends<br>Making a presentation (to a class, a group, or the public)<br>Personal writing/journal (such as a diary)<br>Others   |
| Reading in English              | Reading non-academic text (e.g., novels, comics, news, magazine)<br>Reading an academic article/text<br>Marking<br>Surfing the Internet (e.g., reading updates on facebook, twitter etc)<br>Others  |
| Listening in English            | Listening to a presentation/lecture (by professor, lecturer, classmate)<br>Watching TV/movie<br>Listening to music/radio<br>Others  |
| Using native or other languages | Listening in native or other languages (e.g., watching tv, listening to radio, listening to lecture)<br>Speaking in native or other languages (e.g., having conversation with professor, friends, classmates etc)<br>Reading in native or other languages (e.g., reading novels, papers, web text)<br>Writing in native or other languages (e.g., messages, emails, papers)<br>Others |
| No language use                 | Doing nothing<br>Eating<br>Exercising (ex: jogging/ working out/ walking/ swimming)<br>Thinking (ex: planning/ praying/ problem solving/ remembering)<br>Sleeping   |

---

Chores and daily tasks (ex: cooking/doing laundry/cleaning  
house/office/yard/packing your bag/organizing your materials)  
Collecting data/doing an experiment  
Programming experiments  
Playing computer games  
Solving non-language related problems (e.g., math equation)  
Shopping (online and offline)  
Hobby (arts and crafts, dancing, etc.)  
Others

---

## APPENDIX F. Instructions on the Web-Based Testing Program

The AIED program has two major components: a front end component consisting of HTML files with embedded Javascript code, and a back end component written in Java, which offers an API to store and retrieve data from the server. The both components are organized into a single Webapp project, which is openable with the development tool Eclipse. I will first show you how to import the project and how to make changes to the setup.

### Basic setup

First fire up Eclipse, and choose from **File** → **Import**. In the Import Wizard, select **Maven/Existing Maven Projects** and navigate to the folder containing the code (you will find a **pom.xml** file at the project root) to import the project.

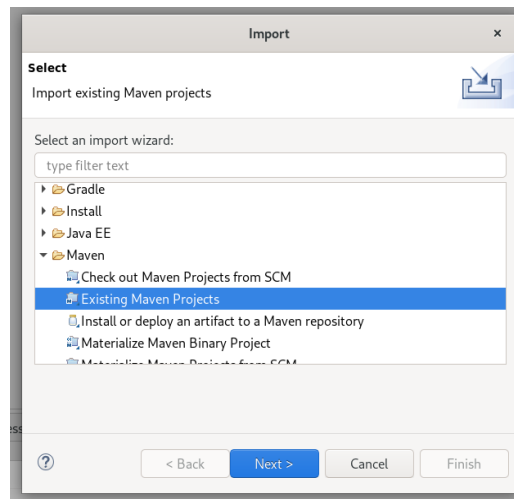


Figure 5.1. Screenshot 1

Once imported, you will see all the files the project requires in the Project Explorer window as shown below. The back-end code which runs on the server is under **Java Resources**, while the front-end code is under **Deployed Resources/webapp**.

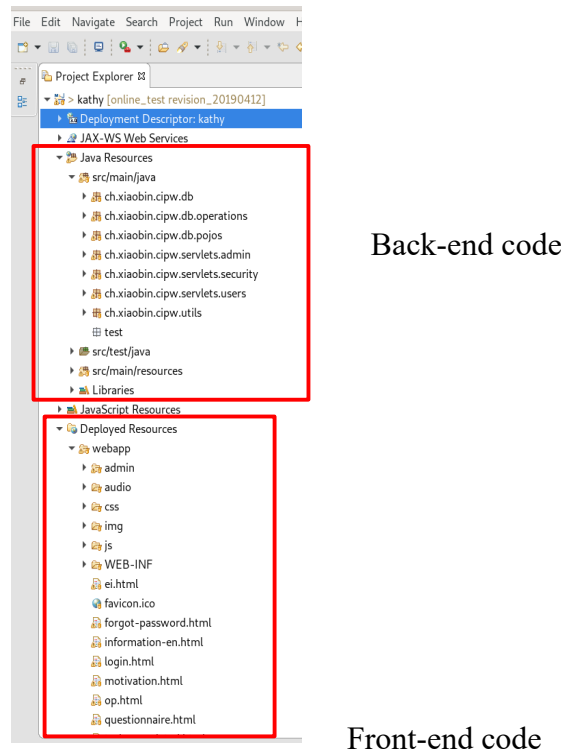


Figure 5.2. Screenshot 2

If you want to change the presentation of the test, you need to change the html files in the “webapp” folder. If you want to change anything related to the back-end, including changing the type of data, number of test items and how they are stored on the server, you would need to change the back-end code.

Before you do any changes, it is better to test if the project already runs in your environment. As said before, the project is a Java Webapp project. As a result, it requires something called a Web Container to run. A Web container is basically an environment to run Java projects for the web. You can think of it as a Web server with Java support. Popular Web containers include Tomcat, Jetty, and so on. We are going to use Jetty to run the project because it is very easy to setup and Jetty has very good documentation.

Before we can run the project under Apache Jetty, we need to install the Jetty plugin for Eclipse. Go to Eclipse's menu **Help** → **Eclipse Marketplace**. Search for “Jetty” in the Marketplace to install it.

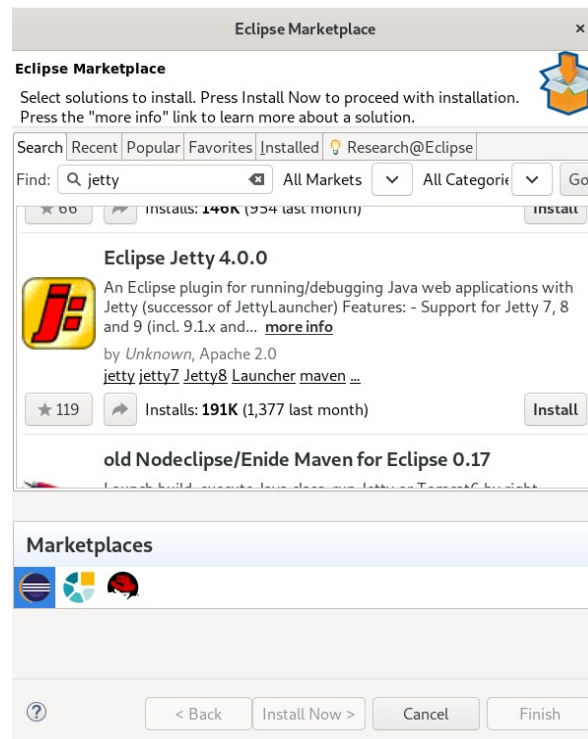


Figure 5.3. Screenshot 3

After installing the Jetty plugin, you can now run the project in Eclipse.

Go to **Run** → **Run Configuration** (if you don't see the menu item, open a random Java file from the Java Resources folder, then select Run → Run Configuration). In the run configuration window, select the Jetty Webapp project, then click the New configuration icon on the top-right corner of the window. Leave everything else by default, and click **Run** to run the project.

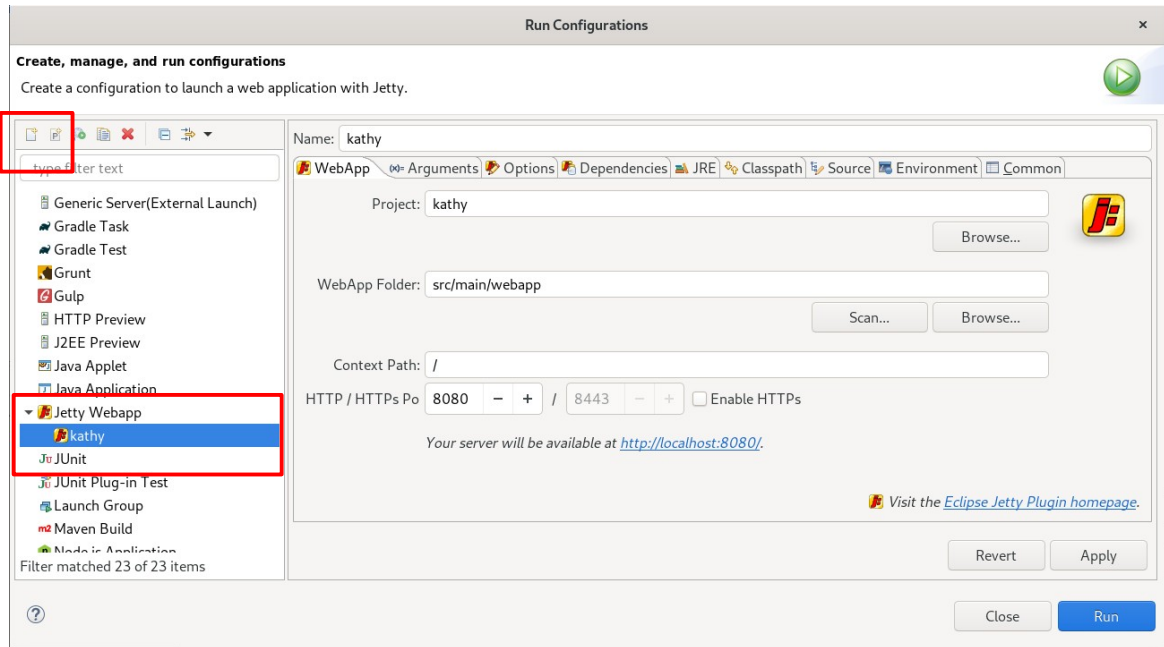


Figure 5.4. Screenshot 4

At this point you should already be able to access the test environment from your Web browser. Try entering the address <http://localhost:8080> in your Web browser. If everything goes well, you will see a login interface to the tests.

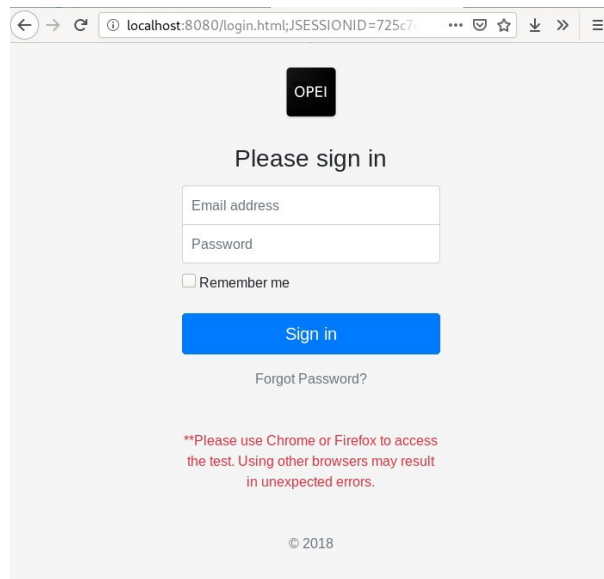
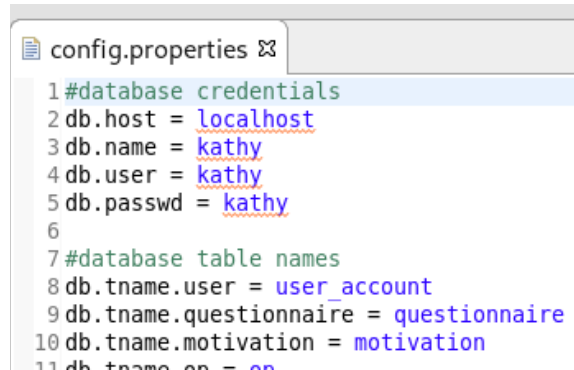


Figure 5.5. Screenshot 5

## Database

In order for the program to run properly, you would also need a database on your server so that the participants' responses to the tests can be stored. The current code uses Postgresql as the database management system. To setup Postgresql server, you can follow the instructions on <https://www.postgresql.org/docs/12/index.html> or ask your system administrator to setup one for you. What you need to get from your system administration is the host address of the database server, the name of the database they created for you, the user and password you can use to access that database.

This information will need to be entered in the **src/main/resources/config.properties** file in the project. In the file, under the “#database credentials” section, enter and replace the values for **db.host**, **db.name**, **db.user**, **db.passwd** with the database credentials you obtained. In the setup shown below, I am accessing my locally installed database, but it can also be a database in another machine or host.



```
config.properties
1 #database credentials
2 db.host = localhost
3 db.name = kathy
4 db.user = kathy
5 db.passwd = kathy
6
7 #database table names
8 db.tname.user = user_account
9 db.tname.questionnaire = questionnaire
10 db.tname.motivation = motivation
11 db.tname...
```

Figure 5.6. Screenshot 6

## Mail server

Another thing to setup for the program to run properly is a mail server with which the program sends the participants credentials for logging into the system to do the tests. Because we want to control who has access to the testing environment, we don't want the system to be open for new-user sign up. The researcher in charge of the experiment will need to gather participant emails and enter them manually into the system to create accounts for the participants. As a result, the system needs a mail server to be able to send notification emails to the participants when an account is setup for them in the system.

The mail server setup is located in the same file as the database setup: **src/main/resources/config.properties**. Under the “#mail server settings”, enter the SMTP



server information as you normally would when setup a mail client. You can find this information from your email service provider.

### Creating participant account

Once everything is setup, restart the application from Eclipse. You may need to stop the previous running instance of the program. Otherwise the internet port would have already been occupied and the new instance will fail to run.

When the program is rebooted, go to <http://localhost:8080/admin/users.html> in your browser. The program would ask you for the admin user name and password to access the admin functions. The

admin credentials are also set in the `src/main/resources/config.properties` file under the section “#Admin credentials”. Enter in the admin page the credentials listed here. Use this admin page to invite participants and later retrieve data for the tests.

To add a new participant, simply enter the participant’s email and click “Add user”. The new user will be recorded in the database and an email will be sent to the participant to inform them about how to participate in the experiment.

Users

| Id | Email | Progress | Group | Created on | Last Login |
|----|-------|----------|-------|------------|------------|
|----|-------|----------|-------|------------|------------|

Figure 5.7. Screenshot 7

The email sent to the participant is customizable. Just search for the file **AdminApiUsersServlet.java** and go to Line 179. This is where the subject and message of the email is set.

The email guides the participants to login the system with the credential and address provided.

## Advanced: Changing the questionnaire and tests

Changing the questionnaire and tests requires changing both the front-end file and the back-end database structure and code to access the database.

For changes to the front-end presentation, open the HTML file you need to change from the **webapp/** folder. Edit the texts or information fields you want to change. This requires knowledge on how to work with HTML and CSS. The pages use the Bootstrap framework for presentation. So some knowledge on how Bootstrap works would be helpful. If you have never heard or worked with these technologies, it is recommended that you ask someone with Web design experience for help. Otherwise you would need to learn at least HTML, CSS, and Bootstrap to be able to make changes to the front-end code. Knowledge about Javascript would also be needed.

The back-end program is responsible for receiving data the participants submit. It is implemented as RESTful APIs under the **servlets/users/** folder. Each test or questionnaire has a corresponding servlet in charge of data recording from and retrieval for the front-end. Change the servlets if you need to change the structure of the data you need to collect.

Data is transferred to and from the servlets in JSON format. The POJOs or data models for each test is located under **db/pojos**. These POJOs are used by the servlets to parse the JSON data passed from the front-end. The **db/operations** folder contains the actual code to handle data transmission between our program and the underlying database management system, namely Postgresql in this case. The operation on each table is implemented as a Java class in the **db/operations** folder. Modifying the operations requires knowledge on SQL. The code in these files are quite self-explanatory and well- documented. Make changes to these files based on your needs.

Again, if any of the technologies sounds unfamiliar to you, ask a Java programmer for help, or learn Java, RESTful API, Servlet, and SQL technologies before you make any changes.

## Deploying the tests

The procedures described above are mostly for testing the program on your own machine. Once you are done making changes and are ready to deploy the tests to your participants, you would need a production server, which runs 24\*7 and is publicly accessible on the Internet so that you participants can do the tests anywhere anytime. Ask your system administrator for the deployment environment if you don't have access to a production server. Once you gained access to such a server, package the program as a WAR file and deploy it in a Web container (usually Tomcat or Jetty).

To package the program, run in a terminal under the root folder of the project files **mvn clean package**, or create an Eclipse **Run Configuration** of Maven Build with the Goals set to “package”. The root folder of the project is where you can find a **pom.xml** file and an **src/** folder.

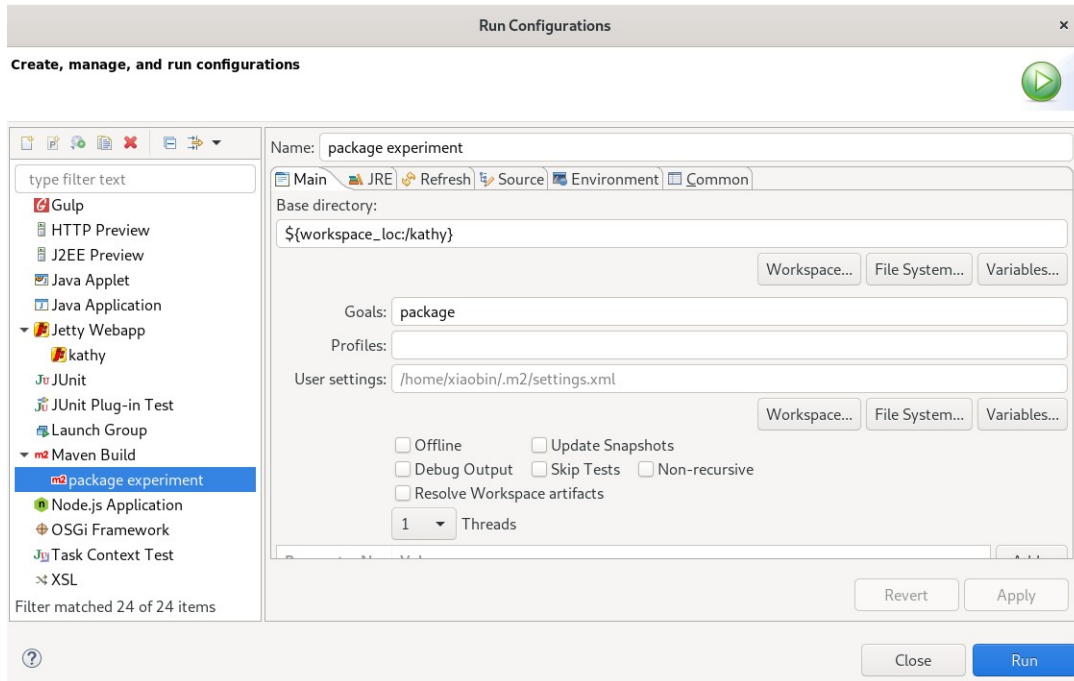
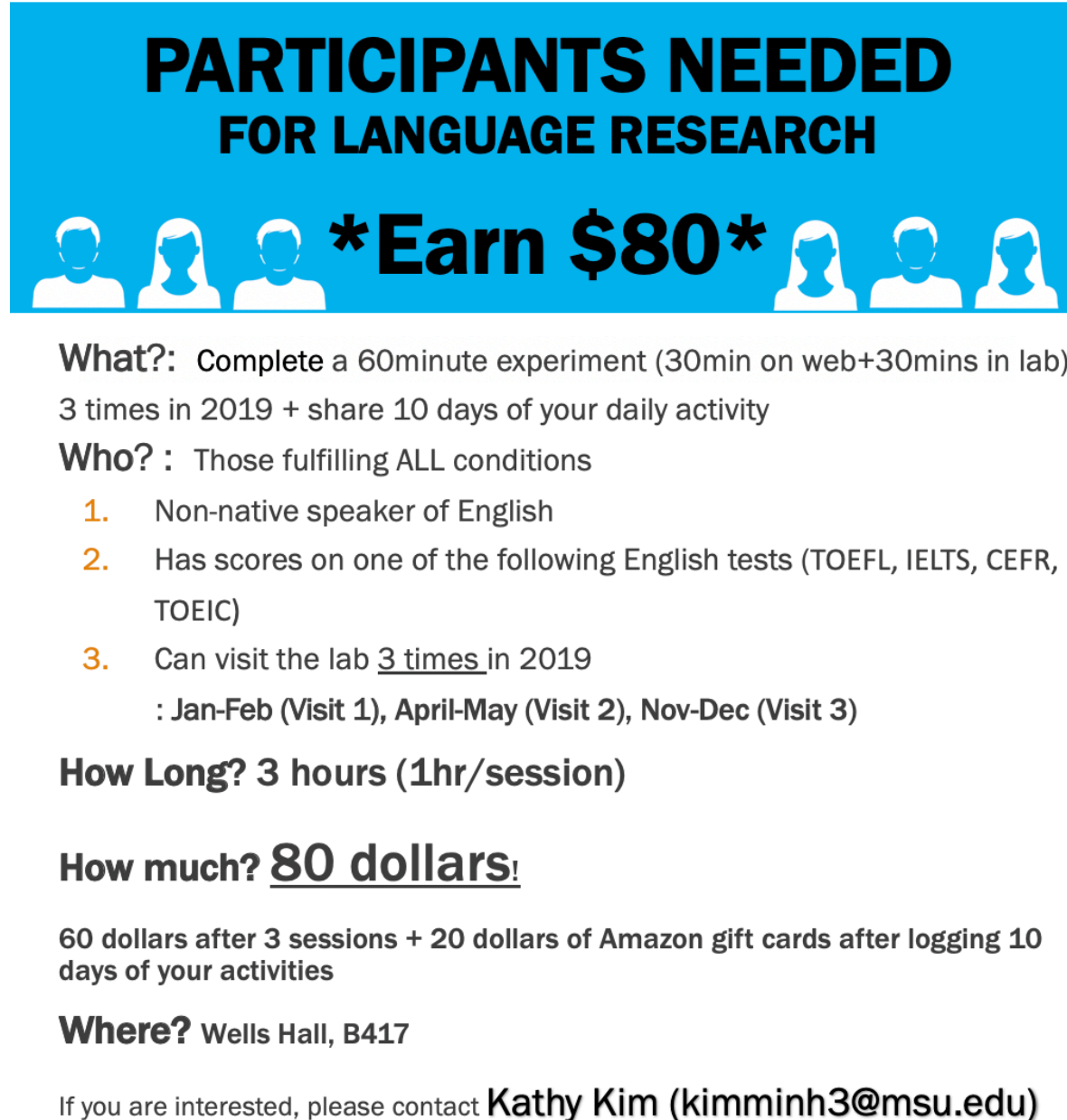


Figure 5.8. Screenshot 8

After running the package command, you get a new folder called **target** in the project root folder under which you will see the packaged WAR file whose filename is suffixed with **.war**. Send this file to your system administrator for deployment or deploy it yourself on the Web container in the production server. This usually is as simple as copying the war file into a folder where the Web container searches for Web applications.

Good luck with your experiments!



**PARTICIPANTS NEEDED  
FOR LANGUAGE RESEARCH**

**\*Earn \$80\***

**What?:** Complete a 60minute experiment (30min on web+30mins in lab)  
3 times in 2019 + share 10 days of your daily activity

**Who? :** Those fulfilling ALL conditions

1. Non-native speaker of English
2. Has scores on one of the following English tests (TOEFL, IELTS, CEFR, TOEIC)
3. Can visit the lab 3 times in 2019  
: Jan-Feb (Visit 1), April-May (Visit 2), Nov-Dec (Visit 3)

**How Long? 3 hours (1hr/session)**

**How much? 80 dollars!**

60 dollars after 3 sessions + 20 dollars of Amazon gift cards after logging 10 days of your activities

**Where?** Wells Hall, B417

If you are interested, please contact **Kathy Kim (kimminh3@msu.edu)**

Figure 5.9. Recruitment Flyer

## APPENDIX H. Oral Production/EI – Coding Guidelines

- Immediate repetitions, including those with repairs, are only coded once. If there is a repair, code the repair rather than the prior attempt.
- Reformulations, using different lexical items, can be coded separately.

The following table lists the target features to be coded, with relevant notes for coders.

Table 5.1 Oral Production/EI – Coding Guidelines

| Feature            | Notes  |
|--------------------|--|
| third-person -s    | Subject-verb agreement for a 3 <sup>rd</sup> person singular subject in the present tense. No copula (e.g., not <i>be</i> ).   |
| mass(/count) noun  | Mass nouns must be in singular form. Mass nouns should not be marked with the -s morpheme that marks plural on countable nouns.  |
| passive            | <i>be</i> -passive construction (ignore get passives, following Spada et al., 2015). Focus is on presence of <i>be</i> and form of verb following <i>be</i> (i.e., past participle).   |
| embedded question  | This is syntactic in nature. In an embedded clause following a <i>wh</i> -word, word order is SVO (e.g., I asked him what he will do with the money); no inversion is permitted. Do not code relative clauses. Ignore issues with complementizer choice (e.g., double complementizer: <i>told his friend that why he...</i> ). Be careful not to code reported speech (e.g., <i>they say what would we do without him</i> . [transcription may not include quotation marks or indicate intonation with a question mark]) |
| comparative adj.   | Adjectives can be marked with the suffix <i>-er</i> (1-2 syllable adjectives) or preceded by <i>more</i> (>2 syllable adjectives).   |
| to-verb complement | Code for the target verbs (need, have, want, ask) which require <i>to</i> -infinitive verb complements. Only code for occasions with verb complements (e.g., wanted to help), ignore noun complements, etc. Focus on the complement, ignore agreement or tense issues on the target (head) verb.   |

## APPENDIX I. Metalinguistic Knowledge Test: Scoring Guide

### General principles

1. To score a point for explanation, participants need to explain why they make such corrections; providing the rule and metalinguistic terms is often evidence of metalinguistic knowledge.
2. Mere description of the correction does not suffice an explanation.
3. If one part of an explanation violates the rule/contradicts another part of the explanation, it should be scored as incorrect.
4. Do not penalize for non-relevant mistakes/errors. For example, if the verb complement error “needs asking” is corrected to “need to ask”, the learner should still get the point, as the original error was corrected, despite the learner introducing a 3<sup>rd</sup>-person singular error. Learners vary in how much of the sentence they reproduce during corrections, which is not strictly relevant to what the item is targeting.

### 1. Third-person – s OR subject-verb agreement

Full explanation for third-person singular should consist of = because the noun/subject/“the name of the word appearing in the sentence” is singular, verb + s should be added to the verb/not plural.

*Participants will have to mention:*

- (1) the (noun / subject) is (singular / third-person / third person singular)

Full explanation for subject-verb agreement should consist of = because the subject is a plural noun, the verb takes a plural form

*Participants will have to mention:*

- (1) the (noun / subject) is (plural / is not singular)

### 2. Mass/count nouns

Full explanation should consist of = mention of specific terminology, such as countable/uncountable or can be plural/cannot be plural.

*Participants will have to mention:*

- (1) the noun (is not countable/ is uncountable /cannot be plural / cannot have -s added to it)

### 3. Be passive

Full explanation should consist of = mention of the subject is a receiver of an action or not an active subject, and thus requires a passive verb form.

*Participants will have to mention:*

- (1) the subject is (the receiver of an action / not an active subject); OR
- (2) passive voice is needed here.

### 5. Embedded questions

Full explanation should consist of = the position of the subject and an auxiliary verb or the verb “to be”/ a verb have to follow affirmative sentence order, not the word order as in direct questions.

*Participants will have to mention:*

- (1) follow word order (subject + [auxiliary] verb) of a statement; OR
- (2) follow SVO [word order]; OR
- (3) inversion is not needed

## **6. Comparatives**

Full/required explanation should consist of = add **-er** to one-syllable adjectives to make a comparison or add **more** in front of adjectives with more than two syllables.

*Participants will have to mention:*

- (1) **-er** is used for (short words / words with one syllables); OR
- (2) **more** is used for (long words / words with two or more syllables); OR
- (3) double marking is not needed

## **7. Gerund/Infinitives**

Full/required explanation should consist of = the infinitive form is needed after the initial verb in the sentence. Only one verb takes inflection in English clauses. A mere statement such as “need to do something” is not qualified; also responses such as “it is a convention” is not qualified.

*Participants will have to mention:*

- (1) [after to / after certain verbs] (infinitive/base/dictionary form) of the verb is needed; OR
- (2) [after to / after certain verbs] (-ing form/gerund form) is not needed

## **Notes:**

(alternatives / other options acceptable)  
[optional part of the response]

## REFERENCES



## REFERENCES

- Andringa, S., & Curcic, M. (2015). How explicit knowledge affects online L2 processing. *Studies in Second Language Acquisition*, 37(2), 237–268. <https://doi.org/10.1017/s0272263115000017>
- Batterink, L., Oudiette, D., Reber, P. J., & Paller, K. A. (2014). Sleep facilitates learning a new linguistic rule. *Neuropsychologia*, 65, 169–179. <https://doi.org/10.1016/j.neuropsychologia.2014.10.024>
- Bentler, P. M., & Chou, C. P. (1987). Practical issues in structural modeling. *Sociological Methods & Research*, 16(1), 78–117. <https://doi.org/10.1177/0049124187016001004>
- Bollen, K. A., & Curran, P. J. (2006). *Latent curve models: A structural equation perspective*. John Wiley & Sons.
- Boomsma A. (1982). Robustness of LISREL against small sample sizes in factor analysis models. In K.G. Joreskog, H. Wold, (Eds.), *Systems under indirection observation: Causality, structure, prediction* (pp.149–173). North Holland.
- Bowles, M. A. (2011). Measuring implicit and explicit linguistic knowledge. *Studies in Second Language Acquisition*, 33(2), 247–271. <https://doi.org/10.1017/s0272263110000756>
- Brown, T. A. (2015). *Confirmatory factor analysis for applied research*. Guilford publications.
- Browne, M. W., & Cudeck, R. (1993). Alternative ways of assessing model fit. In K. A. Bollen & J. S Long (Eds.), *Testing structural equation models* (pp. 136–136), SAGE. <https://doi.org/10.1177/0049124192021002005>
- Cintrón-Valentín, M., & Ellis, N. C. (2015). Exploring the interface: Explicit focus-on-form instruction and learned attentional biases in L2 Latin. *Studies in Second Language Acquisition*, 37(2), 197–235. <https://doi.org/10.1017/s0272263115000029>
- Cleeremans, A. (2007). Consciousness: the radical plasticity thesis. *Progress in Brain Research*, 168, 19–33. [https://doi.org/10.1016/s0079-6123\(07\)68003-0](https://doi.org/10.1016/s0079-6123(07)68003-0)
- Curcic, M., Andringa, S., & Kuiken, F. (2019). The role of awareness and cognitive aptitudes in L2 predictive language processing. *Language Learning*, 69, 42–71. <https://doi.org/10.1111/lang.12321>
- De Bot, K., Lowie, W., & Verspoor, M. (2007). A dynamic systems theory approach to second language acquisition. *Bilingualism: Language and Cognition*, 10, 7–21. <https://doi.org/10.1017/s1366728906002732>

- DeKeyser, R. M. (2003). Implicit and explicit learning. In C. J. Doughty & M. Long (Eds.), *The handbook of second language acquisition* (pp. 313–348). Blackwell. <https://doi.org/10.1002/9780470756492.ch11>
- DeKeyser, R. (2007). *Practice in a second language: Perspectives from applied linguistics and cognitive psychology*. Cambridge University Press. <https://doi.org/10.1017/cbo9780511667275>
- DeKeyser, R. M. (2014). Skill acquisition theory. In B. VanPatten & J. Williams (Eds.), *Theories in second language acquisition* (2nd ed.), (pp. 221–264). Erlbaum.
- DeKeyser, R. (2017). Knowledge and skill in ISLA. In S. Loewen & M. Sato (Eds.), *The Routledge handbook of instructed second language acquisition* (pp. 15–32). Routledge. <https://doi.org/10.4324/9781315676968-2>
- Dewey, D. P., Belnap, R. K., & Hillstrom, R. (2013). Social network development, language use, and language acquisition during study abroad: Arabic language learners' perspectives. *Frontiers: The Interdisciplinary Journal of Study Abroad*, 22, 84–110. <https://doi.org/10.36366/frontiers.v22i1.320>
- Duckworth, A. L., Peterson, C., Matthews, M. D., & Kelly, D. R. (2007). Grit: perseverance and passion for long-term goals. *Journal of Personality and Social Psychology*, 92, 1087–1101. <https://doi.org/10.1037/0022-3514.92.6.1087>
- Ellis, N. C. (2002). Frequency effects in language processing: A review with implications for theories of implicit and explicit language acquisition. *Studies in Second Language Acquisition*, 24(2), 143–188. <https://doi.org/10.1017/S0272263102002024>
- Ellis, N. C. (2003). Constructions, chunking, and connectionism: The emergence of second language structure. In C. J. Doughty & M. Long (Eds.), *The handbook of second language acquisition* (pp. 63–103). Blackwell. <https://doi.org/10.1002/9780470756492.ch4>
- Ellis, N. C. (2005). At the interface: Dynamic interactions of explicit and implicit language knowledge. *Studies in Second Language Acquisition*, 27, 305–352. <https://doi.org/10.1017/S027226310505014x>
- Ellis, N. C. (2006). Cognitive perspectives on SLA: The associative-cognitive CREED. *AILA Review*, 19, 100–121. <https://doi.org/10.1075/aila.19.08ell>
- Ellis, N. C., & Larsen-Freeman, D. (2006). Language emergence: Implications for applied linguistics—Introduction to the special issue. *Applied Linguistics*, 27, 558–589. <https://doi.org/10.1093/applin/aml028>

- Ellis, R. (1990). *Instructed second language acquisition: Learning in the classroom*. Blackwell.
- Ellis, R. (2002). Does form-focused instruction affect the acquisition of implicit knowledge?: A review of the research. *Studies in Second Language Acquisition*, 24, 223–236. <https://doi.org/10.1017/s0272263102002073>
- Ellis, R. (2005). Measuring implicit and explicit knowledge of a second language: A psychometric study. *Studies in Second Language Acquisition*, 27, 141–172. <https://doi.org/10.1017/s0272263105050096>
- Ellis, R. (2008). *The study of second language acquisition* (2nd ed.). Oxford University Press.
- Ellis, R. (2009). Measuring implicit and explicit knowledge of a second language. In R. Ellis, S. Loewen, C. Elder, R. Erlam, J. Philp, & H. Reinders (Eds.), *Implicit and explicit knowledge in second language learning, testing and teaching* (pp. 31–64). Multilingual Matters. <https://doi.org/10.21832/9781847691767-004>
- Ellis, R., & Loewen, S. (2007). Confirming the operational definitions of explicit and implicit knowledge in Ellis (2005): Responding to Isemonger. *Studies in Second Language Acquisition*, 29, 119–126. <https://doi.org/10.1017/s0272263107070052>
- Enders, C. K. (2010). *Applied missing data analysis*. Guilford press.
- Erlam, R. (2006). Elicited imitation as a measure of L2 implicit knowledge: An empirical validation study. *Applied linguistics*, 27(3), 464–491. <https://doi.org/10.1093/applin/aml001>
- Erlam, R. (2009). The elicited oral imitation test as a measure of implicit knowledge. In R. Ellis, S. Loewen, C. Elder, R. Erlam, J. Philp, & H. Reinders (Eds.), *Implicit and explicit knowledge in second language learning, testing and teaching* (pp. 65–93). Multilingual Matters. <https://doi.org/10.21832/9781847691767-005>
- Esser, S., & Haider, H. (2017). The emergence of explicit knowledge in a serial reaction time task: the role of experienced fluency and strength of representation. *Frontiers in Psychology*, 8. <https://doi.org/10.3389/fpsyg.2017.00502>
- Ferman, S., Olshtain, E., Schechtman, E., & Karni, A. (2009). The acquisition of a linguistic skill by adults: Procedural and declarative memory interact in the learning of an artificial morphological rule. *Journal of Neurolinguistics*, 22(4), 384–412. <https://doi.org/10.1016/j.jneuroling.2008.12.002>
- Fischer, S., Drosopoulos, S., Tsen, J., & Born, J. (2006). Implicit learning—explicit knowing: a role for sleep in memory system interaction. *Journal of Cognitive Neuroscience*, 18, 311–319. <https://doi.org/10.1162/jocn.2006.18.3.311>
- Freed, B. F., Dewey, D. P., Segalowitz, N., & Halter, R. (2004). The language contact profile.

- Studies in Second Language Acquisition*, 26(2), 349–356.  
<https://doi.org/10.1017/s027226310426209x>
- Freed, B. F., Segalowitz, N., & Dewey, D. P. (2004). Context of learning and second language fluency in French: Comparing regular classroom, study abroad, and intensive domestic immersion programs. *Studies in Second Language Acquisition*, 26, 275–301.  
<https://doi.org/10.1017/s0272263104262064>
- Frensch, P. A., Haider, H., R nger, D., Neugebauer, U., Voigt, S., & Werg, J. (2003). The route from implicit learning to verbal expression of what has been learned. *Attention and Implicit Learning*, 335–366. <https://doi.org/10.1075/aicr.48.17fre>
- Gass, S. M. (1997). *Input, interaction and output in second language acquisition*. Lawrence Erlbaum Associates.
- Godfroid, A., Loewen, S., Jung, S., Park, J. H., Gass, S., & Ellis, R. (2015). Timed and untimed grammaticality judgments measure distinct types of knowledge: Evidence from eye-movement patterns. *Studies in Second Language Acquisition*, 37(2), 269–297.  
<https://doi.org/10.1017/s0272263114000850>
- Godfroid, A. & Kim, K (Under review). Contributions of explicit-implicit learning aptitudes to explicit-implicit L2 grammar knowledge: An SEM approach.
- Godfroid, A., Kim, K., Hui, B., & Isbell, D. (In preparation). Synthesizing 12 years of validation research on implicit and explicit knowledge.
- Goo, J., Granena, G., Yilmaz, Y., & Novella, M. (2015). Implicit and explicit instruction in L2 learning: Norris & Ortega (2000) revisited and updated. In P. Rebuschat (Ed.), *Implicit and explicit learning of languages* (pp. 443–482). John Benjamins.  
<https://doi.org/10.1075/sibil.48.18goo>
- Goujon, A., Didierjean, A., & Poulet, S. (2014). The emergence of explicit knowledge from implicit learning. *Memory & Cognition*, 42, 225–236. <https://doi.org/10.3758/s13421-013-0355-0>
- Guti rrez, X. (2013). The construct validity of grammaticality judgment tests as measures of implicit and explicit knowledge. *Studies in Second Language Acquisition*, 35(3), 423–449. <https://doi.org/10.1017/s0272263113000041>
- Graham, J. W., Hofer, S. M., & MacKinnon, D. P. (1996). Maximizing the usefulness of data obtained with planned missing value patterns: An application of maximum likelihood procedures. *Multivariate Behavioral Research*, 31, 197–218.  
[https://doi.org/10.1207/s15327906mbr3102\\_3](https://doi.org/10.1207/s15327906mbr3102_3)
- Haider, H., & Frensch, P. A. (2005). The generation of conscious awareness in an incidental

- learning situation. *Psychological Research*, 69, 399–411. <https://doi.org/10.1007/s00426-004-0209-2>
- Hayduk, L., Cummings, G., Boadu, K., Pazderka-Robinson, H., & Boulianne, S. (2007). Testing! testing! one, two, three—Testing the theory in structural equation models!. *Personality and Individual Differences*, 42, 841–850. <https://doi.org/10.1016/j.paid.2006.10.001>
- Hopp, H. (2013). Grammatical gender in adult L2 acquisition: Relations between lexical and syntactic variability. *Second Language Research*, 29, 33–56. <https://doi.org/10.1177/0267658312461803>
- Hopp, H. (2016). Learning (not) to predict: Grammatical gender processing in second language acquisition. *Second Language Research*, 32, 277–307. <https://doi.org/10.1177/0267658315624960>
- Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, 6, 1–55. <https://doi.org/10.1080/10705519909540118>
- Hulstijn, J. H. (2002). Towards a unified account of the representation, processing and acquisition of second language knowledge. *Second Language Research*, 18, 193–223. <https://doi.org/10.1191/0267658302sr207oa>
- Hulstijn, J. H. (2005). Theoretical and empirical issues in the study of implicit and explicit second-language learning. *Studies in Second Language Acquisition*, 27, 129–140. <https://doi.org/10.1017/S0272263105050084>
- Hulstijn, J. H. (2007). Psycholinguistic perspectives on language and its acquisition. In J. Cummins & C. Davison (Eds.), *International handbook of English language teaching* (pp. 701–713). Springer. [https://doi.org/10.1007/978-0-387-46301-8\\_52](https://doi.org/10.1007/978-0-387-46301-8_52)
- Hulstijn, J. H. (2015). Explaining phenomena of first and second language acquisition with the constructs of implicit and explicit learning. In P. Rebuschat (Ed.), *Implicit and explicit learning of languages* (pp. 25–46). John Benjamins. <https://doi.org/10.1075/sibil.48.02hul>
- Institute of International Education. (2017). International Student Enrollment Trends, 1948/49–2016/17. *Open doors: Report on international educational exchange*. Washington, DC.
- Isabelli-García, C., Bown, J., Plews, J. L., & Dewey, D. P. (2018). Language learning and study abroad. *Language Teaching*, 51, 439–484. <https://doi.org/10.1017/s026144481800023x>
- Kenny, D. A. (2015). Measuring model fit. <http://davidakenny.net/cm/fit.htm>
- Kim, J. E., & Nam, H. (2017). Measures of implicit knowledge revisited: Processing modes, time pressure, and modality. *Studies in Second Language Acquisition*, 39(3), 431–457.

<https://doi.org/10.1017/s0272263115000510>

- Kline, R. B. (2016). *Principles and practice of structural equation modeling*. Guilford publications.
- Kramer, B., McLean, S., & Martin, E. S. (2018). Student grittiness: A pilot study investigating scholarly persistence in EFL classrooms. <http://irlib.wilmina.ac.jp/dspace/handle/10775/3498>
- Krashen, S. D. (1981). *Second language acquisition and second language learning*. Pergamon.
- Krashen, S. D. (1985). *The input hypothesis: Issues and implications*. Longman. <https://doi.org/10.2307/414800>
- Lake, J. (2013). Positive L2 self: Linking positive psychology with L2 motivation. In T. M. Apple, D. Da Silva, & T. Fellner (Eds.), *Language learning motivation in Japan* (pp. 71–225). Multilingual Matters. <https://doi.org/10.21832/9781783090518-015>
- Leow, R. P. (2015). *Explicit learning in the L2 classroom: A student-centered approach*. Routledge. <https://doi.org/10.4324/9781315887074>
- Loewen, S. (2020). *Introduction to instructed second language acquisition*. Routledge. <https://doi.org/10.4324/9781315616797>
- Long, M. (1996). The role of the linguistic environment in second language acquisition. In W. Ritchie and T. Bhatia (Eds.), *Handbook of second language acquisition*. (pp. 413–468). Academic Press. <https://doi.org/10.1016/b978-012589042-7/50015-3>
- Lightbown, P. M. (2008). Transfer appropriate processing as a model for classroom second language acquisition. In Z. Han (Ed.), *Understanding second language process* (pp. 27–44). Multilingual Matters. <https://doi.org/10.21832/9781847690159-005>
- Little, T. D., & Rhemtulla, M. (2013). Planned missing data designs for developmental researchers. *Child Development Perspectives*, 7, 199–204. <https://doi.org/10.1111/cdep.12043>
- Little, R. J. A., & Rubin, D. B. (1987). *Statistical analysis with missing data*. John Wiley & Sons.
- Norris, J. M., & Ortega, L. (2000). Effectiveness of L2 instruction: A research synthesis and quantitative meta-analysis. *Language learning*, 50(3), 417–528. <https://doi.org/10.1111/0023-8333.00136>
- Mathews, R. C., Buss, R. R., Stanley, W. B., Blanchard-Fields, F., Cho, J. R., & Druhan, B.



- (1989). Role of implicit and explicit processes in learning from examples: A synergistic effect. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 15, 1083-1100. <https://doi.org/10.1037/0278-7393.15.6.1083>
- Mehl, M. R., Pennebaker, J. W., Crow, D. M., Dabbs, J., & Price, J. H. (2001). The Electronically Activated Recorder (EAR): A device for sampling naturalistic daily activities and conversations. *Behavior Research Methods, Instruments, & Computers*, 33, 517–523. <https://doi.org/10.3758/bf03195410>
- McDonald, R. P. (1999). *Test theory: A unified treatment*. L. Erlbaum Associates.
- McLaughlin, B. (1987). *Theories of second-language learning*. Routledge.
- McManus, K., Mitchell, R., & Tracy-Ventura, N. (2014). Understanding insertion and integration in a study abroad context: The case of English-speaking sojourners in France. *Revue française de linguistique appliquée*, 19, 97–116. <https://doi.org/10.3917/rfla.192.0097>
- Office for International Scholars and Students (2017). <http://oiss.isp.msu.edu/about/statistical-report/>
- Papi, M., Rios, A., Pelt, H., & Ozdemir, E. (2019). Feedback-seeking behavior in language learning: Basic components and motivational antecedents. *The Modern Language Journal*, 103, 205–226. <https://doi.org/10.1111/modl.12538>
- Paradis, M. (1994). Neurolinguistic aspects of implicit and explicit memory: implications for bilingualism. In N. Ellis (Ed.), *Implicit and explicit learning of second languages*. (pp. 393–419). Academic Press.
- Paradis, M. (2009). *Declarative and procedural determinants of second languages*. John Benjamins.
- Pienemann, M. (1989) Is language teachable? Psycholinguistic experiments and hypotheses. *Applied Linguistics*, 10, 52–79. <https://doi.org/10.1093/applin/10.1.52>
- Plonsky, L., & Oswald, F. L. (2014). How big is “big”? Interpreting effect sizes in L2 research. *Language Learning*, 64(4), 878–912. <https://doi.org/10.1111/lang.12079>
- Ranta, L., & Meckelborg, A. (2013). How much exposure to English do international graduate students really get? Measuring language use in a naturalistic setting. *Canadian Modern Language Review*, 69, 1–33. <https://doi.org/10.3138/cmlr.987>
- Rebuschat, P. (2013). Measuring implicit and explicit knowledge in second language research. *Language Learning*, 63, 595–626. <https://doi.org/10.1111/lang.12010>

- Rünger, D., & Frensch, P. A. (2008). How incidental sequence learning creates reportable knowledge: The role of unexpected events. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 34, 1011–1026. <https://doi.org/10.1037/a0012942>
- Satorra, A., & Bentler, P. M. (2010). Ensuring positiveness of the scaled difference chi-square test statistic. *Psychometrika*, 75, 243–248. <https://doi.org/10.1007/s11336-009-9135-y>
- Schoonen, R., van Gelderen, A., Stoel, R. D., Hulstijn, J., & de Glopper, K. (2011). Modeling the development of L1 and EFL writing proficiency of secondary school students. *Language Learning*, 61, 31–79. <https://doi.org/10.1111/j.1467-9922.2010.00590.x>
- Segalowitz, N. (2003). Automaticity and second languages. In C. J. Doughty & M. H. Long (Eds.), *The handbook of second language acquisition* (pp. 382–408). Blackwell. <https://doi.org/10.1002/9780470756492.ch13>
- Segalowitz, N., & Freed, B. F. (2004). Context, contact, and cognition in oral fluency acquisition: Learning Spanish in at home and study abroad contexts. *Studies in Second Language Acquisition*, 26, 173–199. <https://doi.org/10.1017/s0272263104262027>
- Segalowitz, N. & Hulstijn, J. (2005). Automaticity in bilingualism and second language learning. In J. F. Kroll and A. M. B. De Groot (Eds.), *Handbook of bilingualism: psycholinguistic approaches* (pp.371–388). Oxford University Press.
- Selig, J. P., & Little, T. D. (2012). Autoregressive and cross-lagged panel analysis for longitudinal data. In B. Laursen, T. D. Little, & N. A. Card (Eds.), *Handbook of developmental research methods* (pp. 265–278). Guilford publications.
- Sharwood Smith, M (1980). Strategies, language transfer and the simulation of learners' mental operations. *Language Learning*, 29(2), 345–361. <https://doi.org/10.1111/j.1467-1770.1979.tb01074.x>
- Spada, N., Shiu, J. L. J., & Tomita, Y. (2015). Validating an elicited imitation task as a measure of implicit knowledge: Comparisons with other validation studies. *Language Learning*, 65(3), 723–751. <https://doi.org/10.1111/lang.12129>
- Suzuki, Y., & DeKeyser, R. (2015). Comparing elicited imitation and word monitoring as measures of implicit knowledge. *Language Learning*, 65(4), 860–895. <https://doi.org/10.1111/lang.12138>
- Suzuki, Y., & DeKeyser, R. (2017). The interface of explicit and implicit knowledge in a second language: Insights from individual differences in cognitive aptitudes. *Language Learning*, 67(4), 747–790. <https://doi.org/10.1111/lang.12241>
- Swain, M. (1985). Communicative competence: Some roles of comprehensible input and



- comprehensible output in its development. In S. Gass and C. Madden (Eds.), *Input in second language acquisition* (pp. 235–253). Newbury House.
- Vafaei, P., Suzuki, Y., & Kachinske, I. (2017). Validating grammaticality judgment tests: Evidence from two new psycholinguistic measures. *Studies in Second Language Acquisition*, 39(1), 59–95. <https://doi.org/10.1017/s0272263115000455>
- Wagner, U., Gais, S., Haider, H., Verleger, R., & Born, J. (2004). Sleep inspires insight. *Nature*, 427, 352–355. <https://doi.org/10.1038/nature02223>
- Wilhelm, I., Rose, M., Imhof, K. I., Rasch, B., Büchel, C., & Born, J. (2013). The sleeping child outplays the adult's capacity to convert implicit into explicit knowledge. *Nature neuroscience*, 16, 391–393. <https://doi.org/10.1038/nn.3343>
- Widaman, K. F. (1985). Hierarchically nested covariance structure models for multitrait-multimethod data. *Applied Psychological Measurement*, 9, 1–26. <https://doi.org/10.1177/014662168500900101>
- William, J. N. (2018). Implicit and explicit learning: interactions and synergies. Paper presentation at the meeting of LEAD summer school. University of Tübingen, Germany.
- Zhang, R. (2015). Measuring university-level L2 learners' implicit and explicit linguistic knowledge. *Studies in Second Language Acquisition*, 37(3), 457–486. <https://doi.org/10.1017/s0272263114000370>