2017

Adult L2 Processing and Acquisition Of The English Present Perfect

Christopher J. Farina

University of South Carolina

ADULT L2 PROCESSING AND ACQUISITION OF THE ENGLISH PRESENT PERFECT

by

Christopher J. Farina

Honors Bachelor of Arts
Xavier University, 2009

Master of Arts
University of South Carolina, 2013

Submitted in Partial Fulfillment of the Requirements

For the Degree of Doctor of Philosophy in

Linguistics

College of Arts and Sciences

University of South Carolina

2017

Accepted by:

Mila Tasseva-Kurktchieva, Major Professor

Stanley Dubinsky, Committee Member

Paul Malovrh, Committee Member

Robin Morris, Committee Member

Yasuhiro Shirai, Committee Member

Cheryl L. Addy, Vice Provost and Dean of the Graduate School

DEDICATION

This dissertation is dedicated to you as a reader, without whom this labor is for naught, and to my friends and family, especially to my wife Lily, whose support is unyielding, whose wisdom is irreproachable, and whose love is unconditional.

ACKNOWLEDGEMENTS

If I were required to list every individual who has helped me throughout my graduate studies or even throughout the process of completing this dissertation, I would have to compile them using a carefully coded Excel spreadsheet, and it would take days, if not weeks, to accomplish this feat to my own satisfaction. Even without this database, there are several individuals, groups, and organizations whose contributions stand out. Without their individual and combined contributions, this research would not have been possible. I am incredibly grateful to each of you.

Thank you to all of the speakers of English as a second language who have generously donated their time to discuss their experiences as English learners and speakers. You have inspired nearly all of my research projects (including this one), you have consulted on their development, and you have willingly participated in them whenever possible and encouraged others to do so as well.

Thank you to the faculty and staff of English Programs for Internationals and those of the International Accelerator Program at the University of South Carolina for introducing me to several cohorts of English learners and allowing me to utilize valuable class time for my research. Likewise, thank you to all of the professors and instructors of undergraduate linguistics courses who have also allowed me into their classrooms to recruit participants and to present my findings.

Thank you to Mila Tasseva-Kurktchieva for your guidance and steadfast support over the last several years. It was your class in the spring of 2013 that piqued my interest in pursuing my doctorate in second language acquisition, and you have been there guiding me since then, encouraging me to pursue my interests, developing my skills as a social science researcher, and exhorting me to present my ideas simply and clearly. Thank you so much for your time, patience, and emoji-laced comments as you examined several iterations of this thesis paper. I value your expertise, and I appreciate your encouragement and faith during my tenure as a doctoral student. Selecting you as the director of my dissertation committee was the easiest thing I have done in graduate school.

Thank you to Stanley Dubinsky, Paul Malovrh, Robin Morris, and Yasuhiro Shirai for serving on my committee. Courses taught by and conversations with Stan have been fundamental to my understanding of syntax, compositionality, and linguistic theory broadly. Thank you for your willingness to take the time to meet with me to discuss theory, politics, and how to navigate academia. Paul introduced me to applied linguistics and more specifically to language processing and processing-based pedagogy. Thank you for helping me to see beyond linguistic theory to the second language learners who will hopefully benefit from this research. Robin lent me her experience and expertise in investigating real-time language processing. Thank you for helping me to focus my research questions and to develop the methods used in this investigation. Yas, I have been interested in your work since my first exposure to the Aspect Hypothesis four years ago, and it has been foundational to much of my own research since that time. Thank you for your willingness to correspond with me as I grappled with applying some facets of your work within this investigation.

Thank you to my fellow graduate students who have shared all or part of this journey through academia with me. It is only through community, fellowship, and commiseration that we are able to mitigate the stressors that permeate our lives, wrestling with impostor syndrome, maintaining our composure when asked how many languages we speak, biting our lips while listening to self-appointed ‘experts’ lecture us on how nonstandard dialects are somehow less-than, and smiling in the face of being over-worked and under-paid as we balance our dual roles as researcher and educator. You are all intelligent people who deserve to be here even when you are despondently limping through your studies having lost perspective and all self-confidence. You can do this.

This doctoral thesis project was made possible through the generous financial support of the Bilinski Educational Foundation through their Russell J. and Dorothy S. Bilinski Dissertation Fellowship Program, which allowed me to focus on my research for the past year, and through the financial and institutional support of the Linguistics Program at the University of South Carolina.

ABSTRACT

This research investigates the second language (L2) processing and acquisition of the English present perfect via two features: boundedness and current relevance. Boundedness indicates whether an action reaches an endpoint (Smith 1997; Verkuyl 1972); it divides the functions of the present perfect into sets that denote completed situations or ongoing/iterative ones (Bybee et al. 1994; Housen 2002). Current relevance indicates the present importance of a past situation (Siemund 2004); it differentiates the present perfect from the simple past (Bardovi-Harlig 2002). Previous research has relied on offline methods (that evaluate metalinguistic knowledge); no research in SLA has investigated the acquisition of the present perfect using online methods (that measure real-time processing). This investigation addresses this gap using two novel tasks.

In this study, 155 adult L2 English learners of varying proficiency from three first language (L1) backgrounds (Arabic, Chinese & Other) participated; 72 L1 English speakers were controls. Online data were collected using a self-paced reading task wherein participants read sentences manipulated for grammatical tense & boundedness and for grammatical tense & current relevance. Reading time differences for each condition were analyzed by L2 proficiency and L1 group. Offline data were collected using a rating task wherein participants provide judgments concerning the meanings of phrases excerpted from similarly manipulated sentences. Rating differences were analyzed by L2 proficiency and L1 group.

The results show that boundedness and current relevance affect L2 English learners' processing and comprehension of the present perfect; first language and L2 proficiency influence these effects. In boundedness conditions, high-proficiency learners exhibit inhibited reading times in nonbounded contexts, and they more accurately rate boundedness contrasts in the present perfect. These results suggest that advanced learners can distinguish the functions of the present perfect. The Arabic group performs like higher-proficiency learners, which indicates positive L1 transfer. In current relevance conditions, only the highest proficiency group exhibits reading times affected by current relevance marking, and they understand current relevance contrasts marked adverbially and morphosyntactically; less proficient groups only comprehend current relevance contrast marked adverbially. These results suggest that less proficient learners can use adverbially marked current relevance to distinguish the present perfect and simple past; only at higher proficiency do they become sensitive to morphosyntactic current relevance marking. The Chinese group unexpectedly performs like lower proficiency learners, which indicates negative L1 transfer.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Chapter 1 Introduction

One domain of research in second language (L2) acquisition is the development of the temporal system in the *interlanguage grammar* (a learner's grammar that is neither fully that of the first nor of the second language).[1] The acquisition of the temporal system is an especially fruitful field of study for two reasons. First, a comparatively long time is needed to fully acquire the system – approximately 15 years, according to Davydova (2011). For L2 learners of English, this system begins developing very early in the acquisition process, and it likely fossilizes among L2 learners before reaching a targetlike state (§2.3). Second, the process of acquisition is multifaceted. In English, the temporal system is complex (§2.2). Mastering this system requires the acquisition of grammatical properties from multiple linguistic domains[2] and at the interfaces thereof. Thus, research into the acquisition of a language's temporal system provides a rich domain of inquiry: it offers multiple opportunities for investigation across the spectrum of L2 proficiency, and these opportunities incorporate multiple linguistic domains and their interfaces.

The present investigation focuses on a structure that emerges relatively late in the development of the temporal system among both L1 and L2 learners (§2.4), the *English present perfect* (e.g. *I have written*; §2.3.2). The temporal system of English is complex; it expresses meanings and functions that convey *tense* (temporal deixis; §2.2.1) and

[1] See Selinker (1972)'s original proposal and/or VanPatten and Williams (2015) brief explanation of interlanguage grammar for more information.
[2] Linguistic domains include phonetics, phonology, morphology, syntax, semantics, and pragmatics.

aspect (speaker viewpoint; §2.2.2).[3] In early acquisition, learners apply novel morphology to mark meanings that are already present in the grammar but have hitherto been marked via non-grammatical means (Bardovi-Harlig 2000b). Early-emerging structures are often used to investigate formal accuracy. In late acquisition, learners are developing the understanding of the meaning and usage of these forms. Studying a late-emerging structure like the present perfect allows for the investigation of the interlanguage as it adjusts to accommodate novel associations between linguistic forms and their meanings (Bardovi-Harlig 2002). That is, the interlanguage grammar must be restructured so that the known linguistic form can be associated with novel meaning. For example, regarding form, *have* is restructured from a verb that denotes possession to an auxiliary verb that together with the past participle (V-*en/-ed*) forms the present perfect. Further, regarding meaning, the time that the situation occurs is anterior to the present for both the present perfect and the simple past (e.g. *I wrote*; §2.3.1), the time being referenced is the present for both the present perfect and the simple present (e.g. *I write*); these overlapping features (situation time and reference time) must be restructured such that the present perfect becomes distinguished from the simple past and the simple present, which are already present in the interlanguage grammar.

Two meanings associated with the present perfect in the target grammar are central to the present investigation. The first of these is *completedness*. The present perfect can describe either a completed situation with an enduring result or an ongoing/iterative situation. For example, in a sentence with a completed action like *I have broken my arm,* the present perfect describes the present result (my injured arm) of

[3] The temporal system also expresses modality (the expression of possibility and necessity), but these meanings extend beyond the scope of the present investigation.

the completed past action (fracturing my arm). Further, in a sentence with an incomplete action like *I have lived in Detroit,* the present perfect leaves open the possibility that I still live in Detroit (*…for nine years now*) or I may do so again at some point in the future (*…and I'll be moving back soon*). Typologically, these two sets of meanings or *functions* emerge from the semantics of the verb or verb phrase (§2.1). The present investigation operationalizes completedness at the level of the verb phrase via *boundedness* (see §2.1.2). Completedness is an interesting meaning to investigate because it is initially associated with the simple past, but it comes to be associated with the functions of the present perfect that describe completed situations.

The second of these meanings associated with the present perfect is *current relevance* (§2.3.2). Whether completed, iterative, or ongoing, the past situation described by the present perfect is viewed as important to the present moment. For example, when asked whether you would like to go to dinner, you may exploit this feature of the present perfect and respond with *I have eaten*; in this context, the current relevance of the result (my presently sated hunger) allows the sentence to function as a polite refusal. Current relevance is considered a central feature of the present perfect; it distinguishes the present perfect from the simple past and aligns the present perfect with meanings originally associated with the simple present. The present investigation manipulates current relevance overtly using adverbial modifiers (see Table 2.1 & Table 4.3). Completedness and current relevance are certainly not the only important meanings associated with the present perfect; however, investigating how these meanings are represented in the grammar and how they are understood in real time provides insight into the acquisition of the present perfect.

The present investigation examines how instructed adult L2 learners of English acquire the present perfect. These L2 users are divided by English proficiency (§3.2) and by first language (§3.3) via an independent measure of proficiency (§4.3) and a background questionnaire (§4.4), respectively. The use of cross-sectional English proficiency groups allows for the approximation of the process of acquiring the present perfect within a sample of learners. Grouping participants by first language permits the investigation of the effects of L1 transfer, which may otherwise confound the results.

This investigation relies on two reading tasks to explore form-meaning acquisition. The first task measures a participant's L2 *processing*, the real-time procedures that convert linguistic input into a form that is interpretable in the language system (§4.5.1).[4] Broadly, this task captures how L2 users comprehend English in real time in order to study their implicit knowledge of completedness and current relevance as these features emerge in context. This task requires the assumption that how participants process these structures in context is indicative of how the grammar is represented (cf. Slabakova 2008). If this assumption is valid, the first task enables the comparison of two accounts that describe the manner of language acquisition. The complexity account proposes that acquisition occurs in order of increasing semantic-syntactic complexity; that is, less complex structures emerge before more complex ones (§3.1.1). The prototype account suggests that acquisition occurs through bootstrapping based on shared semantic features; that is, learners restrict the usage of structures to those forms with exemplary sets of features before extending the use of these structures to forms with less exemplary

[4] This investigation follows the broad definition of processing used by VanPatten and Jegerski: "processing is the umbrella term for moment-by-moment operations during comprehension… from syntax to morphology to lexicon to interfaces" (2010:5).

features (§3.1.2). The second task measures a participant's associations between forms and their meanings, which probes their explicit knowledge of these two features (§4.5.2). This task relies on metalinguistic judgments to further explore how the grammar is represented.

The results of these two tasks shed light on how learners acquire the present perfect as they distinguish it from the simple past. The first task reveals what effects manipulations in completedness and overt marking of current relevance have on how L2 users differentially process the present perfect and the simple past. Special attention is paid to differential processing between linguistic structures that vary according to their relative complexity and their relative prototypicality. The second task elucidates the associations between form and meaning in the interlanguage grammars of the L2 users for the structures that indicate completedness and those that indicate current relevance. Overall, the results of both tasks reveal how L2 users incorporate completedness and current relevance into their interlanguage grammars as they acquire the meanings associated with the present perfect.

Chapter 2 Foundational tense-aspect theory and acquisition research

The purpose of the present investigation is to shed light on the acquisition and processing of the present perfect, which are studied via tasks that measure how learners read and understand specific linguistic features in context. This chapter provides a review of the literature that serves as the foundation for the present investigation. Section 2.1 concerns how endpoints emerge at the verbal and phrasal levels. The first subsection introduces verb-level aspect with a focus on telicity. The second subsection introduces phrasal aspect with a focus on *boundedness*, one of the central linguistic features in the present investigation. Section 2.2 concerns tense and grammatical aspect, which emerge at the level of the clause. Section 2.3 describes the two English verbal categories relevant to the present investigation, the present perfect and the simple past. The present perfect is the construction of interest and the simple past is the construction with which it is compared. Focus is placed on the tense reference and aspectual character of each construction along with their adverbial modification patterns and functions. The subsection on the present perfect characterizes the second central linguistic feature in this investigation: current relevance. Following a comparative summary of the two grammatical tenses, section 2.4 summarizes prior research concerning the first and second language acquisition of the present perfect. The discussion therein situates the present investigation in the literature on the acquisition of tense-aspect broadly and of the present perfect specifically.

2.1 LEXICAL AND COMPOSITIONAL ENDPOINTS.

Endpoints emerge in language via *aspect*, which can describe the completedness of an occurrence. Completedness is a central feature of the present investigation because of its importance in the English verbal system. Two aspectual features that denote completedness are the focus of the following discussion. *Telicity*, a feature of *lexical aspect*,[5] is generated by the semantics of the verb alone at the syntactic level of the verb (V node; see §2.1.1). *Lexical aspect* is a semantic property inherent in the meaning of the verb that describes a situation and has scope over the verb. Lexical aspect has three common binary feature distinctions: telicity (has/lacks an inherent endpoint), dynamicity (the predicate is/is not changed by the action of the verb), and durativity (the action occurs over time/nearly instantaneously; Vendler 1967). Lexical aspect does not change due to the point of view of the speaker, and it is generally not marked morphologically.[6] *Boundedness*, a feature of *compositional aspect*,[7] is generated by the interaction of the verb and its arguments at the syntactic level of the verb phrase (VP node; see §2.1.2). It can be generated in two ways depending on the semantics of the verb: quantifying the direct object of an activity predicate and specifying the goal of a verb of motion. The

[5] Lexical aspect is the preferred term for the present investigation to describe the aspectual meanings conveyed by the lexical semantics of the verb. Previous analyses have used varying terminology, including *Aktionsart* 'kind of action' (Streitberg 1891), *objective aspect* (Deutschbein 1939 – contra subjective aspect), *inherent or semantic aspect* (Comrie 1976), and *Aspect*$_2$ (Sasse 2002 – contra Aspect$_1$). Although these terms are not entirely synonymous, they are similar enough in character that their differences should not affect the present investigation.
[6] The most notable exception to this occurs in some of the Slavic languages wherein morphological prefixes, called *preverbs*, are used to mark telicity (Mikhaylova 2012).
[7] Compositional aspect is the preferred term for the present investigation to describe the aspectual meanings conveyed by the lexical semantics of the verb and the verbal constellation at the phrasal level. The other term commonly used to describe this kind of aspect is *situation aspect* (Smith 1991, 1997).

below subsections provide a necessary introduction to aspect at the verbal and phrasal levels.

2.1.1 TELICITY.

Telicity concerns whether or not the predicate has an inherent endpoint (Bardovi-Harlig 2000b; Slabakova 2000), and this feature distinguishes two kinds of *occurrences*: (telic) *events* and (atelic) *activities*.[8] A telic predicate describes an occurrence that progresses for some time before reaching its inherent endpoint, at which point, it ends. An atelic predicate describes an occurrence that lacks an inherent endpoint and thus continues indefinitely. This feature distinction is often considered to be the difference between ‘completion’ and ‘finishing’; that is, an occurrence can only be completed if it has an endpoint and that endpoint has been reached, whereas an occurrence can be finished if it lacks an endpoint but the situation denoted by the verb has ceased for any reason (including the reaching of some arbitrary or predetermined endpoint). For example, one can be said to complete or finish *ripening* [+tel] because a predetermined endpoint is reached once the action is finished; however, one can only be said to finish

[8] The present investigation occasionally uses the nomenclature of Vendler (1967) and Mourelatos (1978) to describe different predicates based on their aspectual features as described below.

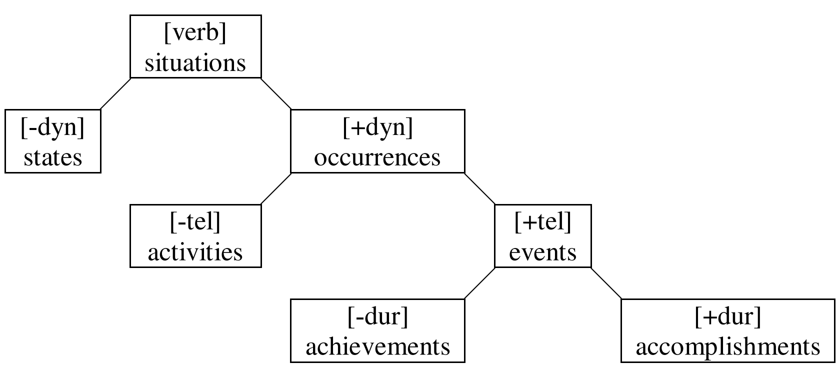

FIGURE 2.1. Verbal classes (Mourelatos 1978; Vendler 1967).

maturing [-tel] as there is no predetermined endpoint to reach, only an arbitrary one assigned in context (such as reaching a developmental milestone). Once a fruit is ripe, the action cannot continue further and would better be called *rotting*, but the action of *mature* may continue indefinitely.

Linguistic research has demonstrated that telicity is unique, psychologically real, and exerts measurable effects. Typological research finds that telicity is the feature most regularly marked using grammatical means cross-linguistically (Slabakova 2008). This research suggests that telicity occupies a privileged position among aspectual features. The trend indicates that telicity is more similar to grammatical markers of temporality than to purely lexical ones. Psycholinguistic research by Zacks et al. (2001) finds that humans identify and hierarchically organize occurrences and events based the presence or absence of endpoints and sub-endpoints within the action. This finding indicates that endpoints are psychologically real and that these endpoints are used to understand linguistic and real-world events. In addition, feature variation has been shown to affect the manner and speed of processing in reading time studies (Coll-Florit & Gennari 2011; Gennari & Poeppel 2003; McKoon & McFarland 2002), behavioral research (Zacks et al. 2001), and magnetoencephalography (MEG; Brennan & Pylkkänen 2010).

2.1.2 Boundedness.

Boundedness is a kind of aspect that often confused with telicity; however, the two differ in syntactic predication and in meaning. Unlike telicity, boundedness emerges through the syntactic composition of a verb and the *verbal constellation* (Depraetere

1995; Tenny 1987; Verkuyl 1972).[9] Further, while telicity concerns an action's inherent endpoint, boundedness concerns whether or not the action actually reaches some inherent or contextually determined endpoint (Depraetere 1995; Smith 1997). That is, telicity is based on an inherent feature of a situation, but it need not be realized when present. Example 1 shows a sentence in which the inherent endpoint of the telic event *die* is not realized as indicated by the acceptability of the *when* clause.

(1) The sick man was dying in the hospital (when the doctor cured him).

Although *die* describes a telic event, the use of the past progressive (*was dying*) predicates that the man does not reach the inherent endpoint (death) in this context. In contrast to telicity, which concerns *inherent* endpoints, boundedness concerns the *actual* spatio-temporal boundaries of a situation as they emerge in context. Example 2 provides two similar sentences with the atelic predicate *treat* wherein 2a is bounded and 2b is nonbounded; note the relative acceptability of the coordinated clause in each sentence.

(2) a. The doctor treated the two infections (and her treatment worked).

b. The doctor treated infections [?](and her treatment worked).

In 2a, the action is bounded because the doctor actually completes the treatments for the two infections; that is, when the two infections are treated, the action of *treat* is complete (Langacker 2008; Niemeier 2013). Only once the course of treatment is complete, can it be faithfully said whether or not the treatment worked. This is not the case for example 2b, which is nonbounded. Without a contextually specified endpoint (a limit on the number of infections to be treated), the action cannot actually complete (Niemeier 2013).

[9] In English, the verbal constellation is composed of the internal arguments (direct and indirect object), external arguments (subject), and adjuncts (adverbial modifiers), but this need not be the case. The constellation can also include verbal morphology that encodes aspectual information.

The fact that the treatments are ongoing leads to the questionability of the coordinated clause. The present investigation assumes a framework in which verbs that can be either nonbounded or bounded are specified for a feature in the lexicon in such a way that boundedness is possible but not required (perhaps, [α bounded]; cf. Pickering et al. 2006; Mikhaylova 2012).[10] In this framework, boundedness, like telicity, is a lexical feature of the verb itself; boundedness differs from telicity in that the actualization of [α bounded] emerges in context.[11]

The present investigation relies on two methods of predicating boundedness in its experimental manipulations: i) quantifying the direct object of an activity predicate and ii) specifying the goal of a verb of motion.

QUANTIFIED OBJECT. The first method concerns the predication of boundedness through the quantification of the direct object.[12] These predicates are first discussed in

---

[10] Horrocks and Stavrou (2003) and others suppose that there is a genuine feature distinction between activities [- telic] and accomplishments [+ telic] in the lexicon, and they explain that activities can act like accomplishments as coercion caused by contextual elements. The distinction here is not trivial. In accordance with Horrocks and Stavrou, other analyses have found that coerced predicates take longer to process (Piñango et al. 1999; Todorova et al. 2000), and the assumption made in the present investigation may be a confounding factor in the analysis. However, the assumption of the present investigation is not unfounded; another investigation by Pickering and colleagues (2006) found that coercion has no effect with differences between conditions correlating with relative feature complexity only. That is, there are no differences between activities and accomplishments, but there are observable differences between states, activity-accomplishments, and achievements.

[11] Although boundedness is compositionally generated aspect, it is not truly grammatical aspect (see: §2.2.2) in that it is not governed by verbal morphology in English and does not have scope over the entire sentence. It is only when the effects of tense-aspect morphology on compositional aspect are considered that the entire IP is considered to have grammatical aspect (Verkuyl 1989).

[12] The present investigation only requires a cursory overview of this rather complex means of predicating boundedness. For additional information on this topic, see Verkuyl (1972, 1989, 1993), Krifka (1992, 1998), Dowty (1979. 1991), and Tenny (1987, 1994) for the foundations of this topic and Borer (1994), Mourelatos (1978), Pustejovsky

depth by Verkuyl (1972).[13] He notes that when an activity verb has a quantized or definite noun as its direct object, the verb phrase is bounded; a stative verb in a similar context remains nonbounded. Sentences 3a and 3b demonstrate how quantification affects the emergence of boundedness in context. The nonquantified determiner phrase in 3a *sonatas* yields a nonbounded predicate, which is indicated by the licit duration adverbial and illicit frame adverbial. In contrast, the quantified determiner phrase in 3b *those sonatas* interacts with the activity verb *play* to form a bounded predicate, which is indicated by the stretched reading of the duration adverbial and the licitness of the frame adverbial.[14] Sentence 3c indicates that, whether or not *sonatas* is quantified, only a nonbounded reading is possible because the state *hate* lacks the [α bounded] feature.

(3) a. James played [$_{DP}$ Ø [$_{NP}$ sonatas]] for/*in one hour. [nonbounded]

b. James played [$_{DP}$ those [$_{NP}$ sonatas]] #for/in one hour. [bounded]

c. James hated [$_{DP}$ (those) [$_{NP}$ sonatas]] for/*in one hour. [nonbounded]

The quantification of the object is the main means by which boundedness is manipulated in the present investigation. Nonbounded predicates are those dynamic predicates that do

(1991), Schmitt (1996), Taylor (1977), Travis (1994), Van Valin (1990), and Van Voorst (1988) for supplemental discussions of this topic.

[13] Verkuyl (1972) uses the term *terminative aspect* to describe boundedness and *durative aspect* to describe nonboundedness. The present investigation does not use the original terminology for the sake of clarity.

[14] These examples make use of Dowty's (1979) adverbial modification test to determine telicity and/or boundedness. Her test is based on the licitness of modification using either *frame* or *duration* adverbials. Frame adverbials denote the span of time in which an action occurs (e.g. *in x time*); the duration of the action cannot exceed this span of time, but it may occur in less time. Duration adverbials denote the span of time over which the action occurs, and it may not occur in less time (e.g. *for x time*). Atelic/nonbounded predicates are licit when modified by duration adverbials but not by frame adverbials. Telic/bounded predicates follow the opposite pattern: they are licit when modified by frame adverbials but not by duration adverbials.

not have quantified direct objects (3a), and bounded predicates are those predicates that have quantified direct objects (3b).

MOTION-GOAL. The second method concerns the predication of boundedness through specification of a goal of motion. These predicates are first discussed in depth by Talmy (1975, 1985) who focuses on two structure types: manner-framed and path-framed.[15] Manner-framed structures (e.g. *walk, run, swim,* & *dance*) are those in which the verb describes the manner of the motion but not the direction of the motion, which may be specified via a locative modifier (Talmy 1991). Path-framed structures (e.g. *enter, exit, ascend,* & *descend*) are those in which the verb describes the direction of the motion but the not manner of the motion, which may be specified using an adverbial modifier (Talmy 1991). The present investigation only uses manner-framed structures, which are bounded when a locative goal is specified via a prepositional phrase or verbal particle (Levin 1993).[16] Example 4 shows how the addition of a goal affects the boundedness of the manner verb *swim* in context. The manner-framed structure *swim in the ocean* in 4a is a nonbounded predicate, which is indicated by the licit duration adverbial and illicit frame adverbial. The same manner verb *swim* specified by the goal *to the opposite shore*

[15] Strictly speaking, Talmy's (1975, 1985) typological studies concern languages, but it is more appropriate to consider the two framing conditions as structure types than representative of any language (cf. Horrocks & Stavrou 2007).

[16] Talmy (1975, 1985) refers to this coercion as a 'lexicalization problem,' and it has since been adapted into the research on unaccusatives following Levin (1993). As is the case for the other means of predicating boundedness, a full characterization of this perspective is beyond the scope of the present investigation. For an introduction into research that considers this method of predicating boundedness, see Beck & Snyder (2001), Chierchia (2004), Dowty (1991), Folli & Ramchand (2005), Hale & Keyser (1993, 1998, 2002), Hoekstra (1988), Hoekstra & Mülder (1990), Horrocks & Stavrou (2007), Levin (1993), Levin & Rappaport (1995), Mateu & Rigau (1999, 2002), and Snyder (1995, 2001).

in 4b forms a bounded predicate, which is indicated by the stretched reading of the duration adverbial and the licit frame adverbial.

(4) a. James swam in the ocean for/*in one hour. [nonbounded]

b. James swam to the opposite shore #for/in one hour. [bounded]

This method of predicating boundedness by providing a goal for a motion predicate is secondary within the present investigation. All motion-goal sentences used in the reading task are bounded.

The present investigation uses boundedness to characterize whether the verb phrase describes a completed event. It does so following two manners of adding an endpoint to an activity verb. This first predicates boundedness by quantifying the direct object the verb, and the second does so by specifying a goal of motion. The two are not directly compared, but the investigation assumes that they require comparable processing effort.

2.2 English tense-aspect system.

In addition to the verb-phrase level distinctions discussed above, sentence-level temporality is expressed via verbal morphology and auxiliaries that encode tense and/or aspect. In English, verbal morphosyntax encodes a combination of tense and grammatical aspect, which are conflated in the paradigm to that point that it is appropriate to use the term ‘tense-aspect’ to refer to the temporal distinctions that grammatical tenses encode. English verbal morphosyntax is sentential in scope, which means that the forms govern the entire clause in which they are embedded and not just the verb phrase (Comrie 1976; Slabakova 2000; Smith 1997). The scope of tense-aspect indicates that the minimal context in which to observe how it is processed is the sentence. Despite the aforementioned conflations in morphosyntax and scope, tense and grammatical aspect encode separate temporal distinctions. Each concept is described in brief below.

2.2.1 TENSE.

Tense concerns the temporality of a situation. It describes when a situation occurs. The function of tense is to position the situation described by the verb in relation to some other time; this other time may be the speech time or some other reference time (Ayoun & Salaberry 2008). Because tense describes the positions of two times in relation to one another, tense is deictic on the time axis (Andersen & Shirai 1996). That is, tense 'points' to the time of the situation with the reference of some other point. Tense is utilized by the speaker to tell the hearer if a situation occurs prior to (past tense), contemporaneously to (present tense), or after another one (future tense). English encodes three tenses: past, present, and future.[17]

2.2.2 GRAMMATICAL ASPECT.

Grammatical aspect concerns 'the internal temporal constituency of a situation' (Comrie 1976:3). Grammatical aspect describes how a situation is composed. This kind of aspect is also known as *viewpoint aspect* because it encodes different ways from which a situation can be seen from the speaker's point of view (Bardovi-Harlig 2000b; Comrie 1976; Klein 1994; Smith 1997). The main distinction is between *perfective* and *imperfective*.[18] The present investigation follows Slabakova's (1997, 2003) conception of

---

[17] The arguments of Declerck (2006) and others that English lacks a future tense (i.e. past & nonpast tenses) are noted, but the present investigation is more concerned with the capability of English to convey 'a mental division of time' after the present than the lack of a unique, analytic verbal morpheme.

[18] There are several different ways to conceptualize this distinction. In the past, linguists have understood the perfective to describe 'momentariness' (J.E. Miller 1970),

the perfective-imperfective distinction that proposes that the function of grammatical aspect is to express whether the clause denotes a 'bounded' (complete and inaccessible) or 'unbounded' (incomplete and interruptible/modifiable) situation. The expression of completedness emerges from the speaker's viewpoint of the situation from the outside as a complete whole, whereas incompleteness emerges from the viewpoint within the situation as it unfolds without definite, internal temporal boundaries (Ayoun & Salaberry 2008; Comrie 1976; Slabakova 2000; Slabakova & Montrul 1999). According to this position, complete occurrences are impenetrable in that they cannot be decomposed into their component sub-occurrences, whereas further decomposition is possible without such internal boundaries. Example 5 illustrates this point. In 5a, *has shaved* is a complete (bounded) occurrence; it cannot be interrupted/divided by the action of the *when* clause. On the contrary, in 5b, *was shaving* is incomplete/ongoing (unbounded) and is, therefore, interruptible/divisible by the action of the *when* clause.

(5) a. *Harry has shaved his beard when Joan wanted to blowdry her hair.

b. Harry was shaving his beard when Joan wanted to blowdry her hair.

The semantic contrast between complete and incomplete is integral to the verbal systems of many languages including English (Slabakova 2001). This contrast is integral despite the fact that English lacks discrete perfective and imperfective verbal morphology. Traditional Standard English grammars indicate that there are four morphosyntactically marked aspectual distinctions (simple, progressive, perfect,[19] and perfect progressive), but the present investigation is only concerned with two of them:

'completedness' (Stilman et al. 1972), 'boundedness' (Stutterheim 1991), 'an unanalyzable whole' (Bybee et al. 1994), or 'totality' (Offord & Gogolitsyna 2005).

[19] See section 2.3.2 below for a more in depth discussion of the present perfect.

simple and perfect. Simple aspect describes whether a situation occurs. Perfect aspect describes an anterior situation that is in some way significant at the reference time. Sentences 6 and 7 give examples of the simple and perfect aspects in the present tense; the verbal morphology and auxiliaries are underlined.

(6)		Jonette studies very hard for her tests.	[simple present]
(7)		Jonette has studied very hard for her tests.	[present perfect]

An interesting feature of these aspects emerges when the situation occurs in the past. Both of these aspects interact with VP-level aspect to generate novel readings: bounded dynamic predicates like *ace the test* generate perfective readings (8a & 9a), and nonbounded and stative predicates like *know the answers* generate imperfective readings (8b & 9b).

(8)	a.	Jonette aced the test.	[bounded & simple past]
	b.	Jonette knew the answers.	[nonbounded & simple past]
(9)	a.	Jonette has aced the test.	[bounded & pres. perfect]
	b.	Jonette has known the answers.	[nonbounded & pres. perfect]

The effects of boundedness on the present perfect are crucial to the present investigation as discussed below in section 2.3.2.

2.3 TENSE-ASPECT CONSTRUCTIONS OF INTEREST.

This investigation focuses on the English present perfect and compares it to the English simple past. Both the present perfect and the simple past refer to some anterior situation, and they are functionally equivalent (Binnick 1991; Comrie 1976; Michaelis 1998); however, they are not identical, and a learner must come to understand the semantic and pragmatic differences between them in order to acquire these categories. To do this, learners must grasp each category's tense reference and aspectual character (McCoard 1978), and they must master the adverbial modifiers that collocate with them (Andersen 2002). The following sections characterize the simple past and the present perfect. As the object of study, the present perfect's description is considerably more in depth than the simple past. The simple past is characterized in brief with emphasis on the similarities and differences between the two constructions.

2.3.1 SIMPLE PAST.

The simple past is formed when the past tense inflectional suffix is attached to the verbal stem (V-*ed*). Following the nomenclature used by Comrie (1985), the English simple past is an *absolute tense*, which means that the event time and the reference time are coincidental. It places the situation denoted by the verb anterior to the present moment on the time axis (Binnick 1991; Petersen 2004). The below figure diagrams the structure of this tense; following Reichenbach (1947), 'E' indicates the time of the action of the verb, 'R' indicates the point of reference from which tense is evaluated, and 'S' indicates the moment of speech.

<-----------E,R--------------S------------------------------>
Past Now Future

FIGURE 2.2. Diagram of the temporal structure of the simple past.

The simple past presents a situation as having occurred at a moment prior to the speech time. Since it separates the point to which it refers from the speech time, it is deictic.

When a situation is predicated in the simple past, the utterance implicates that the situation is complete at the reference point.[20] That is, since the situation is anterior to the

[20] This has led linguists to conclude that the simple past predicates perfective or perfective-like aspect (cf. Anderson et al. 2008; Andersen 2002; Ferretti et al. 2007; Madden & Zwaan 2003).Whether or not the simple past predicates perfective aspect is

present, listeners infer that the situation is not ongoing. As noted by Lyons, 'anteriority is not always distinguishable from completion and termination' (1977:689). This implication is especially strong when the situation is a bounded event. Independent investigations find that English speakers understand bounded events as complete at reference time even though there is no grammatical requirement (Anderson et al. 2008; Ferretti et al. 2007; Madden & Zwaan 2003). The following sentence is an example of a bounded event in the simple past whose completion is contradicted in the coordinated *but* clause.

(10) Lily wrote the essay in an hour, but she didn't finish it by the time that the exam was over.

This example certainly describes an odd scenario; however, the sentence is grammatical, and the described scenario is possible (i.e. Lily wrote her essay in the one hour provided by the exam, but she was not finished when the exam period ended). The licitness of this sentence indicates that it is more appropriate to refer to a bounded event in the simple past as finished than as complete. That is, if this event in the simple past were truly complete, the coordinated clause would make the sentence illicit. Furthermore, although the event is finished at present, the simple past does not implicate that the result of the situation holds at the reference point (Declerck 2006). This is a distinguishing feature between the simple past and the bounded functions of the present perfect (discussed below).

ADVERBIAL MODIFICATION. The simple past can be modified through the addition of adverbial phrases. Since the simple past is used to refer to past time contexts, its usage

---

beyond the scope of the present investigation, but it assumes that the simple past may be used perfectively and imperfectively.

collocates with adverbials that situate the action firmly before the present. For example, adverbs that situate an action in the past include *yesterday, in the past, after the election* and many more (Declerck 2006). Likewise, since the simple past always denotes a definite time span, this span can either be specified using an adverbial phrase or left unspecified. The simple past often collocates with adverbial phrases that overtly indicate definiteness in some past time context (e.g. *on Sunday, in 1900, when I close my eyes*) or it appears in narrative contexts in which the order of presentation covertly suggests definiteness (Davydova 2011).

FUNCTIONS. For the purposes of the present investigation, the simple past is considered to have two functions: past time reference and narrative ordering (cf. Davydova 2011). The past time reference function places a past situation at a definite moment in the past (Elsness 1997).[21] The narrative ordering function places a string of situations on the time axis in the order in which they occurred.

<--E_1,R_1---E_2,R_2----------S------------------------------>
Past Now Future

FIGURE 2.3. Diagram of the temporal structure of the simple past with two events.

This function relies on the definite reference of the simple past to establish order of occurrence. Definite reference is central to both functions of the simple past.

2.3.2 PRESENT PERFECT.

The present perfect is formed using a present tense form of the verbal auxiliary *have* and a past participle (V-*en*/-*ed*). Definitions of the present perfect vary, but they

[21] Anterior reference is not restricted to the simple past. Other past time grammatical tenses (including the past progressive, present perfect, present perfect progressive, past perfect, and past perfect progressive) can predicate the past time reference of this function; however, only the simple past and the past progressive are used to refer to a definite or specific moment in the past.

share the common thread that the category somehow relates a past situation to the present moment (cf. Bardovi-Harlig 2002). Following Reichenbach (1947), this relation between the past and present is symbolized by marking the reference point as coincident with the speech time.

```
<-----------E---------------S,R----------------------------->
  Past                      Now                        Future
```

FIGURE 2.4. Diagram of the temporal structure of the present perfect.

As a present tense structure, the present perfect refers to the present moment; it only indefinitely describes the anterior situation (Binnick 1991). According to Comrie, ‘it tells us nothing directly about the situation in itself, but rather relates some state to a preceding situation’ (1976:52). This relation between the time that the situation occurred and the present time is negotiated through tense and aspect.

The present investigation contends that the present perfect is a combination of the *pre-present tense* and *perfect aspect* (cf. Comrie 1985). The pre-present tense is a subcategory of the present tense that conveys a portion of the time span that is still within the present time but which is located immediately before the present moment (Declerck 2006). This tense denotation separates the time span referred to by the present perfect from the one referred to by the simple past and it explains how certain anterior time adverbs are acceptable while others are not (see Table 2.1 below). Like other aspectual distinctions, perfect aspect indicates the orientation (or viewpoint) of the speaker toward some anterior event (Bardovi-Harlig 2002). Specifically, it indicates that the prior event has *current relevance* for the situation at the reference time (Van Herk 2010). Current relevance is a feature that links a past situation to the reference time, focusing on the state of affairs that persist at this reference time as a consequence of the past situation (Joos

1964; Siemund 2004). For example, in the sentence *The volcano has erupted*, current relevance focuses the interlocutor on the fact that the consequences of the *eruption* persist at present; the past event of *erupting* that caused the present situation is backgrounded. The present investigation understands the present perfect as a *retrospective present* that denotes a relevant present state that exists as a consequence of some past situation (cf. Binnick 1991).[22]

Like the simple past, the present perfect exhibits varying aspectual character when the verb phrase is bounded or nonbounded. Bounded predicates in the present perfect exhibit perfective readings (11a), while nonbounded ones most often exhibit imperfective readings (11b; Smith 1991; Uno 2014).

(11) a. Luisa has found the answer. [perfect & perfective]

b. Luisa has searched for the answer. [perfect & imperfective]

Boundedness is the primary feature that divides the present perfect into perfective and imperfective functions (described below). This division between bounded and nonbounded functions is central to the present investigation.

ADVERBIAL MODIFICATION. The present perfect can be modified by the addition of adverbial phrases although it is most often unmodified (Schlüter 2006).[23] Since the present perfect is used to refer to (pre-)present contexts, its usage collocates with

---

[22] The present investigation is careful to select the word 'consequence' as a conscious rejection of those analyses that would use the word 'result' in its place (cf. Comrie 1976; Fenn 1987; Leech 1969; McCawley 1973). These analyses privilege resultativeness as the central feature of the present perfect. See Declerck (1991) for a compelling refutation of this perspective.

[23] The fact that the present perfect is only overtly modified in 31-45% of instances has led to the argument that null modification is the most typical use of the present perfect, but this argument is flawed. The same argument can be made concerning adverbial co-occurrence with the simple past (Miller 2000).

adverbials that situate the action at or slightly anterior to the present. Adverbs that specify an anterior time in the present (like *already*, *just*, and *today*) are acceptable in standard dialects of English (12a); those that specify an anterior time in the past (like l*ong ago*, and *yesterday*) are unacceptable in most standard and nonstandard dialects of English (12b).

(12) a. Sam has eaten lunch already/today. [pre-present]

b. Sam has eaten lunch ??long ago/*yesterday. [past]

Since the present perfect always denotes an indefinite time span, the present perfect is largely to entirely incompatible with definite adverbial modifiers (Bardovi-Harlig 2002, Inoue 1979; McCormick 2008; Rothstein 2007; Suh 1992). Common indefinite adverbial modifiers include the following: *already, always, ever, just, lately, never, recently, since,* and *(not) yet*.

The present investigation relies on a small set of adverbs in the reading tasks. These adverbs are selected from the work of Davydova (2011). Following McCoard (1978), Davydova schematizes adverbs for their current relevance ([CR]). Research finds that [+CR] adverbials regularly collocate with the present perfect and [–CR] adverbials collocate with the simple past (Davydova 2011; Elsness 1997; McCoard 1978).

TABLE 2.1. Davydova's (2011) classification of adverbials from McCoard (1978).

+ CR	±CR	+CR
long ago	never	at present
five years ago	ever	up till now
once	always	so far
yesterday	just (now)	as yet
the other day	today	not yet
those days	in my life	already
last night	recently	during these five years past
in 1900	lately	herewith
at 3:00	often	since the war
after the war	before (now)	for (three years)
no longer		long since
in the past		

As alluded to above, the present perfect has four functions: resultative, continuative, experiential, and recent past (cf. Comrie 1976). In their unmarked usage, these functions can be divided by their boundedness, which is crucial to the manipulations in the present investigation. The bounded functions are the resultative and recent past functions, and the nonbounded functions are the continuative and experiential functions (see below).

BOUNDED FUNCTIONS. The *resultative* function[24] describes the existence of some state of affairs that emerged as a result of some past occurrence and persists into the present moment (Comrie 1976, 1985).

<-----------E[----result---S,R]------------------------------>
Past Now Future

FIGURE 2.5. Diagram of the temporal structure of the resultative function.

The following are examples of the resultative function. Note how each sentence implies the persistence of a clear result of the action of the verb.

[24] Terminological note: The resultative perfect has also been called the *perfect of result* (Comrie 1976; 1985), the *resultant state perfect* (Brinton 1988; Carey 1990), and the *stative perfect* (McCawley 1971).

(13) I've broken my leg (and it's still broken). (McCawley 1971)

(14) Meg has already washed the elephant. (Declerck 2006)

The crux of this function is that a bounded action that began at an undefined time in the past produced some direct result that persists in the present (cf. Bybee et al. 1994; Declerck 2006). Resultatives delimit the action of the verb, putting the focus on the effect of the occurrence and not on the event itself (Slabakova 2001); that is, the focus is on the fact that the leg is broken not the action of its breaking. This function is predicated by bounded dynamic verbs; the predicate must denote a change of state in order for the resultant state to emerge (Declerck 2006; Huddleston & Pullum 2002).

The *recent past* function[25] describes a past situation while focusing its proximity and relevance to the present moment (Comrie 1985).

<--------------[recent-E]-S,R------------------------------>
Past Now Future

FIGURE 2.6. Diagram of the temporal structure of the recent past function.

The following are examples of the recent past function. Note the presence of lexical markers that indicate recency: *just*, *recently*, and the imperative from of *run*.

(15) Argh! I've (just) been scooped! (McCawley 1971)

(16) I have recently learned that the match is to be postponed (Comrie 1976)

(17) Run! The river has burst its banks. (Declerck 2006)

The crux of this function is to mark the occurrence of some past action as meaningfully recent (Binnick 1991; McCormick 2008). The focus of this function is on the occurrence itself and not on its result, which does not necessarily persist at present (Declerck 2006).

---

[25] Terminological note: The recent past perfect was originally labeled the *perfect of recent* past (Comrie 1976) and has also been called the *hot news perfect* (McCawley 1971).

Although, the recent past is most often associated with bounded predicates, boundedness is not required (Declerck 2006). This function relies on the interpretation that the action being described occurred ‘recently.’ It relies on context and/or modifiers that suggest the importance of recency (Liu 1993; Quirk et al. 1985).

NONBOUNDED FUNCTIONS. The *continuative* function[26] describes a situation that occurred ‘in the past but continues (persists) into the present’ (Comrie 1976:60; Declerk 2006).

```
<-----------[E--------------S,R]----------------------------->
  Past                     Now                     Future
```

FIGURE 2.7. Diagram of the temporal structure of the continuative function.

The following are examples of the continuative function. Note how the action of the verb persists into the present and may even persist into the future.

(18) I’ve lived here (from some point in the past until now). (McCawley 1971)

(19) I've shopped there for years. (Comrie 1985)

With this function, the situation continues into the present and/or is able to be repeated in the future (as seen in the above examples).[27] The continuative focuses on the persistence of the past situation itself (not its result; Liu 1993). This function requires that the situation is nonbounded (Declerck 2006). The continuative requires rich context and/or

---

[26] Terminological note: The continuative perfect has also been called the *perfect of persistent situation* (Comrie 1976), the *extended-now perfect* (McCoard 1978), the *universal perfect* (McCawley 1971), and the *inclusive past-and-present* (Jespersen 1924).

[27] Fenn (1987) labels the two variants of the continuative from the above examples the *actual continuative* and the *recurrent continuative*, respectively. Actual continuatives (18) are continuous throughout the timespan and do not end which makes them non-repeatable. Recurrent continuatives (19) are either i) repeated over a short period of time or ii) repeated over a much longer period and may be considered habits. This second kind of recurrent continuative is often called the *habitual perfect*, but this is not an independent function of the perfect; rather, it is a special reading of the continuative.

overt adverbial modification, which are ‘indispensible’ for generating and interpreting this function (Liu 1993:66; cf. Schlüter 2000).

The *experiential* function[28] describes a situation that has occurred at least once (but possibly many times) at some point in the past up to the present (Comrie 1976). This function is used to describe a situation that occurred at some time or times in the past but does not extend into the present (Brinton 1988; Carey 1990; Declerck 2006).

<-[E_1---(E_2---E_n---E_{n+1})]S,R------------------------------>
Past Now Future

FIGURE 2.8. Diagram of the temporal structure of the experiential function.

The following are examples of the experiential function. Note how these actions must have occurred once but may have occurred on multiple occasions in the span of time before the present.

(20) I’ve eaten mooseburgers (at some point in the past). (McCawley 1971)

(21) I’ve read most of Raymond Chandler’s novels. (Declerck 2006)

The relevance of the pre-present situation to the present makes this function licit. The unmarked interpretation describes a nonbounded situation that occurs over an indefinite time span within the pre-present (Declerk 2006). The experiential function describes a situation that is repeatable (Davydova 2011; Declerck 2006; Fenn 1987).[29] The experiential often requires rich context to be the most favorable interpretation.

---

[28] Terminological note: The experiential perfect has also been called the *existential perfect* (McCawley 1971), the *up-to-now* interpretation (Declerck 2006), and the *indefinite anterior perfect* (Filppula 1999).

[29] Repeatability for the experiential function takes on a modal character that indicates that the action can occur again in the present and future. The continuative can only reoccur in the future because the situation persists through the present (Michaelis 1998).

2.3.3 Summary of constructions of interest.

The previous two subsections discuss how the simple past and present perfect are understood for the purposes of the present investigation. The acquisition of these constructions requires that the learner distinguish them based on certain tense and aspect features. The present investigation focuses on two of these distinguishing features: current relevance and boundedness.

Current relevance characterizes the relation of the anterior situation to the present; it is crucial to the meaning of the present perfect and must be understood for the present perfect to be acquired. The [+CR] feature is certainly carried by the verbal morphology of the present perfect, but the present investigation elects instead to focus on adverbial modifiers, which adult L2 learners tend to rely on before verbal morphology (see §2.4.2 below). It is common to investigate the present perfect through the instantiation of current relevance in the immediate linguistic context (cf. Elsness 1997; Fenn 1987; Tagliamonte 2000; Van Herk 2008; Winford 1993), and this investigation does so using Davydova's (2011) classification of adverbials. The present perfect regularly collocates with adverbials that are listed in Davydova's (2011) [+CR] category, while the simple past regularly collocates with adverbials that are listed in the [–CR] category (Table 2.1 above). The present investigation utilizes a subset of these modifiers to operationalize current relevance to investigate the effects of overtly marking this feature.

Boundedness concerns a predicate's contextually determined endpoint. It is a very important feature to understand the different functions of the present perfect, which either denote a completed or ongoing/iterative action (cf. Bybee et al. 1994; Carey 1994; Housen 2002; Shirai & Andersen 1995; Slobin 1994; Van Herk 2008); however, it is less meaningful in the simple past where even nonbounded predicates are still implicationally completed at present. This feature helps to separate the present perfect from the simple past and to differentiate the functions of the present perfect. The least marked functions of the present perfect are bounded (Dahl & Hedin 2000; Davydova 2001; Declerck 2006; Huddleston & Pullum 2002; Radden & Dirven 2007) and the most marked are nonbounded (Dahl 1985; Davydova 2011; Declerck 2006). The current investigation manipulates boundedness in order to investigate the effects of this feature's effects.

2.4 Acquisition of present perfect.

The purpose of the present investigation is to shed light on the acquisition and processing of the present perfect through two of its most central features: boundedness and current relevance. In order to understand the context of the investigation and some of the assumptions made therein, it is important to be familiar with some of the prior research into acquisition of tense and aspect broadly and of the present perfect specifically.

Before discussing the acquisition of the present perfect, let us define what it means for the category to be 'acquired'. That is, what forms, meanings, and functions must a learner incorporate into their grammar? According to Bardovi-Harlig (2002), acquisition of the present perfect involves the acquisition of three characteristics: the morphosyntactic form, its semantic and pragmatic features, and its (contrastive) usage. In her investigation, she finds that the acquisition of this first characteristic is accomplished rather early in the acquisition process, so this step is glossed over in the present investigation. Second, a learner must acquire its semantic and pragmatic features. Acquisition of these features occurs very late if at all (Bardovi-Harlig 2002; Housen

2002; Lim 2003). The present investigation is focused on these late-acquired features. The third component is a corollary to the second; the focus of this component is on distinguishing the category from 'close neighbors' with similar usage (Bardovi-Harlig 2002:224). In the present investigation, the simple past is the close neighbor from which the present perfect is distinguished.

The following subsections provide a brief overview of research that is especially relevant to the present investigation. First discussed is child L1 acquisition. This research provides the foundation on which the adult L2 acquisition research is based and provides details relevant to the present investigation's conception of the present perfect and the manner in which it is acquired. The discussion continues with adult L2 acquisition. This research provides the foundation of the present investigation and is used to justify many of the assumptions that underlie the investigation itself and the design of the tasks used therein.

2.4.1 Child L1 acquisition.

Information concerning how L1s are acquired provides necessary background information to understand the research that concerns the acquisition of the present perfect. Children learn to mark temporal distinctions very early in their linguistic development. According to Weist's (1986) cross-linguistic review of developmental patterns of tense and aspect, children first learn to mark temporal distinctions using verbal morphology around two years of age. Within the next year, they acquire the ability to mark these distinctions lexically using adverbs (Pawlak et al. 2006; Weist 1986). Newport (1990) suggests that children's limited ability to attend to large spans of information encourages them to preferentially attend to forms that bear the most meaning (i.e. verbs & verbal morphology; Smith 1980; Valian 2006). This developmental pattern is followed by all typically developing children.

During acquisition, there is a clear bias for marking certain predicates with semantically similar grammatical markers. Most relatable to the present investigation, several early studies find that children do not apply past/perfective morphology to all classes of verbs equally (Antinucci & Miller 1976; Bloom et al. 1980; Bronckart & Sinclair 1973; Brown 1973). Instead, children use past/perfective morphology almost exclusively with telic/bounded predicates (Bloom et al. 1980; Bronckart & Sinclair 1973).[30] This usage bias diminishes with age as children approximate adult use, but it indicates that children are sensitive to aspectual distinctions and adult usage patterns. Children associate past/perfective morphology with telicity because situations marked with past/perfective morphology often express an endpoint or result in much the same way that a telic predicate has a clear endpoint (Bickerton 1981; Slobin 1985). A similar association and usage bias emerges with present/imperfective morphology and durative aspect because both express the persistence of the situation.

Contemporary studies on the acquisition of the present perfect take as their starting point seminal L1A studies in the 1970s and 1980s by Nussbaum and Naremore (1975), Fletcher (1981), Johnson (1985), and Gathercole (1986).[31]

---

[30] See Shirai (2009) for an extensive overview of cross-linguistic investigations of this pattern in L1A.

[31] Treating these studies as seminal is not to ignore other works from 1970s. Studies like Brown (1973), Cromer (1974), Lee (1974), and Fletcher (1979) are foundational studies that consider acquisition of present perfect, but they focus more on the age of acquisition than anything else. Two main findings that emerge out of this research: i) the simple past is acquired before the present perfect, and ii) British English learners acquire the present perfect before American English learners. These findings are interesting in their own

Nussbaum and Naremore (1975) investigate the production of L1 learners of American English aged four to six years. Their research demonstrates two important findings. First, the present perfect is one of the last constructions to be acquired by L1 learners of English. All learners used the simple past in a stable manner before the present perfect, which was not stabilized among the oldest participants (Nussbaum & Naremore 1975). Second, none of the L1 learners in their study consistently used the present perfect; they often substituted similar constructions in present perfect contexts (Nussbaum & Naremore 1975). The younger learners most commonly substituted the simple past; the older learners most often substituted *did/*V-*ed* + ADV where the character of the adverb matched the contextually appropriate function of the present perfect (Nussbaum & Naremore 1975). There are two important takeaways from these results: i) the present perfect emerges after the simple past is acquired, and ii) learners use adverbial modification to index functions of the present perfect before targetlike production.

Fletcher (1981) investigates the production of L1 learners of British English aged three years and three months. His investigation yields three findings relevant to the present investigation. First, the present perfect is first used in simple past contexts (Fletcher 1981). This indicates that the present perfect is conflated with the simple past at the time of emergence. Second, his results show clear differences in the verbs most commonly marked with the simple past and the present perfect. The children use more unique verbs in the simple past (48) than in the present perfect (31), and there is very little overlap between these verbs. Fletcher posits that this result is due to an overuse of

---

right concerning the effects of cognitive development and of the learning environment, but they are not directly relevant to the present investigation.

stative verbs in the present perfect (Fletcher 1981). Third, contexts of use indicate that the children use the present perfect to mark recency or stativity (Fletcher 1981). Overall, Fletcher's work captures three facts: i) the present perfect emerges in simple past contexts, ii) children use a small set of aspectually similar verbs for each construction, and iii) children associate the present perfect with recency or stativity.

Johnson (1985) investigates the comprehension and production of the present perfect among preschool-aged L1 learners of English. The critical addition of this study concerns the overt modification of the stimuli sentences with adverbs. The results of her production task demonstrate a clear association between specific functions of the present perfect and telicity and adverbial modification (Johnson 1985). Learners better comprehend and produce when the telicity of the verb and the character of the adverbial modification semantically concord. Johnson's work indicates that learners benefit from a context in which lexical aspect and adverbial modification exhibit similar semantics.

Gathercole (1986) studies the acquisition of the present perfect by L1 learners of Scottish and American English whose ages range from three to approximately six years.[32] Gathercole makes two important findings that are relevant to the present investigation. First, she finds that the four functions of the present perfect emerge productively at different ages. She proposes that the path of acquisition for the four functions begins with the recent past (41.5%, age <3;0), followed by the experiential (15.5%; age 3;0-3;7), followed by the resultative (6.0%; age 3;0-4;4), and finishes with the continuative (2.5%; age 4;4-4;11). Second, she finds that preschool-aged children have the cognitive capacity

[32] Acquisition for the American English children occurs in children older than Gathercole observes; this discussion concern the acquisition of Scottish English with the caveat that this path is not necessary applicable to American English.

to acquire the complex tense-aspect predications of the present perfect (Gathercole 1986; contra Cromer 1976). The limiting factor is not the learner's cognitive ability but a linguistic deficit. Gathercole's work points to two findings that are important for the present investigation: i) the functions of the present perfect emerge individually, and ii) acquisition of the structure is not limited by an inability of the learner to grasp tense-aspect predications.

The research on child L1 acquisition contributes to the present investigation in five ways. First, learners acquire the present perfect after the simple past. At the time of emergence, the present perfect is used in contexts in which the simple past is more appropriate (Nussbaum and Naremore 1975; Fletcher 1981). This suggests that the present perfect emerges out of the simple past ([anteriority] is acquired before [current relevance]). Second, older child learners are sensitive to features within the verbal predicate, and they use them in order to acquire tense-aspect categories (Bloom et al. 1980; Bronckart & Sinclair 1973). If adult learners behave like older child learners, this suggests that manipulations of boundedness may significantly affect learner processing and comprehension. Third, learners associate the present perfect with telicity, boundedness, durativity, and [+CR] adverbial modifiers (Fletcher 1981; Johnson 1985). This suggests that the features being manipulated herein emerge naturally as indicators of the present perfect and they may be similarly important for adult L2 learners. Fourth, these learners rely on adverbial modification to index functions of the present perfect before they are able to accurately produce the structure (Nussbaum and Naremore 1975), which suggests some benefit to their use and presence in context. Finally, the functions of the present perfect emerge at different rates due to a linguistic deficit (Gathercole 1986).

Adult L2 learners may be similarly constrained by linguistic factors when acquiring the present perfect.

2.4.2 Adult L2 acquisition.

Adult learners mark temporal distinctions in their L2 following a path opposite the one used by children. As discussed above in section §2.4.1, children first mark temporality via grammatical means, then via lexical means (e.g. adverbs), and finally via pragmatic means. In her investigation into L2 tense-aspect marking of adult learners, Bardovi-Harlig (2000b) finds that adult L2 learners go through three stages of temporality marking (with some variability). First, they rely on pragmatic means; that is, they use context to provide the information necessary for the hearer to infer the time a situation occurs. Second, they use lexical means (adverbial modifiers). Third, they use grammatical means (verbal morphosyntax). Many L2 learners never reach the third stage because adults tend to over-rely on lexical information (cf. VanPatten 2002). The reason that adults follow the opposite path children do probably stems from differences in cognitive development and the presence of an L1 grammar with which they can establish temporal reference (Shirai 2009). The cognitive differences give adults more means by which they can express temporal reference without relying on grammatical means. The presence of an L1 is often an obstacle to acquisition. In order to acquire a temporal distinction not present in L1, the learner must realize that an existing category in their interlanguage grammar is insufficient to capture the L2 distinction (Hawkins 2009).

As is the case for children, adult learners exhibit biases concerning tense-aspect marking and semantically similar predicates. Focusing on the emergence of past/perfective morphology, adults do not apply past/perfective morphology to all classes of verbs equally in English (Andersen & Shirai 1994, 1996; Bardovi-Harlig 2000b; Li & Shirai 2000).[33] Adult L2 learners selectively mark telic/bounded predicates with past/perfective morphology; however, these associations are much less consistent than they are in L1 acquisition (Lee 2001; Robison 1995; Rohde 1996; Shirai 2002). One potential cause of this less clear association is the fact that adult L2 learners, especially lower proficiency learners, are more likely to use memorized chunks without internalizing them into their grammar (Bates et al. 1988; Bloom et al. 1980; MacWhinney 1978; Shirai 2004). Adult L2 learners approach the associations found among children only as proficiency increases (Bardovi-Harlig 1998; Bardovi-Harlig & Reynolds 1995; Robison 1995). The following is a brief description of several studies that concern the acquisition of tense-aspect morphology that investigate the effects of lexical aspect:

i) Bardovi-Harlig (1992) finds that adult L2 learners of English mark the simple past more accurately on telic predicates and less accurately on atelic predicates.

ii) Bardovi-Harlig and Reynolds (1995) find that learners of English first mark telic predicates with simple past morphology and that learners do so more frequently. Learners at all proficiency levels mark telic and atelic predicates in the simple past consistently.

iii) Bardovi-Harlig and Bergström (1996) find that lexical aspect does not affect the selection of perfective and imperfective morphology among learners of French.

iv) Comajoan (2006) finds that adult L2 learners of Catalan mark telic predicates more frequently and accurately with perfective morphology and they mark atelic

[33] See Shirai (2009) for an extensive overview of cross-linguistic investigations of this pattern in L2A.

predicates more frequently and accurately with imperfective morphology in a verbal narrative task.

v) Ayoun and Salaberry (2008) find adult L2 learners of English associate telicity and simple past morphology in two elicitation tasks. They find a stronger effect in spoken production than in written production, which suggests that learners rely more strongly on the similarity of the two meanings when pressured.

vi) Izquierdo and Collins (2008) find that L2 learners of French are affected by their L1. L1 Spanish speakers transfer the shared features from the L1 verbal system to the L2, while L1 English speakers rely instead on lexical semantics and rote memorization of pedagogical rules, resulting in less accurate production than the L1 Spanish speakers.

vii) Chan et al. (2012) find that adult L2 learners of English have a tendency to use past tense forms with telic predicates and progressive aspect forms with atelic predicates.

With the exception of Bardovi-Harlig and Bergström (1996), these studies indicate that L2 learners show that learners prefer to mark telic events as completed via past tense morphology, and they prefer to mark atelic situations as ongoing via progressive aspect morphology.

For the present investigation, the most critical factors that differentiate child L1 acquisition and adult L2 acquisition are the presence of an L1 and L2 proficiency. A summary of the literature on L1 influence is far beyond the scope of the present investigation, but we note that L1 effects have been found in all subsystems of language: phonology (Eckman 1977; Major 1987), morphology (Andersen 1983), syntax (Gass

1984; Zobl 1980), semantics (Tanaka 1983; Shirai 1989), pragmatics (Silva 1999), and tense-aspect (Gass & Ard 1983). Concerning the acquisition of tense-aspect, research finds limited L1 transfer effects for inflectional morphology but more pronounced effects for the similarity of the tense-aspect system as a whole (Andersen 1983; Shirai 1992, 2004; Terauchi 1994). That is, L2 learners are most affected by tense-aspect similarities in their L1 and L2 verbal systems (Izquierdo & Collins 2008). For example, in the works of Salaberry and Ayoun, L2 learners transfer the tense-aspect distinctions of their L1 directly to their interlanguage grammar. When acquiring the Spanish *preterito* 'the preterit,' L1 English speakers transfer the simple past from the L1 to the interlanguage grammar as the starting point of the L2 category. In so doing, they largely ignore the perfective-imperfective distinction and instead treat it as a past tense marker. L1 French speakers, whose verbal system marks this distinction like Spanish, mark tense and aspect distinctions more accurately and earlier than the L1 English speakers.

The present perfect, if it is acquired at all, is acquired after all of the other non-perfect tenses (Cromer 1968, 1971; Kuczaj & Daly 1979). Typically, it is acquired by high proficiency learners in the final stages of the interlanguage development (Housen 2002). Among instructed and uninstructed adult L2 learners of English alike, nativelike use of the present perfect occurs after approximately 15 years of exposure to the language (Davydova 2011). Given that children around three years of age have the cognitive ability to understand the functions of the present perfect (Gathercole 1986), why do adult L2 learners acquire the present perfect so late when they have the cognitive ability to do so earlier? The answer concerns both linguistic exposure and proficiency. Learners must reach a certain level of English proficiency before they are able to notice the unambiguous evidence needed to generate and acquire the present perfect category and its associated features as distinct from similar categories in the L1 and L2 (Dussias & Piñar 2009; Hawkins 2009). This is true even among instructed learners who must acquire the morphosyntactic and semantic features associated with the present perfect before they benefit from instruction (Slabakova & Montrul 2002; White 2008). Once they acquire these features, accuracy in production and comprehension increases with English proficiency (Bardovi-Harlig 2002).

As is the case for L1 acquisition of the present perfect, L2 researchers have also conducted seminal research that has laid the foundation for the present investigation. These studies include Bardovi-Harlig (2002) and Housen (2002), which concern the acquisition of the present perfect and of tense-aspect constructions more broadly. Recent investigations of the acquisition of the present perfect include McCormick (2008), Davydova (2011), Terán (2014), Uno (2014), and the pilot for the present investigation (Farina 2014, 2015).

Bardovi-Harlig (2002)[34] investigates the written and oral production of instructed adult L2 learners of American English in an intensive English language program. Her research demonstrates several important findings. First, Bardovi-Harlig (2002) observes that L2 learners acquire the present perfect after the simple past has been fully acquired. Second, at the time of emergence, most learners use the present perfect in writing before doing so in speech (Bardovi-Harlig 2002). This finding suggests that the written mode is a prime area to investigate the structure's emergence because the written mode permits

[34] This article is a revised version of Bardovi-Harlig (1997), Another Piece of the Puzzle: The Emergence of the Present Perfect, which was originally published in Language Learning 47(3).

less complex tasks than the spoken mode. Third, she finds that the learners reproduce the adverbial modification patterns present in native speaker production (Bardovi-Harlig 2002). Learners are sensitive to adverbial modification patterns in the input and successfully replicate them by the end of the study. Fourth, the learners' nonnativelike usage of the present perfect indicates that they associate the construction with both past time and present time (Bardovi-Harlig 2002). Overgeneralized use of the present perfect occurs in simple past contexts in 63.1% of cases, which indicates that L2 learners associate the present perfect most strongly with the simple past. Bardovi-Harlig (2002) interprets this finding as an indication that current relevance has not yet been integrated into the interlanguage grammar. Fifth, Bardovi-Harlig notes that proficiency and familiarity with the construction positively correlate with frequency and appropriateness of use (Bardovi-Harlig 2002). This finding only begins to emerge among the higher proficiency learners, which suggests that learners must reach a certain proficiency level before acquisition is measurable. Overall, Bardovi-Harlig (2002) makes five contributions relevant to the present investigation: i) L2 learners are able to distinguish the simple past and present perfect, ii) they can demonstrate nativelike adverbial marking, iii) performance is more accurate in the written mode, iv) L2 learners associate the present perfect with the simple past, and v) proficiency and familiarity affect performance.

Housen (2002) investigates the oral production of instructed late-child L2 learners of British English in a foreign language context. His research demonstrates several important findings concerning the order of acquisition of the English grammatical tenses and of the present perfect specifically. First, the present perfect only emerges in contrastive use after the simple past has been fully acquired (Housen 2002). This late acquisition is despite the fact that the learners receive explicit, contrastive instruction on the differences between the simple past and present perfect (Housen 2002; cf. Buczowska & Weist 1991; Pienemann 1987). Thus, the order of acquisition is stable even among instructed learners. Second, Housen (2002) finds evidence of nonnativelike use of the structure at the time of emergence. Like Bardovi-Harlig (2002), he suggests that these early uses are encoding only select features of the appropriate construction. After this initial stage, the simple past and present perfect are rarely misused (Housen 2002). Third, he finds a distinct contrast between the order of emergence of morphological forms and of functional marking of tense-aspect with said forms (Housen 2002). The morphological and functional developmental paths are summarized in the following tables.[35]

TABLE 2.2. Order of emergence of morphological categories (Housen 2002:158).

Stage	Category	Comment	Example
0	Invariant *V*	esp. base form *Vø*	*see, play*
1	Present Participle *Ving*	initially without Aux. *Be*	*seeing, playing*
	Irregular Past of *Be*		*was*
2	Irregular Past (other verbs)		*had, got*
3	**Regular Past *Ved***		*played, worked*
	Future *Be Going* + *Vinf*	allomorphs: without Aux. *Be, to, -ing; gonna*	*is going married ; are go dancing; am going to take; is gonna happen*
4	**Perfect *Aux + V***	allomorphs: Aux. *Be* and *Have;* initially *V* =*Vø*	*have see, is fall, is fallen, has fall, have fallen*
	Present *Vs*		*goes, comes, does*
	Future *Will* + *V*		*will make, will see*

[35] These tables are influenced most by the work of Dietrich et al. (1995) and Giacalone-Ramat (1995).

This table shows the order in which the verbal morphology first appear among the learners. The present investigation focuses only on stages 3 and 4, which concern the acquisition of the regular past tense and the present perfect.

TABLE 2.3. Order of functional marking of tense-aspect meanings and their respective markers (Housen 2002:162).

Stage	Meaning	Form
0	Pre-functional stage	Random and complementary distribution of forms
1	**Anteriority (past and perfect)**	1. Past of *Be* (*was*)
		2. Irreg. Past (other verbs)
		3. Perfect *Have/Be + V*
		4. Reg. Past *Ved*
	Imperfectivity/progressivity	1. Aux. *Be + Ving*
2	Futurity	1. Aux. *Be + Going + Vinf*
		2. Aux. *Be + Ving*
		3. Aux. *Will + V*
		4. Present *Vø/Vs*
	Habituality	1. Aux. *Be + Ving*
	Present	1. Present *Vø/Vs*
3	Simple past	1. Past of *Be* (*was, were*)
		2. Irreg. Past
		3. Reg. Past *Ved*
	Present perfect	**1. *Have/Has + Ved/Virreg***
	Past perfect	1. *Had + Ved/Virreg*

This table shows the order in which functions emerge and with what verbal morphology they are encoded. The important contrast between these two tables for the present study concerns the discrepancy between the present perfect being used to mark anteriority (stage 1) and the targetlike functions (stage 3). These findings suggest that present perfect morphology first marks anteriority before other target features. Housen (2002) makes three contributions to the present investigation: i) the present perfect emerges after the simple past, ii) present perfect morphology is not fully encoding target features at the time of emergence, and iii) anteriority is encoded before other target features.

McCormick (2008) investigates the written production of instructed adult L2 learners of Canadian English taking college-level courses in an English-speaking university. His research concerns the teachability of the present perfect and focuses on how the first language affects the interlanguage after contrast-focused instruction. McCormick (2008) finds that the first language of the learner strongly affects their success. Chinese speakers, whose L1 overtly marks current relevance, experience the most benefit from instruction and experience virtually no L1 interference. Spanish and Farsi speakers, whose L1s have formally similar but functionally non-equivalent perfect constructions, experience some benefit but are limited by L1 interference. Russian speakers, whose L1 has no formal or functional equivalent to the present perfect, experience only minimal benefit from instruction. McCormick's (2008) research provides clear evidence that features in the learner's L1 affect their acquisition of the present perfect.

Davydova (2011) uses corpora of standard, nonstandard, and learner varieties of English to investigate acquisition through typological variation. Her most crucial additions to the present investigation concern patterns of use for the functions of the present perfect and how L2 learner varieties rely on VP-level aspect and overt markers of current relevance to convey them. First, her results indicate that the learner varieties of English rely on the present perfect for the continuative and resultative function with the continuative function being much more strongly associated with the present perfect (Davydova 2011). These functions are strongly favored in usage over the experiential and recent past functions (Davydova 2011). Learner usage approaches the standard norm as proficiency increases. That is, their bias toward the continuative balances out with the resultative. Second, her results show that learners associate the present perfect with durativity (in general) and with telicity (at higher proficiency). The learner varieties of English show a strong preference for durative, atelic predicates in the present perfect (Davydova 2011). This association makes sense considering that they most often use the present perfect to mark the continuative function, which marks an enduring situation. At higher proficiencies, they associate telic aspect with the resultative function specifically. Third, she finds variation in overt current relevance marking using [+CR] adverbial modifiers. The learner varieties strongly favor overt marking of current relevance (Davydova 2011), which may be due to their strong preference for the continuative function. In general, overt marking of current relevance occurs more often in learner varieties than in Standard English. It seems that these learners benefit from overtly marking this feature of the present perfect. Overall, Davydova (2011) makes three contributions to the present investigation: i) the resultative and continuative functions are most strongly associated with the present perfect, ii) telicity and durativity correlate with the use of the present perfect, and iii) overt markers of current relevance are used by learners more often than by Standard English speakers.

Terán (2014) investigates the effects of lexical aspect on the grammatical ability of a sample of L1 Spanish speakers learning L2 English in an EFL teacher training program. Her research focuses on the effects of lexical aspect on the accuracy with which learners use the present perfect to express the continuative and experiential functions; it yields three important findings. First, it finds that accuracy varies according to proficiency and function (Terán 2014). She finds that accuracy increases with proficiency and that both proficiency groups perform more accurately in continuative contexts than in experiential ones. Second, a qualitative analysis of the contexts of production indicates that all of the continuative contexts contain adverbial modifiers that collocate with this function (Terán 2014). As such, adverbial modification may be the cause of the learners' increased accuracy with the continuative. Third, she finds effects for lexical aspect on accuracy (Terán 2014). Learners perform more accurately on durative verbs for both constructions. Likewise, the continuative is more accurately marked on nondynamic predicates, and the experiential is more accurately marked on dynamic predicates. Overall, Terán (2014) makes three contributions to the present investigation: i) learners more accurately produce the continuative than the experiential function, ii) adverbs in the sentential context correlate with improvements in accuracy, and iii) learners more accurately produce the construction when the predicate is atelic and nonbounded.

Uno (2014) investigates the effects of lexical aspect on the grammatical ability of a sample of native speakers of Japanese learning L2 English. Her research focuses on the effects of lexical aspect on the accuracy with which learners use the present perfect in contexts with and without durative adverbial modifiers. She finds that learners more often and more accurately use nonbounded verbs in the present perfect in contexts with durative adverbs (Uno 2014). In contexts without durative modifiers, learners show no preference for boundedness and are less accurate in their production (Uno 2014). These results contribute the finding that adverbial modification and boundedness affect the accuracy of learner production.

The pilot for the present study investigates the effects of lexical aspect and boundedness on the processing speed and accuracy of instructed adult L2 English learners with high proficiency (Farina 2014, 2015). The investigation used a variation of the self-paced reading task to investigate real-time processing. It finds that learners experience processing facilitation when reading telic/bounded events in the present perfect compared to the same predicates in the simple past due to the shared semantics of boundedness and the present perfect. Likewise, English native speaking controls experience processing inhibition when reading atelic/nonbounded occurrences in the present perfect compared to the same predicates in the simple past due to the contrastive semantics of nonboundedness and the present perfect. Methodologically, this pilot reveals that a self-paced reading task is sensitive enough to measure the factors that contribute to language processing being studied in the present investigation. The pilot makes two contributions to the present investigation: i) L2 learners and native speakers experience processing facilitation and inhibition depending on manipulations in telicity, boundedness, and tense-aspect and ii) a self-paced reading task is capable of capturing subtle processing differences that result from manipulations in tense and aspect.

The research on adult L2 acquisition contributes to the present investigation in seven ways. First, the present perfect emerges after the simple past (Housen 2002), and L2 learners initially associate it with anterior reference alone at the time of emergence (Bardovi-Harlig 2002; Housen 2002). For the present investigation, these findings suggest that the present perfect is best compared with the simple past. Second, the present perfect and its functions emerge at different levels of English proficiency: the resultative and continuative functions are more strongly associated with the present perfect and more accurately produced than are the experiential and recent past functions (Davydova 2011; Terán 2014). The present investigation attempts to clarify the manner in which these functions are differentiated and understood in real time. Third, learners associate the present perfect as a whole with durativity, the resultative and recent past functions with telicity/ boundedness, and the continuative and experiential functions with atelicity nonboundedness (Davydova 2011; Terán 2014; Uno 2014). The present investigation uses these associations to scaffold the learners in the reading tasks. Fourth, adult L2 learners are sensitive to and benefit from adverbial modifiers that are present in the immediate context (Davydova 2011; Terán 2014; Uno 2014), and they can produce these modifiers in a nativelike way (Bardovi-Harlig 2002). For the present investigation, these results suggest that manipulations of adverbial modifiers will affect how the present perfect is processed by L2 learners. Fifth, prior research indicates that L2 learners benefit from having a category in the L1 that is similar (morphosyntactically and semantically) to the one being acquired in the L2 (Andersen 1983; Izquierdo & Collins 2008; Shirai 1992; Terauchi 1994).[36] Sixth, instruction has been shown to improve frequency and accuracy of production (Bardovi-Harlig 2002; McCormick 2008). No data suggest that instruction alters the path through which the present perfect is acquired, but there are data to suggest that instruction increases the speed with which it is acquired (Bardovi-Harlig 2002). These findings suggest that instructed learners are the ideal population to be sampled because they are more likely to acquire the present perfect and they do so more quickly. Finally, the pilot study for the present investigation indicates that the aforementioned factors and effects should be measurable by the online processing task used in the present investigation. The results suggest that subtle manipulations to linguistic context significantly affect the processing strategies of native speakers and L2 learners alike.

[36] The present investigation considers L1 transfer effects by comparing the performance of L2 learners from three linguistic backgrounds: L1 Arabic, L1 Chinese, and Other L1 (see §3.3).

CHAPTER 3 SCOPE OF THE PRESENT INVESTIGATION

The present investigation concerns the processing and acquisition of two features central to the present perfect: boundedness and current relevance. This chapter discusses the three central foci of this investigation. The primary focus is an evaluation of the accounts that describe the acquisition of the present perfect in relation to the simple past; both accounts are briefly described in the first section. The secondary foci concern possible factors that may affect the acquisition of the present perfect: L2 English proficiency and transfer of first language processing strategies or linguistic structures. These secondary foci are addressed in the second and third sections. In each section, these effects are situated in their broader theoretical context and specifically described as they are operationalized in the present investigation. Next, the rationale for the present study is explained. The research questions and hypotheses conclude this chapter.

3.1 COMPETING ACCOUNTS OF ACQUISITION.

The present investigation compares two accounts of acquisition: the complexity account and the prototype account. The complexity account predicts that structures are acquired from least complex to most complex (§3.1.1), and the prototype account predicts that structures are acquired from most exemplary to least exemplary (§3.1.2). These two accounts are incorporated into the present investigation because, in the case of the present perfect, they provide contradictory predictions for the path through which the structure is acquired. Because of this difference in predictions between accounts, the acquisition of present perfect and its associated features is an excellent avenue through which to investigate the descriptive power of each account. The present investigation compares these accounts via the opposing predictions that they make concerning the manner in which two features associated with the present perfect are processed and acquired. Specifically, this thesis project evaluates the ability of each account to describe online processing data investigating the interactions of boundedness and current relevance in the present perfect and the simple past. In addition, it evaluates the L2 users' grammatical (morphosyntactic and semantic) knowledge of boundedness and current relevance in these past-time contexts.

Both accounts share the foundation that information is mapped within the learner's interlanguage grammar using individual features that combine to compose the target category. In order to learn a tense-aspect category, learners must master the morphosyntactic forms used to mark the category (*form-to-form mapping*) and they must master the temporal-semantic relations expressed by that form (*form-to-function mapping*; Housen 2002). For example, acquiring the present perfect may involve associating the form *have* + V-*en/-ed* with reference to functions expressed by the category. In these accounts, functions are represented by features. Learning occurs through the stepwise mapping of features that comprise a category individually (or at least preferentially relying on one feature over another) as opposed to the mapping of the entire category at once (Bardovi-Harlig 2002; Ellis 2008). In the case of the present perfect, a learner needs to associate the verbal morphology with its temporal-semantic relations, which can be understood as the set of features: [+bounded], [+anterior], [+indefinite], and [+current relevance]. Formal accounts hold that a learner acquires each feature individually instead of acquiring the features that comprise present perfect category at the same time.

Now that the foundations of formal, feature-based acquisition have been introduced, the accounts themselves can be considered. The two formal accounts being compared in the present investigation are the complexity account and the prototype account. The following two sections briefly describe these two accounts as they are understood in the present investigation; the final section summarizes each.

3.1.1 Complexity account.

The complexity account operates under the assumption that the characteristics of a category's component features determine its complexity. Complexity can be considered in absolute or relative terms. Absolute approaches to complexity consider complexity to be an objective feature of language. In this approach, complexity is mathematically based (Davydova 2011). Categories with fewer features are less complex than those with more features (Miestamo 2008). Relative approaches to complexity consider complexity from a user-oriented, subjective perspective. Instead of measuring it mathematically, complexity is defined in terms of processing cost to the speaker, which is generally measured using a metric like time-on-task or time-to-response (Miestamo 2008). This perspective shifts the operationalization from number of formal features to number of milliseconds it takes to perform some task. In both perspectives, the complexity of a category determines its relative difficulty for a learner trying to acquire it. Less complex categories are acquired before more complex ones. The present investigation relies on both absolute and relative measures to estimate the complexity of the linguistic forms under investigation. Three levels are treated here: morphology, semantics, and temporal relations.

Morphological complexity concerns the complexity of the inflectional system of the language. For the purposes of the present investigation, morphological complexity is best considered at interfaces: i) morphology & phonology, ii) morphology & syntax, and iii) morphology & semantics. i) At the first interface, complexity concerns the phonetic form of the morpheme as well as how it interacts with the verbal stem phonologically. Within the present investigation of English, this level of complexity can largely be avoided with careful consideration of experimental items. In terms of inflectional morphology alone, the simple past and the present perfect are equally complex. The present investigation eliminates this level of complexity by only using verbs that follow regular strategies of word formation (i.e. weak verbs[37]). ii) At the second interface, complexity concerns the interaction of morphological and syntactic marking. The first meaningful distinction here is the difference in absolute complexity between analytic strategies (relying chiefly on syntax to convey grammatical information) and synthetic strategies (relying chiefly on morphology to convey grammatical information).[38] According to Kusters (2003), analytic strategies are less complex because they only deal with one subsystem of language, syntax; being less complex, analytic languages are easier to process. A system that combines both analytic and synthetic strategies is more complex still because it requires both strategies in concert (Davydova 2011). The present

[37] The dichotomy between *strong* and *weak* verbs is currently used by English grammarians to differentiate the irregular strong verbs from the regular weak verbs. Originally, both labels applied to separate regular systems in the Germanic languages. Grimm (1819) labeled verbs that are inflected for tense-aspect via the Proto-Indo-European system of ablaut (vowel alternation) *starke Verben* 'strong verbs' and those verbs that are inflected in the past via a dental suffix (i.e. /t/, /d/, and /ɪd/) *schwache Verben* 'weak verbs.'

[38] For additional information on analytic and synthetic languages, see Eifring & Theil's (2005) fourth chapter, which discusses linguistic typologies.

perfect is structurally more complex than the simple past because it employs more strategies in its proper formation (cf. Quirk et al. 1985). The present perfect requires two strategies: analytic (*have* + verb) and synthetic (V-*ed/-en*). The simple past only requires one strategy: synthetic (V-*ed*). iii) At the third interface, complexity concerns morphology and semantics. Morphology that conveys grammatical information increases the complexity of a linguistic (sub)system (Kusters 2003; McWhorter 2001; Szmrecsanyi & Kortmann 2009). The present perfect is more complex than the simple past. Present perfect morphology contains tense, aspect, and semantic/pragmatic features at this level while the simple past contains only tense and aspect features.

Semantic complexity concerns the complexity of the meaning(s) denoted by a linguistic category. For the present investigation, semantic complexity is considered through two lenses: i) composition of the category and ii) the relationship between form and function. i) Complexity emerges from how the category is constructed following the principle of compositionality. This principle states that 'the meaning of a grammatically complex form is a compositional function of the meanings of its grammatical constituents' (Cruse 2000:68). That is, the meanings of each form that comprise a category sum to the meaning of the category as a whole, which makes the present perfect more complex than the simple past. The present perfect construction is (minimally) composed of three meaning-bearing constituents (auxiliary + verb + inflectional suffix), whereas the simple past construction is composed of two meaning-bearing constituents (verb + inflectional suffix). The principle of compositionality also helps to distinguish the relative complexity of the functions of the present perfect. Following the assumption within the absolute approach that lacking a feature is less complex than possessing a feature, the nonbounded functions are less complex than the bounded functions in their unmarked usage. ii) Complexity also emerges from the regularity with which and the uniqueness by which a semantic distinction is made. Miestamo (2008) operationalizes the One-Meaning-One-Form Principle as a metric for complexity. This principle states that a linguistic form is ideally matched to a single meaning. That is, a form with a single meaning is less complex than a form that possesses more than one meaning (Miestamo 2008). It follows from this proposal that the present perfect is more complex than the simple past. This is because the present perfect predicates four functions while the simple past only predicates two.

The final factor that affects complexity is the *temporal relation* (or relations) predicated by a category. The important distinction for temporal predication is between *simple tenses* and *complex tenses*. Following Radden and Dirven (2007), simple tenses are grammatical tenses that only express deictic temporal relations in which the point of reference is the present moment. Complex tenses are grammatical tenses that express two temporal relations: i) the relation between the present moment and the time of the situation and ii) the relation between the time of the situation and the speaker's point of reference.[39] Radden and Dirven (2007) propose that simple tenses, which express one relation, are less complex than complex tenses, which express two separate relations. The simple past is an absolute tense that expresses deictic time from the present moment, which makes it a simple tense. The present perfect expresses both a deictic temporal relation and grammatical aspect; therefore, the present perfect is a complex tense. Thus,

[39] Simple tenses are the same as Comrie's (1985) absolute tenses. Complex tenses are not the same as Comrie's (1985) relative tense; rather, they are the combination of an absolute tense and grammatical (viewpoint) aspect.

in terms of the temporal relations they express, the present perfect is considered more complex than the simple past.

Complexity operates at three levels for the purposes of the present investigation. Morphological complexity is largely operationalized within the absolute approach. The phonetic length, the number and kind of morphosyntactic strategies employed, and the number of meanings per form determines comparative complexity of each structure. Semantic complexity is determined within the absolute approach via a form's syntactic-semantic composition and its singularity of meaning/function. Temporal relational complexity concerns the number of relations between three points in time: the event time (E), the speech time (S), and the reference time (R). Structures that mark fewer relations (i.e. E=S=R or E≠S=R) are less complex than those that mark more relations (i.e. E≠S≠R). Overall, these levels indicate that the present perfect is more complex than the simple past, that bounded structures are more complex than nonbounded ones, and that structures with current relevance overtly marked via adverbial modifiers are more complex than those structures that lack the overt marking of this feature.

Following the above discussion, the complexity account posits the following path of acquisition outlined in Figure 3.1:

$$\text{Simple past} \rightarrow \text{Pres. perfect}^{[-\text{BND}]}_{[-\text{CR}]} \rightarrow \text{Pres. perfect}^{[+\text{BND}]}_{[-\text{CR}]} | ^{[-\text{BND}]}_{[+\text{CR}]} \rightarrow \text{Pres. perfect}^{[+\text{BND}]}_{[+\text{CR}]}$$

FIGURE 3.1. Complexity-driven path of present perfect acquisition.

The path begins with the least complex form of the present perfect, the one which is nonbounded and not marked for current relevance in the immediate context. The final step is the most complex form of the present perfect, the one which is both bounded and overtly marked for current relevance. Those forms of the present perfect that are characterized by the presence of one feature and the absence of the other comprise the middle step (or steps) of the path. The present investigation is not concerned with the exact ordering of the structures in the middle step; their order depends on the relative complexity and the concomitant effects on processing of the syntactic-semantic feature boundedness and of the temporal relational feature current relevance.

3.1.2 PROTOTYPE ACCOUNT.

The prototype account operates under the assumption that certain features are more central to a category while others are more peripheral. Central features are considered *prototypical* in this paradigm. Central features are acquired first due to cognitive processing factors (discussed below). As summarized by Andersen (2002),[40] there are five factors that underlie the account and that describe the formation of categories and prototypes. These factors account for how trends between verbs and morphology come about based on universal cognitive-processing strategies (Shirai 2009). The factors are i) prototypes, ii) the one-to-one principle of interlanguage construction, iii) the relevance principle, iv) the congruence principle, and v) the distributional bias principle.

The notion of the *prototype* is the factor most central to this account. It is the factor that each of the others affect. Within the cognitive sciences, prototypes have been conceptualized in several different ways since they were first proposed in the works of Rosch (1973, 1977, 1978; Rosch & Mervis 1975).[41] The present investigation combines

[40] A similar set of factors appears in Andersen (1989, 1993) and Andersen & Shirai (1994, 1996). For the sake of consistency, only Andersen's (2002) labels are used here although information comes from all sources.
[41] A full characterization of the history of prototypes is beyond the scope of the present investigation. For a summary of different perspectives on the topic, see Taylor (2008).

several perspectives to form a single coherent notion.[42] Per Rosch (1973, 1978), prototypes are treated as collections of features. These features are abstracted through induction from focal exemplars (Heider 1971). Following Rosch (1977), the present investigation conceptualizes prototypical features as those features so central to a category so as to act as category defaults (features that are assumed to be present given no additional knowledge). Which features are considered default can vary depending on the knowledge of the speaker (Barasalou 1987). Prototypical features are the most basic or central to the category and features are considered less prototypical if the meaning is extended or motivated in some way by the context of use (cf. Verspoor & Tyler 2009). Categories are defined in contrast to the prototypical features of their most similar neighbors (Taylor 2008). In this study, the present perfect is defined in contrast to the simple past. The present investigation follows Shirai (2009) for the prototype of the simple past; he proposes that the prototypical past describes a [-durative], [+telic], [+anterior], [+unitary] situation wherein some features are more central to the category. The present investigation follows Dahl (1985), Bybee and Dahl (1989), and others for the prototype of the (present) perfect; the researchers propose that the prototypical perfect describes a [+durative], [+telic], [+bounded], [+anterior], [+current relevance], [-definite] situation wherein some features are more central to the category than others (cf. Anderson 1982; Bybee et al. 1994; McCormick 2008). The features of greatest interest for this investigation are boundedness, which distinguishes bounded and nonbounded

[42] This decision is neither taken lightly nor is it particularly unique to this investigation. There is not currently a commonly accepted means by which the structure or nature of a prototype category can be determined. The present investigation follows previous research that determines prototypicality based on corpus data, psycholinguistic investigations, and the researcher's own intuitions (cf. Tanaka 1990 in Andersen & Shirai 1996).

functions of the present perfect, and current relevance, which distinguishes the present perfect and simple past.

The *One-to-One Principle of Interlanguage Construction* leads a learner to assume that each grammatical morpheme has only a single meaning, function, and typological distribution (Andersen 1984, 2002). That is, learners simplify a new form such that polysemy and usage patterns are not considered. This leads to a simpler conception of the morpheme than is present in the target grammar. This principle causes learners to use morphemes much more conservatively than is done by speakers of the target language (Andersen & Shirai 1996). Over time, the learner's interlanguage becomes more targetlike as the learner incorporates semantic and functional complexity to the morpheme.

The *Relevance Principle* follows from the cross-linguistic work of Bybee (1985) and Slobin (1985). In these works, they find that grammatical morphemes are placed closer to the verbal stem when they are more 'relevant' to the meaning of the verb. From this trend, they conclude that aspectual markers are more relevant to the meaning of the verb than are tense or agreement markers. Both learners and native speakers choose which verbal morphology to use based on their relevance to the semantics of the verb (Andersen & Shirai 1996) and based on this typological trend (Bybee 1985; Slobin 1985).

The *Congruence Principle* is a corollary to the Relevance Principle.[43] It helps to explain which relevant verbal morpheme is selected to mark the predicate. This principle describes how learners associate a morphological form with a predicate based on the similarity (or 'congruence') of both their meanings (Andersen 1993). According to this

[43] Giacalone-Ramat (1995) labels this factor the *Principle of Selective Association*.

principle, learners assign the grammatical morpheme to the predicate that is most similar to it in meaning (Andersen 2002).

The *Distributional Bias Principle* explains why learners exhibit a strong bias in their use of verbal morphology.[44] Usage patterns of learners reflect an exaggerated or absolute version of the distributional biases present in the target language (Andersen 2002). That is, when one form appears in a certain context more often than in another context, learners initially restrict its usage to that context (Andersen 1990). For example, if 60% of simple past forms in the input occur with bounded verbs, learners will initially create a nearly 100% restriction of use based on the overextension of typological frequency observed in the input. Specific to the present investigation, the biases in the input of bounded events used in both the present perfect and simple past may lead not to two prototypes but a single one, which only separates with increased proficiency in the target language (Ellis 2002; Wulff et al. 2009).

Prototypicality is rooted in the categorization of forms based on similarities and differences in meaning. Prototypes are composed of the features that are most similar to the meanings of other forms in the same category and most different from those in neighboring categories. Following Andersen (2002), four principles underlie the generation of prototypes during acquisition of tense-aspect markers. First, learners tend to initially over-simplify the morphosyntactic forms, assigning to each only a single meaning/function. Second, learners interpret aspectual features to be more central to the category than tense features, a tendency which contributes to the next principle. Third,

[44] Also called the *Distributional Bias Hypothesis.* For more information on this principle, see Andersen (1990, 1993) and Andersen & Shirai (1994, 1996). For a summary of research in support of this principle, see Andersen & Shirai (1996).

learners assign tense-aspect markers to verbs based on their shared aspectual character; that is, lexical aspect bootstraps assignment of grammatical aspect markers. Finally, learners exaggerate target usage biases, restricting their own usage patterns to only the most typologically frequent contexts. As the learner's interlanguage grammar develops, the effects of the principles diminish and the learner approximates target usage. Overall, this perspective indicates that the prototypical features of the present perfect differentiate it from the simple past. The effects of these features should be observable in understanding and usage, and they should diminish as proficiency increases.

Following the above discussion, the prototype account posits the following path of acquisition outlined in Figure 3.2:

$$\text{Simple past} \rightarrow \text{Pres. perfect}_{[+\text{CR}]}^{[+\text{BND}]} \rightarrow \text{Pres. perfect}_{[+\text{CR}]}^{[-\text{BND}]} \big| {}_{[-\text{CR}]}^{[+\text{BND}]} \rightarrow \text{Pres. perfect}_{[-\text{CR}]}^{[-\text{BND}]}$$

FIGURE 3.2. Prototype-driven path of present perfect acquisition.

The path begins with the most prototypical form of the present perfect, the one which is bounded and marked for current relevance in the immediate context. The final step is the least prototypical form of the present perfect, the one which is nonbounded and not marked for current relevance. As is the case for the complexity account above, the present investigation is not concerned with the middle step, which is not clearly predicted; the order of the two structures in the middle step(s) depends on the comparative centrality of the [+CR] and [+BND] features to the prototype on the whole.

3.1.3 COMPARISON OF ACCOUNTS.

The complexity account characterizes the formal variation between the simple past and the present perfect and between the bounded and nonbounded functions of the present perfect. The comparative complexity of the present perfect and simple past is summarized in Table 3.1; '=' indicates equal complexity, '+' greater complexity, and '-' less complexity. Under the principle of compositionality, the presence of a feature ([+F]) is considered more complex than the absence of a feature ([-F]).

TABLE 3.1. Summary of relative complexity of the present perfect and simple past.

	Morphological			Semantic		
	Phonology	Syntax	Semantics	Composition	Form-meaning	Temporal relational
Simple past	=	-	-	-	-	-
Present perfect	=	+	+	+	+	+

Current relevance is a temporal relational feature. Since each function of the present perfect conveys a different sense of current relevance, it is rather difficult to compare current relevance across functions. As such, the present investigation uses this feature chiefly to distinguish the present perfect from simple past and to evaluate the ability of L2 users to process and comprehend this feature. Boundedness is a semantic-syntactic feature. Unmarked uses of the recent past and resultative functions are [+BND] and unmarked uses of the experiential and continuative functions are [–BND]. Although, boundedness need not be present or absent, the consistency with which it collocates with these functions makes it especially relevant to the composition of the functions among learners (Davydova 2011; Terán 2014; Uno 2014). The present investigation uses this feature to evaluate the ability of L2 users to process and comprehend this feature. Previous research in the complexity account has focused on the production of morphosyntax and semantics; online investigations of adult native speakers have shown that participants take more time to process complex constructions than their less complex equivalents (Coll-Florit & Gennari 2011). The present investigation seeks to evaluate the complexity account by measuring the qualitative and quantitative processing strategies of native speakers and adult L2 users for indications processing is affected by the manipulation of features that affect complexity. It follows the assumption that less complex present perfect structures ([–BND] & [–CR]) are processed more readily than more complex ones ([+BND] & [+CR]).

The prototype account characterizes the cognitive processing strategies by which learners come to associate form and meaning. The prototypical features of the simple past and the present perfect are summarized in Table 3.2; features manipulated in the present investigation are in bold: boundedness and current relevance.

TABLE 3.2. Summary of prototypical features of the present perfect and simple past.

	[Durative]	[Anterior]	[Definite]	[Telic]	**[Bounded]**	**[Current Rel.]**
Simple past	-	+	+	+	**+**	**-**
Present perfect	+	+	-	+	**+**	**+**

The present perfect prototype is theoretically [+BND] as is well-justified in the literature summarized in the previous chapter.[45] However, this prototype may not be contrastive enough to be successfully distinguished from the simple past without sufficient understanding of the [CR] feature (cf. Clark 1990; Taylor 2008). That being said, the bounded prototype is assumed for the present investigation. The [CR] feature is believed to be more important for the learner. Knowing that category prototypes develop in opposition to neighboring categories (Taylor 2008), the present perfect category develops based on prototypical features that are most different from the prototypes of the simple

[45] Although Andersen and Shirai (1996) propose that child and adult learners need not share the same prototype, the surveyed child L1 and adult L2 acquisition research indicates that learners probably share prototypes. Both sets of learners associate the resultative and recent past functions with bounded predicates and associate the continuative and experiential functions with nonbounded predicates (Johnson 1985; Davydova 2011; Terán 2014; Uno 2014). Corpus-based research finds the same pattern among native speakers (Davydova 2011; Declerck 2006).

past (Verspoor & Tyler 2009). The more a feature reliably predicts category membership, the more learners rely on this feature (Ellis 2008; Verspoor & Tyler 2009). The most reliable feature for learners is current relevance (Dahl & Hedin 2000; Davydova 2011; Lindstedt 2000); once it is acquired, child L1 and adult L2 learners alike rely on it to successfully parse present perfect constructions (Bardovi-Harlig 2002; Comajoan 2006; Wulff et al. 2009). Previous research in the prototype account has focused on the production of forms based on lexical aspect and adverbial context; the present investigation seeks to evaluate the prototype account by measuring the qualitative and quantitative processing strategies of native speakers and adult L2 users for indications processing is affected by the manipulation of prototypical features. It follows the assumption that present perfect prototypes ([+BND & [+CR]) are processed more readily than nonprototypes ([–BND] & [–CR]).

3.2 L2 PROFICIENCY EFFECTS.

One of the more stable findings in acquisition research is the positive correlation between L2 proficiency and L2 performance. That is, as one's L2 proficiency increases so too does their performance on tasks in the L2. The offline production study by Bardovi-Harlig (2002) reveals that performance in a grammatical tense assignment task varies by proficiency. Offline comprehension studies suggest that adult L2 learners are capable of accurately comprehending differences in tense-aspect morphology regardless of their proficiency level (Slabakova 2003; Slabakova & Montrul 1999). Offline research specific to the present perfect, such as Bardovi-Harlig (2002), Terán (2014), and Uno (2014), indicates that L2 users have the ability to integrate this information, but it is not clear that this behavior will be measurable in an online task. That is, the results of the previous research may be more the result of metalinguistic knowledge than the state of interlanguage grammar. Online research minimizes the effects of metalinguistic knowledge so that the state of the grammar can be studied. In online research, increased L2 proficiency has been found to free up cognitive resources used in language processing (Dussias & Piñar 2009). It has also been shown that processing patterns vary by L2 proficiency level among upper intermediate and advanced L2 learners (Farina 2014, 2015; Liszka 2002). It is expected that proficiency will significantly affect the performance of the adult L2 English learners in the present investigation as well. The present investigation relies on proficiency-based grouping in order to investigate the effects of proficiency on language processing and metalinguistic knowledge. The L2 English user participants are divided into four groups (described below) determined by their performance on an independent measure of English proficiency (§4.3). Based on the previous research, there should be measurable improvements in performance as proficiency increases.

3.2.1 Low proficiency.

The lowest proficiency group of English learners in the present investigation is the Low L2 proficiency group (Low). Learners in this group have been exposed to the English simple past and have been instructed in its prescriptive use in their English language courses. These learners have also been exposed to the English present perfect in their environment but have not necessarily been instructed in its prescriptive use in their English language courses.

3.2.2 Lower intermediate proficiency.

The next most proficient group is the Lower Intermediate L2 proficiency group (Int.-Low). Learners in this group have been exposed to both the English simple past and the English present perfect and have been instructed in their prescriptive use in their English language courses.

3.2.3 Higher intermediate proficiency.

The second most proficient group is the Higher Intermediate L2 proficiency group (Int.-High). Learners in this group have been exposed to both the English simple past and the English present perfect, have been instructed in their prescriptive use in their English language courses, and have had the opportunity for practice with both forms.

3.2.4 Advanced proficiency.

The highest proficiency group is the Advanced L2 proficiency group (Adv.). These learners have been exposed to both the English simple past and the English present perfect, have been instructed in their prescriptive use in their English language courses, and have had the opportunity to practice both forms inside and outside of the classroom.

3.3 First language effects.

One factor that separates adult L2 learners from child L1 learners is that adults approach learning the L2 with a fully formed L1, with all of its feature distinctions fully developed in the linguistic system (Lardiere 2003). That is, adult L2 learners approach the task of learning a language at a different initial state than do child L2 learners. Components of the L1 are *transferred* into the developing L2 or interlanguage (Corder

1971). This is also true of the features that organize tense and aspect within the L1.[46] The present investigation is not able to fully incorporate the aspect systems of the participants' L1. Specifically, it does not account for the lexical aspect of the translational equivalent verbs in each L1. This failure may introduce some uncertainty into the results; however, it is not expected to confound the data because boundedness, unlike telicity, is a syntactic and semantic relation. The semantic portion (i.e. which verbs can be bounded) varies from language to language, but the syntactic portion (i.e. quantification of a direct object) is comparatively stable across languages, given the semantic requirements are satisfied. The present investigation does attempt to incorporate the tense-aspect systems of the participants' L1 grammar. It is through this grammatical system that adult L2 learners of English approach the tense-aspect system of English, which is the main focus of this investigation. The L1-influenced interlanguage grammar of an adult L2 learner affects L2 development both through the transfer of linguistic categories and features and through the transfer of entrenched processing patterns. Prior research indicates that L2 development is more inhibited by L1 transfer than it is facilitated by it; this tendency holds even when the L1 and L2 constructions are similar (Bybee 2008). This inhibition emerges most when learners associate an L2 construction with an L1 construction with which it has a 'misleading similarity' (Spada et al. 2005:201). It is expected that L1 transfer effects will measurably affect the results of the present investigation.

The potential effects caused by L1 influence are investigated through comparisons of three L2 English user groups divided by first language. The languages selected for

[46] For more information on L1 lexical aspectual propertied transferring into the L2, see Nishi and Shirai (2007).

further investigation are Arabic, Chinese, and Other. It is expected that differences in the L1 grammars will result in measurable differences between groups.

3.3.1 ARABIC.

The Arabic verbal system focuses on the perfective-imperfective distinction. That is, for past tense situations, the system can express aspectual meanings similar to those expressed via the (perfective) simple past and the (imperfective) past progressive grammatical tenses in English. The system does not express aspectual meanings that are directly comparable to the English present perfect. The present investigation is concerned primarily with the effects of the Arabic perfect and the past continuous grammatical tenses. The Arabic perfect describes a completed situation that is usually in the past (Mace 1998; Schulz 2004). Depending on context, the perfect corresponds with either the English simple past or bounded present perfect functions (Badawi et al. 2004). The following example demonstrates a sentence with a bounded predicate in the Arabic perfect that can be accurately translated into English via either the simple past or the present perfect.

(1)	rafaḍat	al	-	wizārah	al	-	khutah
	reject.PERF	the	-	ministry	the	-	plan

‘The ministry (has) rejected the plan.’

The Arabic past continuous describes an ongoing or iterative past situation. It is used to express the nonbounded perfects, and it is formed via periphrasis: *kāna* ‘to be.PERF’ + V.IMPF (Badawi et al. 2004). This example demonstrates a sentence with a nonbounded predicate that can be accurately translated into English via the simple past, the present perfect, or the past progressive.

(2) kāna yata‘ahhaduhuhā bi al - ri‘āyati ṭuwāla wujūdihā ma‘ahu

was.PERF promise.IMPF for the - care-take throughout existence with her

‘He (has) looked after her all of the time she was with him.’

There is no grammatical marker that indicates current relevance in Arabic; the morpheme with the function most similar to the feature is the actualization particle *qad* ‘actually/already/yet’ (or one of its variants; Badawi et al. 2004). The Arabic perfect and past continuous can express aspectual relations similar to the present perfect; however, they are more similar to the English simple past and past progressive than to the present perfect. Without the benefit of an aspectual contrast similar to the English simple past and present perfect in the L1, Arabic L1 learners of L2 English will not experience positive L1 transfer of tense-aspect expression, and they will need to acquire the distinction for the first time in the L2.

3.3.2 CHINESE.

The Mandarin Chinese aspectual system expresses aspectual meanings that are similar to those expressed via the English present perfect, but it does so in a manner that is not directly analogous formally and functionally. The present investigation is concerned primarily with the effects of the aspectual marker *le*, which marks current relevance, and is secondarily concerned with the perfective post-verbal particles -*le* and the experiential particle -*guo*. The marker *le* is a sentence-final particle that indicates ‘the state of affairs has special current relevance with respect to some particular situation’ (Li & Thompson 1981:240; Li et al. 1982). This marker is not functionally equivalent to the present perfect; it only expresses the [+CR] feature. Example 3 from Li & Thompson (1981) illustrates the difference in meaning between an unmarked sentence (3a) and one marked with *le* (3b).

(3)	a.	zhèi	-	ge	dìfang	hěn	ānjìng	
		this	-	CL	place	very	peaceful	
		'This place is very peaceful'						
	b.	zhèi	-	ge	dìfang	hěn	ānjìng	le
		this	-	CL	place	very	peaceful	CR
		'This place has become very peaceful' / 'This place is very peaceful now'						

The perfective post-verbal particle *-le* indicates that an occurrence is bounded or completed (Bayley 2013); combined with *le*, it generates interpretations nearly equivalent to a bounded present perfect. The experiential post-verbal particle *-guo* indicates that 'an event has been experienced with respect to some reference time' (Li & Thompson 1981:226). The unmarked interpretation is similar to an iterative nonbounded present perfect, which can be strengthened further through the use of *le* (Li & Thompson 1981). Chinese L1 speakers have the current relevance feature in their L1 grammar, and the L1 grammar has the capability to separate perfective and imperfective aspect in the L1 via the verb-final particle *-le*. That is, the L1 grammar has the features that are required to distinguish the English simple past from the present perfect, and it has the morphosemantics features to mark and semantically distinguish the bounded and nonbounded functions of the English present perfect.

3.3.3 Other.

The Other group is composed of speakers from varying L1 backgrounds. Their results are chiefly used as a L2 English user baseline against which the other two groups are measured.

3.4 Rationale for the study.

The present investigation seeks to address the lack of online research on the acquisition of the present perfect. Prior research into the acquisition of the present perfect has focused on production, comprehension, and use of the construction using offline measures. These measures capture the learner's application of memorized metalinguistic rules as much as they capture the state of the learner's interlanguage grammar. This research has hitherto failed to capture L2 users' interpretations of features used to compose the present perfect. Research that uses online procedures to examine the acquisition and processing of aspect have focused on the perfective-imperfective distinction (Anderson et al. 2008; Baggio et al. 2008; Ferretti et al. 2007; Madden & Zwaan 2003; Stutterheim et al. 2009), and it leaves the present perfect underrepresented in the data. That being said, offline research into the acquisition of the present perfect has demonstrated that manipulations in boundedness and adverbial modification affect how learners produce and comprehend the construction in context. More specifically, offline research has shown that the present perfect emerges from the simple past and that manipulations to boundedness and current relevance affect the nativelikeness of production and comprehension. Online research has demonstrated that manipulations in boundedness affect how native speakers process the aspectual character of the phrase and the aspectual character of the clause as a whole. Together, the results of prior investigations indicate that manipulations in adverbial modification are likely to produce valid and significant effects during real-time tense-aspect processing and that these effects are indicative of the developmental acquisition of interlanguage processing strategies (cf. VanPatten & Jegerski 2010). The present investigation investigates two features hypothesized to affect processing using online and offline measures: boundedness and current relevance.

Except for the pilot study of the procedures used herein (Farina 2014, 2015), this investigation is the first to use online and offline methods to investigate the processing and acquisition of the English present perfect by instructed, adult nonnative speakers of English. This investigation advances the work of prior research by evaluating its findings and implications using online procedures, which has been requested by researchers over the last decade (Felser 2005; Jegerski 2014; Leeser 2014; Marinis 2010; Papangeli 2010; Roberts & Liszka 2013). Moreover, researchers have specifically requested that online procedures be used to address the development of the interlanguage grammar and to justify competing paths of acquisition (Dussias & Piñar 2009; Jegerski 2014; Papangeli 2010; Roberts 2009). Processing data are crucial because they are indicative of how language is represented in the minds of native speakers and L2 learners alike (Slabakova 2008); these data are further validated using offline measures. This research uses online and offline data in an attempt to investigate the potential effects of L2 proficiency and L1 background on the processing strategies used by instructed adult L2 learners. The online processing task used in this investigation is a self-paced reading task. This task was selected because it minimizes the effects of the learner's metalinguistic knowledge and more accurately captures acquisition beyond rote memorization (Leeser 2014), and it has been used in a similar investigation by Roberts (2009) to study the influence of L1 patterns on L2 processing of tense-aspect agreement. The offline task is a Likert-style rating task that was selected to validate the online results.

The present investigation evaluates the respective abilities of two opposing accounts to describe how manipulations of central features affect how the present perfect is processed and understood by instructed adult L2 learners: i) the *complexity account* under which less semantic-syntactically complex structures are processed more readily than more complex ones, and ii) the *prototype account* under which more semantically central structures are processed more readily than peripheral structures. Offline data from prior research provide evidence for both accounts; however, no known research tests the validity of these accounts in a cross-sectional analysis across proficiency levels using both online and offline tasks. The present investigation addresses this gap in the literature. The online and offline data obtained during this investigation are used to analyze these two competing perspectives of acquisition while accounting for potential L2 proficiency and L1 transfer effects.

3.5 Research questions and hypotheses.

The following research questions guide the present investigation. Relevant hypotheses and brief justifications for these hypotheses follow the questions and subquestions.

The first research question concerns the effects of boundedness on tense-aspect processing and, by extension, the degree to which the processing of manipulations in boundedness follows a pattern predicted by either the complexity or prototype account.

RQ1: Do manipulations in boundedness affect how instructed adult L2 learners of English process past time constructions following either the complexity or prototype account?

It is hypothesized that nonbounded present perfects will be processed more quickly than bounded ones, which would support the predictions of the complexity account. This hypothesis stems from the results of several studies in L2 production of the present perfect that find that nonbounded functions of the present perfect are produced earlier, more often, and more accurately (Davydova 2011; Terán 2014; Uno 2014).

There are three subquestions for each research question. The first two questions have the following three corollaries concerning adult L2 learner processing:

(a) Is L2 learner processing qualitatively similar to that of adult L1 speakers?

(b) Is L2 learner processing affected by differences in L2 English proficiency?

(c) Is L2 learner processing affected by first language?

Concerning (a) qualitative processing similarities, it is hypothesized that the NS controls processing will be too rapid to distinguish between conditions and the L2 English users will respond to grammatical tense more than boundedness. These predictions are based on the results of the pilot of the present investigation, which examines both L1 and L2 English processing (Farina 2014, 2015), and based on prior research that suggests that acquiring the present perfect overall is more effortful than acquiring its functions (Bardovi-Harlig 2002; Housen 2002). Concerning (b) L2 proficiency effects, it is predicted that higher proficiency learners (Int.-High & Adv. groups) will perform in a more nativelike manner than lower proficiency learners (Low & Int.-Low). The prediction is based on several studies that find a positive correlation between L2 proficiency and nativelike performance (Bardovi-Harlig 2002; Davydova 2011; Dussias & Piñar 2009; Housen 2002). Finally, concerning (c) L1 effects, it is hypothesized that the Arabic L1 group will experience no transfer effects due to a lack of similar L1 structures; the Chinese L1 group will experience positive transfer caused by the functionally similar post-verbal *-le* particle. These hypotheses parallel the results in McCormick (2008): Chinese L1 users' performance is positively affected by the presence of a functionally similar current relevance marker and L1 users of a language without a formally or functionally equivalent structure to the present perfect exhibit only minimal benefits to performance.

The second research questions concerns the effects of adverbial marking of current relevance on tense-aspect processing and, by extension, the degree to which the processing of manipulations in current relevance follows a pattern predicted by either the complexity or prototype account.

RQ2: Do manipulations in adverbial modifiers that overtly mark current relevance affect how instructed adult L2 learners of English process past time constructions following either the complexity or prototype account?

It is hypothesized that [+CR] modifiers will facilitate processing of the present perfect and inhibit processing of the simple past. This result would indicate that manipulations in current relevance affect processing in a manner following the prototype account. This hypothesis emerges from L2 research that notes the importance of adverbial modification and context to the acquisition of the present perfect (Bardovi-Harlig 2002; Davydova 2011; Farina 2014, 2015; Terán 2014). That is, specific contexts and modifiers positively or negatively affect performance in present perfect contexts.

The corollaries to the second research question are the same as those for the first question, but they apply to current relevance per RQ2. Concerning (a) processing similarities, it is expected that the NS controls processing will again be too rapid to distinguish between conditions and that the L2 English users will respond to current relevance as it interacts with grammatical tense. The first prediction is again based on the pilot of this investigation, which finds that L1 processing is not measurably affected by similar manipulations (Farina 2014, 2015). The second prediction is rooted in the L2 acquisition research noted above in the prior paragraph (Bardovi-Harlig 2002; Terán 2014) as well as typological and corpus-based research that note strong interactions between adverbial modification and tense-aspect predication (Declerk 2006; Elsness 1997; Liu 1993; McCoard 1978). Concerning (b) L2 proficiency effects, it is hypothesized that the Intermediate (Int.-Low & Int.-High) and Advanced groups will respond to manipulations in CR, and their performance will become more nativelike as proficiency increases. This prediction is based on L2 research that finds that proficiency positively correlates with performance as well as interactions between current relevance and grammatical tense (Bardovi-Harlig 2002; Davydova 2011; Terán 2014). Lastly, concerning (c) L1 effects, it is predicted that the Arabic L1 group will again experience no transfer effects due to a lack of similar L1 structures and the Chinese L1 group will experience positive transfer caused by the functionally similar sentence-final *le* particle. The rationale for this prediction is roughly the same as that of the L1 effects discussed above; the difference being that, instead of a post-verbal perfective marker *-le*, the Chinese L1 users will likely benefit from the sentence-final current relevance marker *le* (McCormick 2008).

The third research question concerns the metalinguistic understanding of the manner in which boundedness interacts with these tense-aspect constructions.

RQ3: Are the grammars of instructed adult L2 learners of English sensitive to contrasts in boundedness?

It is hypothesized that the grammars of the L2 English users will be sensitive to contrasts in boundedness and they will also be sensitive to its interactions with grammatical tense. This hypothesized result indicates that the L2 users' interlanguage grammars are sensitive to the feature that delimits the functions of the present perfect into those that describe a completed situation and those that describe an ongoing or iterative one. This hypothesis is generated based on the work of Slabakova who finds that L2 learners are able to acquire compositional aspect and to understand its interactions with tense-aspect (Slabakova 2008, 2003, 2000; Slabakova & Montrul 2002). It is further supported by research that finds effects for boundedness in acquisition of the present perfect specifically (Bardovi-Harlig 2002; Davydova 2011; Terán 2014; Uno 2014)

The third and fourth research questions have three subquestions that differ slightly from those that apply to the first two questions. The second two research questions have the following three corollaries concerning adult L2 metalinguistic judgments:

(a) Are L2 learner metalinguistic judgments qualitatively similar to those of adult L1 speakers?

(b) Are L2 learner metalinguistic judgments affected by differences in L2 English proficiency?

(c) Are L2 learner metalinguistic judgments affected by first language?

Concerning (a) qualitative similarities of metalinguistic judgments, it is hypothesized that the metalinguistic judgments of the L2 English users will be qualitatively similar to those of the NS controls once they understand the semantics of boundedness. This hypothesis is based on the work of Wulff and colleagues (2009) who find that aspect ratings by native speakers correlate with patterns of use among L2 learners; further, it is also supported in the work of Slabakova (2000), who finds that nativelike aspect judgments are attainable. Concerning (b) L2 proficiency effects, it is predicted that the Intermediate (Int.-Low & Int.-High) and Advanced groups will be sensitive to contrasts in boundedness, but only the Int.-High and Adv. groups will exhibit nativelike ratings. Effects for proficiency are well attested in the literature, but this hypothesis is specifically generated by research by Bardovi-Harlig (2002) and Slabakova (2000) who find meaningful differences among the interlanguage grammars of L2 learners of varying proficiency. Concerning (c) L1 effects, it is hypothesized that the Arabic L1 group will be sensitive to contrasts in boundedness, but they will not exhibit nativelike ratings, and that the Chinese L1 group will exhibit nativelike ratings (or at least ratings beyond their expected average proficiency) due to the functionally similar post-verbal *-le* particle. These hypotheses are based on McCormick (2008) and Slabakova (2000); both of these studies find variation in performance that is mediated by the presence or absence of functionally and formally similar structures in the first language.

The fourth research question concerns the metalinguistic understanding of current relevance in these past time contexts.

RQ4: Are the grammars of instructed adult L2 learners of English sensitive to contrasts in current relevance marked via i) adverbial modifiers and ii) verbal morphosyntax?

It is hypothesized that the grammars of the L2 English users will be sensitive to contrasts in current relevance marked in both manners. Such results would indicate that L2 learners use the [+CR] feature to help differentiate the present perfect from the simple past. This hypothesis is rooted in works that find effects for adverbial, morphosyntactic, and contextual current relevance markers on the acquisition and use of the present perfect (Davydova 2011; Terán 2014; Uno 2014).

The corollaries to the fourth research question are the same as those for the third question, but they apply to current relevance per RQ4. Concerning (a) qualitative similarities of metalinguistic judgments, it is expected that the metalinguistic judgments of the L2 English users will be qualitatively similar to those of the NS controls i) in the adverbially marked conditions and ii) in the verbal morphologically marked conditions. These results are expected because many of these learners have been instructed within their English program on the semantics of the present perfect, which has been shown to improve performance to a degree (Bardovi-Harlig 2002; McCormick 2008). Furthermore, L2 English learners have been observed marking current relevance in a nativelike manner (Davydova 2011). Concerning (b) L2 proficiency effects, it is hypothesized that i) the Intermediate (Int.-Low & Int.-High) and Advanced groups will be sensitive to contrasts in current relevance marked using adverbial modifiers and that ii) only the higher proficiency learners (Int.-High & Adv. groups) will be sensitive to contrasts in current relevance marked using verbal morphology. Both of these hypotheses stem from the accuracy of production results in Bardovi-Harlig's (2002) study wherein she observed that performance dramatically improves among higher proficiency learners. Lastly, concerning (c) L1 effects, it is predicted that i) the Arabic L1 group will be sensitive to contrasts in current relevance marked adverbially morphology due to the similar use of extra-grammatical adverbial markers in the L1, but ii) they will not be sensitive to contrasts marked via verbal morphology due to the lack of a grammatical feature that marks current relevance in the L1. Likewise, it is also predicted that i) the Chinese L1 group will not be sensitive to contrasts in current relevance marked adverbially due to negative L1 transfer of the post-sentence particle *le*, ii) but they will be sensitive to contrasts marked via verbal morphology due to positive transfer of the same. Both of these predictions are extrapolations of McCormick's (2008) findings concerning the effects of formal a functional similarities of L1 structures on L2 performance; however, the second prediction is strongly influenced by the expectation that the function of the singular Chinese current relevance particle may not be successfully transferred to the dozens of forms that indicate current relevance in English (Bybee 2008; Spada et al. 2005).

CHAPTER 4 METHODOLOGY

In order to address the research questions and hypotheses stated at the end of the previous chapter (§3.5), the present investigation utilizes three types of tasks: two reading tasks (administered separately), a background questionnaire, and an independent measure of proficiency (IMP). After being briefed on the investigation (Appendix A), the instructed adult L2 learners of English, who are divided by English proficiency and by first language, perform the three tasks. The IMP is an offline cloze-style passage that gathers English proficiency information from both L1 and L2 users of English (§4.3). The questionnaire is an offline survey that collects basic personal and language history information (§4.4). The reading tasks measure real-time language processing in the first administration and offline metalinguistic knowledge in the second administration (§4.5). The first reading task sheds light on the factors that inhibit and facilitate processing of the English present perfect. Specifically, this task compares and contrasts the effects of boundedness and current relevance on language processing in two tense-aspect constructions: the simple past and the present perfect. The second task investigates L2 English users' intuitions concerning the meaning of boundedness and current relevance in the same tense-aspect contexts.

4.1 PARTICIPANTS.

The participants for this study comprise two groups: 155 instructed, adult L2 speakers of English and 72 adult L1 English-speaking controls. Data were collected from the L1 English group to serve as a baseline for comparison with the data from the L2

English group *in toto*, by English proficiency (defined by proficiency score on IMP: low, lower intermediate, upper intermediate & advanced), and by subgroup (defined by L1).[47] English native-speaking controls were undergraduate students at the University of South Carolina, recruited through introductory-level linguistics courses. Non-native speaker participants were students at the University of South Carolina's affiliated intensive English Program, English Programs for Internationals (EPI), or students at USC's International Accelerator Program (IAP). Students at EPI were approached through their reading and vocabulary course. Students at IAP were approached through their linguistics for nonnative speakers course. Instructed L2 English learners were selected due to the difficulty of the target structure; uninstructed learners rarely become proficient users of the present perfect (Davydova 2011). In the first administration using the SPRT, adult L2 English learners have the following native languages: Arabic (11), Chinese (42), French (2), Hindi (2), Japanese (5), Korean (2), Spanish (3), and Ukrainian (1). In the second administration using the rating task, they have the following native languages: Arabic (11), Chinese (44), French (2), Hindi (2), Japanese (10), Korean (4), Spanish (3), Thai (1), and Vietnamese (1).

4.2 MATERIALS AND PROCEDURES.

The investigation consists of three procedures: i) completion of an independent measure of English proficiency, ii) collection of a background questionnaire, and iii) administration of two reading tasks. Participants complete each of the three tasks on two separate occasions. These two administrations are separated by approximately ten weeks,

[47] These group variables were selected in order to answer the specific research questions of the present investigation. See Chan and colleagues (2012) for a detailed discussion of other individual differences among language learners.

which should prevent any priming effects but may not prevent minor practice effects among those participants who complete the tasks twice. The same proficiency measure and questionnaire are used in both administrations. The reading task differs by administration. During each administration, the participants complete the reading task first. They complete the questionnaire and the proficiency task separately to prevent any priming effects and to inhibit fatigue. All tasks are completed on an internet-connected computer in a classroom laboratory setting.

4.3 Independent measure of proficiency.

The offline independent measure of proficiency (IMP) gathers comparable proficiency data for both L1 and L2 English speakers (Appendix B). The instrument is a cloze-style passage from Slabakova (2000). This proficiency measure was selected for three reasons: i) it is a consistent and reliable measure, ii) it has an attested normal distribution of responses, and iii) it can be administered in a comparatively short time. All of the vocabulary items in the text are frequently occurring and known to the L2 English users as determined by pilot testing and post-testing interviews. After the first sentence, which is shown uninterrupted, there is a blank every seventh word where a specific word was removed from the story. There are 40 total blanks. Participants are evaluated following an all-or-nothing method of coding, in which they are given 1 point for writing in the exact same word that was removed (with correct inflections when appropriate) and 0 points for any other response, following Slabakova (2000). The highest possible score is 40 points. The results of this task are reported below in sections 5.1.1 and 5.2.1.

4.4 Background questionnaire.

The background questionnaire is an adapted version of the Bilingual Language Profile (BLP; Appendix C; Birdsong et al. 2012). The BLP is a questionnaire designed to assess language dominance in bilingual adults through self-report data. The BLP captures personal history and language use data that fall into five categories: i) biographical information, ii) language history, iii) language use, iv) language proficiency, and v) language attitudes. The adapted version of the BLP used in the present investigation captures additional biographical information that may be relevant to performance on the SPRT and language history information that accommodates a population that is not exclusively bilingual. The information obtained from this questionnaire is used to sort the participants into groups and L1 subgroups, to ensure that these groups are comparable, and to investigate meaningful correlations regarding performance. The results of this task are reported below in sections 5.1.2 and 5.2.2.

4.5 Reading tasks.

The two reading tasks chiefly address the research questions. They are administered through the web-based experiment manager Ibex (spellout.net/ibexfarm). The first administration of this research uses a self-paced reading task (SPRT) with a comprehension question where focus is placed on the reading times; the second administration relies on a SPRT with a rating task where focus is placed on the ratings. The reading component of each task is an online measure of linguistic ability that investigates language processing in real time, and it is assumed to minimize the effects of metalinguistic knowledge, which affects the validity of offline measures (Leeser 2014). The task that occurs after the reading portion (either comprehension or rating task) is an offline measure of metalinguistic knowledge. SPRT is considered the fundamental experimental measure used to study sentence-level processing (Jegerski 2014; McDonough & Trofimovich 2012). Recall from section 2.2 that tense-aspect is sentential in scope (Comrie 1976; Smith 1997); therefore, the sentence is the minimal experimental domain to measure tense-aspect processing. SPRT is most often used to compare processing facilitation and inhibition patterns caused by ambiguity, anomaly, or distance dependency (Jegerski 2014), and it is a tool that is commonly used to compare native L1 and nonnative L2 processing (Jackson 2010; Jegerski 2014; Koda 2012; Papangeli 2010).

Using SPRT to investigate tense-aspect processing is not without precedent. Roberts (2009) and Roberts and Liszka (2013) successfully use SPRT to investigate the influence of L1 inflectional patterns on L2 processing and comprehension of tense-aspect agreement and adverbial modification in the simple past and present perfect. Further, Farina (2014, 2015) successfully uses SPRT to investigate the effects of lexical and compositional aspect on tense-aspect processing in the simple past and present perfect. More generally, much of the research in L2 processing concerns whether speakers use parsing strategies particular to their L1, L2, or neither language; the learnability of linguistic features; and differences between competence and performance (Jegerski 2014; Koda 2012; McDonough & Trofimovich 2012). Processing studies have hitherto not been used to justify paths of acquisition as has been called for in the literature (Dussias & Piñar 2009; Jegerski 2014; Papangeli 2010; Roberts 2009). This research uses processing data to investigate any effects of L2 proficiency and L1 background on the acquisition of the present perfect.

The SPRT used for this investigation follows the non-cumulative, moving window paradigm. This paradigm was selected following the recommendations of Jegerski (2014) and McDonough and Trofimovich (2012) who suggest that this method of presentation offers the ability to capture very precise data with minimal task effects. First, the participant sees a series of dashes that match the length of each word in the sentence. When the participant taps the spacebar, the first word appears and replaces the first set of dashes. As the participant continues to tap the spacebar, words are presented linearly. When each word appears, the one previous to it is replaced with dashes. Participants are shown a simple comprehension question following each sentence (Jegerski 2014). This comprehension question is a distractor task with a two-fold purpose: i) it ensures that the participant understands the content of the sentence and ii) it directs the attention of the participants away from the variables being manipulated (Jegerski 2014). The question and answer choices appear at the same time after the participant has completed reading the sentence. Figure 4.1 approximates the presentation of this paradigm.

```
--- --- --- - ----.
The --- --- - ----.                Who ran a race?
--- man --- - ----.                (1) The man
--- --- ran - ----.                (2) The boy
--- --- --- a ----.
--- --- --- - race.
```

FIGURE 4.1. Sentence presentation in non-cumulative, moving window SPRT.

Instead of focusing on the entire sentence presented, researchers usually focus on a region of interest. Per the discussion of boundedness in section 2.1.2, this research captures a five-word region of interest. [Table 4.1 & Table 4.2 below include sentences with marked regions of interest.] The task captures the reading time (RT) variable, which is operationalized as the time each word is displayed on the screen. To allow for valid between-condition comparisons, each critical word is balanced for syllable length, lexical content, syntactic category, token frequency, and region in the sentence (Jegerski 2014; McDonough & Trofimovich 2012). The region of interest is always at least five words from the end of the sentence so as to avoid the effects of whole-sentence meaning calculation, which result in inconsistent RTs (McDonough & Trofimovich 2012). The SPRT's comprehension questions evaluate whole-sentence understanding. Only RTs from sentences with correctly answered comprehension questions are analyzed in the results. This procedure is valid following the assumption that, if the participants do not understand the basic meaning of the sentence, it is very unlikely that their processing will be meaningfully affected by the experimental manipulations (McDonough & Trofimovich 2012). The ratings used in the rating task evaluate metalinguistic knowledge; all ratings are considered in the analysis.

The character and focus of the reading task differs by administration. The SPRT captures online processing data and focuses on a participant's reading times in a region of interest. The rating task captures offline judgments and focuses on metalinguistic ratings. The composition and organization of the sentences that comprise each reading task are as follows.

In the SPRT, participants read 104 sentences, and they read 56 sentences in the rating task (see below for a full description of the conditions and examples). After being presented with instructions, participants practice the task for eight (8) sentences. Once they understand the procedure, they each read the task sentences. Half of the 96 manipulated sentences (48) are experimental items and half (48) are distractors. Half (24) of the experimental items test the effects of boundedness on tense-aspect processing, and half (24) of them test the effects of current relevance on tense-aspect processing. A full description of these conditions is found below in section 4.5.1.

In the rating task, two thirds of the 48 manipulated sentences (32) are experimental items and one third (16) are distractors (see below). Half (16) of the experimental items evaluate the participants' understanding of boundedness in these contexts, and half (16) evaluate their understanding of current relevance. A full description of these conditions is found below in section 4.5.2.

The distractors and fillers are the same for both administrations of the reading tasks. The distractors, which are of comparable length and are superficially similar to the experimental items, are also subdivided but consistent across administrations. Following Jegerski (2014), half (24 & 8) involve a manipulated non-experimental variable, and half (24 & 8) are non-manipulated fillers. The manipulated variable in these distractors is the definiteness of the NP subject. Definiteness of the NP subject was selected as the distractor manipulation for four reasons: i) the manipulation is straight-forward (definite *the* & indefinite *a(n)*) and discrete; ii) the manipulation does not affect the licitness of the sentence, which may affect processing strategies used throughout the task; iii) the manipulation is outside of the target and manipulated regions of the other conditions; and iv) the new region of interest is already balanced according to the characteristics mentioned in the previous paragraph.

Presentation is counter-balanced and pseudorandom to ensure that each item is only shown in one condition and that sentences in the same condition are not displayed consecutively. This method of presentation is recommended to prevent participants from becoming conscious of the experimental manipulations, which may cause them to process the sentences differently and confound the data (Jegerski 2014).

The manipulated items are presented following 2x2 Latin square designs wherein the SPRT tests the effects of each boundedness and current relevance separately. That is, the SPRT manipulates boundedness and grammatical tense separately from its manipulation of overtly marked current relevance and grammatical tense, and the rating task quantifies intuitions concerning boundedness and overtly marked current relevance. The rating task queries relevant metalinguistic knowledge concerning the how the participants understand boundedness and current relevance.

4.5.1 Self-paced reading task.

The self-paced reading task investigates the effects on tense-aspect processing of two variables: boundedness and overtly marked current relevance (Appendix D). Boundedness is manipulated via the presence or absence of quantization of the direct object of the verb (e.g. *I shot* [$_{\text{DP}}$ Ø [$_{\text{NP}}$ *sheriffs*]] [–BND] v. *I shot* [$_{\text{DP}}$ *the* [$_{\text{NP}}$ *sheriffs*]] [+BND]). Overtly marked current relevance is manipulated using adverbial modifiers in first position that are [+/-current relevance] and [-anterior]. The following two subsections describe in more detail the specific manipulations used in the SPRT.

Boundedness, SPRT. Conditions A through D in the SPRT investigate the effects of manipulations in boundedness on the reading times of the participants. These conditions address the first research question.

RQ1: Do manipulations in boundedness affect how instructed adult L2 learners of English process past time constructions?

The effects of boundedness on tense-aspect processing are examined using 24 sentences designed to determine speaker sensitivity to the interactions between boundedness and tense-aspect (Appendices D.4 & D.5). As mentioned previously, boundedness is manipulated through the quantification of the direct object of an activity predicate; such manipulations generate both bounded and nonbounded predicates for comparison. Table 4.1 schematizes the conditions; verbs and critical words are underlined and annotated with numbers that indicate the location within the region of interest.

TABLE 4.1. Boundedness x grammatical tense conditions – SPRT.

Cdn.	Boundedness	Grammatical tense
A	+BND – *her theory*	Present Perfect – *has tested*
B	+BND – *her theory*	Simple Past – *tested*
C	–BND – *theories*	Present Perfect – *has tested*
D	–BND – *theories*	Simple Past – *tested*

A. Deliberately, the researcher has <u>tested</u>$_V$ her <u>theory</u>$_1$ <u>on</u>$_2$ <u>the</u>$_3$ <u>circus</u>$_4$ <u>monkeys</u>$_5$ who had to identify colors.

B. Deliberately, the researcher <u>tested</u>$_V$ her <u>theory</u>$_1$ <u>on</u>$_2$ <u>the</u>$_3$ <u>circus</u>$_4$ <u>monkeys</u>$_5$ who had to identify colors.

C. Deliberately, the researcher has <u>tested</u>$_V$ <u>theories</u>$_1$ <u>on</u>$_2$ <u>the</u>$_3$ <u>circus</u>$_4$ <u>monkeys</u>$_5$ who had to identify colors.

D. Deliberately, the researcher <u>tested</u>$_V$ <u>theories</u>$_1$ <u>on</u>$_2$ <u>the</u>$_3$ <u>circus</u>$_4$ <u>monkeys</u>$_5$ who had to identify colors.

The effects of boundedness on tense-aspect processing are expected to emerge in two ways. First, these effects will emerge as faster RTs for bounded events in the present perfect (A) than in the simple past (B) because bounded events should facilitate processing of the present perfect through either semantic-syntactic bootstrapping (complexity account) or shared semantic features (prototype account). Second, they will emerge as faster RTs for nonbounded activities in the simple past (D) than in the present perfect (C) because nonbounded predicates should inhibit processing of the present perfect by clashing against processing trends based on bootstrapping or shared features. These trends will be more apparent in the L2 English user groups than the NS one (Farina 2014, 2015).

CURRENT RELEVANCE, SPRT. Conditions E through H in the SPRT investigate the effects of manipulations in overtly marked current relevance on the reading times of the participants. These conditions address the second research question.

RQ2: Do manipulations in adverbial modifiers that overtly mark current relevance affect how instructed adult L2 learners of English process past time constructions?

The effects of overtly marked current relevance on tense-aspect processing are examined using 24 sentences designed to determine speaker sensitivity to the interactions between current relevance and tense-aspect (Appendices D.6 & D.7). These sentences rely on motion predicates so as not to also manipulate boundedness through quantization while still having an endpoint; motion predicates with a specified goal and bounded predicates are processed similarly among native speakers (Hodgson 2010; Wagner 2010), which permits their indirect comparison. Current relevance is manipulated using nine adverbial modifiers that are [-anterior] (i.e. they are licit with both grammatical tenses). Five of these modifiers are [+current relevance] (*since* V-*ing, for* [x time], *at present, often,* & *sometimes*), and four of these modifiers are [-current relevance] (*after* V-*ing, in* [x time], *at some point, until* [x event]). Table 4.2 schematizes the conditions; the same annotation procedures apply as in Table 4.1. Within these conditions, there are four subconditions within which adverbial markers indicate current relevance in different manners. Table 4.3 schematizes these subconditions.

TABLE 4.2. Current relevance x grammatical tense conditions – SPRT.

Cdn.	Current relevance	Grammatical tense
E	+CR – *for two minutes*	Present Perfect – *has chased*
F	+CR – *for two minutes*	Simple Past – *chased*
G	–CR – *in two minutes*	Present Perfect – *has chased*
H	–CR – *in two minutes*	Simple Past – *chased*

E. For two minutes, the dog has chased$_V$ the foxes$_1$ to$_2$ their$_3$ nearby$_4$ burrow$_5$ while barking to alert the hunters.

F. For two minutes, the dog chased$_V$ the foxes$_1$ to$_2$ their$_3$ nearby$_4$ burrow$_5$ while barking to alert the hunters.

G. In two minutes, the dog has chased$_V$ the foxes$_1$ to$_2$ their$_3$ nearby$_4$ burrow$_5$ while barking to alert the hunters.

H. In two minutes, the dog chased$_V$ the foxes$_1$ to$_2$ their$_3$ nearby$_4$ burrow$_5$ while barking to alert the hunters.

TABLE 4.3. Adverbially marked current relevance subconditions – SPRT.

	Current relevance		
Sub.	[+CR]	[–CR]	Category
1	*since* V-*ing*	*after* V-*ing*	Durative vs framing phrases
2	*for* [x time]	*in* [x time]	Duration vs frame AdvPs
3	*at present*	*at some point*	(Pre-)present temporal reference PPs
4	*often* & *sometimes*	*until* [x event]	Iterative vs noniterative phrases

The effects of current relevance on tense-aspect processing are expected to emerge in two ways. First, these effects will emerge as faster RTs for [+CR] clauses in the present perfect (E) than in the simple past (F) because [+CR] adverbials should facilitate processing of the present perfect through either bootstrapping or shared features.

Second, they will emerge as faster RTs for [–CR] clauses in the simple past (H) than in the present perfect (G) because [–CR] adverbials should inhibit processing of the present perfect by clashing against processing trends based on bootstrapping or shared features. These trends will again be more apparent in the L2 English user groups. The prototype account predicts that these trends emerge at lower proficiencies, and the complexity account predicts they emerge at higher proficiencies.

4.5.2 RATING TASK.

The rating task investigates how the participants interpret the meanings of boundedness and current relevance in past time contexts using a rating task that follows each sentence (Appendix E). This task differs from the SPRT chiefly in that the participants are asked to rate their agreement for a statement that describes the content of a phrase using a six-point Likert-style scale (1-6). A six-point scale was chosen in order to remove the neutral option that is present in odd-numbered scales (e.g. *Neither agree nor disagree*). When present, the neutral option has been shown to shift participant answers away from the neighboring options on the scale (Bishop 1987; Kalton et al. 1980). Likewise, Carter and colleagues (2012) observe significant cultural differences in the use of the middle option that would likely affect a cross-cultural investigation like the present one. The scale used in the present investigation attempts to prevent these effects by forcing all participants to indicate even slight preferences in judgment. The rating portion of the task occurs after the participant reads the sentence one word at a time in the same manner as in the SPRT. In the sentences, boundedness is manipulated in the same manner as in the SPRT; overtly marked current relevance is manipulated in two ways: i) as in the SPRT, using adverbial modification in first position that are [+/–CR] and [-

anterior], and ii) using verbal morphology (simple past [–CR] & present perfect [+CR]). The following two subsections describe in more detail the specific manipulations used in the rating task.

BOUNDEDNESS, RATING TASK. Conditions A through D in the rating task investigate the effects of manipulations in boundedness on the continuability ratings of the participants. These conditions address the third research question.

RQ3: Are the grammars of instructed adult L2 learners of English sensitive to contrasts in boundedness?

How L2 English users understand the interactions of boundedness and grammatical tense is examined through the rating of 16 sentences for the perceived continuability of the tensed verb phrases. The sentences and phrases to be rated are taken from the pool of sentences used in the conditions that investigate boundedness in the SPRT (Appendix E.4).

The statements that the participants rate in the boundedness conditions are designed to evaluate knowledge of contextually assigned endpoints in the past. To make the task easier for the participants to understand, participants are asked to rate continuability (i.e. whether or not an action can continue or be restarted) rather than boundedness itself. Continuability is selected as the opposite of boundedness following definitions of Vendler (1967) and Verkuyl (1989). It is operationalized using a scalar representation of the binary continuable-noncontinuable distinction used in Shirai (2013) following the methodology of Wulff and colleagues (2009). That is, continuability ratings are made for particular phrases from the sentences on a scale from 1 to 6. A score of ‘1’ indicates that the participant strongly disagrees that the action of the verb can continue; a score of ‘6’ indicates that the participant strongly agrees that the action can continue. Instructions and examples shown to the participants can be found in Appendices E.2 and E.3. Since continuability is treated in opposition to boundedness, a lower continuability rating indicates that the participant views the phrase as more bounded than a higher rating. Table 4.4 schematizes the conditions; tensed verb phrases are underlined. The underlined portions of these sentences are used in the construction of the rating statements, wherein they appear in bold.

TABLE 4.4. Boundedness x grammatical tense – Rating task.

Cdn.	Boundedness	Grammatical tense
A	+BND – *her theory*	Present Perfect – *has tested*
B	+BND – *her theory*	Simple Past – *tested*
C	–BND – *theories*	Present Perfect – *has tested*
D	–BND – *theories*	Simple Past – *tested*

A. Deliberately, the researcher <u>has tested her theory</u> on the circus monkeys who had to identify colors.

Rating: When the action **has tested her theory** is finished, it can be continued.

Strongly Disagree 1 2 3 4 5 6 Strongly Agree

B. Deliberately, the researcher <u>tested her theory</u> on the circus monkeys who had to identify colors.

Rating: When the action **tested her theory** is finished, it can be continued.

Strongly Disagree 1 2 3 4 5 6 Strongly Agree

C. Deliberately, the researcher <u>has tested theories</u> on the circus monkeys who had to identify colors.

Rating: When the action **has tested theories** is finished, it can be continued.

Strongly Disagree 1 2 3 4 5 6 Strongly Agree

D. Deliberately, the researcher <u>tested theories</u> on the circus monkeys who had to identify colors.

Rating: When the action **tested theories** is finished, it can be continued.

Strongly Disagree 1 2 3 4 5 6 Strongly Agree

The participants' understandings of boundedness and its interaction with tense-aspect are expected to emerge in two ways. First, participants will rate phrases in the nonbounded conditions (C & D) more continuable than those in the bounded conditions (A & B). This result is expected because bounded events should, by definition, be considered less continuable than nonbounded activities. Second this difference in ratings between bounded and nonbounded conditions will be more exaggerated in the present perfect than in the simple past. That is, the differences in continuability ratings between Conditions A and C (present perfect) will be greater than the differences between Conditions B and D. This is because boundedness has a more salient function in the present perfect (distinguishing between functions) than in the simple past. These differences will be most observable in the NS group, but will also be observable among higher proficiency L2 English user groups (Farina 2014, 2015).

CURRENT RELEVANCE, RATING TASK. Conditions E through H in the rating task investigate the effects of manipulations of adverbial modification and of tense-aspect construction on the current relevance ratings of the participants. These conditions address the fourth research question.

RQ4: Are the grammars of instructed adult L2 learners of English sensitive to contrasts in current relevance marked i) adverbially or ii) using verbal morphology?

How L2 English users understand two instantiations of current relevance is examined through the rating of 16 sentences, 8 of which mark current relevance via first-position adverbial phrases and 8 of which do so via verbal morphology, The sentences for Conditions E and F are adapted from the pool of sentences used in the conditions that investigate current relevance in the SPRT, and the sentences for Conditions G and H are adapted from the pool of sentences used in the conditions that investigate boundedness in the SPRT (Appendix E.6).

The statements that the participants rate concerning current relevance are designed to evaluate knowledge of current relevance in two instantiations: adverbial phrases and verbal morphology. The participants rate the adverbial phrases or tensed verb phrases for their perceived ability to indicate the contemporary importance of the action of the verb or of the verb's resultant state. As is the case for the other rating conditions, current relevance ratings are made for particular phrases from the sentences on a scale from 1 to 6. A score of '1' indicates that the participant strongly disagrees that the action or resultant state of the verb is more important to the present than to the past; a score of '6' indicates that the participant strongly agrees that the action or resultant state of the verb is more important to the present than to the past. Instructions and examples shown to the participants can again be found in Appendices E.2 and E.3. Table 4.5 schematizes the two conditions where current relevance is marked adverbially; adverbial modifiers are underlined. The same subconditions apply in this task as apply in the SPRT; it is presented again here as Table 4.6. Table 4.7 schematizes the conditions for two conditions where current relevance is marked via verbal morphology (i.e. grammatical tense); tensed verb phrases are underlined. The underlined portions of these sentences are used in the construction of the rating statements, wherein they appear in bold.

TABLE 4.5. Adverbially marked current relevance conditions – Rating task.

Cdn.	Current relevance	Grammatical tense
E	+CR – *at present*	Present Perfect – *has hiked*
F	–CR – *at some point*	Present Perfect – *has hiked*

E. At present, Robert has hiked with Emma to the mountain summit in order to see the beautiful view.

Rating: The phrase **at present** indicates that the action or its consequences are still relevant.

Strongly Disagree 1 2 3 4 5 6 Strongly Agree

F. At some point, Robert has hiked with Emma to the mountain summit in order to see the beautiful view.

Rating: The phrase **at some point** indicates that the action or its consequences are still relevant.

Strongly Disagree 1 2 3 4 5 6 Strongly Agree

TABLE 4.6. Adverbially marked current relevance subconditions – Rating task.

	Current relevance		
Sub.	[+CR]	[–CR]	Category
1	*since* V-*ing*	*after* V-*ing*	Durative vs framing phrases
2	*for* [x time]	*in* [x time]	Duration vs frame AdvPs
3	*at present*	*at some point*	(Pre-)present temporal reference PPs
4	*often* & *sometimes*	*until* [x event]	Iterative vs noniterative phrases

TABLE 4.7. Morphologically marked current relevance conditions – Rating task.

Cdn.	Grammatical tense	Boundedness
G	Present Perfect [+CR] – *has painted*	+BND – *a picture*
H	Simple Past [–CR] – *painted*	+BND – *a picture*

G. Cheerfully, the artist has painted a picture of the splendid mountains when the sun was rising because it was so beautiful.

Rating: The fact that **a picture** has been **painted** is relevant at the present time.

Strongly Disagree 1 2 3 4 5 6 Strongly Agree

H. Cheerfully, the artist painted a picture of the splendid mountains when the sun was rising because it was so beautiful.

Rating: The fact that **a picture** was **painted** is relevant at the present time.

Strongly Disagree 1 2 3 4 5 6 Strongly Agree

The participants' understanding of current relevance as marked using adverbial phrases and tensed verb phrases are expected to emerge in two ways. First, in the adverbial marking conditions, participants will rate [+CR] markers in Condition E as more indicative of current relevance than [–CR] markers in Condition F. This result is expected because of the aspectual character of these adverbial modifiers and because of the collocation patterns observed for the [+CR] markers in the present perfect and the [–CR] markers in the simple past. Second, in the morphological marking conditions, participants will rate the present perfect verb phrases in Condition G as more indicative of current relevance than the simple past verb phrases in Condition H. This result is expected by definition under the assumption that the grammars of the participants associate the present perfect with the [+CR] feature. Both of these differences should be observable in the NS group and among the L2 English user groups who are acquiring/have acquired the features of the present perfect.

Chapter 5 Results

The following sections describe the results of the tasks utilized by the present investigation. They are divided by reading task for each administration: the SPRT from the first administration and the rating task from the second administration. Three tasks are shown in each section: the independent measure of proficiency, the modified Bilingual Language Profile, and the appropriate reading task. The results of each task are summarized in three ways: divided by L1 and L2 English language use, by English proficiency score, and by first language.

Per Larson-Hall's (2010:100-3) recommendation, the present investigation considers statistical significance to be found when the p-value is less than .10. She argues that a p-value of .10 decreases the probability of Type II errors (false negative), which are usually between 20% and 33% due to the limitations of statistical power in social science research while still balancing the need to prevent Type I errors (false positives).[48]

5.1 Self-paced reading task – First administration

The SPRT is the principle task of the first administration. Conditions A through D investigate the L2 users' online processing of boundedness in context through individual reading times. Conditions E through H investigate the L2 users' online processing of current relevance in context through individual reading times as well. Below are the summary results of the three tasks used in this administration.

[48] See Kline (2004) as well as Murphy and Myors (2004) for additional arguments concerning raising the p-value threshold to .10.

5.1.1 INDEPENDENT MEASURE OF PROFICIENCY, SPRT.

Scores were assigned using an all-or-nothing coding method. That is, the word inputted in the blank had to be the same word that was originally removed. Points were awarded for misspelled words so long as the intended entry was apparent. The highest possible score for the task is 40. The results of this task are presented in the following two tables. Relevant inferential statistics for group difference (ANOVAs and posthoc tests) are presented after each table.

TABLE 5.1. Descriptive statistics for proficiency score for groups by L2 English proficiency.

	n	$\bar{x}$	sd
Adv.	11	16.091	3.534
Int.-High	11	11.000	0.632
Int.-Low	11	8.000	0.775
Low	11	3.182	2.442
All L2 Users	68	9.567	5.196
English NS	39	20.462	4.129

An ANOVA revealed that there is a statistically significant difference for proficiency scores between the L2 user and English NS groups ($F_{1,81} = 109.88$, $p < .001$). An additional ANOVA revealed that there are statistically significant differences for proficiency scores among the proficiency groups ($F_{4,78} = 79.17$, $p < .001$). A Tukey-Kramer posthoc test revealed significant differences pairwise between each group except between the two Intermediate groups at an alpha of .05.

TABLE 5.2. Descriptive statistics for proficiency score for groups by L1.

	n	$\bar{x}$	sd
Arabic	7	7.857	4.375
Chinese	24	7.958	4.248
Other	13	13.462	5.410
English NS	39	20.462	4.129

An ANOVA revealed that there are statistically significant differences for proficiency scores among the L1 groups ($F_{3,79} = 47.08$, $p < .001$). A Tukey-Kramer posthoc test revealed significant differences pairwise between each group except between the Arabic and Chinese groups at an alpha of .05. The descriptive statistics indicate that the Other group performs more accurately than the Arabic and Chinese groups.

5.1.2 MODIFIED BILINGUAL LANGUAGE PROFILE QUESTIONNAIRE, SPRT.

This modified version of the Bilingual Language Profile (BLP) contains responses divided into five categories: i) biographical information, ii) language history, iii) language use, iv) language proficiency, and v) language attitudes. The items in these five categories capture participant data that have previously been shown to affect L2 acquisition (Birdsong et al. 2012). Any signification differences between groups captured by this task introduce the potential that a resultant difference in the reading task may be due to this confounding difference and not the group difference. This section departs from the format used elsewhere in the chapter in an effort to focus the reader on specific significant differences captured by the BLP that have the potential to confound the main analyses. These factors are included in the statistical models used to analyze the SPRT data (Appendix G.1). A complete presentation of the BLP results headed by an initial discussion of all significant differences between groups can be found in Appendix F.

The English NS group and the L2 English users *in toto* differ by gender ($\chi^2_3 = 5.649$, $p = .130$). The L2 English users are much more evenly balanced than the English NS group, which is female-dominant. Although females have been shown to be more successful language learners than males, this difference is not expected to cause any meaningful differences in the results, but it is factored into the statistical models on the chance that it does.

The English NS group and L2 English groups divided by proficiency exhibit two potentially meaningful differences. First, the Int.-Low and the English NS groups differ by gender (DSCF = 3.841, $p = .052$); the Int.-Low group has a higher proportion of males to females than does the English NS group. This difference is not expected to affect the more meaningful comparisons between the L2 user groups. Second, the Low group is significantly more formally educated than the Int.-Low and the English NS groups (DSCF = 3.941, $p = .043$). Level of education has been demonstrated to correlate with performance on certain linguistic tasks, but this difference is also not expected to affect the more meaningful comparisons between the L2 user groups. Both gender and highest level of education are factored into the statistical models to account for these differences, but neither is expected to affect the results.

The L2 English groups divided by L1 exhibit two potentially meaningful differences. First, the Arabic and Chinese groups have significantly lower English proficiency scores than the Other group, but the Arabic and Chinese groups are not different from each other ($p < .05$). This difference should not affect the results of the reading task because the critical comparisons are between the Arabic and Chinese groups, which are matched for proficiency. Second, the values for overall use of the second language are significantly lower for the Chinese group than for the Other group (DSCF = 3.612, $p = .029$). This finding is also not expected to significantly affect the below analysis because the Arabic and Chinese groups, which are the groups being most directly compared, do not differ. Both L2 English proficiency and overall use of the second language are factored into the statistical models to account for these differences, but neither is expected to affect the results.

5.1.3 SELF-PACED READING TASK.

The data considered here and in the following chapter are the individual reading times within the regions of interest from the sentences from Conditions A-H whose comprehension question was answered correctly. Reading time data are trimmed using the following procedures: any value below 100ms is deleted and excluded from the analysis, and any value above 2sd from the group mean for that condition is replaced with the cutoff value ($\bar{x}$+2sd; cf. Jegerski 2014; Marinis 2010). All remaining NS data are incorporated into the analyses; NNS data are divided by response on the comprehension questions (correct/incorrect; cf. Shirai 2003). Only the data associated with correct responses are analyzed (Jegerski 2014). These data are evaluated through the use of generalized linear modeling (GLM; cf. Bayley 2013; Jegerski 2014), which creates regression models that can then be used to investigate group differences in the data using analyses of variance (ANOVAs). In order to satisfy the assumptions of the GLM, the reading time data were transformed using the box-cox power transformation, which normalizes the reading time data so that the data can be compared using parametric and linear statistical models. This transformation was applied to the data separately based on the set of conditions being compared and based on how the L2 English users were divided. Reported here are the results of the ANOVAs and Tukey posthoc tests obtained from minimalist descriptive GLMs. The descriptive and inferential statistics of participant reading times are reported below. Factors entered into the models and additional inferential statistics for transformed reading times by condition and by group can be found in Appendix G.

BOUNDEDNESS CONTRASTS, SPRT. Conditions A through D in the first administration investigate the effects of manipulations in boundedness on the reading times of the participants. Table 5.3 schematizes these conditions.

TABLE 5.3. Boundedness x grammatical tense – SPRT.

Cdn.	Boundedness	Grammatical tense
A	+BND	Present Perfect
B	+BND	Simple Past
C	–BND	Present Perfect
D	–BND	Simple Past

These conditions address the first research question.

RQ1: Do manipulations in boundedness affect how instructed adult L2 learners of English process past time constructions?

The results in this section demonstrate how the variations in compositional and grammatical aspect affect the participants' ability to process the sentences while reading. The comparative speed with which each group reads words in the region of interest indicate the amount of processing required to parse these words. As discussed in section 3.5, it was predicted that L2 English users will be sensitive to boundedness contrasts in a way that supports the complexity account. Int.-High and Advanced groups were expected to exhibit more nativelike processing patterns than lower proficiency groups, and the L1 Chinese group was expected to perform in a more nativelike manner than the L1 Arabic group.

IN TOTO. When comparing L2 learners to native speakers, the L2 English users demonstrate slower average reading times than the English NS controls. This is the case for all reading time data together (Table 5.4).

TABLE 5.4. Descriptive statistics for reading times (ms) for L2 and native English speakers.

	n	$\bar{x}$	sd
L2 Users	4440	869	576
English NS	4890	442	191

This group difference also emerges between groups by condition (Table 5.5).

Comparisons between conditions within groups are evaluated using linear modeling.

TABLE 5.5. Descriptive statistics for reading times (ms) for L2 and native English speakers.

	Condition	n	$\bar{x}$	sd
L2 Users	A. PPerf [+BND]	1125	889	615
	B. SPast [+BND]	1080	856	556
	C. PPerf [–BND]	1115	914	637
	D. SPast [–BND]	1120	818	477
English NS	A. PPerf [+BND]	1220	438	184
	B. SPast [+BND]	1220	434	189
	C. PPerf [–BND]	1225	442	193
	D. SPast [–BND]	1225	455	195

A GLM was constructed in order to investigate potential differences between the groups ($F_{41,8353} = 44.36$, $r^2 = .179$, $p < .001$). This GLM uses normalized values for reading time following a box-cox power transformation ($\lambda = 0$). As expected, an ANOVA indicates that there is a significant difference for overall reading times between the groups by condition ($F_3 = 2.47$, $p = .060$). The same ANOVA indicates that there is no significant difference for reading times at each location between the groups by condition ($F_{28} = 1.01$, $p = .448$).

A Tukey posthoc test was run to further investigate the former significant difference. It finds that the reading times of the L2 English users as a whole are significantly different from those of the English speaking controls in all four conditions (A-D: $p < .001$).[49]

The same posthoc test finds no significant differences in reading times between conditions within groups for either the L2 English users or the English NS group.

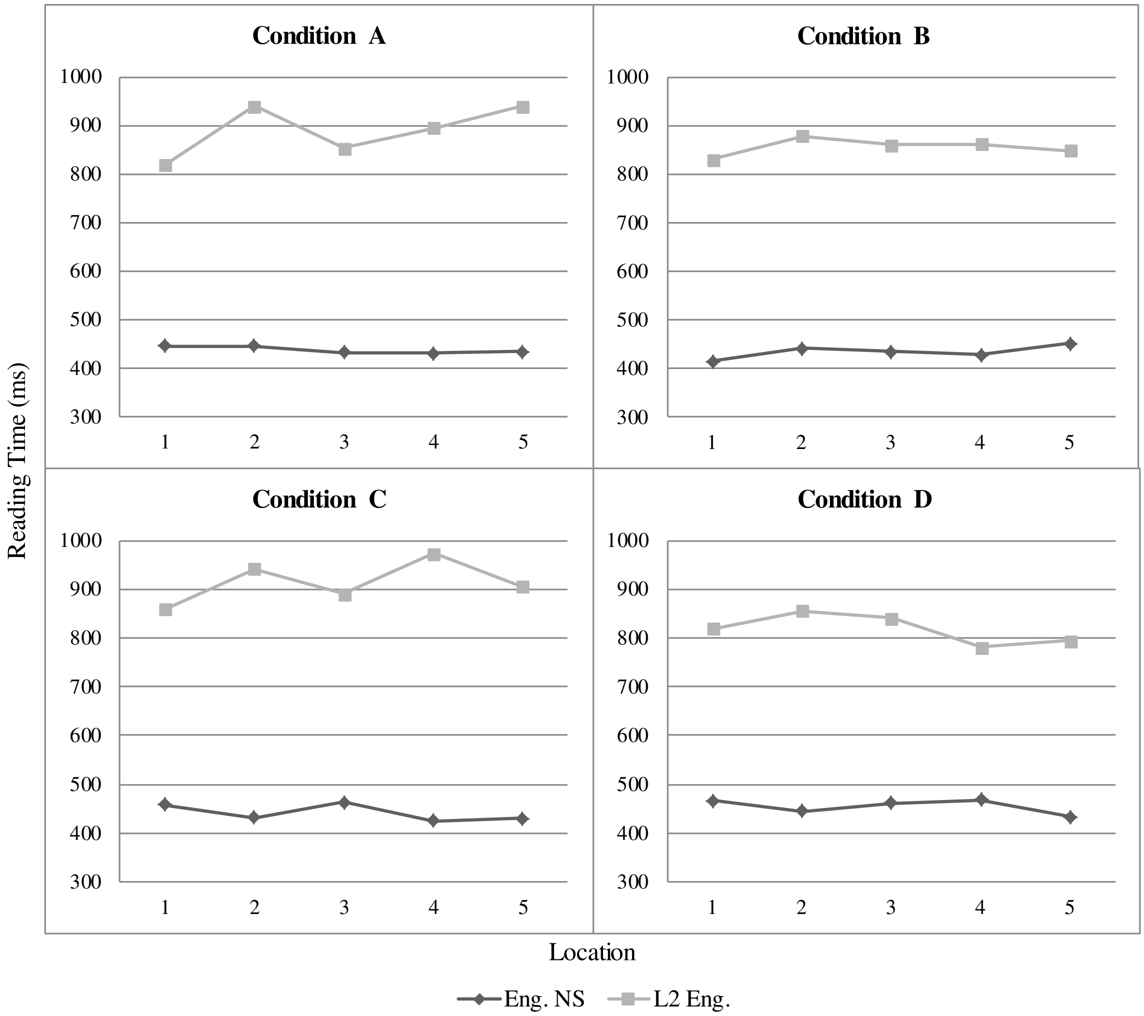

FIGURE 5.1. Response time by condition at each location, *In toto*.

ENGLISH PROFICIENCY. When divided by English proficiency, the L2 English users appear to demonstrate an effect for proficiency, showing faster reading times as

[49] The p-values from this and all other Tukey posthoc tests' pairwise comparisons can be found in Appendices G.4 and G.5.

proficiency increases. The Advanced and Int.-High groups have the fastest reading times followed by the Int.-Low group and, finally, the Low group (Table 5.6).

TABLE 5.6. Descriptive statistics for reading times (ms) for groups by L2 English proficiency.

	n	$\bar{x}$	sd
Adv.	795	670	416
Int.-High	890	633	375
Int.-Low	1315	900	532
Low	1440	1097	687
English NS	4890	442	191

This group difference appears to emerge between groups by condition also (Table 5.7), and is further investigated using linear modeling.

TABLE 5.7. Descriptive statistics for reading times (ms) for groups by L2 English proficiency.

	Condition	n	$\bar{x}$	sd
Adv.	A. PPerf [+BND]	190	617	393
	B. SPast [+BND]	215	647	365
	C. PPerf [–BND]	180	775	571
	D. SPast [–BND]	210	651	302
Int.-High	A. PPerf [+BND]	205	619	430
	B. SPast [+BND]	235	664	362
	C. PPerf [–BND]	240	611	378
	D. SPast [–BND]	210	639	326
Int.-Low	A. PPerf [+BND]	335	912	554
	B. SPast [+BND]	310	894	483
	C. PPerf [–BND]	340	951	605
	D. SPast [–BND]	330	838	464
Low	A. PPerf [+BND]	395	1139	715
	B. SPast [+BND]	320	1099	719
	C. PPerf [–BND]	355	1153	730
	D. SPast [–BND]	370	996	567
English NS	A. PPerf [+BND]	1220	438	184
	B. SPast [+BND]	1220	434	189
	C. PPerf [–BND]	1225	442	193
	D. SPast [–BND]	1225	455	195

A GLM was constructed in order to investigate potential differences between the groups divided by L2 English proficiency ($F_{83,7741} = 21.74$, $r^2 = .222$, $p < .001$). This GLM uses normalized values for reading time following a box-cox power transformation ($\lambda = 0.25$). An ANOVA indicates that there is a significant difference for overall reading times between the groups by condition ($F_{12} = 3.\ 21$, $p < .001$). The same ANOVA indicates that there is no significant difference for reading times at each location between the groups by condition ($F_{76} = 1.17$, $p = .144$).

A Tukey posthoc investigating the former contrast finds that the reading times of each group in most cases do not meaningfully differ between groups. That is, when comparing two groups, there are only a few instances in which the groups' performance in a given condition is different from the groups' performance in the other conditions. For example, if a group performs like the English NS group in one condition but not like the English NS group in the other conditions, it can be said that this group is experiencing some beneficial effects for the manipulations in the former condition. This test indicates that, with a few exceptions, the only meaningful differences occur between groups across conditions. These exceptions concern the Conditions C [PPerf –BND] and D [SPast – BND] between the Advanced & the Low groups (C: p = .268; D: p = .187) and between the Advanced & Int.-High groups (C: p = .003).

The same posthoc test finds no significant differences in reading times between conditions within groups for any of the groups.

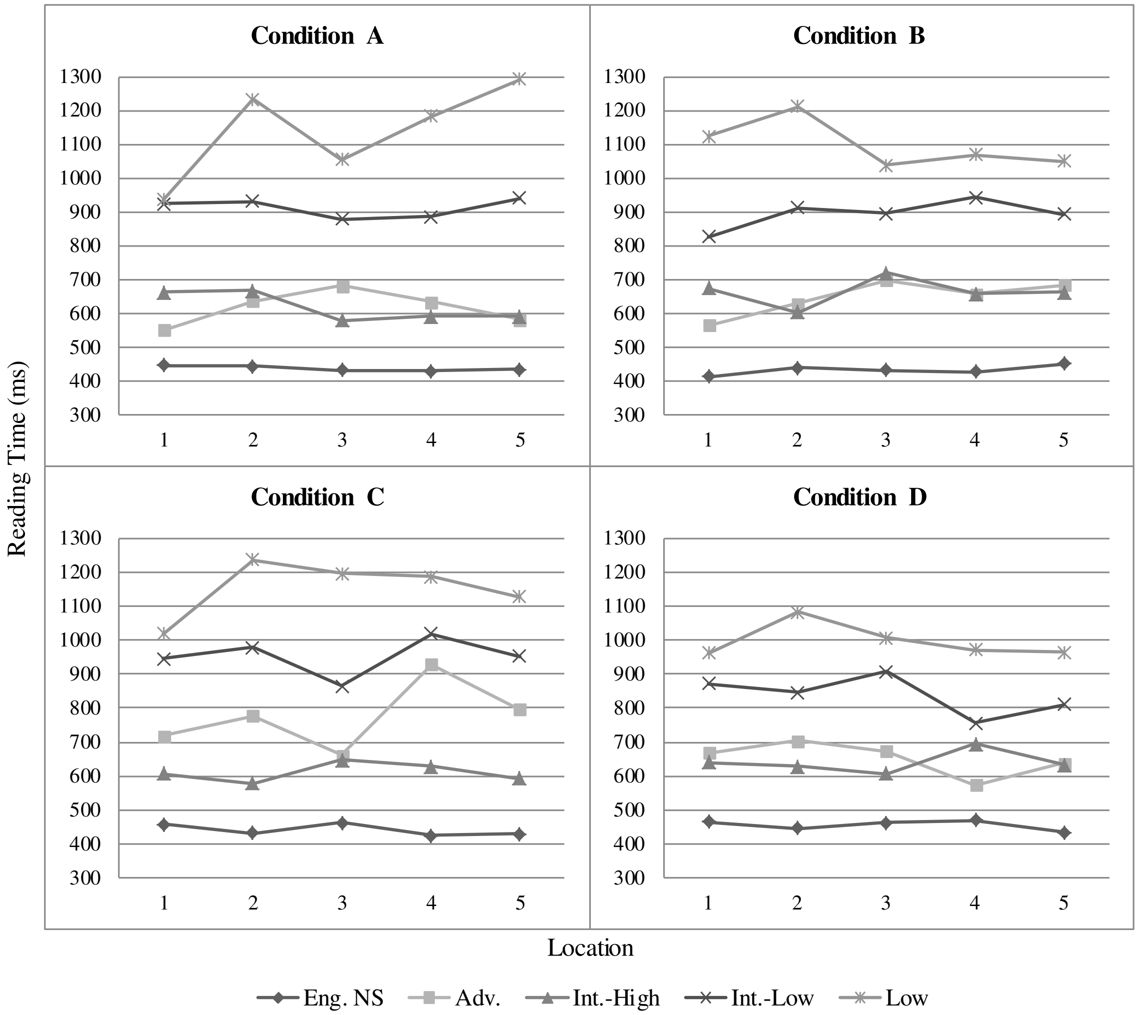

FIGURE 5.2. Response time by condition at each location, English proficiency.

FIRST LANGUAGE. When divided by first language, the L2 English users appear to demonstrate an effect for proficiency but not language. The Other group, which has the highest proficiency, exhibits faster reading times than the Arabic and the Chinese groups (Table 5.8).

TABLE 5.8. Descriptive statistics for reading times (ms) for groups by L2 English proficiency.

	n	x̄	sd
Arabic	640	952	555
Chinese	2635	908	600
Other	1165	736	503
English NS	4890	442	191

This group difference appears stable between groups by condition (Table 5.9), and is again further investigated using linear modeling.

TABLE 5.9. Descriptive statistics for reading times (ms) for groups by L2 English proficiency.

	Condition	n	$\bar{x}$	sd
Arabic	A. PPerf [+BND]	175	988	576
	B. SPast [+BND]	160	939	501
	C. PPerf [–BND]	150	1026	667
	D. SPast [–BND]	155	855	446
Chinese	A. PPerf [+BND]	665	930	643
	B. SPast [+BND]	595	896	593
	C. PPerf [–BND]	705	933	633
	D. SPast [–BND]	670	869	521
Other	A. PPerf [+BND]	285	731	539
	B. SPast [+BND]	325	740	491
	C. PPerf [–BND]	260	796	612
	D. SPast [–BND]	295	683	347
English NS	A. PPerf [+BND]	1220	438	184
	B. SPast [+BND]	1220	434	189
	C. PPerf [–BND]	1225	442	193
	D. SPast [–BND]	1225	455	195

A GLM was constructed in order to investigate potential differences between the groups divided by first language ($F_{83,7741} = 22.84$, $r^2 = .197$, $p < .001$). This GLM uses normalized values for reading time following a box-cox power transformation ($\lambda = 0$). An ANOVA indicates that there is no significant difference for overall reading times between the groups by condition ($F_9 = 0.93$, $p = .500$). The same ANOVA indicates that there is no significant difference for reading times at each location between the groups by condition ($F_{60} = 1.15$, $p = .201$).

A Tukey posthoc finds that the reading times of each group differs from the English NS controls, but they are not all significantly different from each other. Specifically, the Other group's reading times are not significantly different from Arabic group in Condition D [SPast –CR] ($p = .231$), and they are not significantly different from the Chinese group in all conditions (A: p =.948; B: p =.843; C: p = .210; D: p = .730).

The same posthoc test finds no significant differences in reading times between conditions within groups for any of the groups.

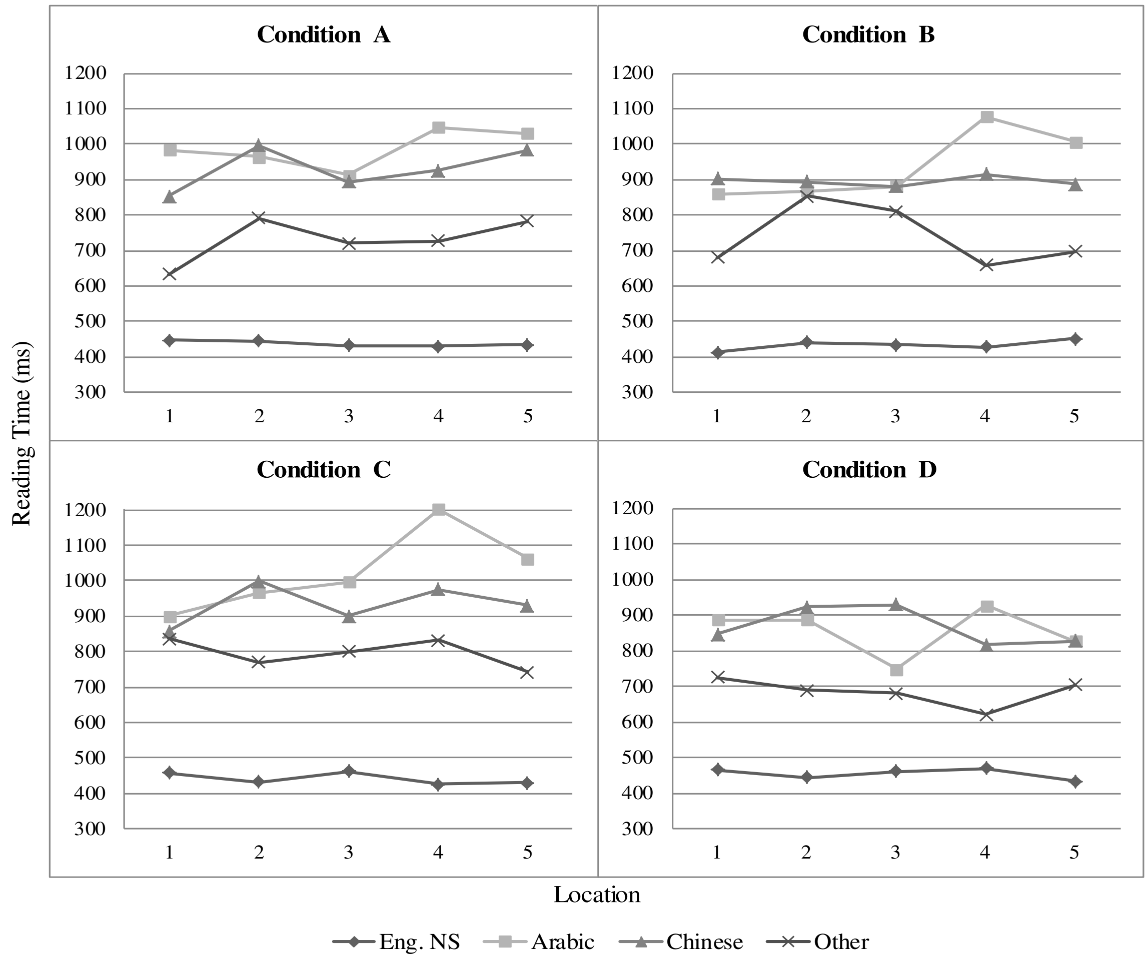

FIGURE 5.3. Response time by condition at each location, First language.

CURRENT RELEVANCE CONTRASTS, SPRT. Conditions E through H in the first administration investigate the effects of manipulations in overtly marked current relevance on the reading times of the participants. Table 5.10 schematizes these conditions.

TABLE 5.10. Current relevance x grammatical tense conditions – SPRT.

Cdn.	Current relevance	Grammatical tense
E	+CR	Present Perfect
F	+CR	Simple Past
G	–CR	Present Perfect
H	–CR	Simple Past

Overtly marked current relevance is manipulated using four categories of adverbial modifiers, which comprise the subconditions (Table 5.11).

TABLE 5.11. Adverbially marked current relevance subconditions – SPRT.

	Current relevance		
Sub.	[+CR]	[–CR]	Category
1	*since* V-*ing*	*after* V-*ing*	Durative vs framing phrases
2	*for* [x time]	*in* [x time]	Duration vs frame AdvPs
3	*at present*	*at some point*	(Pre-)present temporal reference PPs
4	*often & sometimes*	*until* [x event]	Iterative vs noniterative phrases

These conditions and subconditions address the second research question.

RQ2: Do manipulations in adverbial modifiers that overtly mark current relevance affect how instructed adult L2 learners of English process past time constructions?

The results in this section demonstrate how the variations in adverbial modification and grammatical aspect affect the participants' ability to process the sentences while reading. The comparative speed with which each group reads words in the region of interest indicate the amount of processing required to parse these words. It was hypothesized that L2 English users will be sensitive to current relevance contrasts in a way that support the prototype account (§3.5). Intermediate and Advanced groups were expected to exhibit more nativelike processing in prototypical sentences as proficiency increases, and the L1 Chinese group was expected to perform more natively than the L1 Arabic group in these conditions.

In toto. When comparing L2 learners to native speakers, the L2 English users demonstrate slower average reading times than the English NS controls. This is the case for all reading time data together (Table 5.12).

TABLE 5.12. Descriptive statistics for reading times (ms) for L2 and native English speakers.

	n	x̄	sd
L2 Users	4492	604	349
English NS	4895	385	170

This group difference also emerges between groups by condition (Table 5.13). Comparisons between conditions within groups are evaluated using linear modeling as in the previous set.

TABLE 5.13. Descriptive statistics for reading times (ms) for L2 and native English speakers.

	Condition	n	x̄	sd
L2 Users	E. PPerf [+CR]	1161	595	337
	F. SPast [+CR]	1121	609	352
	G. PPerf [–CR]	1080	601	374
	H. SPast [–CR]	1130	612	332
English NS	E. PPerf [+CR]	1230	386	158
	F. SPast [+CR]	1220	386	158
	G. PPerf [–CR]	1215	379	154
	H. SPast [–CR]	1230	389	205

A GLM was constructed in order to investigate potential differences between the groups ($F_{41,8474} = 45.94$, $r^2 = .182$, $p < .001$). This GLM uses normalized values for reading time following a box-cox power transformation ($\lambda = -0.25$). Contrary to expectations, an ANOVA indicates that there is no significant difference for overall reading times between the groups by condition ($F_3 = 0.64$, $p = .589$). The same ANOVA indicates that there is a significant difference for reading times at each location between the groups by condition ($F_{28} = 4.18$, $p < .001$).

A Tukey posthoc test was run to further investigate the former failure to find a significant difference. It finds that the reading times of the L2 English users as a whole are significantly different from those of the English NS controls in all four conditions (E-H: $p < .001$).

The same posthoc test finds no significant differences in reading times between conditions within groups for either the L2 English users or the English NS group.

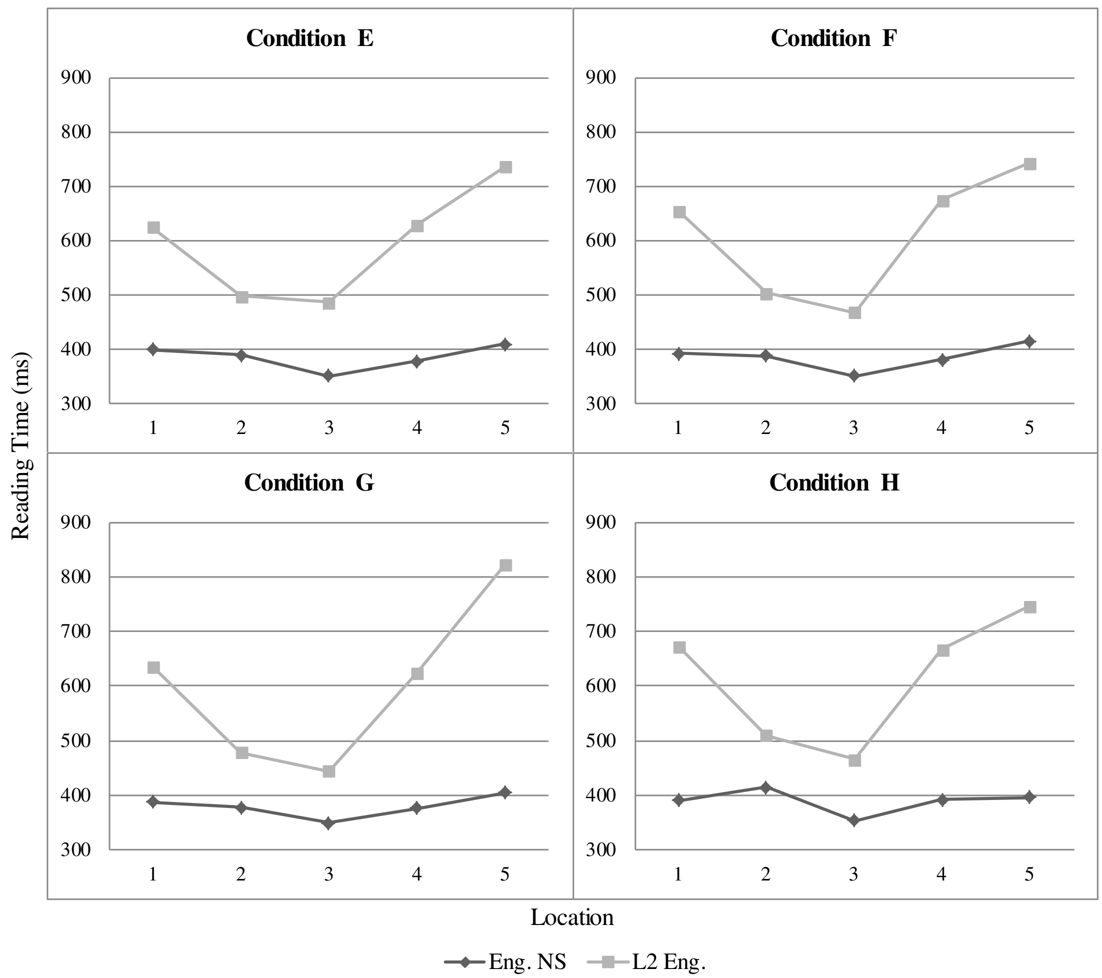

FIGURE 5.4. Response time by condition at each location, *In toto*.

When these data are further subdivided by subcondition, the same lack of meaningful differences is maintained. (See Appendices G.2 & G.3 for a tabular representation of the reading times for all subconditions.) These comparisons are again performed using linear modeling. A second GLM was constructed to investigate the subconditions ($F_{131,8354} = 12.38$, $r^2 = .193$, $p < .001$). This GLM uses normalized values for reading time following the same box-cox power transformation ($\lambda = -0.25$). Contrary to expectations again, an ANOVA indicates that there is no significant difference for overall reading times between the groups by subcondition ($F_{15} = 1.13$, $p = .327$). The same ANOVA indicates that there is a significant difference for reading times at each location between the groups by subcondition ($F_{124} = 1.20$, $p = .062$).

A Tukey posthoc test was run to further investigate the former failure to find a significant difference. It finds that the reading times of the L2 English users as a whole are significantly different from those of the English NS controls in all subconditions (E1-H4: $p < .001$).

The same posthoc test finds no significant differences in reading times between subconditions within groups for either the L2 English users or the English NS group.

ENGLISH PROFICIENCY. When divided by English proficiency, the L2 English users appear to demonstrate an effect for proficiency, showing faster reading times as proficiency increases. The Advanced and Int.-High groups have the fastest reading times followed by the Int.-Low and the Low groups (Table 5.14).

TABLE 5.14. Descriptive statistics for reading times (ms) for groups by L2 English proficiency.

	n	$\bar{x}$	sd
Adv.	810	503	268
Int.-High	895	515	280
Int.-Low	1411	670	361
Low	1376	653	391
English NS	4895	385	170

This group difference appears to emerge between groups by condition also (Table 5.15), and is again investigated using linear modeling.

TABLE 5.15. Descriptive statistics for reading times (ms) for groups by L2 English proficiency.

	Condition	n	x̄	sd
Adv.	E. PPerf [+CR]	205	522	284
	F. SPast [+CR]	215	515	263
	G. PPerf [–CR]	195	470	264
	H. SPast [–CR]	195	504	260
Int.-High	E. PPerf [+CR]	220	538	293
	F. SPast [+CR]	250	507	272
	G. PPerf [–CR]	210	530	298
	H. SPast [–CR]	215	488	253
Int.-Low	E. PPerf [+CR]	400	625	323
	F. SPast [+CR]	301	700	391
	G. PPerf [–CR]	325	656	400
	H. SPast [–CR]	385	705	332
Low	E. PPerf [+CR]	336	642	392
	F. SPast [+CR]	355	659	383
	G. PPerf [–CR]	350	664	416
	H. SPast [–CR]	335	648	372
English NS	E. PPerf [+CR]	1230	386	158
	F. SPast [+CR]	1220	386	158
	G. PPerf [–CR]	1215	379	154
	H. SPast [–CR]	1230	389	205

A GLM was constructed in order to investigate potential differences between the groups ($F_{109,8406} = 28.82$, $r^2 = .272$, $p < .001$). This GLM uses normalized values for reading time following a box-cox power transformation ($\lambda = -0.25$). An ANOVA indicates that there is a significant difference for overall reading times between the groups by condition ($F_{12} =$ 3. 21, $p < .001$). The same ANOVA indicates that there is a significant difference for reading times at each location between the groups by condition ($F_{76} = 2.40$, $p < .001$).

A Tukey posthoc investigating the former contrast finds that the reading times of each group in most cases do not meaningfully differ between groups by condition. This test indicates that, with a few exceptions, the only meaningful differences occur between groups across conditions. That is, the groups differ according to proficiency and not according to condition. The p-values of these potential exceptions are close to the significance threshold. The Low group is significantly different from the Int.-High and Advanced groups, as expected, except for Condition E [PPerf +CR] (Int.-H: $p = .104$; Adv.: $p = .153$) and Condition H [SPast –CR] (Adv.: $p = .102$). If the p-values on the cusp are removed, then the only significant difference is between the Low and Advanced groups in Condition E [PPerf +CR].

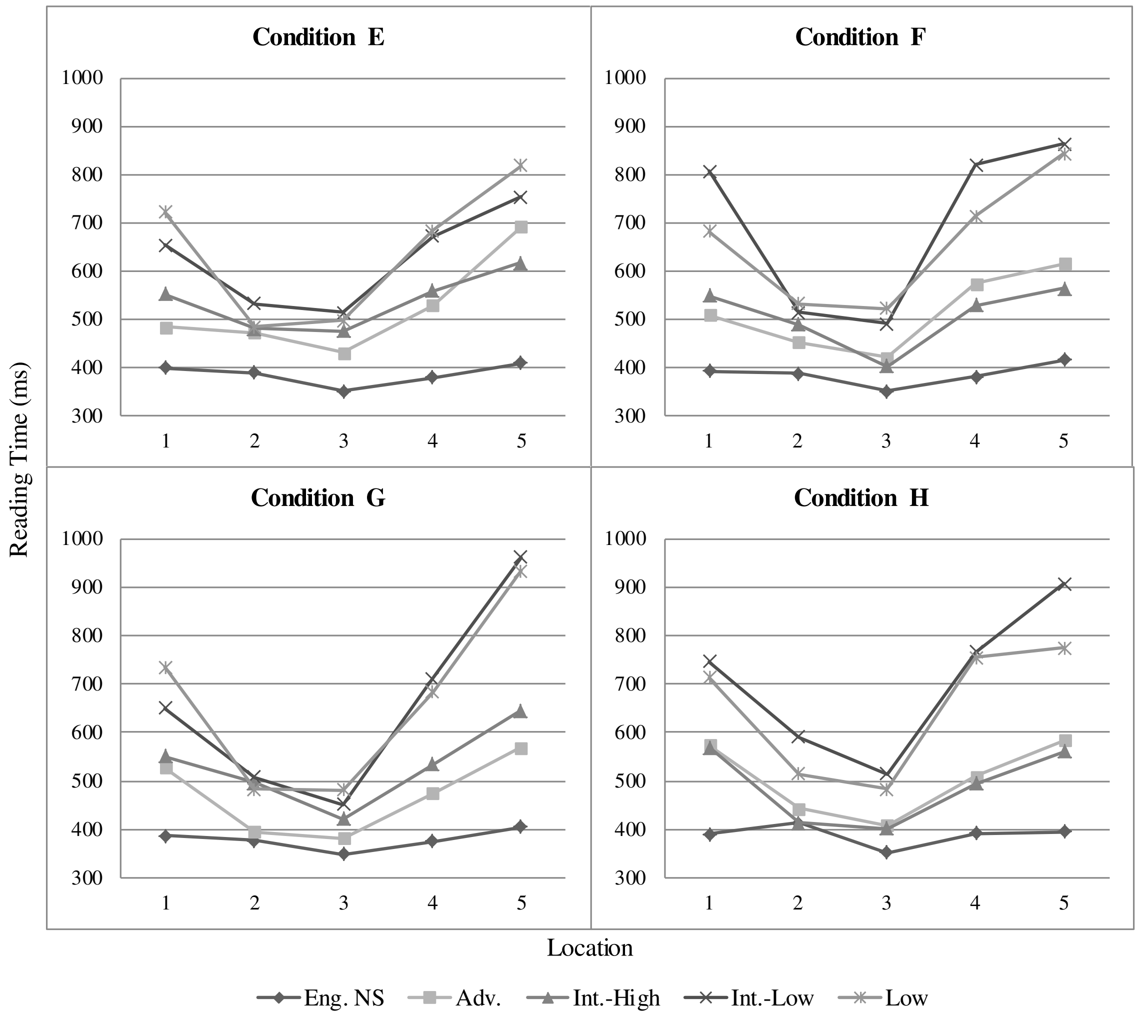

FIGURE 5.5. Response time by condition at each location, English proficiency.

When these data are further divided into subconditions, the same general effect for proficiency is maintained. This effect is evaluated using linear modeling. A second GLM was constructed to investigate the subconditions ($F_{409,8106} = 8.29$, $r^2 = .295$, $p <$

.001). This GLM uses normalized values for reading time following the same box-cox power transformation (λ = -0.25). An ANOVA indicates that there is a significant difference for overall reading times between the groups by subcondition (F_{60} = 2.52, p < .001). The same ANOVA indicates that there is no significant difference for reading times at each location between the groups by subcondition (F_{316} = 0.90, p = .905).

A Tukey posthoc test was run to further investigate the former significant difference between subconditions. It finds significant differences in the overall reading times of the L2 English users divided by proficiency (Table 5.16). The differences between the L2 English users and the English NS group are more numerous; as such, it is more parsimonious to focus on the subconditions in which no differences between groups are found (Table 5.17).

TABLE 5.16. Abbr. Tukey posthoc results for subcondition between proficiency groups.

Prof. Group	Comparison Group	Comparison Subcondition	p-value
Low	Int.-Low	E3. PPerf [+CR] – *at present*	.088
	Int,-High	G1. PPerf [–CR] – *after V-ing*	.003
		F3. SPast [+CR] – *at present*	<.001
	Adv.	H1. SPast [–CR] – *after V-ing*	.019
Int.-Low	Int.-High	G1. PPerf [–CR] – *after V-ing*	.005
		E2. PPerf [+CR] – *for* [x time]	.030
		F2. SPast [+CR] – *for* [x time]	.002
		H2. SPast [–CR] – *in* [x time]	<.001
		F3. SPast [+CR] – *at present*	<.001
		G3. PPerf [–CR] – *at some point*	.051
		H3. SPast [–CR] – *at some point*	.060
		H4. SPast [–CR] – *until* [x event]	<.001
	Adv.	F1. SPast [+CR] – *since V-ing*	<.001
		F2. SPast [+CR] – *for* [x time]	.005
		G2. PPerf [–CR] – *in* [x time]	.047
		H2. SPast [–CR] – *in* [x time]	.042
		H3. SPast [–CR] – *at some point*	<.001
		G4. PPerf [–CR] – *until* [x event]	<.001
Int.-High	Adv.	G3. PPerf [–CR] – *at some point*	.004

TABLE 5.17. Abbr. Tukey posthoc results for subcondition between English NS and prof. groups.

Control Group	Comparison Group	Comparison Subcondition	p-value
Eng. NS	Low	E3. PPerf [+CR] – *at present*	.390
	Int.-Low	Ø	
	Int.-High	F1. SPast [+CR] – *since V-ing*	.581
		G1. PPerf [–CR] – *after V-ing*	1.00
		H1. SPast [–CR] – *after V-ing*	.424
		F2. SPast [+CR] – *for* [x time]	.471
		H2. SPast [–CR] – *in* [x time]	.496
		F3. SPast [+CR] – *at present*	.231
		G3. PPerf [–CR] – *at some point*	.280
		F4. SPast [+CR] – *often & sometimes*	.448
		H4. SPast [–CR] – *until* [x event]	1.00
	Adv.	E1. PPerf [+CR] – *since V-ing*	.416
		H1. SPast [–CR] – *after V-ing*	1.00
		G2. PPerf [–CR] – *in* [x time]	.876
		G3. PPerf [–CR] – *at some point*	.569
		H3. SPast [–CR] – *at some point*	.989
		G4. PPerf [–CR] – *until* [x event]	1.00

The same posthoc test finds a few significant differences in overall reading times between subconditions within the proficiency groups of the L2 English users. The Low group's reading times significantly differ between subconditions E3 [+CR *at present*] & F3 [–CR *at some point*] (p = .007), which vary in grammatical tense. The Int.-High group's reading times significantly differ between subconditions G4 [PPerf –CR *until [x event]*] & H4 [SPast –CR *until [x event]*] (p = .048), which vary in grammatical tense. The Advanced group's reading times significantly differ between subconditions F4 [SPast +CR *often & sometimes*] & G4 [PPerf –CR *until [x event]*] (p = .051), which vary in grammatical tense and current relevance, and G4 & H4 (p = .002), which vary in grammatical tense.

FIRST LANGUAGE. When divided by first language, the L2 English users appear to demonstrate a minimal effect for language. The Arabic group has reading times that are significantly slower than the Chinese and the Other groups. (Table 5.18).

TABLE 5.18. Descriptive statistics for reading times (ms) for groups by L2 English proficiency.

	n	$\bar{x}$	sd
Arabic	700	810	417
Chinese	2622	564	320
Other	1170	571	324
English NS	4895	385	170

This group difference appears stable between groups by condition (Table 5.19), and is investigated using linear modeling.

TABLE 5.19. Descriptive statistics for reading times (ms) for groups by L2 English proficiency.

	Condition	n	$\bar{x}$	sd
Arabic	E. PPerf [+CR]	170	817	407
	F. SPast [+CR]	165	811	395
	G. PPerf [–CR]	165	812	491
	H. SPast [–CR]	200	802	376
Chinese	E. PPerf [+CR]	701	547	299
	F. SPast [+CR]	636	579	348
	G. PPerf [–CR]	635	556	324
	H. SPast [–CR]	650	575	307
Other	E. PPerf [+CR]	290	582	327
	F. SPast [+CR]	320	562	299
	G. PPerf [–CR]	280	577	360
	H. SPast [–CR]	280	563	310
English NS	E. PPerf [+CR]	1230	386	158
	F. SPast [+CR]	1220	386	158
	G. PPerf [–CR]	1215	379	154
	H. SPast [–CR]	1230	389	205

A GLM was constructed in order to investigate potential differences between the groups ($F_{83,7791} = 29.89$, $r^2 = .241$, $p < .001$). This GLM uses normalized values for reading time following a box-cox power transformation ($\lambda = -0.25$). An ANOVA indicates that there is no significant difference for overall reading times between the groups by condition ($F_9 =$

0.80, p = .6181). The same ANOVA indicates that there is a significant difference for reading times at each location between the groups by condition ($F_{60} = 3.06$, $p < .001$).

A Tukey posthoc finds that the reading times of each group are significantly different except for between two groups. The Chinese and the Other groups' reading times are not significantly different in all four conditions (E: p = .996; F: p = 1.00; G: p = .999; H: p = 1.00).

The same posthoc test finds no significant differences in reading times between conditions within groups for any of the groups.

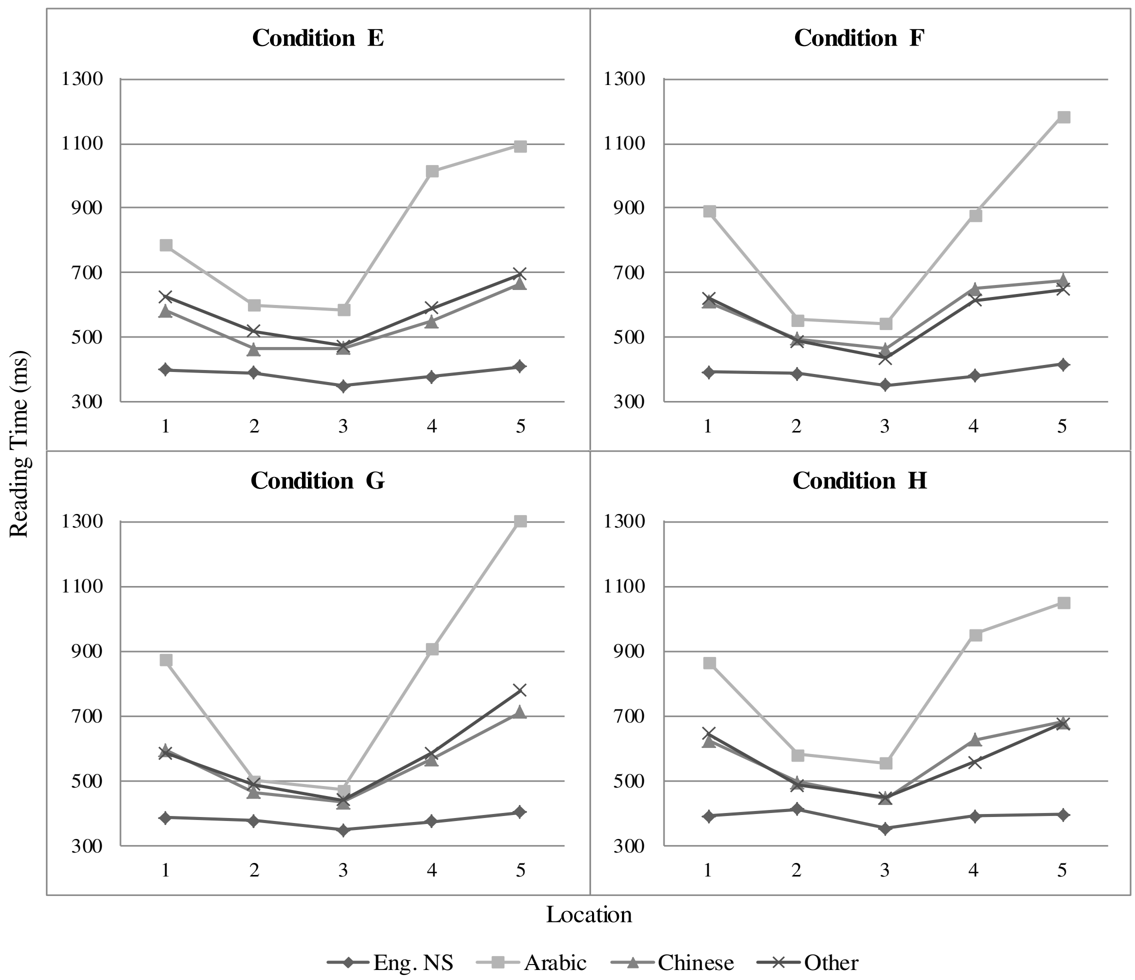

FIGURE 5.6. Response time by condition at each location, First language.

When these data are further divided into subconditions, the same general differences between the Arabic group and the Chinese and Other groups are maintained. These differences are investigated using linear modeling.

A second GLM was constructed to investigate the subconditions ($F_{323,7551} = 8.16$, $r^2 = .259$, $p < .001$). This GLM uses normalized values for reading time following the same box-cox power transformation ($\lambda = -0.25$). An ANOVA indicates that there is a significant difference for overall reading times between the groups by subcondition ($F_{45} = 1.36$, $p = .053$). The same ANOVA indicates that there is no significant difference for reading times at each location between the groups by subcondition ($F_{252} = 1.02$, $p = .405$).

A Tukey posthoc test was run to further investigate the former significant difference. It finds significant differences in the reading times of the L2 English users divided by first language (Table 5.20). The differences between the L2 English users and the English NS group are more numerous; as such, it is more constructive to focus on the subconditions in which no differences between groups are found (Table 5.21).

TABLE 5.20. Abbr. Tukey posthoc results for subcondition between L1 Groups.

L1 Group	Comparison Group	Comparison Subcondition	p-value
Arabic	Chinese	G1. PPerf [–CR] – *after V-ing*	<.001
		H1. SPast [–CR] – *after V-ing*	.064
		E2. PPerf [+CR] – *for* [x time]	<.001
		F2. SPast [+CR] – *for* [x time]	<.001
		H2. SPast [–CR] – *in [x time]*	.059
		E3. PPerf [+CR] – *at present*	<.001
		F3. SPast [+CR] – *at present*	.005
		E4. PPerf [+CR] – *often* & *sometimes*	.094
		F4. SPast [+CR] – *often* & *sometimes*	.011
		G4. PPerf [–CR] – *until* [x event]	.003
		H4. SPast [–CR] – *until* [x event]	.002

L1 Group	Comparison Group	Comparison Subcondition	p-value
Arabic	Other	E1. PPerf [+CR] – *since V-ing*	.006
		G1. PPerf [–CR] – *after* V-*ing*	<.001
		H1. SPast [–CR] – *after* V-*ing*	<.001
		E2. PPerf [+CR] – *for* [x time]	.002
		F2. SPast [+CR] – *for* [x time]	<.001
		G2. PPerf [–CR] – *in* [x time]	.036
		H2. SPast [–CR] – *in* [x time]	.066
		F3. SPast [+CR] – *at present*	.027
		G4. PPerf [–CR] – *until* [x event]	<.001
		H4. SPast [–CR] – *until* [x event]	.091
Chinese	Other	Ø	

TABLE 5.21. Abbr. Tukey posthoc results for subcondition between English NS and L1 Groups.

Control Group	Comparison Group	Comparison Subcondition	p-value
Eng. NS	Arabic	Ø	
	Chinese	E1. PPerf [+CR] – *since V-ing*	.525
		F1. SPast [+CR] – *since* V-*ing*	.544
		G1. PPerf [–CR] – *after V-ing*	1.00
		E2. PPerf [+CR] – *for* [x time]	.636
		F2. SPast [+CR] – *for* [x time]	.637
		H2. SPast [–CR] – *in* [x time]	594
		E3. PPerf [+CR] – *at present*	.997
		G3. PPerf [–CR] – *at some point*	.881
		F4. SPast [+CR] – *often & sometimes*	1.00
		G4. PPerf [–CR] – *until* [x event]	.349
		H4. SPast [–CR] – *until* [x event]	.939
	Other	E1. PPerf [+CR] – *since* V-*ing*	1.00
		F1. SPast [+CR] – *since* V-*ing*	.917
		G1. PPerf [–CR] – *after* V-*ing*	1.00
		H1. SPast [–CR] – *after* V-*ing*	1.00
		E2. PPerf [+CR] – *for* [x time]	.604
		F2. SPast [+CR] – *for* [x time]	.991
		G2. PPerf [–CR] – *in* [x time]	1.00
		H2. SPast [–CR] – *in* [x time]	.667
		G3. PPerf [–CR] – *at some point*	.149
		F4. SPast [+CR] – *often & sometimes*	.916
		G4. PPerf [–CR] – *until* [x event]	1.00
		H4. SPast [–CR] – *until* [x event]	.864

The same posthoc test finds no significant differences in overall reading times between subconditions within the first language groups of the L2 English users.

5.2 Rating task – Second administration.

The rating task is the principle task of the second administration. Conditions A through D investigate the L2 users' understanding of boundedness in context through continuability ratings. Conditions E through H investigate the L2 users' understanding of current relevance in context through current relevance ratings based on first-position adverbial modifiers or verbal morphosyntax. Below are the summary results of the three tasks used in this administration.

5.2.1 Independent measure of proficiency, Rating task.

The same scoring method is used for the second administration as the first. The results of this task are presented below.

Table 5.22. Descriptive statistics for proficiency score for groups by L2 English proficiency.

	n	$\bar{x}$	sd
Adv.	13	17.538	4.115
Int.-High	24	11.708	1.122
Int.-Low	14	8.214	0.893
Low	16	4.250	1.949
All L2 Users	67	10.328	5.037
English NS	30	21.733	3.039

An ANOVA revealed that there is a statistically significant difference for proficiency scores between the L2 user and English NS group ($F_{1,95} = 131.81$, $p < .001$). An additional ANOVA revealed that there are statistically significant differences for proficiency scores among the proficiency groups ($F_{4,91} = 165.00$, $p < .001$). A Tukey posthoc test revealed significant differences pairwise between each group at an alpha of .05.

TABLE 5.23. Descriptive statistics for proficiency score for groups by L1.

	n	$\bar{x}$	sd
Arabic	7	11.714	3.402
Chinese	44	9.727	4.722
Other	16	11.375	6.323
English NS	30	21.733	3.039

An ANOVA revealed that there are statistically significant differences for proficiency scores among the groups ($F_{3,93} = 44.84$, $p < .001$). A Tukey posthoc test revealed no significant differences pairwise between each L2 English group at an alpha of .05.

5.2.2 MODIFIED BILINGUAL LANGUAGE PROFILE QUESTIONNAIRE, RATING TASK.

This modified BLP collects the same data as the questionnaire from the first administration. Recall from section 5.1.2 that the BLP captures data that have been shown to affect L2 acquisition; it is important that any differences between groups be considered in the analyses in order to prevent confounds. This section departs from the format used elsewhere in the chapter in an effort to focus the reader on specific significant differences captured by the BLP. A complete presentation of the BLP results headed by an initial discussion of all significant differences between groups can be found in Appendix H.

The English NS group and the L2 English users *in toto* differ by age ($F_{1,94} = 3.76$, $p = .056$). The L2 English users are on average slightly older than the native speakers; the difference in age is minimal, and the English NS group is already expected to read more quickly than the older L2 English users. This difference is not expected to affect the results

The L2 English groups divided by proficiency differ by the number of years spent immersed in a school/work environment in which English is spoken ($p < .05$). The pairwise analysis reveals that the members of Int.-Low group have spent significantly less time than their peers in such an environment on average. If exposure to English specifically at school/work affects the ratings in a way dissimilar to the other measures of immersion, then it is expected that the Int.-Low group will perform more poorly than another proficiency-matched sample of participants would in a future running of this task. This difference is not expected to affect the results because the groups do not significantly differ for any of the other measures of immersion.

The L2 English groups divided by L1 differ by several factors. First, the members of the Arabic group are generally older than the members of the Chinese and Other groups ($F_{3,92} = 5.550$, $p = .002$); age is not expected to meaningfully affect ratings. Second, the Arabic group has an older age of English exposure than the Chinese group (DSCF = 4.982, p = .001) and the Other group (DSCF = 3.765, p = .021). Third, and related to the second difference, the Arabic group has also received less formal education in English than the other two groups (Chinese: DSCF = 4.593, p = .003; Other: DSCF = 4.126, p = .010). Fourth, the members of the Other group have significantly higher values for overall use of English than the Arabic group (DSCF = 3.024, p = .082) and the Chinese group (DSCF = 3.548, p = .033); if the Other group performs in a more nativelike manner than the Arabic or Chinese groups, this difference may indicate that frequency of language use leads to improved accuracy. All four of these factors are considered in the analyses, but none are expected to affect the results.

5.2.3 Rating task.

The data considered here and in the following chapter are the individual ratings from the judgment task that follows all of the sentences from Conditions A-H. These data are evaluated through the use of nonparametric group comparison statistics (Wilcoxon Rank Sum Test, Kruskal-Wallis Test, and Dwass, Steel, and Critchlow-Fligner pairwise comparisons) and nonparametric Spearman correlations (whenever meaningful). Reported here are the results of the group comparison statistics obtained from the appropriate analyses. The descriptive and inferential statistics of participants' metalinguistic ratings are reported below.

BOUNDEDNESS CONTRASTS, RATING TASK. Conditions A through D in the second administration investigate the effects of manipulations in boundedness on the continuability ratings of the participants. Table 5.24 schematizes these conditions.

TABLE 5.24. Boundedness x grammatical tense conditions – Rating task.

Cdn.	Boundedness	Grammatical tense
A	+BND / –CONT	Present Perfect
B	+BND / –CONT	Simple Past
C	–BND / +CONT	Present Perfect
D	–BND / +CONT	Simple Past

In this task, continuability is operationalized as the opposite of boundedness (cf. Shirai 2013:283). These conditions address the third research question.

RQ3: Are the grammars of instructed adult L2 learners of English sensitive to contrasts in boundedness?

The results in this section demonstrate how the variations in compositional and grammatical aspect affect how completed or continuable participants rate the tensed verb phrase. The comparative continuability ratings that are assigned to the tensed verb phrases characterize the boundedness feature's meaning in the grammar. That is, these ratings indicate how the English users understand the boundedness feature and its interactions with the simple past and the present perfect. As discussed in section 3.5, it was predicted that L2 English users will be sensitive to boundedness contrasts in a way that is mediated by proficiency and first language. Intermediate and Advanced groups were expected to exhibit nativelike continuability ratings, and the L1 Chinese group was expected to perform in a more nativelike manner than the L1 Arabic group.

IN TOTO. When comparing L2 learners to native speakers, the L2 English users rate the continuability of the predicates differently than the English NS controls. This is demonstrated in the present perfect conditions (A [PPerf +BND] & B [SPast +BND]), which are given higher continuability ratings by the L2 English group (Table 5.25).

TABLE 5.25. Descriptive statistics for continuability rating for L2 and native English speakers.

	Condition	n	$\bar{x}$	sd
L2 Users	A. PPerf [+BND]	344	3.785	1.944
	B. SPast [+BND]	340	3.529	1.987
	C. PPerf [–BND]	339	3.956	1.897
	D. SPast [–BND]	340	3.535	2.020
English NS	A. PPerf [+BND]	122	3.164	2.206
	B. SPast [+BND]	124	3.137	2.150
	C. PPerf [–BND]	123	3.650	2.199
	D. SPast [–BND]	124	3.556	2.142

A Wilcoxon Rank Sum Test revealed that there is a statistically significant difference for continuability ratings between the groups ($\chi^2_1 = 6.678$, $p = .010$).

A series of Kruskal-Wallis Tests that there are statistically significant differences for continuability rating between the groups for Condition A [PPerf +BND] ($\chi^2_1 = 6.678$, $p = .010$) and Condition B [SPast +BND] ($\chi^2_1 = 2.968$, $p = .085$), but there are no significant differences between the groups for continuability rating for Condition C [PPerf –BND] ($\chi^2_1 = 1.043$, $p = .307$) and Condition D [SPast –BND] ($\chi^2_1 = 0.036$, $p = .849$).

A pair of Kruskal-Wallis Tests revealed that there is a statistically significant difference for continuability rating within the L2 user group between conditions (χ^2_3 = 10.615, p = .014), but there are no significant differences within the English NS group (χ^2_3 = 5.830, p = .120). A pairwise two-sided multiple comparison analysis using the Dwass, Steel, Critchlow-Fligner method revealed that this significant finding emerges from the differences in the proportions of ratings between Conditions B [SPast +BND] & C [PPerf –BND] (DSCF = 4.034, p = .023) and between Conditions C [PPerf –BND] & D [SPast –BND] (DSCF = 3.728, p = .042).

ENGLISH PROFICIENCY. When divided by English proficiency, the L2 English users appear to demonstrate an effect for proficiency. Ratings appear to be most affected by tense-aspect construction, but the effect of boundedness on these ratings appears to increase with higher proficiency (Table 5.26). A Kruskal-Wallis Test revealed that there is a statistically significant difference for continuability rating between the groups (χ^2_4 = 23.265, p < .001).

A series of Kruskal-Wallis Tests showed that there are statistically significant differences for continuability rating between the proficiency groups for Condition A [PPerf +BND] (χ^2_4 = 16.514, p = .002), but there are no significant differences for continuability rating between the groups for Condition B [SPast +BND] (χ^2_4 = 7.710, p = .103), Condition C [PPerf –BND] (χ^2_4 = 5.175, p = .270), and Condition D [SPast –BND] (χ^2_4 = 5.393, p = .249). A pairwise two-sided multiple comparison analysis using the Dwass, Steel, Critchlow-Fligner method revealed that this significant finding in Condition A [PPerf +BND] emerges from the differences in the proportions of ratings between the Advanced group and the two lowest proficiency groups (Int.-Low: DSCF =

3.795, p = .056; Low: DSCF = 4.417, p = .015) and between the English NS group and these same groups (Int.-Low: DSCF = 3.695, p = .068; Low: DSCF = 4.507, p = .013).

TABLE 5.26. Descriptive statistics for continuability rating for groups by L2 English proficiency.

	Condition	n	x̄	sd
Adv.	A. PPerf [+BND]	60	3.083	2.053
	B. SPast [+BND]	60	3.083	2.053
	C. PPerf [–BND]	60	3.450	2.111
	D. SPast [–BND]	60	3.117	2.051
Int.-High	A. PPerf [+BND]	114	3.623	2.058
	B. SPast [+BND]	115	3.452	2.014
	C. PPerf [–BND]	112	4.000	1.941
	D. SPast [–BND]	113	3.593	2.099
Int.-Low	A. PPerf [+BND]	65	4.123	1.781
	B. SPast [+BND]	66	3.894	1.906
	C. PPerf [–BND]	65	4.292	1.792
	D. SPast [–BND]	67	3.290	1.946
Low	A. PPerf [+BND]	105	4.152	1.731
	B. SPast [+BND]	99	3.646	1.939
	C. PPerf [–BND]	102	3.990	1.743
	D. SPast [–BND]	100	3.880	1.919
English NS	A. PPerf [+BND]	122	3.164	2.206
	B. SPast [+BND]	124	3.137	2.150
	C. PPerf [–BND]	123	3.650	2.199
	D. SPast [–BND]	124	3.556	2.142

A pair of Kruskal-Wallis Tests revealed that there is statistically significant difference for continuability rating within the Int.-Low group between conditions (χ^2_3 = 9.945, p = .019), but there are no significant differences for any other group (Adv.: χ^2_3 = 1.458, p = .692; Int.-High: χ^2_3 = 4.117, p = .249; Low: χ^2_3 = 3.460, p = .326; English: χ^2_3 = 5.830, p = .120). A pairwise two-sided multiple comparison analysis using the Dwass, Steel, Critchlow-Fligner method revealed that this significant finding emerges from the differences in the proportions of ratings between Conditions A [PPerf +BND] & D [SPast –BND] (DSCF = 3.359, p = .082) and Conditions C [PPerf –BND] & D [SPast –BND] (DSCF = 4.176, p = .017).

FIRST LANGUAGE. When divided by first language, the L2 English users appear to demonstrate an effect for first language, with each group exhibiting different rating preferences (Table 5.27).

TABLE 5.27. Descriptive statistics for continuability rating for groups by L1.

	Condition	n	$\bar{x}$	sd
Arabic	A. PPerf [+BND]	44	3.159	2.057
	B. SPast [+BND]	44	3.045	1.988
	C. PPerf [–BND]	44	3.841	2.112
	D. SPast [–BND]	44	3.136	2.131
Chinese	A. PPerf [+BND]	210	3.852	1.869
	B. SPast [+BND]	205	3.766	1.880
	C. PPerf [–BND]	206	4.044	1.809
	D. SPast [–BND]	207	3.710	1.962
Other	A. PPerf [+BND]	90	3.933	2.021
	B. SPast [+BND]	91	3.231	2.155
	C. PPerf [–BND]	89	3.809	1.994
	D. SPast [–BND]	89	3.326	2.071
English NS	A. PPerf [+BND]	122	3.164	2.206
	B. SPast [+BND]	124	3.137	2.150
	C. PPerf [–BND]	123	3.650	2.199
	D. SPast [–BND]	124	3.556	2.142

A Kruskal-Wallis Test revealed that there is a statistically significant difference for continuability rating between the groups ($\chi^2_3 = 16.428$, $p < .001$).

A series of Kruskal-Wallis Tests revealed that there are statistically significant differences between the proficiency groups for continuability rating for Condition A [PPerf +BND] ($\chi^2_3 = 11.559$, $p = .009$) and for Condition B [SPast +BND] ($\chi^2_3 = 9.890$, $p = .020$), but there are no significant differences for continuability rating between the groups for Condition C [PPerf –BND] ($\chi^2_3 = 1.552$, $p = .670$) and Condition D [SPast –BND] ($\chi^2_4 = 3.156$, $p = .368$). A pairwise two-sided multiple comparison analysis using the Dwass, Steel, Critchlow-Fligner method revealed that the former significant finding in Condition A [PPerf +BND] emerges from the differences in the proportions of ratings between the Chinese and the English NS groups (DSCF = 3.583, p = .055) and between the Other and the English NS groups (DSCF = 3.692, p = .045). The same analysis revealed that the latter significant finding in Condition B [SPast +BND] emerges from the differences in the proportions of ratings between the Chinese and the English NS groups (DSCF = 3.618, p = .052).

A series of Kruskal-Wallis Tests revealed that there is statistically significant difference for continuability rating within the Other group between conditions (χ^2_3 = 8.032, p = .045), but there are no significant differences for any other group (Arabic: χ^2_3 = 3.778, p = .287; Chinese: χ^2_3 = 3.324, p = .344; English: χ^2_3 = 5.830, p = .120). A pairwise two-sided multiple comparison analysis using the Dwass, Steel, Critchlow-Fligner method revealed that this significant finding emerges from the differences in the proportions of ratings between Conditions A [PPerf +BND] & B [SPast +BND], which is significant one-way (DSCF = 3.216, p = .052).

ADVERBIALLY MARKED CURRENT RELEVANCE CONTRASTS, RATING TASK. Conditions E & F and G & H in the second administration investigate the effects of manipulations of adverbial modification and of tense-aspect construction on the current relevance ratings of the participants. Conditions E [+CR] & F [–CR] investigate the perceived importance of the adverbial modifiers in conveying current relevance (Table 5.28), and Conditions G [PPerf] & H [SPast] investigate the understood current relevance of the predicates themselves (Table 5.29). In Conditions E [+CR] & F [–CR], an adverbial modifier indicates current relevance when it characterizes the action or its consequences as more important to the present than to the past.

TABLE 5.28. Adverbially marked current relevance (sub)conditions – Rating task.

Cdn.	Grammatical tense	Current Relevance	Subcondition
E	Present Perfect	+CR	1 – *since* V-*ing*
			2 – *for* [x time]
			3 – *at present*
			4 – *often & sometimes*
F	Present Perfect	–CR	1 – *after V-ing*
			2 – *in [x time]*
			3 – *at some point*
			4 – *until [x event]*

TABLE 5.29. Morphologically marked current relevance conditions – Rating task.

Cdn.	Grammatical tense	Boundedness
G	Present Perfect [+CR]	+BND
H	Simple Past [–CR]	+BND

Conditions E [+CR] & F [–CR] address the fourth research question, focusing on the effect of i) overt adverbial modification, and Conditions G [PPerf] & H [SPast] focus on the effect of ii) verbal morphology.

> RQ4: Are the grammars of instructed adult L2 learners of English sensitive to contrasts in current relevance marked i) adverbially or ii) using verbal morphology?

The results in this section demonstrate how the participants perceive the relations between adverbial modifiers and current relevance. The comparative current relevance ratings that are assigned to each adverbial modifier characterize its association to current relevance in the English user's lexicon. That is, these ratings indicate how the English users understand how current relevance can be marked overtly in certain adverbial phrases that collocate with the present perfect. It was hypothesized that L2 English users will be sensitive to current relevance contrasts in a way that is mediated by proficiency and first language (§3.5). Intermediate and Advanced groups were expected to exhibit nativelike current relevance ratings for adverbially marked current relevance, and the L1 Arabic group was expected to perform more natively than the L1 Chinese group in these conditions.

IN TOTO. When comparing L2 learners to native speakers, the L2 English users rate the current relevance of the adverbial modifiers differently than the English NS controls. The L2 English users appear to show no effect for condition, whereas the controls seem to do so (Table 5.30).

TABLE 5.30. Descriptive statistics for CR rating for L2 and native English speakers.

	Condition	n	$\bar{x}$	sd
L2 Users	E. [+CR]	336	4.464	1.659
	F. [–CR]	338	4.320	1.673
English NS	E. [+CR]	124	4.815	1.727
	F. [–CR]	122	4.049	2.036

A Wilcoxon Rank Sum Test revealed that there is a statistically significant difference between the groups for current relevance rating ($\chi^2_1 = 2.863$, $p = .091$).

A pair of Kruskal-Wallis Tests that there is a statistically significant difference between the groups for current relevance rating for Condition E [+CR] ($\chi^2_1 = 7.904$, $p = .004$), but there are no significant differences between the groups for current relevance rating for Condition F [–CR] ($\chi^2_1 = 0.229$, $p = .632$). The former significant difference probably emerges due to the differences in the proportions of ratings for Subcondition E3 [+CR *at present*] ($\chi^2_1 = 10.198$, $p = .001$) and, possibly, Subcondition E4 [+CR *often* & *sometimes*] ($\chi^2_1 = 2.284$, $p = .131$); both of which are [+CR] marking. (See Appendix I for additional descriptive and inferential statistics for Subconditions E & F.)

A pair of Kruskal-Wallis Tests revealed that there is statistically significant difference within the English NS group for current relevance rating between conditions

(χ^2_1 = 8.999, p = .003), but there are no significant differences within the L2 user group (χ^2_1 = 1.790, p = .181). This significant difference probably emerges due to the differences in the proportions of ratings for concerning Subconditions E3 [+CR *at present*] & F3 [–CR *at some point*] that is reported in the previous paragraph.

ENGLISH PROFICIENCY. When divided by English proficiency, the L2 English users appear to demonstrate an effect for proficiency. Ratings appear to approach those of native speakers only at the highest proficiency level (Table 5.31).

TABLE 5.31. Descriptive statistics for CR rating for groups by L2 English proficiency.

	Condition	n	x̄	sd
Adv.	E. [+CR]	60	4.033	1.947
	F. [–CR]	60	3.667	1.847
Int.-High	E. [+CR]	112	4.607	1.700
	F. [–CR]	113	4.602	1.650
Int.-Low	E. [+CR]	65	4.415	1.713
	F. [–CR]	65	4.508	1.522
Low	E. [+CR]	99	4.596	1.332
	F. [–CR]	100	4.270	1.595
English NS	E. [+CR]	124	4.815	1.727
	F. [–CR]	122	4.049	2.036

A Kruskal-Wallis Test revealed that there is a statistically significant difference between the groups for current relevance rating (χ^2_4 = 17.452, p = .002).

A pair of Kruskal-Wallis Tests that there are statistically significant differences between the proficiency groups for current relevance rating for Condition E [+CR] (χ^2_4 = 11.849, p = .019) and Condition F [–CR] (χ^2_4 = 12.478, p = .014). A pairwise two-sided multiple comparison analysis using the Dwass, Steel, Critchlow-Fligner method revealed that the former significant finding emerges from the differences in the proportions of ratings between the Advanced group and the English NS group (DSCF = 4.083, p = .032). The same test revealed that the latter significant finding emerges from the differences in the proportions of ratings between the Advanced group and the two intermediate proficiency groups (Int.-High: DSCF = 4.819, p = .006; Int.-Low: DSCF = 3.620, p = .078).

A series of Kruskal-Wallis Tests revealed that there is statistically significant difference within the English NS control group for current relevance rating between conditions ($\chi^2_1 = 8.999$, p = .003), but there are no significant differences for any other group (Adv.: $\chi^2_1 = 1.629$, p = .202; Int.-High: $\chi^2_1 = 0.056$, p = .813; Int.-Low: $\chi^2_1 = 0.000$, p = .996; Low: $\chi^2_1 = 1.507$, p = .220). This significant difference probably emerges due to the differences in the proportions of ratings for concerning Subconditions E3 [+CR *at present*] & F3 [–CR *at some point*], which a pairwise two-sided multiple comparison analysis using the Dwass, Steel, Critchlow-Fligner method revealed to be the strongest difference between subconditions (DSCF = 7.879, p < .001).

FIRST LANGUAGE. When divided by first language, the L2 English users appear to demonstrate an effect for first language, with the Arabic group appearing to demonstrate more nativelike ratings than the other two groups (Table 5.32).

TABLE 5.32. Descriptive statistics for CR rating for groups by L1.

	Condition	n	$\bar{x}$	sd
Arabic	E. [+CR]	44	4.818	1.769
	F. [–CR]	44	4.364	1.906
Chinese	E. [+CR]	203	4.419	1.637
	F. [–CR]	206	4.398	1.592
Other	E. [+CR]	89	4.393	1.649
	F. [–CR]	88	4.114	1.738
English NS	E. [+CR]	124	4.815	1.727
	F. [–CR]	122	4.049	2.036

A Kruskal-Wallis Test revealed that there is a statistically significant difference between the groups for current relevance rating ($\chi^2_3 = 0.619$, p = .062).

A series of Kruskal-Wallis Tests that there is a statistically significant difference between the proficiency groups for current relevance rating for Condition E [+CR] (χ^2_3 = 12.592, p = .006), but there are no significant differences between the groups for current relevance rating for Condition F [–CR] (χ^2_3 = 1.685, p = .640). A pairwise two-sided multiple comparison analysis using the Dwass, Steel, Critchlow-Fligner method revealed that this significant finding emerges from the differences in the proportions of ratings between the Chinese and the English NS groups (DSCF = 4.069, p = .021) and between the Other and the English NS groups (DSCF = 3.806, p = .036). This significant difference between the Chinese and English NS groups emerges due to the differences in the proportions of ratings concerning Subconditions E3 [+CR *at present*] (DSCF = 4.842; p = .004).

A series of Kruskal-Wallis Tests revealed that there is statistically significant difference within the English NS group for current relevance rating between conditions (χ^2_1 = 8.999, p = .003), but there are no significant differences for any other group (Arabic: χ^2_1 = 2.023, p = .155; Chinese: χ^2_3 = 0.148, p = .701; Other: χ^2_1 = 1.129, p = .288). The significant difference within the English NS group emerges due to the differences in the proportions of ratings for concerning Subconditions E3 [+CR *at present*] & F3 [–CR *at some point*]; this significant difference is noted above in the previous subsection.

GRAMMATICALLY MARKED CURRENT RELEVANCE CONTRASTS, RATING TASK. Conditions G [PPerf] & H [SPast] in the second administration investigate the effects of manipulations of tense-aspect construction on the current relevance ratings of the participants. In these conditions, the only variation is in the predicate: simple past &

present perfect. These conditions address the fourth research question, focusing on the effect of ii) tense-aspect morphology.

> RQ4: Are the grammars of instructed adult L2 learners of English sensitive to contrasts in current relevance marked i) adverbially or ii) using verbal morphology?

The results in this section demonstrate how the participants perceive the relations between the two tense-aspect constructions and current relevance. The comparative current relevance ratings that are assigned to each predicate characterize its association to current relevance in the English user's grammar. That is, these ratings indicate how the English users understand how current relevance can be marked overtly using the present perfect. It was hypothesized that L2 English users will be sensitive to current relevance contrasts in a way that is mediated by proficiency and first language (§3.5). Only the most advanced groups (Int.-High & Adv.) were expected to exhibit nativelike current relevance ratings for current relevance marked via verbal morphology, and the L1 Chinese group was expected to perform more natively than the L1 Arabic group in these conditions.

IN TOTO. When comparing L2 learners to native speakers, the L2 English users rate the current relevance of the predicates similarly to the English NS controls. Neither the L2 English users nor the controls appear to show an effect for condition (Table 5.33).

TABLE 5.33. Descriptive statistics for CR rating for L2 and native English speakers.

	Condition	n	$\bar{x}$	sd
L2 Users	G. PPerf	339	4.212	1.761
	H. SPast	335	4.131	1.839
English NS	G. PPerf	123	4.740	1.778
	H. SPast	124	4.532	1.828

A Wilcoxon Rank Sum Test revealed that there is a statistically significant difference between the groups for current relevance rating ($\chi^2_1 = 18.100$, $p < .001$).

A pair of Kruskal-Wallis Tests that there is a statistically significant difference between the groups for current relevance rating for Condition G [PPerf] ($\chi^2_1 = 13.018$, $p < .001$) and Condition H [SPast] ($\chi^2_1 = 5.899$, $p = .015$).

A pair of Kruskal-Wallis Tests revealed that there are no statistically significant differences for current relevance rating between conditions within the English NS group ($\chi^2_1 = 1.364$, $p = .243$) or within the L2 user group ($\chi^2_1 = 0.120$, $p = .729$).

ENGLISH PROFICIENCY. When divided by English proficiency, the L2 English users appear to demonstrate an effect for proficiency. Ratings appear to begin to separate only at the highest proficiency level (Table 5.34).

TABLE 5.34. Descriptive statistics for CR rating for groups by L2 English proficiency.

	Condition	n	x̄	sd
Adv.	G. PPerf	60	4.567	1.760
	H. SPast	60	4.333	1.829
Int.-High	G. PPerf	114	4.342	1.799
	H. SPast	112	4.250	1.948
Int.-Low	G. PPerf	64	3.984	1.750
	H. SPast	64	3.859	1.789
Low	G. PPerf	101	4.000	1.703
	H. SPast	99	4.051	1.752
English NS	G. PPerf	123	4.740	1.778
	H. SPast	124	4.532	1.828

A Kruskal-Wallis Test revealed that there is a statistically significant difference between the groups for current relevance rating ($\chi^2_4 = 29.784$, $p < .001$).

A pair of Kruskal-Wallis Tests revealed that there are statistically significant differences between the proficiency groups for current relevance rating for Condition G [PPerf] ($\chi^2_4 = 21.325$, $p < .001$) and Condition H [SPast] ($\chi^2_4 = 10.164$, $p = .038$). A pairwise two-sided multiple comparison analysis using the Dwass, Steel, Critchlow-Fligner method revealed that the former significant finding emerges from the differences in the proportions of ratings between the English NS group and the two lowest proficiency groups (Int.-Low: DSCF = 4.782, $p = .007$; Low: DSCF = 5.661, $p < .001$). The same test revealed that the latter significant finding emerges from the differences in the proportions of ratings between the English NS and the Int.-Low group (DSCF = 3.935, $p = .043$).

A series of Kruskal-Wallis Tests revealed that there are no statistically significant differences within the groups for current relevance rating between conditions (Adv.: $\chi^2_1 = 0.788$, $p = .375$; Int.-High: $\chi^2_1 = 0.003$, $p = .987$; Int.-Low: $\chi^2_1 = 0.156$, $p = .693$; Low: $\chi^2_1 = 0.099$, $p = .753$; English: $\chi^2_1 = 1.364$, $p = .243$).

FIRST LANGUAGE. When divided by first language, the L2 English users appear to demonstrate a small effect for first language, with the Arabic group appearing to demonstrate slightly more nativelike ratings than the other two groups (Table 5.35).

TABLE 5.35. Descriptive statistics for CR rating for groups by L1.

	Condition	n	$\bar{x}$	sd
Arabic	G. PPerf	43	4.256	1.853
	H. SPast	43	4.000	1.976
Chinese	G. PPerf	206	4.243	1.641
	H. SPast	204	4.206	1.769
Other	G. PPerf	90	4.122	1.988
	H. SPast	88	4.023	1.942
English NS	G. PPerf	123	4.740	1.778
	H. SPast	124	4.532	1.828

A Kruskal-Wallis Test revealed that there is a statistically significant difference between the groups for current relevance rating ($\chi^2_3 = 18.127$, $p < .001$).

A series of Kruskal-Wallis Tests that there is a statistically significant difference between the proficiency groups for current relevance rating for Condition G [PPerf] (χ^2_3 = 13.252, p = .004) and for Condition H [SPast] (χ^2_3 = 6.365, p = .095). A pairwise two-sided multiple comparison analysis using the Dwass, Steel, Critchlow-Fligner method revealed that the former significant finding emerges from the differences in the proportions of ratings between the Chinese and the English NS groups (DSCF = 5.086, p = .002) and between the Other and the English NS groups (DSCF = 3.556, p = .0578). The same analysis revealed that the latter significant finding does not obtain under greater scrutiny.

A series of Kruskal-Wallis Tests revealed that there are no statistically significant differences within the groups for current relevance rating between conditions (Arabic: χ^2_1 = 0.412, p = .521; Chinese: χ^2_3 = 0.017, p = .896; Other: χ^2_1 = 0.173, p = .678; English: χ^2_1 = 1.364, p = .243).

CHAPTER 6 DISCUSSION

The following sections discuss the results of the self-paced reading task (SPRT) and the rating task utilized in the first and second administration, respectively, of the present investigation. These discussions are divided by task and further divided by sets of conditions. The discussion begins with first two research questions, which concern the online processing of boundedness and current relevance and are answered through analysis of the SPRT results. The discussion continues with the final two research questions, which concern the metalinguistic knowledge of boundedness and current relevance and are answered through analysis of the rating task results. The final section summarizes and concludes the discussion.

6.1 SELF-PACED READING TASK – FIRST ADMINISTRATION.

The following discussion concerns the results of the reading times from the self-paced reading task (SPRT) for conditions A-D and E-H. Conditions A-D concern boundedness in RQ1: *Do manipulations in boundedness affect how instructed adult L2 learners of English process past time constructions?* Conditions E-H concern RQ2: *Do manipulations in adverbial modifiers that overtly mark current relevance affect how instructed adult L2 learners of English process past time constructions?* In each of the following sections, the results are discussed in reference to the classification metrics used throughout the present investigation: native-nonnative, L2 English proficiency, and first language of the L2 English learners.

6.1.1 BOUNDEDNESS CONTRASTS, SPRT.

As discussed previously in section 4.5.1, Conditions A-D collect reading times at five locations within the region of interest of sentences manipulated for boundedness (via object quantification, see §2.1.2) and for tense-aspect construction. These conditions capture reading times that indicate the relative facility of processing for variations in boundedness in both the simple past and the present perfect. Refer to the below table for a summary of the manipulations in these conditions.

TABLE 6.1. Manipulations in boundedness and tense-aspect for Conditions A-D.

Cdn.	Boundedness	Grammatical Tense
A	+BND	Present Perfect
B	+BND	Simple Past
C	–BND	Present Perfect
D	–BND	Simple Past

Recall from section 3.5 that it is expected that manipulations in boundedness and grammatical tense will affect reading times in the regions of interest. It is hypothesized that bounded and nonbounded predicates are not read with equal facility in the simple past and the present perfect, and it is further hypothesized that such variation emerges following prototypical clusters of features. That is, bounded predicates in the present perfect and the simple past (Conditions A & B) will be processed more quickly than nonbounded ones (Conditions C & D), and this trend will especially be true of the present perfect conditions (A & C), following the assumption that boundedness is a prototypical feature of the present perfect. Likewise, it is hypothesized that the effects of the interactions will be affected by differences in L2 English proficiency and by first language.

IN TOTO. The first set of comparisons concerns the L2 English users as a whole and the English NS group. The first meaningful comparison concerns the bound-nonbound distinction. The reading times of both the L2 English users and the English NS controls do not show any significant differences between the bounded predicates (A & B) and the nonbounded predicates (C & D). Likewise, the trends observed for reading times by location provide no clear evidence of differential processing obscured by the condition means. These trends exhibited by the groups indicate that, on the whole, the boundedness of the predicates does not meaningfully affect the processing speed of the L2 English users or the English NS controls.

The second meaningful comparison concerns the tense-aspect contrast. Again, the reading times of both the L2 English users and the English NS controls do not show any meaningful differences between the present perfect predicates (A & C) and the simple past ones (B & D) when they are matched for boundedness (A & B and C & D). As is the case for the previous comparisons, the reading times by location provide no evidence of differential processing. However, the descriptive statistics indicate that there may be a meaningful difference in reading times for grammatical tense among the L2 English users. The reading times for the present perfect ($\bar{x}_A = 889$, $\bar{x}_C = 914$) are slower than the reading times for the simple past ($\bar{x}_B = 856$, $\bar{x}_D = 818$). This result is expected because the simple past is learned before and is used more often than the present perfect. As such, the L2 English users should demonstrate greater facility when processing the simple past than when processing the present perfect, which is suggested by these reading times. This facility is more apparent in nonbounded contexts where the simple past predicates (D) are read much more quickly than the present perfect ones (C). This is a surprising result in the prototype perspective because the nonbounded feature and the simple past clash, which should inhibit processing. It is the expected result in the complexity perspective because Condition D [SPast –BND] is less complex and should require less processing effort.

Overall, the failure to find meaningful differences between conditions indicates that there is not sufficient evidence to suggest that the processing speed of either group is significantly affected by manipulations in boundedness or grammatical tense in the sentences used in the present investigation. However, the descriptive statistics seem to support some meaningful effect for grammatical tense among the L2 English users. This effect seems to derive from more exposure to or experience with the simple past; this possibility is investigated further when the L2 users are divided by English proficiency.

ENGLISH PROFICIENCY. The second set of comparisons concerns the L2 English users divided by proficiency and the English NS group with a focus on the L2 English users. The Advanced group is considered separately from the others, which are treated together. The Advanced group is the only group to show an effect for boundedness and grammatical tense; the lower proficiency groups show an effect for changes in grammatical tense only or they show no effects at all.

The reading times of the Advanced group exhibit meaningful differences between conditions. These differences chiefly concern Condition C [PPerf –BND] ($\bar{x}_C = 775$), whose reading times are slower than the reading times of the other three conditions. The graphed reading times by location indicate that this slower reading time is due to slower reading times at Locations 4 and 5 ($\bar{x}_{C.L4} = 927$, $\bar{x}_{C.L5} = 795$). The most informative comparisons for Condition C [PPerf –BND] concern Condition A [PPerf +BND] ($\bar{x}_A$ =

617) and Condition D [SPast –BND] ($\bar{x}_D$ = 651), which are matched for grammatical tense and for boundedness, respectively. The reading times in Condition B [SPast +BND] ($\bar{x}_B$ = 647) cast doubt on the meaningfulness of the comparison between Conditions C [PPerf –BND] and D [SPast –BND]. Alone, the comparison between Conditions C [PPerf –BND] and D [SPast –BND] would indicate that the Advanced group reads the simple past more quickly than the present perfect in nonbounded contexts. The fact that the reading times for Conditions B [SPast +BND] and D [SPast –BND] are so similar emphasizes the effect of grammatical tense over any effect of boundedness. The comparison between the two present perfect conditions (A [PPerf +BND] & C [PPerf –BND]) indicates that the Advanced group processes the bounded predicates more rapidly than nonbounded ones in the present perfect. In order to determine whether or not this trend is one of facilitation or inhibition, comparisons are made between the Advanced and Int.-High groups. For Condition A [PPerf +BND], the two groups' reading times are very similar (Adv.: $\bar{x}_A$ = 617; Int.-H: $\bar{x}_A$ = 619); however, in Condition C [PPerf –BND], their reading times are significantly different (Adv.: $\bar{x}_C$ = 775; Int.-H: $\bar{x}_C$ = 611). That is, the two groups perform similarly in Condition A [PPerf +BND] and differently in Condition C [PPerf –BND]. The deviation from this trend emerges within the Advanced group, who perform significantly more slowly in Condition C [PPerf –BND], which suggests that the nonbounded feature inhibits processing in the present perfect. This finding supports the prototype account, following the assumption that bounded perfects are more prototypical (and thus processed more easily) than nonbounded ones. It contradicts the complexity account because the bounded predicates, which are

syntactically and semantically more complex, are processed more quickly than the nonbounded ones.

There are two additional trends that are potentially meaningful in the reading times for the non-Advanced proficiency groups. First, the Int.-Low and the Low groups exhibit reading times for the present perfect conditions that are slower than they are for the simple past. This finding is expected because the members of both groups either have no formal exposure to the present perfect or have minimal instruction on and practice with the construction. Having only minimal exposure, one would expect some inhibition when these groups try to process the less familiar construction. Second, the reading times of the Int.-High group are largely unaffected by both boundedness and grammatical tense. This is an unexpected finding, and it can be reconciled in three ways: either i) these learners are so familiar with the constructions that they are behaving like native speakers, processing each construction with similar ease, ii) there are task effects associated with the complexity of the SPRT that obscure their processing behavior, or iii) these learners have enough experience with the present perfect for the cost of processing it to approximate that of the simple past but not enough experience with the construction to meaningfully process the differences in boundedness that affect its meaning. The present investigation favors the third possibility because this interpretation situates the performance of the Int.-High group meaningfully between the lower proficiency and higher proficiency groups on the following cline of relative processing from fastest to slowest:

(1) Simple past → [+BND] Present perfect → [–BND] Present perfect.

This cline supports the prototype account as well as the complexity account. The difference in processing effort between bounded and nonbounded predicates in the present perfect indicates inhibition caused by processing nonprototypical features; however, the fact that boundedness appears to not affect processing effort as the present perfect diverges from the simple past suggests that the processor fails to notice VP-level complexity when acquiring syntactically higher TP-level complexity.

Taken as a whole, these findings indicate how L2 English proficiency affects the relative processing effort required for L2 users to read sentences manipulated for boundedness and grammatical tense. First, boundedness only affects performance at the highest proficiency level. Specifically, nonbounded predicates have an inhibitory effect in the present perfect, which suggests that processing the nonbounded predicates is more costly for the processor. On the one hand, these predicates are syntactically and semantically less complex than the bounded ones, so the complexity account fails to capture this phenomenon. On the other hand, the prototype account does successfully capture this phenomenon, predicting that the present perfect would emerge first in the more prototypical bounded context.

FIRST LANGUAGE. The third set of comparisons concerns the L2 English users divided by first language and the English NS group with a focus on the L2 English users. The results are discussed in two clusters based on similarities exhibited by the groups. The Arabic and Other groups are analyzed first, followed by the Chinese group.

The reading times of the Arabic and the Other groups exhibit similar trends despite the fact that the Other group's reading times are significantly faster than the Arabic group's times in all but one condition (D: $p = .231$). The two groups exhibit different trends in bounded contexts only. The Arabic group processes the present perfect ($\bar{x}_A$ = 988) more slowly than the simple past ($\bar{x}_B$ = 939), whereas these two conditions are nearly equal in the Other group ($\bar{x}_A$ = 731; $\bar{x}_B$ = 740). The likely explanation for this difference is that the Other group has a significantly higher average proficiency than the Arabic group, and, as such, the Other and Arabic groups exhibit the patterns of the Advanced or Int.-High groups and the Int.-Low group, respectively. Likewise, a comparison of the reading times by location for these conditions indicate that the Arabic group exhibits a spike in reading times in Condition B [SPast +BND] at Location 4 not replicated by either L1 group ($\bar{x}_{B.L4}$ = 1078). In nonbounded contexts, the two groups are very similar. Both have slower reading times in Condition C [PPerf –BND] (Arabic: $\bar{x}_C$ = 1026; Other: $\bar{x}_C$ = 796) than they have in Condition D [SPast –BND] (Arabic: $\bar{x}_D$ = 855; Other: $\bar{x}_D$ = 683). The reading times by location indicate that the groups follow the same general processing pattern in Condition C [PPerf –BND], exhibiting slower reading times at Location 4 (Arabic: $\bar{x}_{C.L4}$ = 1202; Other: $\bar{x}_{C.L4}$ = 833). In Condition D [SPast –BND], the Other group exhibits no significantly slower reading time at any location, but the Arabic group still exhibits a slightly slower-than-average reading time at Location 4 ($\bar{x}_{D.L4}$ = 928), a fact which is obscured in the average reading time by the faster reading times in Location 3 ($\bar{x}_{D.L3}$ = 749). Together, these trends exhibited in Conditions C [PPerf –BND] and D [SPast –BND] indicate that reading the nonbounded present perfect sentences requires more processing effort than reading the nonbounded simple past ones. The facilitation that the Arabic group experiences in Condition D [SPast –BND] is enough to cause the aforementioned only nonsignificant difference between the two groups. As is the case with the Advanced group above, this trend in the nonbounded conditions supports the prototype account and contradicts the complexity account.

The reading times of the Chinese group exhibits different enough trends that it seems justified to separate them from the other two groups. Like the Arabic group but unlike the Other group, the Chinese group processes the present perfect ($\bar{x}_A = 930$) more slowly than the simple past ($\bar{x}_B = 896$). Like both groups, the Chinese group also has slower reading times in Condition C [PPerf –BND] ($\bar{x}_C = 933$) than in Condition D [SPast –BND] ($\bar{x}_D = 869$); however, this difference is much smaller than it is for the other two groups, and the Chinese group's reading times by location are more irregular in Condition D [SPast –BND] than C [PPerf –BND]. Taken together, these trends exhibited between A [PPerf +BND] & B [SPast +BND] and C [PPerf –BND] & D [SPast –BND] indicate that the Chinese group is chiefly affected by manipulations in grammatical tense. The fact that these two trends are similar indicates that they are less affected by boundedness than the other two groups and are equally affected by grammatical tense. In this respect, they appear to behave like the Int.-Low and Low groups above, and thus provide some support for the complexity account over the prototype account. This similarity in trends should not be considered to be the effect of proficiency. The Chinese group's reading times are significantly faster than the proficiency-matched Arabic group and are not significantly different from the Other group, which has higher average proficiency.

Overall, these findings provide evidence of meaningful difference between the L1 groups that cannot be explained solely by differences in proficiency. In bounded contexts, the Arabic and the Chinese groups exhibit different readings times for each tense-aspect construction while the Other group exhibits no such difference. It should be noted that the difference between Conditions A [PPerf +BND] and B [SPast +BND] is much smaller for the Chinese group than for the Arabic group, which situates the Chinese group between the Arabic and the Other group in these contexts. The nonbounded contexts seem to be more volatile between conditions for all L2 users; each group has greater differences between Conditions C [PPerf –BND] & D [SPast –BND] than between A [PPerf +BND] & B [SPast +BND]. In these conditions, the Arabic and the Other groups perform similarly, exhibiting larger differences between conditions than the Chinese group. Taken together, these findings indicate three aspects of processing: i) the present perfect requires more processing effort than the simple past, ii) the present perfect and the simple past are more distinguished in nonbounded contexts, and iii) the present perfect requires more processing effort in nonbounded contexts than in bounded ones. Overall, the groups' reading times seem to be more closely adhering to the prototype account when boundedness is a factor.

RQ1, CONCLUSIONS. The data from conditions A-D from the first administration support the following conclusions to RQ1 and its subquestions.

Main question: Yes, manipulations in boundedness affect how some instructed adult L2 learners of English process past time constructions.

Subquestion (a): It is unclear whether the processing of the adult learners is qualitatively similar to the processing of the adult native speakers because the reading times of the native speaker controls do not exhibit any significant or meaningful differences between conditions. This may either be the result of a task not sensitive enough to capture so thoroughly entrenched and rapid native speaker processing, or it may be an indication of qualitative differences between the groups.

Subquestion (b): Yes, the processing of the adult learners is affected by differences in L2 English proficiency. The Advanced group is the only group to show an effect for boundedness; the lowest two proficiency groups only show an effect for grammatical tense. The Int.-High group exhibits trend that appears to be between these two extremes wherein changes in grammatical tense no longer inhibit processing but changes in boundedness do not yet measurably affect the processor.

Subquestion (c): Yes, the processing of the adult learners is affected by the L1 processor. There are certainly meaningful differences in processing exhibited by the three groups, but the most meaningful comparisons in the present investigation concern the Arabic and Chinese groups. Contrary to predictions, the Arabic group appears to perform more like the Advanced group than their proficiency would predict, which suggests that they are experiencing some benefit from their L1 processors. The Chinese group performs in the manner expected by their average proficiency, which suggests that their L1 processors do not measurably affect their L2 processing in this task. These findings may be the result of *learned attention*, a phenomenon by which L2 learners preferentially attend either to structures from the L1 or to structures acquired earlier in the L2 (Ellis & Sagarra 2010; Ellis et al. 2012; Sagarra & Ellis 2013). For the former finding that is indicative of positive L1 transfer, attention required to process syntactic-semantic cues in the L1 may have led the Arabic group to process L2 structures that also employ syntactic-semantic cues with greater-than-expected facility. As to the latter finding, since their L1 does not employ comparable syntactic-semantic cues, the Chinese group's processors may ignore or be blind to these cues in the L2, which manifests in these results as reading times that are not indicative of either positive or negative L1 transfer.

6.1.2 CURRENT RELEVANCE CONTRASTS, SPRT.

As discussed previously in section 4.5.1, Conditions E-H collect reading times at five locations within the region of interest of sentences manipulated for overtly marked current relevance (via [+/–CR] adverbial modifiers in first position, see Table 4.3) and for tense-aspect construction. These conditions capture reading times that indicate the relative facility of processing for variations in overtly marked current relevance in both the simple past and the present perfect. Refer to the below table for a summary of the manipulations in these conditions.

TABLE 6.2. Manipulations in overtly marked CR and tense-aspect for Conditions E-H.

Cdn.	Current Rel.	Grammatical Tense
E	+CR	Present Perfect
F	+CR	Simple Past
G	–CR	Present Perfect
H	–CR	Simple Past

Recall from section 3.5 that it is expected that manipulations in overtly marked current relevance and grammatical tense will affect reading times in the regions of interest. That is, it is hypothesized that sentences with the combination of [+CR] adverbial modifiers and the present perfect (Condition E [PPerf +CR]) will be processed more rapidly than sentences with [–CR] adverbial modifiers in the present perfect (Condition G [PPerf –CR]). Likewise, it is hypothesized that sentences with the combination of [–CR] adverbial modifiers and the simple past (Condition H [SPast –CR]) will be processed more rapidly than sentences with [+CR] adverbial modifiers in the simple past (Condition F [SPast +CR]). Further, it is expected that the effects of the interactions between current relevance and grammatical tense will be affected by differences in L2 English proficiency and by first language.

IN TOTO. The first set of comparisons concerns the L2 English users as a whole and the English NS group. The first meaningful comparison concerns the [+/–CR] distinction. The reading times of both the L2 English users and the English NS controls do not show any significant differences between the [+CR]-modified predicates (E & F) and the [–CR]-modified predicates (G & H). These trends exhibited by the groups indicate that, on the whole, overtly marked current relevance does not meaningfully affect the processing speed of either the L2 English users or the English NS controls. Visual analysis of the four graphs for reading times by location indicates that the English NS controls have stable reading times across locations for all conditions, whereas the L2 English users exhibit similar subtle U-shaped pattern that is likely due to word length. This subtle U-shape captures the observed differences between groups at Locations 1, 4, and 5.

The second meaningful comparison concerns the tense-aspect contrast. The reading times of both the L2 English users and the English NS controls do not show any meaningful differences between the present perfect predicates (E [PPerf +CR] & G [PPerf –CR]) and the simple past ones (F [SPast +CR] & H [SPast –CR]) when they are matched for current relevance (E & F and G & H). This is a rather unexpected result considering the results of Conditions A-D, which suggest that L2 users are affected by manipulations in grammatical tense. Although unexpected, this result is readily explained: L2 users have a tendency to over-rely on lexical markers of temporality whenever they are present, largely ignoring the often redundant grammatical markers of tense-aspect (cf. Bardovi-Harlig 2000b).

Overall, there is no evidence that L2 English users as a group are affected by manipulations in overtly marked current relevance and grammatical tense. The English NS controls, likewise, exhibit no observed effects for the manipulations. This same lack of evidence is observed in the subconditions, the results of which yield no useful information for discussion.

ENGLISH PROFICIENCY. The second set of comparisons concerns the L2 English users divided by proficiency and the English NS group with a focus on the L2 English users. Each group is considered separately with the exception of the Low group, whose data show no evidence of differential processing in the conditions or subconditions.

The reading times of the Advanced group exhibit one notable trend between conditions. This trend emerges from the reading times for Condition G [PPerf –CR] ($\bar{x}_G$ = 470), whose reading times are faster than the reading times for the other three conditions. The most informative comparisons concern Condition E [PPerf +CR] ($\bar{x}_E$ = 522) and Condition H [SPast –CR] ($\bar{x}_H$ = 504), which are matched for grammatical tense and for adverbial modification, respectively. A visual comparison of reading times by location for these conditions indicate that the differences are probably due to more stable processing in Condition G [PPerf –CR] and, possibly, due to some facilitation at Locations 2 and 3 in the region of interest ($\bar{x}_{G.L2}$ = 395, $\bar{x}_{G.L3}$ = 383). Condition E [PPerf +CR] exhibits spikes indicative of inhibition at Locations 4 and 5 ($\bar{x}_{E.L4}$ = 530, $\bar{x}_{E.L5}$ = 692), and Condition H [SPast –CR] exhibits a pattern similar to Condition G [PPerf –CR] but without facilitation. The comparison between Condition E [PPerf +CR] and G [PPerf

–CR] indicates that the addition of current relevance inhibits processing of the verb phrase, and the comparison between G and H indicates that the present perfect phrase is processed more rapidly than the simple past one, which is due to facilitation in Condition G [PPerf –CR]. The former result supports the complexity account under the assumption that the [+CR] adverbial modifiers are adding additional information to be processed that are not added by the [–CR] modifiers. That is, the less semantically complex clause is processed more rapidly than the more semantically complex one. The latter result is unexpected and is not describable using either account. In the prototype account, the [–CR] modifiers should inhibit the processing of the present perfect causing Condition G [PPerf –CR] to be read more slowly than Condition H [SPast –CR]. In the complexity account, the more complex structure (the present perfect [G]) should be read more slowly than the less complex structure (the simple past [H]). Neither of which captures these comparisons.

The final interesting comparison between conditions concerns the difference between Conditions F [SPast +CR] and G [PPerf –CR], which approaches significance. Like the latter comparison, it is unclear what this difference means in the context of the present investigation. The conclusion most supported by the data suggests that lexical-semantic complexity ([+CR] in Condition F [SPast +CR]) is more costly to the processor than tense-aspectual complexity (present perfect in Condition G [PPerf –CR]). This conclusion seems unlikely in the context of the literature that overwhelmingly supports the reverse trend among adult L2 learners. This difference between conditions may be a result of the fourth subcondition, which shows that the [–CR] modifier that does not encode iterativity leads to a faster reading time in this context (Subconditions F4 [SPast

+CR *often* & *sometimes*] & G4 [PPerf –CR *until [x event]*]) and leads to a faster reading time for the present perfect than the simple past (Subconditions G4 [PPerf –CR *until [x event]*] & H4 [SPast –CR *until [x event]*]). That is, these results may be more an effect for manipulations in the current relevance indicated by the presence of iterativity than for current relevance as a whole.

The reading times of the Int.-High group exhibit comparatively faster processing in the less complex conditions than the more complex conditions. This trend emerges from an effect for grammatical tense wherein the two simple past conditions ($\bar{x}_F$ = 507; $\bar{x}_H$ = 488) are processed more rapidly than the present perfect conditions ($\bar{x}_E$ = 538; $\bar{x}_G$ = 530). The difference between Conditions G [PPerf –CR] and H [SPast –CR] is significant, which indicates a clear difference in processing speed between the two grammatical tenses when modified by [–CR] adverbial modifiers. An analysis of the reading times by location data indicates that this difference is due to facilitation in Condition H [SPast –CR] at Locations 2 and 4 ($\bar{x}_{H.L2}$ = 414, $\bar{x}_{H.L4}$ = 495). This group shows no effect for overtly marked current relevance in the main conditions, especially in the present perfect conditions (E [PPerf +CR] & G [PPerf –CR]) where the reading times are approximately equal. However, in the subconditions, there is a significant difference in the fourth subcondition between Subconditions G4 [PPerf –CR *until [x event]*] and H4 [SPast –CR *until [x event]*], which may help to explain the significant difference in the conditions. Overall, the results of this group indicate an effect for complexity in tense-aspect and show no indication of an effect for prototypicality of features.

The reading times of the Int.-Low group exhibit the opposite trend of the one exhibited by the Int.-High group: comparatively faster processing in the more complex conditions than the less complex conditions. This trend emerges from an effect for grammatical tense wherein the two present perfect conditions ($\bar{x}_E = 625$; $\bar{x}_G = 656$) are processed more rapidly than the simple past conditions ($\bar{x}_F = 700$; $\bar{x}_H = 705$). A comparison of the reading times by location in Condition E [PPerf +CR] and F indicates that the difference may be due to the facilitative effect of the prototypical combination of features, as exhibited in reading times of Locations 1 and 5 (compare $\bar{x}_{E.L1} = 654$ & $\bar{x}_{F.L1} = 807$ and $\bar{x}_{E.L5} = 754$ & $\bar{x}_{F.L5} = 864$). The meaningful differences in reading time between Conditions E [PPerf +CR] and G [PPerf –CR] suggest that there may be a facilitative effect for [+CR] modifiers in the present perfect, but there is not sufficient evidence to conclude this definitely. The results of the subconditions present a mix of results that indicate that the present perfect and simple past sentences are processed differently across subconditions between the two intermediate groups. Divergent from the trend exhibited by the two more advanced groups, the Int.-Low group's performance in the subconditions was not markedly different in the fourth set than in was in the first three. Overall, the results of this group are contrary to the predictions of the complexity account as it pertains to tense-aspect. Likewise, the results for the two present perfect conditions suggest a facilitative effect for prototypicality of features.

Taken together, these findings provide conflicting evidence. The two intermediate groups provide the clearest examples of these conflicts, suggesting different processing preferences mediated by proficiency. Exhibiting faster reading times in the less complex (and more familiar) simple past conditions (F & H) than in the more complex present perfect conditions (E & G), the Int.-High group's results indicate an effect for complexity in grammatical tense but no effect for manipulations in overt adverbial modification. In opposition, the reading times of the Int.-Low group are faster in the more complex present perfect conditions (E & G) than in the simple past ones (F & G). When combined with the fact that the group also exhibits faster reading times for present perfect sentences modified by [+CR] phrases (Condition E [PPerf +CR]; a prototypical set of features) than those modified with [–CR] phrases (Condition G [PPerf –CR]), the Int.-Low group's results support the prototype account. As a set, the results of these two groups indicate an effect for proficiency wherein the lower proficiency L2 users are more affected by prototypical relations and the higher proficiency L2 users are more affected by complexity of the grammatical tense. These results can be interpreted as a result of the general trend of temporal marking noted in Bardovi-Harlig (2000b): adult L2 learners rely on lexical means of marking time before using grammatical ones. Through this lens then, the results of the Advanced group indicate that lexical-semantic complexity is more taxing on the L2 processor than tense-aspectual complexity at very high proficiency levels. Perhaps, this incongruity is reconcilable with the existing literature if the L2 processor is capable of processing both sets of information together only at this high proficiency level.

FIRST LANGUAGE. The third set of comparisons concerns the L2 English users divided by first language and the English NS group with a focus on the L2 English users. The results are discussed in two meaningful comparisons based on trends exhibited by the groups: i) the differences between the proficiency-matched Arabic and Chinese groups and ii) the potential trends observed between the non-proficiency-matched Chinese and Other groups. The L2 English groups all exhibit some variation of the same pattern for reading times by location with the Arabic group exhibiting a more exaggerated pattern. The results of the subconditions do not add any meaningful data for discussion not already present in the discussion of the results of the conditions.

The reading times of the Arabic and Chinese groups are far more different than is expected. The Arabic group's reading times in these conditions (E-H) are approximately 15 percent faster than those in Conditions A-D ($\bar{x}_{E\text{-}H} = 810$; $\bar{x}_{A\text{-}D} = 952$), which is certainly comparable; however the Chinese group's reading times in these conditions (E-H) are approximately 40 percent faster than they are in Conditions A-D ($\bar{x}_{E\text{-}H} = 564$; $\bar{x}_{A\text{-}D} = 908$), which is a significant difference in reading time. Due to the structural variation between the two sets of conditions, it is expected that there will be some variation, but the amount of variation in the results of the Chinese group suggests that there is an effect for first language that facilitates the processing of this group to the extent that their reading times are not significantly different from the higher proficiency Other group ($\bar{x}_{E\text{-}H} = 571$). The lack of significant variation of reading times within the conditions suggests that this effect is not due to the specific manipulations of the present investigation. Instead, the data suggest it is the result of varied processing of the two forms of boundedness assignment: object quantification and motion-goal (§2.1.2). Additional study is necessary to determine if these two methods of boundedness assignment are processed differently among this group in both their first and second languages.

The comparison between the Chinese and Other groups concerns the relationship between proficiency and comparative processing speed of the two grammatical tenses. Recall from the relevant discussion of English proficiency above that the two intermediate groups exhibit different preferences concerning grammatical tense: the Int.-Low group exhibits faster reading times in the present perfect whereas the Int.-High group exhibits faster reading times in the simple past. The average proficiency scores of the Chinese and Other groups correspond to those of the two proficiency groups. The groups divided by L1 still exhibit traces of this trend, but the decrease in the difference between the reading times of the present perfect (E & G) and simple past conditions (F & H) indicate that the observed trends are affected by proficiency more than first language processing strategies.

Overall, these comparisons fail to provide evidence of meaningful differences between the L1 groups. The results provide no indication that manipulations in tense-aspect and overt adverbial modification affect the L2 processing of the L2 English users in any meaningful way. The only potentially interesting finding in these data pertains to the overall reading times of the Chinese group between these conditions (E-H) and those that investigate the effects of boundedness (A-D). The significantly lower reading times in the former conditions indicate that Chinese group's processing is facilitated by the motion-goal structure and/or inhibited by the quantified object structure. More research needs to be done to investigate the potential variations in the processing of these two structures.

RQ2, CONCLUSIONS. The data from conditions E-H from the first administration support the following conclusions to RQ2 and its subquestions.

Main question: No, manipulations in adverbial modifiers that overtly mark current relevance do not affect how some instructed adult L2 learners of English process past time constructions.

Subquestion (a): It is unclear whether the processing of the adult learners is qualitatively similar to the processing of the adult native speakers because the reading times of the native speaker controls do not exhibit any significant or meaningful differences between conditions. This may either be a result of a task not sensitive enough to capture so thoroughly entrenched and rapid native speaker processing, or it may be an indication of qualitative differences between the groups.

Subquestion (b): Yes, the processing of the adult learners is affected by differences in L2 English proficiency. Although no group shows a clear effect for manipulations in current relevance on its own, the Advanced group shows a small inhibitory effect in [+CR] conditions. The two intermediate groups exhibit the counterintuitive result that the high proficiency group processes the simple past more rapidly than the simple past while the lower proficiency group processes the present perfect more rapidly than the simple past. This result violates the assumption that the learners acquire the present perfect after the simple past.

Subquestion (c): No, the processing of the adult learners is not affected by the L1 processor in a manner that is meaningful for the present investigation. There is no evidence to suggest that processing is affected by the manipulations in the conditions; however, there is evidence to suggest that the Chinese group's processing is affected by how boundedness is predicated in this task.

6.2 Rating task – Second administration.

The following discussion concerns the results of the metalinguistic ratings for conditions A-D, E & F, and G & H. Conditions A-D concern boundedness in RQ3: *Are the grammars of instructed adult L2 learners of English sensitive to contrasts in boundedness?* Conditions E-H concern RQ4: *Are the grammars of instructed adult L2 learners of English sensitive to contrasts in current relevance marked i) adverbially or ii)* *using verbal morphology?* Conditions E [+CR] & F [–CR] concern overtly marked current relevance, and conditions G [PPerf] & H [SPast] concern grammatically marked current relevance. Each set of comparisons is treated by research question. In each of the following sections, the results are discussed in reference to the classification metrics used throughout the present investigation: native-nonnative, L2 English proficiency, and first language of the L2 English learners.

6.2.1 BOUNDEDNESS CONTRASTS, RATING TASK.

As discussed previously in section 4.5.2, Conditions A-D collect speaker judgment data concerning boundedness (operationalized as possibility for continuation) in the present perfect and the simple past. Specifically, these conditions capture speaker understanding of the relation between boundedness and continuability in both the simple past and the present perfect. Refer to the below table for a summary of the manipulations in these conditions.

TABLE 6.3. Manipulations in boundedness and tense-aspect for Conditions A-D.

Cdn.	Boundedness	Grammatical Tense
A	+BND	Present Perfect
B	+BND	Simple Past
C	–BND	Present Perfect
D	–BND	Simple Past

Recall from section 3.5 that it is expected that manipulations in boundedness and grammatical tense affect the perception of continuability. That is, it is hypothesized that bounded predicates (Conditions A & B) are rated as less able to be continued than nonbounded ones (Conditions C & D). Likewise, it is hypothesized that this effect will be more pronounced in the present perfect (Conditions A & C) than in the simple past (Conditions B & D).

IN TOTO. The first set of comparisons concerns the L2 English users as a whole and the English NS group. The first meaningful comparison concerns the bound-nonbound distinction. The continuability ratings of the L2 English users do not show any meaningful differences between the bounded predicates (A & B) and the nonbounded predicates (C & D). That is, there are no significant differences between the condition pairs that vary according to boundedness but maintain the same tense-aspect construction (A & C and B & D). This trend is opposed to that exhibited by the English NS group whose ratings appear to differ in the environments that differ according to boundedness but maintain tense-aspect construction (A & C and B & D); bounded predicates (A & B) are rated approximately half of a point less continuable than nonbounded ones (C & D). These different trends exhibited by the groups indicate that, on the whole, the L2 English users do not interpret boundedness as a factor that meaningfully relates to the continuability of a predicate, whereas the English NS group associates boundedness with noncontinuability and nonboundedness with continuability.

The second meaningful comparison concerns the distinction between present perfect and simple past. The continuability ratings of the L2 English users exhibit one significant difference between the present perfect phrases (A & C) and the simple past phrases (B & D). The nonbounded predicates in the present perfect (C) are rated significantly more continuable than in the simple past (D). This trend is also present in the bounded conditions: the bounded predicates in the present perfect (A) are rated more continuable than in the simple past (B), but this difference is not significant. Among the L2 English users, present perfect phrases are rated as more continuable than those in the simple past by approximately one third of a point for both bounded and nonbounded predicates. This trend within the L2 English group is not exhibited by the English NS group who seem to disregard the tense-aspect construction when making decisions concerning continuability. Instead, the English NS group relies entirely on boundedness. These different trends exhibited by the groups indicate that, on the whole, the L2 English users interpret tense-aspect as a factor that meaningfully relates to the continuability of a phrase, especially when the predicate is nonbounded. The more nativelike performance in Condition C [PPerf –BND] may emerge from an association or familiarity with the nonbounded functions of the present perfect.

Overall, these comparisons indicate that each group associates continuability with a different grammatical feature. The L2 English group associates the tense-aspect construction with changes in continuability. Specifically, they rate the present perfect more continuable than the simple past regardless of boundedness. The fact that the L2 English users associate the present perfect with continuability suggests that the continuative function is more central to the meaning of the present perfect for these L2 learners than, for example, the resultative function is. That is, if the L2 English users associate the present perfect with the resultative function, then it is likely that they would rate it equally continuable or less continuable than the simple past because the resultative function denotes a completed event. Instead, they associate the present perfect with ongoing or continuable occurrences; they understand the present perfect as less completive than the simple past. This association is qualitatively different from the one exhibited by the English NS group whose members associate the boundedness distinction with variation in continuability ratings. They show a preference for using boundedness to discern continuability, rating nonbounded predicates as more continuable than bounded ones regardless of the tense-aspect construction used.

ENGLISH PROFICIENCY. The second set of comparisons concerns the L2 English users divided by proficiency and the English NS group with a focus on the L2 English users. The Advanced and Int.-High groups are considered together; they are compared and contrasted as necessary. With one exception, the Int.-Low and Low groups did not show evidence that they understand the relationship between continuability and boundedness or tense-aspect. As such, their results are not discussed here under the assumption that they have not acquired either relationship noted above.

The continuability ratings of the Advanced group exhibit a meaningful difference within the bounded predicates, and these ratings indicate that there is a distinction forming between the bounded (A & B) and nonbounded predicates (C & D). The ratings of the Advanced group significantly differ from the lowest two proficiency groups in Condition A [PPerf +BND], but do not differ from the English NS group. This finding is partially indicative of a nativelike preference for associating continuability with boundedness, but there is no clear contrast between the simple past conditions (B & D) to clearly justify such a conclusion. If this finding is indicative of a shift in preferences from associating continuability with tense-aspect to associating it with boundedness, the only evidence to support this proposal is the finding that the Int.-High group's ratings for Condition A [PPerf +BND] do not differ from any other group. That is, three stages might be visible: nonnativelike (Low & Int.-Low), partially nativelike (Int.-High), and fully nativelike (Adv.). The emergence of three stages indicates that the L2 users' understanding of boundedness becomes more nativelike as their proficiency increases, but it is not distinctly measurable until advanced proficiency. This finding concerning Condition A [PPerf +BND] is not an especially illuminating finding in itself, but an interesting trend emerges when considering boundedness and tense-aspect together. The Advanced group rates continuability in a more nativelike manner in the present perfect than in the simple past. That is, the nativelike contrast between bounded and nonbounded predicates only emerges within this group in the present perfect conditions (A & C); no contrast is present in the simple past conditions (B & D). The Int.-High group seems to also exhibit this tendency but to a lesser extent. The finding that continuability rating are more nativelike in the present perfect than the simple past indicates that the more advanced groups parse the relation between boundedness and continuability more accurately in the present perfect than they do in the simple past. It is likely that this is due to typological effects. On the one hand, the simple aspect of the simple past very often yields a completive or perfective reading in context. The dominance of this reading may lead L2 learners to not attend to boundedness distinctions. On the other hand, accurately interpreting the function of the present perfect requires that the learner attend to the boundedness distinctions. As such, they will attend to and find use for these distinctions in the present perfect while ignoring them in the simple past.

The continuability ratings for the Int.-Low group present one significant finding. The members of this group rate the continuability of Condition D [SPast –BND] as significantly lower than the other conditions. The ratings for this condition indicate that these L2 users understand that a nonbounded predicate in the simple past is less able to be continued than a nonbounded predicate in the present perfect. Considering the observation L2 users overall rely on tense-aspect to determine continuability, this difference is not surprising; however, it is surprising that this group rates the simple past phrases (D) as less continuable than the present perfect ones (C). This is opposite the trend that emerges among the L2 English users as a whole.

These findings begin to clarify how proficiency affects how the L2 English users understand continuability. The first clarification concerns at what proficiency the L2 English users show any association between continuability and one of the features. Only the Int.-High and the Advanced groups exhibit any meaningful associations. This finding indicates that there is a considerable proficiency threshold that learners have to cross before they are associate continuability with a grammatical feature. Once they reach that threshold, they associate continuability first with the present perfect, rating it more continuable than the simple past regardless of the boundedness of the verb phrase. When their proficiency increases further, they continue to associate continuability with the present perfect, but they also begin to distinguish bounded and nonbounded predicates for continuability in the present perfect. This trend suggests that the boundedness distinction is unnoticed or not attended to in the simple past, but it becomes meaningful for them in the present perfect. It is more meaningful in the present perfect because it can be used to distinguish the function of the present perfect in that context.

FIRST LANGUAGE. The third set of comparisons concerns the L2 English users divided by first language and the English NS group with a focus on the L2 English users. Each linguistic group is treated separately.

The ratings of the Arabic group indicate a preference for associating continuability with tense-aspect and not boundedness. The Arabic group's ratings for Conditions A [PPerf +BND] and B [SPast +BND] are nativelike. That is, they rate bounded predicates as continuable in a manner similar to the English NS group. Although this finding provides initial support for the claim that the Arabic group makes the bound-nonbound distinction in a nativelike way, this claim is not fully supported. For this claim to be fully supported, Conditions C [PPerf –BND] and D [SPast –BND] would need to be different from Conditions A [PPerf +BND] and B [SPast +BND], respectively. The ratings for Conditions A [PPerf +BND] and C [PPerf –BND] are different, but Conditions B [SPast +BND] and D [SPast –BND] are not. Specifically, the Arabic group rates nonbounded predicates in the present perfect as more continuable than bounded predicates in the present perfect, but they rate predicates similarly in the simple past. Taken together, these findings indicate that the bound-nonbound distinction is only understood in a nativelike manner in the present perfect. This suggests that the Arabic group does not understand boundedness as a meaningful feature in the simple past but they use it to accurately parse the present perfect.

The ratings of the Chinese group give no indication that these speakers associate continuability with either boundedness or tense-aspect. The Chinese group's ratings for Conditions A [PPerf +BND] and B [SPast +BND] are not nativelike. Specifically, they are significantly higher than those of the English NS group. This finding could indicate that the Chinese group understands bounded predicates in both tense-aspect constructions as more continuable than native speakers do; however, this finding is probably vacuous because the Chinese group does not measurably distinguish boundedness or tense-aspect in their ratings. As such, although the difference is significant, it may not be meaningful. That is, the data either suggest that the Chinese group rates bounded predicates as more continuable than nonbounded ones with no effect for tense-aspect construction, or they indicate that the Chinese group fails to see any meaningful relations for continuability herein. The present investigation favors the latter explanation due to the preponderance of evidence in this task and in the SPRT that indicates that the Chinese group is affected by manipulations in tense-aspect construction; the former explanation would be viable if this were not the case.

The ratings of the Other group indicate a clear preference for associating continuability with tense-aspect and not boundedness. The Other group's ratings for Condition A [PPerf +BND] are significantly higher than those of the English NS group. This indicates that the Other group understands bounded predicates in the present perfect as more continuable than native speakers do. The fact that the Other group rates bounded predicates in the simple past (B) in a nativelike way indicates that the distinction found in Condition A [PPerf +BND] may be one of tense-aspect and not of boundedness. Support for the possibility that the Other group is rating tense-aspect and not boundedness is found in the significant difference in rating between Conditions A [PPerf +BND] and B [SPast +BND] and the meaningful difference between Conditions C [PPerf –BND] and D [SPast –BND]. That is, the Other group rates predicates in the present perfect as more continuable than those in the simple past regardless of boundedness. This association with continuability and the present perfect is nonnativelike, but it causes their ratings in Condition C [PPerf –BND] to be similar to those of the English NS group despite a qualitatively different manner of reaching similar ratings. Overall, these findings indicate that the Other group understands continuability in relation to tense-aspect and not boundedness. That is, they understand predicates in the present perfect as more continuable than those in the simple past.

These findings provide evidence of some meaningful differences between the L1 groups. Despite being no more proficient than the other two groups, the Arabic group behaves like the Advanced group above, rating bounded and nonbounded predicates in a nativelike manner in the present perfect. This suggests that the L1 Arabic users are either i) noticing and attending to boundedness more than the other proficiency-matched groups or ii) they are positively transferring some feature of the L1 grammar. This case of positive transfer was not hypothesized (§3.5). It is possible that the L1 distinction between the Arabic perfect, which is used to translate the bounded functions of the English present perfect, and the Arabic past continuous, which translates the nonbounded functions. Such a positive transfer would result in ratings like those exhibited by the Arabic group, but it is unclear why a similar positive transfer would not also affect the ratings in the simple past. As proposed above, it is possible that the boundedness distinction goes unnoticed in the simple past because of its minimal effect on the character of the verb phrase, but there is no clear evidence to suggest this in the present investigation. The ratings of the Chinese group indicate no associations between continuability and either set of grammatical features. This result was hypothesized based on the optionality present in the aspectual marking of Chinese (§3.5). The Other group provides a general baseline for comparison. The members of this group associate continuability with the respective tense-aspect constructions as discussed above. This finding may be an aggregated effect caused by positive L1 transfer from similar language systems, but the present investigation does not have the data required to make such a conclusion. Overall, each group associates a different grammatical feature with continuability: the Arabic group associates it with changes in boundedness in the present perfect, the Other group associates it with changes in tense-aspect construction, and the Chinese group has no clear association.

RQ3, CONCLUSIONS. The data from conditions A-D from the second administration support the following conclusions to RQ3 and its subquestions.

Main question: Yes, the grammars of instructed adult L2 learners are sensitive to contrasts in boundedness.

Subquestion (a): No, the metalinguistic judgments of adult learners are not qualitatively similar to the judgments of the adult native speakers: the L2 users associate continuability with tense-aspect, rating the present perfect as more continuable and less complete than the simple past, while the English native speakers associate this distinction with boundedness, rating nonbounded predicates more continuable and less complete than nonbounded ones.

Subquestion (b): Yes, the metalinguistic judgments of the adult learners are affected by differences in L2 English proficiency: the threshold at which L2 users become sensitive to such contrasts is upper-intermediate proficiency. Only the Int.-High and Adv. groups indicated meaningful understanding of the relation between phrasal or aspectual bounding and continuability.

Subquestion (c): Yes, the metalinguistic judgments of the adult learners are affected by first language: the Arabic group exhibits a nativelike association between boundedness and continuability in the present perfect, the Chinese group demonstrates no clear associations, and the Other group exhibits a nonnativelike association between tense-aspect construction and continuability.

6.2.2 CURRENT RELEVANCE CONTRASTS, RATING TASK.

ADVERBIAL MODIFIERS. As discussed previously in section 4.5.2, Conditions E & F collect speaker judgment data concerning current relevance overtly marked by an adverbial phrase in the present perfect and the simple past. Refer to the below table for a summary of the manipulations in these conditions. There is only one meaningful comparison for current relevance ratings at the level of the condition: [+CR] & [–CR] adverbial phrases. Each condition is broken down into four subconditions based on the specific adverbial phrases used in the manipulations; refer to Table 4.6 above for a breakdown of these subconditions.

TABLE 6.4. Manipulations in adverbially marked current relevance for Conditions E & F.

Cdn.	Current Rel.	Grammatical Tense
E	[+CR]	Present Perfect
F	[–CR]	Present Perfect

Recall from section 3.5 that it is expected that adverbial manipulations of relevance affect the perception of current relevance. That is, it is hypothesized that [+CR] adverbial phrases (Condition E [+CR]) are rated more highly for current relevance than [–CR] ones (Condition F [–CR]).

IN TOTO. The first set of comparisons concerns the L2 English users as a whole and the English NS group. The current relevance ratings of the L2 English users do not differ by condition (E & F), which indicates that the L2 English users overall do not associate these adverbials phrases with current relevance. The English NS group's ratings for Condition E [+CR] are significantly higher than those for Condition F [–CR], which indicates that the native speakers associate these adverbials phrases with current relevance. The analysis for subconditions indicates that this difference between conditions is likely due to the rating differential for Subcondition E3 [+CR *at present*] & E4 [+CR *often* & *sometimes*] and F3 [–CR *at some point*] & F4 [–CR *until [x event]*], which manipulate temporal reference within an indefinite (pre-)present frame and iterativity, respectively. The ratings of the L2 English users are not significantly different in any subcondition, but the difference between Subconditions E3 [+CR *at present*] and F3 [–CR *at some point*] is the greatest at approximately half of a point.

Overall, there is no evidence that L2 English users as a group associate the manipulated adverbial phrases with current relevance. This finding is clearly opposed to the ratings of the English NS group, who rate the [+CR] adverbial phrases in more indicative of current relevance in Subconditions E2 [+CR *for [x time]*], E3 [+CR *at present*], and E4 [+CR *often* & *sometimes*] than the [–CR] adverbial phrases in the matching subconditions.

ENGLISH PROFICIENCY. The second set of comparisons concerns the L2 English users divided by proficiency and the English NS group with a focus on the L2 English users. The Advanced group is the only L2 English group to rate [+CR] and [–CR] adverbial phrases differently. As such, this groups results are the only results discussed in detail. Other proficiency groups are mentioned as necessary.

The current relevance ratings of the Advanced group exhibit a meaningful difference between [+CR] and [–CR] adverbial phrases. Of all the L2 English groups, the Advanced group is the only one whose scores indicate any association between adverbial phrase and current relevance. This difference in ratings is due to the significant difference between Subconditions E3 [+CR *at present*] and F3 [–CR *at some point*], which manipulate temporal reference within an indefinite (pre-)present frame. This finding indicates that, like the English NS group, the Advanced group understands the relation between certain adverbial phrases and current relevance. It is not surprising that the Advanced group first understands the association between these adverbial phrases and current relevance because phrases within this category make more explicit the currency of the action or its resultant state. It is, however, surprising that no other categories of adverbial phrases exhibited a significant effect. This is true even among the English NS group. Their ratings for the [+CR] adverbial phrases in Subconditions E2 [+CR *for [x time]*], E3 [+CR *at present*], and E4 [+CR *often* & *sometimes*] follow the predicted association pattern, but only the scores for Subconditions E3 [+CR *at present*] and F3 [–CR *at some point*] are significantly different. Perhaps, one of the more counter-intuitive findings emerges from the results of Subconditions E1 [+CR *since V-ing*] and F1 [–CR *after V-ing*], which contrast *since* and *after* phrases. Both the Advanced and the English NS groups rate the [–CR] *after* phrase as more indicative of current relevance than the [+CR] *since* phrase, which is both contrary to the predictions based on typological frequency and contrary to the ratings of all of the other L2 English groups. The latter contradiction is probably due to the effect of instruction; the contrasting usage of these two phrases is explicitly taught at the IEP from which these students were selected; however, it is unclear at what proficiency the former contradiction begins to emerge.

These results indicate two things. First, only the most advanced L2 English users associate [+CR] adverbial phrases with current relevance. Second, the ratings for the subconditions are not consistent within or between groups. That is, certain subconditions show greater contrasts than others within the English NS and Advanced groups, and which subconditions exhibit the greatest contrasts is not consistent between groups. The most consistent finding is that Subcondition E3 [+CR *at present*], which manipulates temporal reference within an indefinite (pre-)present frame, is assigned the highest rating for current relevance.

FIRST LANGUAGE. The third set of comparisons concerns the L2 English users divided by first language and the English NS group with a focus on the L2 English users. Only the Arabic and Other groups are treated here. The ratings of the Chinese group show no meaningful distinctions between conditions or subconditions.

The current relevance ratings of the Arabic group exhibit a meaningful difference between [+CR] and [–CR] adverbial phrases. Of all the L2 English groups, the Arabic group is the only one whose scores indicate significant understanding of the contrasts in current relevance between the manipulated adverbial phrases. Their current relevance ratings for [+CR] adverbial phrases are significantly higher than their ratings for [–CR] adverbial phrases as a whole. Unlike in the previous sets of comparisons, this trend does not appear to be due to the contrasts in Subconditions E3 [+CR *at present*] and F3 [–CR *at some point*]; rather, Subconditions E1 [+CR *since V-ing*], E3 [+CR *at present*], and E4 [+CR *often* & *sometimes*] are all rated as more indicative of current relevance than the matching [–CR] subconditions. These findings suggest that the Arabic group understands the associations of the [+CR] and [–CR] adverbial phrases. This indicates that they are using the adverbial phrases as cues for tense-aspect. Such a use of adverbial phrases may be a strategy transferred into English from Arabic, which uses adverbial markers to capture relations not explicitly marked on the inflected verbal roots.

Like the current relevance ratings of the Arabic group, those of the Other group exhibit a meaningful difference between [+CR] and [–CR] adverbial phrases. The Other group's current relevance ratings do not provide clear evidence that they understand the contrast in Subconditions E3 [+CR *at present*] and F3 [–CR *at some point*], but there is a meaningful contrast between the scores for these subconditions that is approaching significant ($p = .121$). As this group is composed of L2 English users of varying L1 backgrounds, it is difficult to generalize this finding; however, it does indicate that the contrast observed in this subcondition is not unique to the Arabic group. As such, the finding that both groups are sensitive to this contrast may indicate that the Arabic speakers are not transferring some strategy from the L1 grammar but that adverbial phrases of this type are especially associated with the perception of current relevance.

These findings on the whole divide the L2 English users into two groups: those who notice a meaningful difference between the [+CR] and [–CR] adverbial phrases and those who do not. The former group is composed of the Arabic and Other groups, whose rating indicate a nativelike (or approaching nativelike) understanding of the contrast between the [+CR] and [–CR] adverbial phrases. The latter group is composed of the Chinese group, whose ratings indicate no such understanding. It is possible that the difference between these two groups is one of negative L1 transfer. The former group may associate a broader range of adverbial phrases with current relevance than the latter group because the first languages of the former group are less restrictive with this meaning than is Chinese for the latter group. That is, the relatively restrictive use of *le* as a marker of current relevance in Chinese may negatively affect this group's ability to generalize the concept to a wider range of adverbial markers that indicate varying aspects of current relevance in English. In other words, L1 Chinese users may have difficulty generalizing one form with one function to many forms with many functions, and this difficulty may result in their comparatively performance on this task.

TENSE-ASPECT CONSTRUCTIONS. As discussed previously in section 4.5.2, Conditions G & H collect speaker judgment data concerning current relevance indicated by the tense-aspect construction used. Refer to the below table for a summary of the manipulations in these conditions. There is only one meaningful comparison for current relevance ratings at the level of the condition: present perfect and simple past.

TABLE 6.5. Manipulations in tense-aspect construction for Conditions G & H.

Cdn.	Boundedness	Grammatical Tense
G	+BND	Present Perfect
H	+BND	Simple Past

Recall from section 3.5 that it is expected that manipulations in tense-aspect construction affect the perception of current relevance. That is, it is hypothesized that the sentences in the present perfect (Condition G) are rated more highly for current relevance than those in the simple past (Condition H).

IN TOTO. The first set of comparisons concerns the L2 English users as a whole and the English NS group. The current relevance ratings of the L2 English users do not differ by condition (G & H), which indicates that the L2 English users overall do not associate either tense-aspect construction with current relevance. The English NS group's ratings for Condition G [PPerf] are not significantly higher than those for Condition H [SPast], but these ratings do approach significance ($p = .122$). The lack of a significant difference in the English NS group makes the L2 English users' ratings difficult to parse, and suggest that the task may not adequately measure perceptions of current relevance in the verb phrase. A potentially interesting difference between the two groups is the

English NS group's significantly higher current relevance ratings than the L2 English users. This finding may indicate either that the L2 English group does not associate either construction with current relevance or the English NS group associates both constructions with current relevance (i.e. past actions and their results meaningfully affect the present regardless of their enduring character).

Overall, these results suggest that either this task is not sensitive enough to capture any meaningful differences between these conditions if any differences are indeed present or there is not sufficiency statistical power for the appropriate nonparametric statistics to capture such a minute difference.

ENGLISH PROFICIENCY. The second set of comparisons concerns the L2 English users divided by proficiency and the English NS group with a focus on the L2 English users. Within groups, no group rates the manipulated verb phrases differently; however, there are several interesting between group comparisons in Condition G [PPerf] especially that are worthy of further discussion.

The current relevance ratings of the L2 English groups in Condition G [PPerf] exhibit a positive trend as proficiency increases. The differences between the groups set them into three levels. The lowest level group comprises the Low and Int.-Low groups, whose ratings are significantly lower than those of the Advanced (Low: $p = .054$; Int.-Low: $p = .107$) and English NS groups (Low: $p < .001$; Int.-Low: $p = .072$). The middle level group comprises the Int.-High group, whose ratings are significantly lower than the English NS group alone ($p = .072$). The highest level group then comprises the Advanced and English NS groups, whose scores do not significantly differ from each other but are higher than the other proficiency groups. Maintaining the original four proficiency groups and treating the English NS group as the one with the highest proficiency, a Spearman correlation analysis reveals a weak, positive monotonic correlation between proficiency and current relevance rating ($\rho = .212$, $p < .001$). That is, the current relevance rating for the present perfect increases as proficiency increases.

These results indicate that English proficiency is an important factor in the separation of the present perfect from the simple past through the addition of the [current relevance] feature or, at least, with the association of the present perfect with this feature. The current relevance ratings for the present perfect slowly separate from those of the simple past as proficiency increases.

FIRST LANGUAGE. The third set of comparisons concerns the L2 English users divided by first language and the English NS group with a focus on the L2 English users. All groups are treated together.

The current relevance ratings of the L2 English groups do not provide clear evidence that can be used to meaningfully distinguish them based on first language. In Condition G [PPerf], the Chinese and Other groups are significantly different from the English NS group, but the Arabic group is not significantly different from any group. This provides some evidence that the ratings of the Arabic group are more nativelike than those of the other two L2 English groups. However, this evidence is not sufficient to suggest that these differences between groups and the lack thereof are meaningful.

RQ4, CONCLUSIONS. The data from Conditions E [+CR] & F [–CR] from the second administration support the following conclusions to RQ4 part i) and its subquestions.

Question 4(i): Yes, the grammars of instructed adult L2 learners of English are sensitive to contrasts in adverbially marked current relevance, but this sensitivity is not visible when the data are treated *in toto*.

Subquestion (i.a): Yes, the metalinguistic judgments of the adult learners are qualitatively similar to the judgments of the adult native speakers: within certain groups, the L2 users associate current relevance with the same kinds of adverbial phrases that native speakers do. This is especially true for adverbial modifiers that manipulate temporal reference within an indefinite (pre-)present frame and those that mark iterativity.

Subquestion (i.b): Yes, the metalinguistic judgments of the adult learners are affected by differences in L2 English proficiency: the threshold at which L2 users become sensitive to such contrasts is advanced proficiency. Only the Adv. group indicated meaningful understanding of the relation between the adverbial modifiers and current relevance.

Subquestion (i.c): Yes, the metalinguistic judgments of the adult learners are affected by first language: the Arabic and Other groups exhibit a meaningful association between adverbial modifiers and current relevance, whereas the Chinese group exhibits no clear association between the modifiers and current relevance.

The data from Conditions G [PPerf] & H [SPast] from the second administration support the following conclusions to RQ4 part ii) and its subquestions.

Question 4(ii): It is unclear whether the grammars of instructed adult L2 learners are sensitive to contrasts in current relevance marked using tense-aspect morphology.

Subquestion (ii.a): There is no evidence to indicate that the metalinguistic judgments of the adult learners are qualitatively similar to the judgments of the adult native speakers: neither the English NS group nor the L2 English users clearly differentiate current relevance marking between the simple past and the present perfect.

Subquestion (ii.b): Yes, the metalinguistic judgments of the adult learners are affected by differences in L2 English proficiency: there is a weak positive correlation between English proficiency and current relevance ratings in the present perfect.

Subquestion (ii.c): It is unclear whether the metalinguistic judgments of the adult learners are affected by first language: although the Arabic group slightly differs from the Chinese and Other groups, there is not sufficient evidence to indicate that these differences are motivated by L1 differences.

6.3 CONCLUSIONS.

The following sections summarize the above discussion and propose final conclusions that are justified based on the results of the present investigation. Each section addresses boundedness or current relevance individually. Within these sections, the discussion is focused on the effects of English proficiency and first language on the processing and acquisition of each feature.

6.3.1 BOUNDEDNESS.

The results of the present investigation support the following final conclusions concerning the real-time processing of and metalinguistic understanding of boundedness in past time constructions. Both the first and the second administration yield results indicating that adult L2 English users are sensitive to manipulations in boundedness;

furthermore, these results indicate that performance for both the SPRT and rating task is meaningfully affected by English proficiency and by L1 patterns.

ENGLISH PROFICIENCY. In the SPRT, the two intermediate groups and the Advanced group responded to the manipulations in grammatical tense. More pertinent to the present investigation, only the reading times of the Advanced group were affected by manipulations in boundedness as well. Specifically, this group experienced inhibition in reading times for nonbounded predicates in the present perfect. This finding indicates that only the processors of the Advanced group are able to incorporate both manipulations in boundedness and grammatical tense, and further that processing both manipulations is costly in nonbounded contexts in which the meaning of the present perfect and simple past differ aspectually.

In the rating task, the two intermediate groups and the Advanced group again responded to the manipulations in grammatical tense; however, the continuability ratings of the Int.-High group and the Advanced group were affected by manipulations in boundedness as well. Specifically, these groups rated nonbounded predicates in the present perfect as much more continuable than either bounded predicates in the present perfect or any predicates in the simple past. This finding demonstrates that these groups understand the relation between boundedness and continuability and that they understand how it interacts with tense-aspect in these two past time contexts.

The findings from the two tasks justify two conclusions. First, contexts in which nonbounded predicates are in the present perfect are more marked than the other contexts in the present investigation are. This markedness is observable in both processing and metalinguistic knowledge. It explains the meaningful differences in reading times and the variations in continuability ratings between this marked context and those matched for boundedness and grammatical tense in both tasks. The present investigation presumes that this markedness is due to the contrasts in meaning between the bounded and nonbounded functions of the present perfect. Similar contrasts do not manifest in the simple past, and as such the learner must come to separate these two sets of functions in the present perfect as clearly distinct from what is already present in the interlanguage grammar. Second, explicit grammatical knowledge of the interactions between boundedness and grammatical tense are observable at a lower proficiency than they are in real-time processing. That is, interlanguage development is observable offline before it is observable online as would be expected following any of the entrenchment-driven, usage-based language learning models.[50]

FIRST LANGUAGE. In the SPRT, there were clear differences in reading times between the Arabic and Chinese groups that cannot be explained by the captured personal history information or by English proficiency. The Arabic group performed more like the Advanced group than their average proficiency score would suggest (i.e. their reading times were affected by both manipulations in boundedness and grammatical tense). Conversely, the Chinese group performed just as their average proficiency score would predict (i.e. their reading times were affected only by manipulations in grammatical tense). Under the assumption that these two groups are matched for proficiency as the statistical results indicate, the groups should perform comparably; however, they exhibit quite different trends. These different trends indicate that the Arabic group experience some facilitation or benefit, which seems to be rooted in the transfer of L1 processing

[50] Examples include MacWhinney's (1999) Emergentism, Pienemann's (1998) Processability Theory, Ullman's (2004) Declarative/Procedural Model, among others.

strategies. Arabic marks boundedness semantically and does not have an English-like present perfect, so it is not clear what feature of the processor is being transferred. Perhaps, the Arabic group is more accustomed to the compositional construction of tense-aspect over the entire verb phrase, or they benefit from the transfer of their oppositional perfect and past continuous grammatical tenses, which roughly correspond to the bounded simple past & bounded present perfect and nonbounded present perfect, respectively.

In the rating task, the Arabic group again assessed continuability in a more nativelike manner than the Chinese group, even out-performing the higher-proficiency Other group. The Arabic group rates continuability based on both manipulations in boundedness and grammatical tense in a manner that is again similar to the Advanced group. Assuming that the descriptive knowledge of the interaction of these two features precedes any observable manifestation in processing, it should come as no surprise that the Arabic group's continuability ratings are the most nativelike. This finding is likely the result of L1 transfer as all other captured variation between the groups is accounted for in the statistical model.

The findings from the two tasks justify two conclusions concerning L1 transfer. First, the Arabic group shows some indication of a benefit as a result of L1 transfer, but it is unclear what exactly is being transferred. It seems most likely that the Arabic perfect and past continuous are being functionally mapped onto the English bounded present perfect and nonbounded present perfect, respectively, and that the processing of and metalinguistic knowledge of boundedness exhibited in the tasks may be incidental to this form-function mapping that captures the perfective-imperfective distinction rather than an understanding of the semantic-syntactic composition of boundedness. Second, the Chinese group shows no evidence of transferring the resultative verb compound structure (which bounds activity predicates and coerces them to exhibit the characteristics of accomplishments) to the English sentences in this investigation. That is, although functionally similar to boundedness generation in English, the lack of structural similarity may inhibit the transfer of this L1 phenomenon to the interlanguage.

6.3.2 CURRENT RELEVANCE.

The results of the present investigation support the following final conclusions concerning the real-time processing of and metalinguistic understanding of current relevance (marked via adverbial phrase and via verbal morphology) in past time constructions. Both the first and the second administration yield results indicating that adult L2 English users are only somewhat sensitive to manipulations in current relevance. These results indicate that performance for both the SPRT and rating task is meaningfully affected by proficiency and by L1 patterns but that the majority of these effects are observable only in the adverbially marked conditions.

ENGLISH PROFICIENCY. In the SPRT, only the Advanced group responded to the manipulations in current relevance. Their reading times were slower in the [+CR] conditions than in the [–CR] conditions regardless of grammatical tense. That is, the presence of [+CR] modifiers appears to slow down or inhibit processing. This tendency is likely to be due to the complexity of incorporating the additional aspectual properties of the [+CR] modifiers. It seems likely that the processor of Advanced group is the first to successfully integrate phrasal aspect and adverbially marked aspect. In addition, their reading times showed unexpected facilitation in the [–CR] present perfect condition when compared to the [–CR] simple past condition. This finding is unexpected for two reasons. First, the present perfect is more complex than the simple past and it is learned later, which both suggest that the present perfect would be read more slowly than the simple past. Second, [–CR] modifiers clash with the semantics of the present perfect, which should inhibit processing like other semantic mismatches do.

In the rating task, the Advanced group is again the only group to respond to these manipulations. In those conditions wherein the participants rated adverbial modifiers, they accurately associated higher current relevance ratings with the predicted [+CR] modifiers and lower ratings with the [–CR] adverbial modifiers. In those conditions wherein they rated tensed verb phrases, they associated the present perfect with a higher current relevance rating than the simple past. Interestingly, there is a small correlation between English proficiency and current relevance ratings for the present perfect in this second task, which indicates that the interlanguage grammars of the L2 users are developing toward a more nativelike rating as proficiency increases.

The findings from both tasks support two conclusions. First, only high proficiency L2 English users are able incorporate manipulations in current relevance during processing, and they are the only ones whose grammars understand the relation between current relevance and both adverbial modifiers and verbal morphology. Second, there is evidence to suggest that this understanding is a gradual process that emerges as proficiency increases.

FIRST LANGUAGE. In the SPRT, there is no evidence to suggest that first language affects the processing of the manipulations in current relevance in the present investigation. Instead, the evidence suggests that the Chinese group processes the two manners in which boundedness is composed (object quantification and motion-goal) in different ways.

In the rating task, the Arabic and the Other groups respond to the manipulations in adverbially marked current relevance, but the Chinese group does not; no group indicates a clear association between the verbal morphology and current relevance rating. The Arabic and Other groups demonstrated associated higher current relevance ratings with the [+CR] adverbial modifiers and lower ratings with the [–CR] adverbial modifiers. Due to the vast differences between Arabic and the languages captured in the Other group, it is impossible to definitively know whether these results are due to positive L1 transfer for the Arabic group, which helped them to perform better than expected, or negative L1 transfer for the Chinese group, which caused them to perform worse than the proficiency-matched Arabic group. Due to the similarity between the Arabic and Other groups, it is assumed that it is more likely that the Chinese group's results are due to negative transfer. It is possible that their performance is due to English's more numerous adverbial modifiers that indicate current relevance. That is, the lack of a single form or set of forms onto which the function of *le* (the CR marker) can be singularly mapped may inhibit the acquisition of adverbial current relevance marking in English.

The findings from both tasks combined justify two conclusions. First, only the most advanced proficiency L2 English users showed variations in performance that are due to these manipulations in current relevance. This finding indicates that the association between current relevance and the present perfect is acquired later than the association between boundedness and the functions of the present perfect. Second, the relation between adverbial modifiers and current relevance is more transparent to L2 users than the relation between verbal morphology and current relevance. This finding supports the well-noted trend that adult learners rely on adverbial modifiers to mark tense-aspect before they do so using verbal morphology.

CHAPTER 7 RECOMMENDATIONS

This chapter briefly reflects on the present investigation. The first section considers some of the limitations of this investigation (§7.1). It also discusses how these limitations may have affected the study and its results. The second section offers some recommendations for future research (§7.2). These recommendations address the limitations set out in the prior section, and they offer avenues of research that might be fruitful based on the results of this investigation. The chapter concludes with brief suggestions concerning how the results of this research may be applied in English as a second language teaching (§7.3).

7.1 LIMITATIONS OF THE PRESENT STUDY.

The number and background of participants are prominent limitations of this investigation, as is often the case for research in second language acquisition. When the study was proposed, the population from which the samples were to be drawn was considerably larger and more linguistically diverse; however, this population changed dramatically before the administration of the tasks occurred. This change required that all L2 English participants be divided into subgroups based on both first language and English proficiency, which resulted in 'double dipping' participants in two subgroups and a reliance on statistical matching of proficiency and first language. This introduced potential confounds due to the balancing of subgroups. For example, if the High-Int. group has 20% more L1 Chinese users than any other group, it is difficult to extricate the effects of L2 proficiency on the results from those effects rooted in L1 transfer. The statistics indicated that the subgroups in this investigation were comparable, but the possibility of such confounds remains.

The reading tasks used in the present investigation were designed to provide context rich enough that the manipulations in each sentence would evoke variable responses. Although there was a great deal of effort exerted to make these sentences lexically and syntactically comparable, the sentences were long for a reading task of this kind (17 words, on average). As such, individual differences that affect the comprehension of long strings of written text may have affected the results. The two that seem most critical for this investigation are working memory capacity and L1 reading ability. Although the investigation assumes that these and similar IDs are normally distributed throughout the L1 and proficiency subgroups, this need not be the case. It may be reasonably expected that participants with higher L2 proficiency also have better working memory and L1 reading ability, which facilitated their language development. It is unclear what effect, if any, individual differences not captured by the modified BLP had on the results of this study.

The SPRT used in the first administration measures sentence-level processing. SPRTs best capture inhibited processing caused by a noticeable linguistic feature (i.e. slower reading times or ‘huh?’ responses resulting from some linguistic anomaly). The present investigation compared SPRT data between conditions in order to investigate both facilitation and inhibition. The pilot for this investigation indicated that a SPRT would be sensitive enough to measure comparative processing given these manipulations; however, it is very possible that the manipulations may not cause enough of a change in reading times to be statistically validated. This possibility was considered in the previous chapter whenever *meaningful* results were treated, but several failures to capture variations in processing caused by these manipulations are likely present in the previous analyses.

The rating task used in the second administration measures metalinguistic knowledge. The post-sentence rating statements were based on the task used by Wulff and colleagues (2009), which was designed to capture metalinguistic telicity judgments. The present investigation used this task to capture boundedness ratings of tensed predicates and current relevance ratings of adverbial modifiers and tensed predicates. The rating task was piloted on trained native English speakers and intermediate proficiency L2 English users to ensure that the task was internally valid and that the instructions were comprehensible. The participants of the pilot worked with the researcher individually or in small groups. They were highly motivated to understand the directions and examples and to complete the task to the best of their abilities. The results of this pilot were not replicated in the full-scale investigation. It is likely that the testing circumstances used in the present investigation did not appropriately scale up the pilot. Participants were given the same instructions in both instances, but there were several indications during administration that these instructions were not attended to, especially among some of the less motivated participants. A failure to fully understand the instructions and examples may have led to the failure to replicate the findings of the rating task when it was piloted.

Each administration of the study required the participants to complete three tasks. Depending on their motivation, ability to focus, and English proficiency, these tasks took between 45 and 75 minutes to complete. Several intermediate and low proficiency participants showed signs of fatigue during the respective reading tasks for each administration. Although the statistical analyses showed no effect for sentence order, it is still possible that participants were fatigued or overwhelmed by the considerable length of the tasks. Fatigue and the length of the task may have discouraged many of the participants, preventing them from putting forth their best effort.

Finally, the selection of statistical analyses (and selection of the data analyzed by them) is a limitation of this investigation. The present investigation relied on generalized linear modeling for between-group and within-group comparisons of parametric data. The models used were minimalist: they were generated using the fewest necessary variables. This method was chosen for its theoretical defensibility, not for an individual model's descriptive capacity (i.e. r^2 and BIC). As this research was largely exploratory, it was crucial to approach the results with specific meaningful comparisons as described in the proposal so as to avoid *p-hacking*, a practice of reanalyzing data to yield a target result (Simonsohn et al. 2014). In addition to more complex models, these results may be well analyzed using mixed linear modeling, survival analysis,[51] or nonlinear analysis.[52] In order to encourage principled research using these results, the data have been posted to the author's user profile on ResearchGate, the most active research-sharing platform for scientists and researchers at the time of writing.

7.2 Directions for future research.

This section continues the discussion of the limitations above while emphasizing possible directions for future research. These proposals address methodological

[51] Survival analysis is widely used in the social sciences where researchers are interested in measuring the time until an event of interest occurs.
[52] Nonlinear methods, and fractal analysis in particular, have been used to investigate qualitative differences in SPRT data without making assumptions about the system that caused these differences (Wallot & Van Orden 2011).

limitations of the present investigation, and they address results worthy of further investigation.

As mentioned in the previous section, the number and background of participants limited the investigation. Additional research into L1 effects may seek to evaluate participants from language groups that have structures that vary from the target structure by degree. For example, to investigate the present perfect, it may be beneficial to sample participants from the following L1 groups due to the functional and formal similarities and differences between L1 structures and the English present perfect.

TABLE 7.1. Example L1 groups used to investigate L1 effects.

	Functionally similar	Formally similar
Spanish – *pretérito perfecto* ‘present perfect’	+	+
French – *passé compose* ‘compound past’	-	+
Chinese – *le* [CR]	+	-
Arabic – perfect	-	-

Incorporating these or similar L1 groups permits finer-grained analysis of L1 transfer based on the similarities and differences noted above. Such a change in sampling would clarify the effects of L1 transfer observed in the present study.

There are two ways we recommend to further investigate the effects of L2 proficiency in similar studies that seek to investigate the effects of L1 transfer and L2 proficiency. Both of which prevent assigning participants to multiple groups. First, if there are sufficient participants in each of the L1 groups, then each L1 group can be further subdivided by L2 proficiency. This method allows for very precise comparisons of the effects of proficiency within a language group and the effect of first language among participants with the same proficiency. For example, the performance of the L1 French group can be investigated by L2 proficiency, and the intermediate L2 proficiency groups can be compared across first languages. Comparisons like these are likely to be generalizable. Second, if the sampling population prohibits this method, a set of participants can be reserved for proficiency-based comparisons. This group can either be selected from L2 users whose first languages are not captured for the L1-based comparisons or can be selected and removed from the L1 groups. Given sufficient variation in the first languages of the members of these proficiency groups, the results should be generalizable with the caveat that L1 effects may be present but unaccounted for if certain language families are over-represented in one of more of the proficiency groups.

Boundedness and current relevance were selected as the main features to be investigated due to the importance afforded them in the literature on the English present perfect. This selection is not meant to imply that these are the only two critical features that may be used to distinguish the present perfect from the simple past. Another aspectual feature that this and prior research suggest may be important is *durativity* (the situation occurs over some length of time). Durativity was not used due to potential confounds in this feature's manipulation, but the facilitation observed in the nonbounded present perfect condition indicates that durativity may be central to the meaning of the present perfect for adult L2 English learners. Although the pilot for this investigation found no data to support this possibility, the effect of durativity on the processing and understanding of the present perfect should be investigated further.

In this investigation, the present perfect was only used in its simple form (*have* V-*en/-ed*). Although the vast majority of present perfect usages occur in this form (Schlüter 2000), there is some interesting variation between the bounded and nonbounded functions and marking with the progressive form (*have been* V-*ing*). The nonbounded functions are marked with progressive morphology more often (at a ratio of approximately 5 simple forms for every 1 progressive form) than the bounded functions are (at a ratio of approximately 40 simple forms for every 1 progressive form). This noticeable difference in progressive marking between the two sets of functions seems worthy of further investigation given the fact that boundedness alone only affected processing among the higher proficiency participants; the rest were most affected by grammatical tense-aspect marking. Comparing the processing of the present perfect and the present perfect progressive as well as that of the present perfect progressive and the past progressive would yield results that she light on the possible contribution of progressive aspect to acquisition of the semantics of the present perfect.

As mentioned in the previous section, the tasks used in the present investigation were time-consuming and caused the participants to become fatigued. Future investigations may benefit from encoding L2 proficiency into the task such that lower proficiency L2 users are exposed to approximately half of the items, intermediate users are exposed to between two-thirds and three-fourths of the items, and advanced users alone are exposed to all of the items. This method would require additional participants, but it would likely limit task effects associated with fatigue.

The choice of online reading measures used in the present investigation weighed the costs and benefits of using these tasks, but they are not the only tasks that lend themselves to the study of tense-aspect acquisition and processing. Another prominent task measuring online processing receptively is eye tracking. An eye-tracking study using similar sentences would capture both fixations and regressions, which would provide more detailed data concerning how tense and aspect are being composed in each predicate and in the sentence as a whole. Two online measures of production are type-capturing and timed oral production. An online type-capturing task or a timed oral production task based on those used by Bardovi-Harlig (2002), Terán (2014), or Uno (2014) would capture the speed and accuracy with which L2 English users assign the present perfect in rich contexts in written or oral modalities. Tasks using either methodology would be especially interesting if research of this kind incorporates into the contexts adverbial modifiers that collocate with individual functions of the present perfect instead of modifiers that denote just durativity and iterativity. Research of this kind would shed light on how L2 learners produce in contexts in which a single function of the construction is strongly preferred by native English speakers.

The results of the investigation suggest several fruitful avenues of research in the field of second language acquisition. One avenue of research emerges from the Arabic group's performance on both reading tasks that investigated boundedness. Recall that they performed in a more nativelike manner than their proficiency predicts in the SPRT, and recall that they exhibited more nativelike continuability ratings than the Chinese group, which has comparable English proficiency, in the rating task. These results are suggestive of positive L1 transfer from Arabic; however, as noted in chapter six, it is unclear what feature or features were transferred to the interlanguage that generated these results. As such, more research is encouraged to determine what formal or functional features of Arabic present in the interlanguage grammar benefit the acquisition and processing of boundedness in past time contexts in L2 English.

Another set of results that merit further inquiry concerns the performance of the Chinese group. The Chinese group performed in a less nativelike manner than anticipated considering their average proficiency. In the previous chapter, it was suggested that this performance is the result of negative transfer of the singular morpheme by which current relevance is marked in Chinese (sentence final *le*) to the multiple phrases by which it is indicated in English. It may be fruitful to investigate this suggestion via a task that queries understanding of select categories of adverbial modifiers that collocate with the present perfect (i.e. actualization, iterative, durative, etc.). Based on research concerning the semantics of *le*, the task could ask L2 English users to rate to what degree exemplar phrases from these categories exhibit semantic features similar to that of *le* (i.e. contemporaneousness, relevance, and durativity; cf. Li et al. 1982). Results of an inquiry like this one would help determine to what phrases, if any, L1 Chinese learners of English are transferring the features of *le*. Once this information is known, they could be instructed to apply these features to additional phrases and to the perfect system in English.

One set of curious results from the Chinese group in the SPRT fell outside the scope of the present investigation but may yield interesting results concerning the processing of boundedness in English and Chinese. The Chinese group exhibited different processing for quantified object predicates and motion-goal predicates. Recall that boundedness was manipulated in two manners in the SPRT (§4.5.1). All conditions that addressed the effects of boundedness were manipulated using object quantification, and all conditions that addressed current relevance were motion predicates with a defined locative goal. The Chinese group performed significantly faster in the current relevance conditions (E-H) than in the boundedness conditions (A-D). It is unclear whether or not these processing differences are due to the manner in which boundedness was manipulated or due to the effects of the adverbial modifiers. The Arabic and Other groups did not exhibit such a quantitative difference between the two sets of conditions, which suggest that the difference observed within the Chinese group is due to the former possibility. Further research is necessary to disentangle the manipulations in boundedness from the effects of adverbial modification. We recommend that this research be done in both the first and second language. If differences in processing were found, research from both languages would be necessary to determine whether the two manners of bounding predicates are processed differently in both languages or only in L2 English. The former indicates that the speakers are likely transferring some processing strategy from the L1 grammar. The latter indicates some specific inhibition when processing motion-goal structures in the second language whose cause is not apparent.

7.3 Pedagogical applications.

This section concerns how the results of this investigation may benefit the field of English as a second language pedagogy. The results suggest two possible lessons that may improve learner outcomes.

This investigation found that nonbounded predicates facilitate the processing of the present perfect and its separation from the simple past. This finding suggests that it may be beneficial to use the nonbounded functions to help learners to differentiate the present perfect from the simple past. Practically, this means using lessons that incorporate nonbounded predicates (activities and states) to encourage learners to associate the present perfect with the consequent state and not the anterior situation. Bounded predicates need not be eschewed entirely. This research does not indicate that bounded predicates are more complex or would make learning more difficult; bounded predicates just seem to obscure the differences of the present perfect and the simple past for learners who less familiar with the structure. Using nonbounded predicates in the classroom may enable learners to distinguish these grammatical tenses more effectively than they do without nonboundedness being used to emphasize the stativity of the structure.

The investigation found that participants over-relied on the first-position adverbial modifiers to the detriment of their tense-aspect processing. This finding mostly emerges in the SPRT results in which participants exhibited no differences in processing between tense-aspect constructions. Although the L2 users seem to over-rely on the adverbial modifiers in the SPRT, the results of the rating task indicate that they do not yet associate these modifiers with the current relevance. Together, the results of both tasks suggest that the L2 English users are trying to use the modifiers to assign temporality to the sentence but are unable to do so. The results from Bardovi-Harlig (2002), Terán (2014), and Uno (2014) suggest that English learners benefit from exposure to the present perfect in adverbially modified contexts. The results of the present investigation support these prior findings, and they indicate that modifiers that denote (pre-)present temporal reference and those that indicate iterativity may be especially beneficial since L2 users already seem to associate them with current relevance.

REFERENCES

Andersen, Roger W. 1983. Pidginization and creolization as language acquisition. Rowley, MA: Newbury House.

Andersen, Roger W. 1984. The One to One Principle of interlanguage construction. *Language Learning* 34. 77-95.

Andersen, Roger W. 1989. La adquisicion de la morfologia verbal [The acquisition of verbal morphology]. *Linguistica* I. 89-141.

Andersen, Roger W. 1990. Models, processes, principles, and strategies: Second language acquisition in and out of the classroom. In Bill VanPatten & James F. Lee (Eds.), *Second language acquisition—Foreign language learning*, 45-68. Philadelphia: Multilingual Matters.

Andersen, Roger W. 1991. Developmental sequences: The emergence of aspect marking in second language acquisition. In T. Huebner & C. A. Ferguson (Eds.), *Crosscurrents in second language acquisition and linguistic theories (Language acquisition and language disorders 2)*, 305-324. Philadelphia: John Benjamins.

Andersen, Roger W. 1993. Four operating principles and input distribution as explanations for underdeveloped and mature morphological systems. In K. Hyltenstam and Ä. Viborg (Eds.), *Progression and regression in language*, 309-339. New York: Cambridge University Press.

Andersen, Roger W. 2002. The dimensions of "pastness." In Rafael Salaberry & Yasuhiro Shirai (Eds.), *The L2 acquisition of tense-aspect morphology*, 79-106. Philadelphia: John Benjamins.

Andersen, Roger W. & Yasuhiro Shirai. 1994. Discourse motivation for some cognitive acquisition principles. *Studies in Second Language Acquisition* 16(2). 133-156.

Andersen, Roger W. & Yasuhiro Shirai. 1996. The primacy of aspect in first and second language acquisition: The pidgin-creole connection. In William C. Ritchie & Tej K. Bhatia (Eds.), *Handbook of second language acquisition*, 527-270. San Diego: Academic Press.

Anderson, Lloyd B. 1982. The "perfect" as a universal and as a language-specific category. In Paul J. Hopper (Ed.), *Tense-aspect: Between semantics & pragmatics*, 227-265. Philadelphia: John Benjamins.

Anderson, Sarah E., Teenie Matloc, Caitlin M. Fausey & Michael J. Spivey. 2008. On the path to understanding the on-line processing of grammatical aspect. In *Proceedings of the 30th CogSci conference*. 2253-2258.

Antinucci, Francesco & Ruth Miller. 1976. How children talk about what happened. *Journal of Chile Language* 3. 169-189.

Ayoun, Dalila & M. Rafael Salaberry. 2008. Acquisition of English tense-aspect morphology by advanced French instructed learners. *Language Learning* 58(3). 555-595.

Badawi, Elsaid, Michael G. Carter & Adrian Gully. 2004. *Modern written Arabic: A comprehensive grammar.* New York: Routledge.

Baggio, Giosuè, Michiel van Lambalgen & Michiel Van Lambalgen. 2008. Computing and recomputing discourse models: An ERP study. *Journal of Memory and Language* 59(1). 36-53.

Barsalou, Lawrence W. 1987. The instability of graded structure: Implications for the nature of concepts. *Concepts and conceptual development: Ecological and intellectual factors in categorization* 1987. 101-140.

Bardovi-Harlig, Kathleen. 1992. The relationship of form and meaning: A cross-sectional study of tense and aspect in the interlanguage of learners of English as a second language. *Applied Psycholinguistics* 13. 253-278.

Bardovi-Harlig, Kathleen. 1998. Narrative structure and lexical aspect: Conspiring factors in second language acquisition of tense-aspect morphology. *Studies in Second Language Acquisition* 20. 471-508.

Bardovi-Harlig, Kathleen. 2000a. The aspect hypothesis. *Language Learning* 50. 191-275.

Bardovi-Harlig, Kathleen. 2000b. Tense and aspect in second language acquisition: Form, meaning, and use. *Language Learning Monograph Series.* Malden, MA: Blackwell.

Bardovi-Harlig, Kathleen. 2002. Another piece of the puzzle: The emergence of the present perfect. *Language Learning* 47(3), 375-422.

Bardovi-Harlig, Kathleen & Anna Bergström. 1996. The acquisition of tense and aspect in SLA and FLL: A study of learner narratives in English (SL) and French (FL). *Canadian Modern Language Review.*

Bardovi-Harlig, Kathleen & Dudley W. Reynolds. 1995. The role of lexical aspect in the acquisition of tense and aspect. *TESOL Quarterly* 29. 107-131.

Bates, Elizabeth A., Inge Bremerton & Lynn Snyder. 1988. *From first words to grammar: individual differences and dissociable mechanisms.* New York: Cambridge University Press.

Bayley, Robert. 2013. Data analysis: Quantitative approaches. In M. Rafael Salaberry & Llorenç Comajoan (Eds.), *Research design and methodology in studies on L2 tense and aspect*, 357-390. Boston: De Gruyter Mouton.

Beck, Sigrid & William Snyder. 2001. Complex predicates and goal PP's: evidence for a semantic parameter. In: Anna H.-J. Do, Laura Dominguez & Aimee Johansen (Eds.), *Proceedings of the 25th Annual Boston University Conference on Language Development* 1, 114-122. Somerville, MA: Cascadilla Press.

Bickerton, Derek. 1981. *Roots of language (Classics in linguistics 3)*. Berlin: Language Science Press.

Binnick, Robert I. 1991. *Time and the verb: A guide to tense and aspect*. Oxford: Oxford University Press.

Birdsong, David, Libby M. Gertken & Mark Amengual. 2012. *Bilingual language profile: An easy-to-use instrument to assess bilingualism.* COERLL, University of Texas at Austin: https://sites.la.utexas.edu/bilingual/

Bishop, George F. 1987. Experiments with the middle response alternative in survey questions. *Public Opinion Quarterly* 51. 220-232.

Bloom, Lois, Karin Lifter & Jeremie Hafitz. 1980. Semantics of verbs and the development of verb inflection in child language. *Language* 56. 386-412.

Borer, Hagit. 1994. The projection of arguments. In E. Benedicto and J. Runner (Eds), *Functional projections (University of Massachusetts Occasional Papers in Linguistics 17)*, 19-47. Amherst, MA: GLSA.

Brennan, Jonathan & Liina Pylkkänen. 2010. Processing psych verbs: Behavioral and MEG measures of two different types of semantic complexity. *Language and Cognitive Processes* 25(6). 777-807.

Brinton, Laurel J. 1988. *The development of English aspectual systems*. New York: Cambridge University Press.

Bronckart, Jean-Paul & Hermine Sinclair. 1973. Time, tense, and aspect. *Cognition* 2. 107-130.

Brown, Roger. 1973. *A first language: The early stages*. Cambridge, MA: Harvard University Press.

Buczowska, Ewa & Richard M. Weist. 1991. The effects of formal instruction on the second-language acquisition of temporal location. *Language Learning* 41. 535-554.

Bybee, Joan. 1985. *Morphology: A study of the relation between meaning and form*. Philadelphia: John Benjamins.

Bybee, Joan. 2008. Usage-based grammar and second language acquisition. In Peter Robinson & Nick C. Ellis (Eds.), *Handbook of cognitive linguistics and second language acquisition*, 216-236. New York: Routledge.

Bybee, Joan & Östen Dahl. 1989. The creation of tense and aspect systems in the languages of the world. *Studies in Language* 13. 51-103.

Bybee, Joan, Revere Perkins & William Pagliuca. 1994. *The evolution of grammar: Tense, aspect and modality in the languages of the world*. Chicago: University of Chicago Press.

Carey, Kathleen. 1990. The role of conversational implicature in the early grammaticalization of the English perfect. *Proceedings of the Sixteenth Annual Meeting of the Berkeley Linguistics Society.* Berkeley, CA: Berkeley Linguistics Society.

Carter, N. T., R. P. Griffith, J. Feitosa, R. Moukarzel, M.-C. Kung, A. D. Lawrence & M. O'Connell. 2012. Explaining and predicting non-invariance across cultures using cultural uncertainty avoidance. In N. T. Carter & A. D. Mead (Chairs), *Recent developments in personality measurement invariance: Time, culture, and forms*. Symposium presented at the 27th Annual Meeting of the Society for Industrial and Organizational Psychology, San Diego, CA.

Chan, Wai Meng, Kwee Nyet Chin, Sunil Kumar Bhatt & Izumi Walker. 2012. *Perspective on individual differences and foreign language education (Studies in second and foreign language education 6).* Boston: De Gruyter Mouton.

Chan, Ho Leung, Jessica Finberg, Willie Costello & Yasuhiro Shirai. 2012. L2 acquisition of tense-aspect morphology: Lexical aspect, morphological regularity and L1 transfer. In Luna Filipovi & Katarzyna Jaszczolt (Eds.), *Space and time in languages and cultures: Linguistic diversity*, 181-204. Philadelphia: John Benjamins.

Chierchia, Gennaro. 2004. Scalar Implicatures, Polarity Phenomena, and the Syntax/Pragmatics Interface. In Adriana Belletti (Ed.), *Structures and Beyond,* 39-103. Oxford: Oxford University Press.

Clark, Eve V. 1990. On the pragmatics of contrast. *Journal of Child Language* 17. 417-431.

Comrie, Bernard. 1976. *Aspect.* Cambridge: Cambridge University Press.

Coll-Florit, Marta & Silvia P. Gennari. 2011. Time in language: Event duration in language comprehension. *Cognitive Psychology* 62(1). 41-79.

Comajoan, Llorenç. 2006. The aspect hypothesis: Development of morphology and appropriateness of use. *Language Learning* 56(2). 201-268.

Comrie, Bernard. 1976. *Aspect: An introduction to the study of verbal aspect and related problems*. New York: Cambridge University Press.

Comrie, Bernard. 1985. *Tense*. New York: Cambridge University Press.

Corder, S. P. 1971. Idiosyncratic dialects and error analysis. *IRAL - International Review of Applied Linguistics in Language Teaching* 9(2). 147-160.

Cromer, Richard F. 1968. *The development of temporal reference during the acquisition of language*. Harvard University PhD dissertation.

Cromer, Richard F. 1971. The development of the ability to decentre in time. *The British Journal of Psychiatry* 62. 353-365.

Cromer, Richard F. 1974. The development of language and cognition: The cognition hypothesis. In B. Foss (Ed.), *New perspectives in child development*, 184-252. Harmondsworth: Penguin.

Cromer, Richard F. 1976. The acquisition of syntactic structures by children. *Neurolinguistics* 5. 52-55.

Cruse, Alan D. 2000. *Meaning in language: An introduction to semantics and pragmatics*. Oxford: Oxford University Press.

Dahl, Östen. 1981. On the definition of the telic-atelic (bounded-nonbounded) distinction. In Philip Tedeschi & Annie Zaenen (Eds.), *Tense and aspect*, 79-90. New York: Academic Press.

Dahl, Östen. 1985. *Tense and aspect systems*. Oxford: Basil Blackwell.

Dahl, Östen & Eva Hedin. 2000. Current relevance and event reference. In Östen Dahl (Ed.), *Tense and aspect in the languages of Europe: Empirical approaches to language typology 20-6*, 385-401. Berlin: De Gruyter.

Davydova, Julia. 2011. *The present perfect in non-native Englishes: A corpus-based study in variation*. Boston, MA: Walter de Gruyter.

Declerck, Renaat. 1991. *Tense in English: Its structure and use in discourse*. New York: Routledge.

Declerck, Renaat. 2006. *The grammar of the English tense system: A comprehensive analysis (Topic in English linguistics 60.1)*. New York: Mouton de Gruyter.

Depraetere, Ilse. 1995. On the necessity of distinguishing between (un)boundedness and (a)telicity. *Linguistics and Philosophy* 18. 1-19.

Deutschbein, Max. 1939. Aspekte und aktionsarten im neuenglischen [Aspect and aktionsart in Modern Englishes]. *Neuphilologische Monatsschrift* 10. 129-148, 190-201.

Dietrich, Rainer, Wolfgang Klein, & Colette Noyau. 1995. *The acquisition of temporality in a second language (Studies in bilingualism 7).* Philadelphia: John Benjamins.

Dowty, David R. 1979. *Word Meaning and Montague Grammar*. Boston: Kluwer.

Dowty, David R. 1991. Thematic proto-roles and argument selection. *Language* 67(3). 547-619.

Dussias, Paola E. & Pilar Piñar. 2009. Sentence parsing in L2 learners: Linguistic and experience-based factors. In William C. Ritchie, & Tej K. Bhatia (Eds.), *The new handbook of second language acquisition*, 296-318. Bingley, UK: Emerald Group.

Eckman, Fred R. 1977. Markedness and the contrastive analysis hypothesis. *Language Learning* 27. 315-330.

Eifring, Halvor & Rolf Theil. 2005. *Linguistics for Students of Asian and African Languages*. University of Oslo.

Ellis, Nick 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.

Ellis, Nick C. 2008. Usage-based and form-focused language acquisition: The associative learning of constructions, learned attention, and the limited L2 endstate. In Peter Robinson & Nick C. Ellis (Eds.), *Handbook of cognitive linguistics and second language acquisition*, 372-405. New York: Routledge.

Ellis, Nick C., Kausar Hafeez, Katherine I. Martin, Lillian Chen, Julie Boland & Nuria Sagarra. 2012. An eye-tracking study of learned attention in second language acquisition. *Applied Psycholinguistics*. 1-33.

Ellis, Nick C. & Nuria Sagarra. 2010. The bounds of adult language acquisition: Blocking and learned attention. *Studies in Second Language Acquisition* 32. 553-580.

Elseness, Johan. 1997. *The present perfect and preterite in contemporary and earlier English: Topics in English linguistics 21.* New York: De Gruyter.

Farina, Christopher. 2014. Typological frequency and prototypicality: Processing English past time tense-aspect morphology (Poster). *Second Language Research Forum (SLRF) 33.* Columbia, SC.

Farina, Christopher. 2015. Processing endpoint prototypes in English past time tense-aspects. *Second Language Research Forum (SLRF) 34.* Atlanta, GA.

Felser, Claudia. 2005. Experimental psycholinguistic approaches to second language acquisition. *Second Language Research* 21(2). 95-97.

Fenn, Peter. 1987. *A semantic and pragmatic examination of the English perfect.* Tubingen: Gunter Narr Verlag.

Ferretti, Todd R., Marta Kutas & Ken McRae. 2007. Verb aspect and the activation of event knowledge. *Journal of Experimental Psychology, Learning, Memory, and Cognition* 33(1). 182-196.

Filppula, Markku. 1999. *The grammar of Irish English: Language in Hibernian style.* New York: Routledge.

Fletcher, Paul. 1979. The development of the verb-phrase. In Paul Fletcher & Michael Garman (Eds.), *Language acquisition: Studies in first language development.* New York: Cambridge University Press.

Fletcher, Paul. 1981. Description and explanation in the acquisition of verb-forms. *Journal of Child Language* 8. 93-108.

Folli, Raffaella & Gillian Ramchand. 2005. Prepositions and results in Italian and English: An analysis from event decomposition. In H. Verkuyl, H. De Swart, & A. van Hout (Eds.), *Perspectives on aspect*, 81-105. Boston: Kluwer.

Gass, Susan M. 1984. A review of interlanguage syntax: Language transfer and language universals. *Language Learning* 34. 115-132.

Gass, Susan M. & Josh Ard. 1983. Second language acquisition and the ontology of language universals. In W. E. Rutherford (Ed.), *Language universals and second language acquisition*, 33-67. Philadelphia: John Benjamins.

Gathercole, Virginia C. 1986. The acquisition of the present perfect: explaining differences in the speech of Scottish and American children. *Journal of Child Language* 13. 537-560.

Gennari, Silvia & David Poeppel. 2003. Processing correlates of lexical semantic complexity. *Cognition* 89(1), B27-B41.

Giacalone-Ramat, Anna. 1995. Tense and aspect in learner Italian. In Pier Marco Bertinetto, Valentina Bianchi, Östen Dahl & Marco Squartini (Eds.), *Temporal reference, aspect and actionality (Typological perspectives 2)*, 289-309. Turin: Rosenberg & Sellier.

Grimm, Jacob. 1819. *Deutsche Grammatik* [German grammar]. Göttingen: Bei Dieterich.

Hale, Kenneth & Samuel Jay Keyser. 1993. On argument structure and the lexical expression of syntactic relations. In K. Hale and J. Keyser (Eds), *The View from Building 20*, 53-109. Cambridge, MA: MIT Press.

Hale, Kenneth & Samuel Jay Keyser. 1998. The basic elements of argument structure. In *Papers from the UPenn/MIT Roundtable on Argument Structure and Aspect 32*, 73-118. Cambridge, MA: MIT Working Papers in Linguistics.

Hale, Kenneth & Samuel Jay Keyser. 2002. *Prolegomenon to a theory of argument structure*. Cambridge, MA: MIT Press.

Hawkins, Roger. 2009. Second language acquisition of morphosyntax. In William C. Ritchie & Tej K. Bhatia (Eds.), *The new handbook of second language acquisition*, 211-236. Bingley, UK: Emerald Group.

Heider, Eleanor R. 1971. "Focal" color areas and the development of color names. D*evelopmental Psychology* 4. 447-455.

Hodgson, Miren. 2010. Locatum structures and the acquisition of telicity. *Language Acquisition* 17, 155-182.

Hoekstra, Teun. 1988. Small clause results. *Lingua* 74. 101-139.

Hoekstra, Teun & René Mülder. 1990. Unergatives as copular verbs. *The Linguistic Review* 7. 1-79.

Horrocks, Geoffrey & Melita Stavrou. 2003. Actions and their results in Greek and English: The complementarity of morphologically encoded (viewpoint) aspect and syntactic resultative prediction. *Journal of Semantics* 20. 297-327.

Horrocks, Geoffrey & Melita Stavrou. 2007. Grammaticalized aspect and spatio-temporal culmination. *Lingua* 117. 605-644.

Housen, Alex. 2002. The development of tense-aspect in English as second language and the variable influence of inherent aspect. In Rafael Salaberry & Yasuhiro Shirai

(Eds.), *The L2 acquisition of tense-aspect morphology*, 155-197. Philadelphia: John Benjamins.

Huddleston, Rodney D. & Geoffrey K. Pullum. 2002. *The Cambridge grammar of the English language*. New York: Cambridge University Press.

Kalton, G., J. Roberts & D. Holt. 1980. The effects of offering a middle response option with opinion questions. *The Statistician* 29. 65-78.

Inoue, Kyoko. 1979. An analysis of the English present perfect. *Linguistics* 17. 561-589.

Izquierdo, Jesús & Laura Collins. 2008. The facilitative role of L1 influence in tense–aspect marking: A comparison of Hispanophone and Anglophone learners of French. *The Modern Language Journal* 92(3). 350-368.

Jackson, Carrie N. 2010. The processing of subject-object ambiguities by English and Dutch L2 learners of German. In Bill VanPatten & Jill Jegerski (Eds.), *Second language processing and parsing: Issues in theory and research*, 207-230. Philadelphia: John Benjamins.

Jegerski, Jill. 2014. Self-paced reading. In Jill Jegerski & Bill VanPatten (Eds.), *Research methods in second language psycholinguistics*, 20-49. New York: Routledge.

Jespersen, Otto H. 1924. *Language: Its nature, development and origin*. New York: Henry Holt and Co.

Johnson, Cynthia. 1985. The emergence of present perfect verb forms: Semantic influences on selective imitation. *Journal of Child Language* 12. 325-352.

Joos, Martin. 1964. *The English verb: Forms and meanings.* Madison, WI: The University of Wisconsin Press.

Klein, Wolfgang. 1994. *Time in language*. New York: Routledge.

Kline, Rex B. 2004. *Beyond significance testing: Reforming data analysis methods in behavioral research*. Washington, DC: American Psychological Association.

Koda, Keiko. 2012. How to do research on second language reading. In Alison Mackey & Susan M. Gass (Eds.), *Research methods in second language acquisition: A practical guide*, 158-179. Malden, MA: Blackwell.

Krifka, Manfred. 1992. Thematic relations as links between nominal reference and temporal constitution. In Ivan Sag & Anna Szabolcsi (Eds.), *Lexical matters,* 29-53. Palo Alto, CA: CSLI.

Krifka, Manfred. 1998. The origins of telicity. In Susan Rothstein (Ed.), *Events and grammar*, 197-235. Boston: Kluwer.

Kuczaj, Stan A. & Mary J. Daly. 1979. The development of hypothetical reference in the speech of young children. *Journal of Child Language* 6. 563-579.

Kusters, Christiaan Wouter. 2003. *Linguistic complexity: The influence of social change on verbal inflection*. Utrecht: LOT.

Kusters, Christiaan Wouter. 2008. Complexity in linguistic theory, language learning and language change. In Matti Miestamo, Kaius Sinnemaki & Fred Karlsson (Eds.), *Language complexity, typology, contact, change (Studies in Language Companion Series 94)*, 3-21. Philadelphia: John Benjamins.

Langacker, Ronald W. 2008. *Cognitive grammar: A basic introduction*. Oxford: Oxford University Press.

Lardiere, Donna. 2003. Revisiting the comparative fallacy: a reply to Lakshmanan and Selinker, 2001. *Second Language Research* 19. 129-43.

Larson-Hall, Jennifer. 2010. Changing the way we do statistics: Hypothesis testing, power, effect size, and other misunderstood issues. In *A guide to doing statistics in second language research using SPSS*. New York: Routledge.

Lee, Laura L. 1974. *Developmental sentence analysis: A grammatical assessment procedure for speech and language clinicians*. Evanston: Northwestern University Press.

Lee, Eun-Joo. 2001. Interlanguage development by two Korean speakers of English with a focus on temporality. *Language Learning* 51. 591-633.

Leech, Geoffrey N. 1969. *Meaning and the English Verb*. London: Longman.

Leeser, Michael J. 2014. On psycholinguistic methods. In Jill Jegerski & Bill VanPatten (Eds.), *Research methods in second language psycholinguistics*, 231-251. New York: Routledge.

Levin, Beth. 1993. *English verb classes and alternations: A preliminary investigation.* Chicago: University of Chicago Press.

Levin, Beth & Malka Rappaport Hovav. 1995. *Unaccusativity*. Cambridge, MA: MIT Press.

Li, Charles N. & Sandra A. Thompson. 1981. *Mandarin Chinese: A functional reference grammar.* Berkeley: University of California Press.

Li, Charles N., Sandra A. Thompson & R. McMillan Thompson. 1982. The discourse motivation for the perfect aspect: The Mandarin particle *le.* In Paul J. Hopper (Ed.), *Tense-aspect: Between semantics and pragmatics*, 19-44. Philadelphia: John Benjamins.

Li, Ping & Yasuhiro Shirai. 2000. *The acquisition of lexical and grammatical aspect*. New York: Mouton de Gruyter.

Lindstedt, Jouko. (2000). The perfect—aspectual, temporal and evidential. In Östen Dahl (Ed.), *Tense and aspect in the languages of Europe: Empirical approaches to language typology 20-6*, 365-383. Berlin: De Gruyter.

Liszka, Sarah A. 2002. *The development of tense and the present perfect in second language English*. University of Essex PhD dissertation.

Lim, Jayleon. 2003. *The developmental process of English simple past and present perfect by adult Korean learners*. University of Oregon PhD dissertation.

Liu, Ching Kang. 1993. *Complexity of the English present perfect: A comparative study between the English present perfect and the Mandarin guo and le*. University of Southern California, Santa Barbara PhD dissertation.

Lyons, John. 1977. *Semantics*. New York: Cambridge University Press.

Mace, John. 1998. *Arabic grammar: A revision guide*. Edinburgh: Edinburgh University Press.

MacWhinney, Brian. 1978. Processing a first language: The acquisition of morphophonology. *Monographs of the society for research in child development* 43(174).

MacWhinney, Brian (Ed.). 1999. *The Emergence of Language.* Mahwah, NJ: Lawrence Erlbaum.

Madden, Carol J. & Rolf A. Zwaan. 2003. How does verb aspect constrain event representations? *Memory & Cognition* 31(5). 663-672.

Major, Roy C. 1987. A model for interlanguage phonology. In Georgette Ioup & Steven H. Weinberger (Eds.), *Interlanguage phonology: The acquisition of a second language sound system*, l01-124. Cambridge, MA: Newbury House.

Marinis, Theodoros. 2010. Chapter 7: Using on-line processing methods in language acquisition research. In Elma Blom & Sharon Unsworth (Eds.) *Experimental methods in language acquisition research*, 139-162. Philadelphia: John Benjamins.

Mateu, Jaume & Gemma Rigau. 1999. Universals of lexico-syntactic typology and parametric variation. Paper presented at GLOW 1999. Berlin: ZAS.

Mateu, Jaume & Gemma Rigau. 2002. A minimalist account of conflation processes: Parametric variation at the lexicon-syntax interface. In Artemis Alexiadou (Ed.), *Theoretical approaches to universals*, 211-236. Philadelphia: John Benjamins.

McCawley, James D. 1971. Tense and time reference in English. In Charles J. Fillmore & D. Terence Langèndoen (Eds.), *Studies in linguistic semantics*, 96-113. New York: Holt, Rinehart & Winston.

McCawley James D. 1973. *Grammar and Meaning*. Tokyo: Taishukan.

McCoard, Robert W. 1978. *The English perfect: Tense choice and pragmatic inference*. Amsterdam: New Holland.

McCormick, David C. 2008. *Teaching the English present perfect in the second language classroom*. University of Toronto PhD Dissertation.

McDonough, Kim & Pavel Trofimovich. 2012. How to use psycholinguistic methodologies for comprehension and production. In Alison Mackey & Susan M.

Gass (Eds.), *Research methods in second language acquisition: A practical guide*, 117-138. Malden, MA: Blackwell.

McKoon, Gail & Talke Macfarland. 2002. Event templates in the lexical representations of verbs. *Cognitive Psychology* 45. 1-44.

McWhorter, John H. 2001. The world's simplest grammars are creole grammars. *Linguistic Typology* 6. 125-166.

Michaelis, Laura A. 1998. *Aspectual grammar and past-time reference.* New York: Routledge.

Miestamo, Matti. 2008. Grammatical complexity in a cross-linguistic perspective. In Matti Miestamo, Kaius Sinnemäki & Fred. Karlsson (Eds.), *Language complexity, typology, contact, change: Studies in language companion* 94, 22-41. Philadelphia: John Benjamins.

Mikhaylova, Anna. 2012. *(In)complete acquisition of aspect in second language and heritage Russian.* University of South Carolina, Columbia PhD dissertation.

Miller, J.E. 1970. *Tense and aspect in Russian*. University of Edinburgh PhD dissertation.

Miller, Jim. 2000. The perfect in spoken and written English. *Transactions of the Philological Society* 98. 323-52.

Mourelatos, Alexander P. D. 1978. Events, processes, and states. *Linguistics and Philosophy* 2. 415-434.

Murphy, Kevin R. & Brett Myors. 2004. *Statistical power analysis: Simple and general model for traditional and modern hypothesis tests.* Mahwah, NJ: Lawrence Erlbaum.

Newport, Elissa L. 1990. Maturational constraints on language learning. *Cognitive Science* 14. 11-28.

Niemeier, Susanne. 2013. A Cognitive Grammar perspective on tense and aspect. In M. Rafael Salaberry & Llorenç Comajoan (Eds.), *Research design and methodology in studies on L2 tense and aspect*, 11-56. Boston: De Gruyter Mouton.

Nishi, Yumiko & Yasuhiro Shirai. 2007. Where L1 semantic transfer occurs: The significance of cross-linguistic variation in lexical aspect in the L2 acquisition of aspect. In Yoshiko Matsumoto, David Y. Oshima, Orin Robinson & Peter Sells (Eds.), *Diversity in language: Perspectives and implications*. Stanford, CA: CSLI Publications.

Nussbaum, Jo N. & Rita C. Naremore. 1975. On the acquisition of present perfect "have" in normal children. *Language and Speech* 18. 219-226.

Offord, Derek & Natalia Gogolitsyna. 2005. *Using Russian: A guide to contemporary* usage, 2nd edn. London: Cambridge University Press.

Papangeli, Angeliki. 2010. *Language development and on-line processing in L1 and L2 children.* Reading, UK: University of Reading PhD dissertation.

Pawlak, Aleksandra, Jessica S. Oehlrich & Richard M. Weist. 2006. Reference time in child English and Polish. *First Language* 26. 281-297.

Petersen, Christian Heyde. 2004. *Formal and functional perspectives on tense and related categories (European university studies, series 21 linguistics)*. Bern: Lang.

Pickering, Martin J., Brian McElree, Steven Frisson, Lillian Chen & Matthew J. Traxler. 2006. Underspecification and aspectual coercion. *Discourse Processes* 42(2). 131-155.

Pienemann, Manfred. 1987. Psychological constraints on the teachability of languages. In Carol W. Pfaff (Ed.), *First and second language acquisition processes*, 143-168. Rowley, MA: Newbury House.

Pienemann, Manfred. 1998. *Language processing and second language development: Processability theory (Studies in bilingualism 15)*. Philadelphia: John Benjamins.

Piñango, Maria M., Edgar Zurif & Ray Jackendoff. 1999. Real-time processing implications of enriched composition at the syntax-semantics interface. *Journal of Psycholinguistic Research* 28(4). 395-414.

Pustejovsky, James. 1991. The syntax of event structure. *Cognition* 41. 47-81.

Quirk, Randolph, Sydney Greenbaum, Geoffrey Leech & Jan Svartvik. 1985. *A comprehensive grammar of the English language*. London: Longman.

Radden Günter & René Dirven. 2007. *Cognitive English grammar*. Philadelphia: Benjamins.

Reichenbach, Hans. 1947. The tenses of verbs. In *Elements of symbolic logic*, 287-298. New York: Macmillan Company.

Roberts, Leah. 2009. The L1 influences the on-line comprehension of tense/aspect in L2 English, (Paper presented at the L2 Processing and Parsing Conference). Lubbock, TX.

Roberts, Leah & Sarah Ann Liszka. 2013. Processing tense/aspect-agreement violations on-line in the second language: A self-paced reading study with French and German L2 learners of English. *Second Language Research* 29(4). 413-439.

Robison, Richard E. 1995. The aspect hypothesis revisited: A cross-sectional study of tense and aspect marking in interlanguage. *Applied Linguistics* 16. 344-370.

Rohde, Andreas. 1996. The aspect hypothesis and the emergence of tense distinctions in naturalistic L2 acquisition. *Linguistics* 34. 1115-1137.

Rosch, Eleanor H. 1973. On the internal structure of perceptual and semantic categories. In T. E. Moore (Ed.), *Cognitive development and the acquisition of language*, 111-144. New York: Academic Press.

Rosch, Eleanor H. 1977. Human categorization. In N. Warren (Ed.), *Advances in cross-cultural psychology* 1, 1-72. New York: Academic Press.

Rosch, Eleanor H. 1978. Principles of categorization. In: Eleanor H. Rosch and Barbara B. Lloyd (Eds.), *Cognition and categorization.* Hillsdale, NJ: Lawrence Erlbaum.

Rosch, Eleanor H. & Carolyn Mervis. 1975. Family resemblances: Studies in the internal structure of categories. *Cognitive Psychology* 7. 573-605.

Rothstein, Björn. 2007. A modified *ExtendedNow* for the present perfect. In Louis de Saussure, Jacques Moeschler & Genoveva Puskás (Eds.), *Recent advances in the syntax and semantics of tense, aspect and modality*, 89-108. New York: Mouton de Gruyter.

Rothstein, Susan. 2008. Chapter 1: Verb classes and aspectual classification. In *Structuring events: A study in the semantics of lexical aspect*. Oxford: Blackwell.

Sagarra, Nuria & Nick C. Ellis. 2013. From seeing adverbs to seeing verbal morphology: Language experience and adult acquisition of L2 tense. *Studies in Second Language Acquisition* 35. 261-290.

Sasse, Hans-Jürgen. 2002. Recent activity in the theory of aspect: Accomplishments, achievements, or just non-progressive state? *Linguistic Typology* 6. 199-221.

Schlüter, Norbert. 2000. The present perfect in British and American English: Selected results of an empirical analysis. In Christian Mair & Marianne Hundt (Eds.), *Corpus linguistics and linguistic theory*, 313-320. Atlanta: Rodopi.

Schlüter, Norbert. 2006. How reliable are the results? Comparing corpus-based studies of the present perfect. *Zeitschrift fur Anglistik und Amerikanistik* 54. 135-48.

Schmitt, Cristina. 1996. *Aspect and the syntax of noun phrases*. University of Maryland, College Park PhD dissertation.

Schulz, Eckehard. 2004. *A student grammar of modern standard Arabic*. New York: Cambridge University Press.

Selinker, Larry. 1972. Interlanguage. *International Review of Applied Linguistics* 10. 209-231.

Shirai, Yasuhiro. 1989. *The acquisition of the basic verb PUT: Prototype and transfer.* University of California, Los Angeles MA thesis.

Shirai, Yasuhiro. 1992. Conditions on transfer: A connectionist approach. *Issues in Applied Linguistics* 3. 91-120.

Shirai, Yasuhiro. 2002. The prototype hypothesis of tense-aspect acquisition in second language. In Rafael Salaberry & Yasuhiro Shirai (Eds.), *The L2 acquisition of tense-aspect morphology*, 455-478. Philadelphia: John Benjamins.

Shirai, Yasuhiro. 2003. The Aspect Hypothesis, the comparative fallacy and the validity of obligatory context analysis: A reply to Lardiere, 2003. *Second Language Research* 23(1). 51-64.

Shirai, Yasuhiro. 2004. A multiple-factor account for form-meaning connections in the acquisition of tense-aspect morphology. In Bill VanPatten, Jessica Williams, Susanne Rott & Mark Overstreet (Eds.), *Form-meaning connections in second language acquisition*, 91-112. Mahwah, NJ: Lawrence Erlbaum.

Shirai, Yasuhiro. 2009. Temporality in first and second language acquisition. In Wolfgang Klein & Ping Li (Eds.), *The expression of time*, 167-193. New York: Mouton de Gruyter.

Shirai, Yasuhiro. 2013. Defining and coding data: Lexical aspect in L2 studies. In M. Rafael Salaberry & Llorenç Comajoan (Eds.), *Research design and methodology in studies on L2 tense and aspect*, 271-308. Boston: De Gruyter Mouton.

Shirai, Yasuhiro & Roger W. Andersen. 1995. The acquisition of tense-aspect morphology: A prototype account. *Language* 71. 743-762.

Siemund, Peter. 2004. Substrate, superstrate and universals: Perfect constructions in Irish English. In Bernd Kortmann (Ed.), *Dialectology meets typology*, 401-434. New York: Mouton de Gruyter.

Silva, Rosangela Souto. 1999. *Pragmatic competence and transferability: Native and non-native speakers of Portuguese*. University of Illinois, Urbana-Champaign PhD dissertation.

Simonsohn, Uri, Leif D. Nelson & Joseph P. Simmons. 2014. *P*-curve and effect size: Correcting for publication bias using only significant results. *Perspectives on Psychological Science* 9(6). 666-681.

Slabakova, Roumyana. 1997. *Zero acquisition: Second language acquisition of the parameter of aspect*. McGill University: PhD dissertation.

Slabakova, Roumyana. 2000. L1 transfer revisited: The L2 acquisition of telicity marking in English by Spanish and Bulgarian native speakers. *Linguistics*. 38(4). 739-770.

Slabakova, Roumyana. 2001. Aspectual constraints in the mental lexicon. *Journal of Psycholinguistic Research* 30(2). 197-217.

Slabakova, Roumyana. 2003. Semantic evidence for functional categories in interlanguage grammars. *Second Language Research* 19(1). 42-75.

Slabakova, Roumyana. 2008. Aspectual shifts: Grammatical and pragmatic knowledge in L2 acquisition. In Juana M. Liceras, Helmut Zobl & Helen Goodluck (Eds.), *The role of formal features in second language acquisition*, 455-483. New York: Routledge.

Slabakova, Roumyana & Silvina Montrul. 1999. Aspectual tenses in Spanish L2 acquisition: A UG perspective. In Yasuhiro Shirai & Rafael Salaberry (Eds.), *Tense-aspect morphology in L2 acquisition*, 1-25. Philadelphia: John Benjamins.

Slabakova, Roumyana & Silvina Montrul. 2002. On viewpoint aspect interpretation and its L2 acquisition: A UG perspective. In Rafael Salaberry & Yasuhiro Shirai (Eds.), *The L2 acquisition of tense-aspect morphology*, 363-396. Philadelphia: John Benjamins.

Slobin, Dan Isaac. 1985. Crosslinguistic evidence for the Language-Making Capacity. In Dan. I. Slobin (Ed.), *The crosslinguistic study of language acquisition: Vol.* 2. *Theoretical issues*, 1157-1249. Hillsdale, NJ: Erlbaum.

Slobin, Dan Isaac. 1994. Talking perfectly. Discourse origins of the present perfect. In W. Pagliuca (Ed.), *Perspectives on grammaticalization*, 119-133. Philadelphia: John Benjamins.

Smith, Carlota S. 1980. The acquisition of time talk: Relations between child and adult grammars. *Journal of Child Language* 7. 263-278.

Smith, Carlota S. 1991. *The Parameter of Aspect (Studies in linguistics and philosophy 43)*. Dordrecht, NL: Springer.

Smith, Carlota S. 1997. *The Parameter of Aspect (Studies in linguistics and philosophy 43),* 2nd edn. Dordrecht, NL: Springer.

Snyder, William. 1995. *Language acquisition and language variation: The role of morphology*. Cambridge, MA: MIT Working papers in Linguistics.

Snyder, William. 2001. On the nature of syntactic variation: Evidence from complex predicates and complex word-formation. *Language* 77(2). 324-342.

Spada, Nina, Patsy M. Lightbown & Joanna L. White. 2005. The importance of form/meaning mappings in explicit form-focused instruction. In Alex Housen & Michel Pierrard (Eds.), *Investigations in instructed second language learning*, 199-234. New York: Mouton De Gruyter.

Stilman, Galina, Leon Stilman & William E. Harkness. 1972. *Introductory Russian grammar,* 2nd edn. Toronto: Xerox College Publishing.

Streitberg, Wilhelm. 1891. Perfective and imperfective Actionsart im Germanischen […in Germanic]. *Beitraege zur Geschichte der deutschen Sprache* 15. 70-177 (tr.).

Suh, Kyung-Hee. 1992. *A discourse analysis of the English tense-aspect-modality system.* University of California, Los Angeles: PhD Dissertation.

Stutterheim, Christiane von. 1991. Narrative and description: Temporal reference in second language acquisition. In Thom Huebner & Charles A. Ferguson (Eds.), *Crosscurrents in second language acquisition and linguistics theories*, 385-403. Philadelphia: John Benjamins.

Stutterheim, Christiane von, Mary Carroll & Wolfgang Klein. 2009. New perspectives in analyzing aspectual distinctions across languages. In Wolfgang Klein & Ping Li (Eds.), *The expression of time*, 195-216. New York: Mouton de Gruyter.

Szmrecsanyi, Benedikt & Bernd Kortmann. 2009. Between simplification and complication: Non-standard varieties of English around the world. In Geoffrey Sampson, David Gill & Peter Trudgill (Eds.), *Language complexity as an evolving variable (Studies in the Evolution of Language.)*, 64-80. Oxford: Oxford University Press.

Tagliamonte, Sali. 2000. The grammaticalization of the present perfect in English. In Olga Fischer, Anette Rosenbach & Dieter Stein (Eds.), *Pathways of change: Grammaticalization in English*, 329-354. Philadelphia: Benjamins.

Talmy, Leonard. 1975. Semantics and syntax of motion. In J. P. Kimball (Ed.), *Syntax and semantics 4*, 181–238. New York: Academic Press.

Talmy, Leonard. 1985. Lexicalization patterns: Semantic structure in lexical forms. In Timothy Shopen (Ed.), *Language typology and lexical descriptions 3 (Grammatical Categories and the Lexicon)*, 36-149. New York: Cambridge University Press.

Talmy, Loenard. 1991. *Path to realization: A typology of event conflation*. Berkeley: Berkeley Linguistic Society.

Tanaka, S. 1983. *Language transfer as a constraint on lexico-semantic development in adults learning a second language in acquisition-poor environments.* Columbia University, Teachers College PhD dissertation.

Tanaka, S. 1990. *Ninchi imiron: Eigo doshi no tagi no koza* [Cognitive Semantics: The structure of polysemy of English verbs]. Tokyo: Sanyusha.

Taylor, Barry. 1977. Tense and continuity. *Linguistics and Philosophy* 1. 199-220.

Taylor, John R. 2008. Prototypes in Cognitive Linguistics. In Peter Robinson & Nick C. Ellis (Eds), *Handbook of cognitive linguistics and second language acquisition*, 29-65. New York: Routledge.

Tenny, Carol L. 1987. *Grammaticalizing aspect and affectedness*. PhD Dissertation. Cambridge, MA: MIT University Press.

Tenny, Carol L. 1994. *Aspectual roles and the syntax-semantics interface.* Boston: Kluwer.

Terauchi, M. 1994. Keitaiso no syuutoku [The acquisition of morphology]. In SLA Kenkyulcai (Ed.), *Daini gengo syuutoku riron ni motozulcu saisin no eigo kyooiku* [Current English language teaching based on second language acquisition theory], 24-48. Tokyo: Taishukan.

Terán, Virginia. 2014. *Telicity and the developmental acquisition of the English present perfect by L1 Spanish speakers*. Southern Illinois University, Carbondale MA thesis.

Travis, Lisa. 1994. Event Phrase and a theory of functional categories. In P. Koskinen (Ed.), *Proceedings of the 1994 Canadian Linguistic Association Meeting at the University of Calgary,* 559–570. Toronto Working Papers in Linguistics.

Todorova, Marina, Kathy Straub, William Badecker & Robert Frank. 2000. Aspectual coercion and the online computation of sentential aspect. *Proceedings of the CogSci conference*.

Ullman, Michael T. 2004. Contributions of memory circuits to language: The declarative/procedural model. *Cognition* 92. 231-270.

Uno, Mariko. 2014. Lexical aspect in the use of the present perfect by Japanese EFL learners. *IRAL - International Review of Applied Linguistics in Language Teaching* 52(1). 31-57.

Valian, Virginia. 2006. Young children's understanding of present and past tense. *Language Learning and Development* 2. 251-276.

Van Herk, Gerard. 2008. Letter perfect: The present perfect in early African American correspondence. *English World-Wide* 29(1). 45-69.

Van Herk, Gerard. 2010. Aspect and the English present perfect: What can be coded?. In James A. Walker (Ed.), *Aspect in grammatical variation*, 49-64. Philadelphia: John Benjamins.

van Valin, Robert D. 1990. Semantic parameters of split intransitive. *Language* 66(2). 221-260.

Van Voorst, Jan. 1988. *Event structure*. Philadelphia: John Benjamins.

VanPatten, Bill. 2002. Processing instruction: An update. *Language Learning* 52(4). 755-803.

VanPatten, Bill & Jill Jegerski. 2010. Second language processing and parsing: The issues. In Bill VanPatten & Jill Jegerski (Eds.), *Research in second language processing and parsing*, 3-26. Philadelphia: John Benjamins.

VanPatten, Bill & Jessica Williams (Eds.). 2015. *Theories in second language acquisition: An introduction*. New York: Routledge.

Vendler, Zeno. 1967. Verbs and times. In *Linguistics in Philosophy*, 97-121. Ithaca, NY: Cornell University Press.

Verkuyl, H. J. 1972. On the compositional nature of the aspects. Dordrecht, NL: Reidel.

Verkuyl, H. J. 1989. Aspectual classes and aspectual composition. *Linguistics and Philosophy* 12. 39-94.

Verkuyl, H. J. 1993. *A theory of aspectuality*. New York: Cambridge University Press.

Verspoor, Marjolijn & Andrea Tyler. 2009. Cognitive linguistics and second language learning. In William C. Ritchie & Tej K. Bhatia (Eds.), *The new handbook of second language acquisition*, 160-180. Bingley, UK: Emerald Group.

Wagner, Laura. 2010. Inferring meaning from syntactic structures in acquisition: The case transitivity and telicity. *Language and Cognitive Processes* 25(10). 1354-1379.

Wallot, Sebastian & Guy Van Orden. 2011. Nonlinear analyses of self-paced reading. *The Mental Lexicon* 6(2). 245-274.

Weist, Richard M. 1986. Tense and aspect: Temporal systems in child language. In Paul Fletcher & Michael Garman (Eds.), *Language acquisition: Studies in first language development*, 356-374. New York: Cambridge University Press.

White, Lydia. 2008. Some puzzling features of L2 features. In Juana M. Liceras, Helmut Zobl & Helen Goodluck (Eds.), *The role of formal features in second language acquisition*, 301-326. New York: Routledge.

Winford, Donald. 1993. Variability in the use of perfect have in Trinidadian English: A problem of categorical and semantic mismatch. *Language Variation and Change* 5. 141-188.

Wulff, Stefanie, Nick C. Ellis, Ute Römer, Kathleen Bardovi-Harlig & Chelsea J. Leblanc. 2009. The acquisition of tense-aspect: Converging evidence from corpora and telicity ratings. *The Modern Language Journal* 93(3). 354-369.

Zacks, Jeffrey M., Barbara Tversky & Gowri K. Iyer. 2001. Perceiving, remembering and communicating structure in events. *Journal of Experimental Psychology*. 29-58.

Zobl, Helmut. 1980. Developmental and transfer errors: Their common bases and (possibly) differential effects on subsequent learning. *TESOL Quarterly* 14. 469-479.

APPENDIX A

EXPLANATION OF RESEARCH

CHRISTOPHER FARINA, PI; MILA TASSEVA-KURKTCHIEVA, ADVISOR

Introduction and Purpose

You are invited to participate in a research study conducted by Christopher Farina, who is a doctoral student in the Linguistics Program at the University of South Carolina.

The purpose of this study is to investigate how learners of English as a second language process the meaning and grammar of their second language when asked to read and comprehend sentences in that language. To find this out, I will be asking you to read sentences one word at a time and to answer comprehension questions about these sentences. To collect all of the information that I need, you will also fill out a short survey about your personal and language history and you will complete a brief English proficiency task. This form explains what you will be asked to do if you choose to participate in this study. Please read it carefully and feel free to ask any questions before you make a decision about participating. The study will take about 75 minutes to complete depending on your English reading ability.

Description of Study Procedures

There are three tasks in this study; all of them are completed using a computer. This study i) captures real-time reading and comprehension data, ii) collects relevant personal and language history information, and iii) collects English proficiency data. The reading task asks you to read 56 sentences one word at a time and then indicate how much you agree with a statement about each sentence. Tell the computer to show you the next work by hitting the "Space Bar." Read as quickly as you can. Change your reading speed as needed. The survey asks about your personal history as a language learner, how you use these languages, and your attitudes about them. You may skip any question you do not feel comfortable answering, so please answer each question honestly. A 40-question fill in the blank task determines your English proficiency.

Benefits, Risks and Research Related Injury

The study is designed so that it will not pose risks to the participants other than being under the stress of regular classroom testing. However, if you know you experience anxiety during regular classroom tests, I would encourage you to withdraw your participation. No research-related injuries are expected in relation with your participation with the study.

Costs and Payments

Participation in the study will not cost you anything. We will reimburse you 5USD for your time and participation. Your instructor may also choose to give you a grade or extra credit for your participation. How much and what kind of credit you may receive is entirely up to your instructors. Ask your instructor what credit you will earn for participating.

Confidentiality of Records

To guard your personal information and privacy, the computer will give each participant a code number. I will use this code number on all materials related to your participation in the project such as language background questionnaire, recordings of the actual tasks, and coding of the data. No personal information such as name, age, or your answers will be revealed to anybody but the primary research personnel. Your instructors will not know your individual results at any time.

All data collected during the experiments will be kept for future use. The data will be stored in a password-protected format and will be available to the principal investigator only. It will be used for comparison with future groups of participants. None of your personal information will ever be revealed to third parties and nobody except the principal investigator will have access to any identifying information.

Voluntary Participation

Your participation in this study is strictly voluntary. Your decision not to participate will not affect your standing with your current or future classes. You may withdraw from this study at any time. No written notice or approval is required. You may also refuse to answer any questions you don't want to answer and still remain in the study. The investigator may withdraw you from this research if circumstances arise which warrant doing so.

Contact Persons

If you have any questions or concerns about the research, please feel free to contact the principal investigator, Mr. Christopher Farina via email at farinacj@email.sc.edu or his advisor Mila Tasseva-Kurktchieva via email at tassevak@mailbox.sc.edu.

If you have questions regarding your rights as a research subject, you may contact Thomas Coggins, Office of Research Compliance, University of South Carolina, Columbia, SC 29208, Phone (803) 777-7095, Fax (803) 576-5589, Email tcoggins@mailbox.sc.edu.

APPENDIX B

INDEPENDENT MEASURE OF PROFICIENCY (IMP)

Directions: Please fill in the blanks in the following passage with the word that you believe makes the most sense there. Each blank must have one and only one word.

Joe came home from work on Friday. It was payday, but he wasn't __*too*__ excited about it. He knew that __*when*__ he sat down and paid his __*bills*__ and set aside money for groceries, __*some*__ for the car and a small __*amount*__ in his savings account, there wasn't __*too*__ much left over for a good __*time*__.

He thought about going out for __*dinner*__ at his favorite restaurant, but he __*just*__ wasn't in the mood. He wandered __*about*__ his apartment and ate a sandwich. __*For*__ a while, he couldn't stop himself __*from*__ worrying about the money situation. Finally, __*he*__ got into his car and started __*driving*__. He didn't have a destination in __*mind*__, but he knew that he wanted __*to*__ be far away from the city __*where*__ he lived.

He drove onto a quiet country __*road*__. The country sights made him feel __*good*__. His mind wandered as he drove __*along*__ small farms and he began to __*imagine*__ living on his own piece of __*land*__ and becoming self-sufficient. It had always __*been*__ a dream of his, but he __*had*__ never done anything to make it __*a*__ reality. Even as he was thinking, __*his*__ logical side was scoffing at his __*wild*__ imaginings. He debated the advantages and __*disadvantages*__ of living in the country and __*growing*__ his own food. He imagined his __*farmhouse*__ equipped with a solar energy panel __*on*__ the roof to heat the house __*in*__ winter and power a water heater. __*He*__ envisioned fields of vegetables for canning __*and*__ preserving to last through the winter. __*If*__ the crops had a good yield, __*then*__ he could sell the surplus and __*buy*__ some farming equipment with the extra __*cash*__.

Suddenly, Joe stopped thinking and laughed __*out*__ loud, ‘‘I’m really going to go __*ahead*__ with all this?’’

APPENDIX C

ADAPTED BILINGUAL LANGUAGE PROFILE (BLP)

C.1 BLP INSTRUCTIONS.

Information Survey

Please answer the following questions about your language history, use, attitudes, and proficiency. This survey was created with support from the Center for Open Educational Resources and Language Learning at the University of Texas at Austin to better understand the profiles of bilingual speakers in diverse settings with diverse backgrounds. It has been adapted for research conducted at the University of South Carolina.

The survey consists of 19 questions. It will take approximately 10 to 20 minutes to complete. This is not a test, so there are no right or wrong answers. Please answer every question honestly.

Thank you very much for your help.

C.2 BIOGRAPHICAL INFORMATION QUESTIONS.

I. Biographical Information

Please answer the following questions honestly. If you are uncomfortable responding to certain questions, you may skip them and continue. No one other than the researchers will see your responses. Your responses are linked to your participant number NOT your name.

Age: ________ years

Sex:
□ Male
□ Female

Which hand do you write with?
□ Right
□ Left
□ I use both hands equally

Where do you currently live?
City: ________________________________
State/Province: ________________________
Country: _____________________________

Highest level of formal education:

□ Less than high school	□ Some graduate school
□ High school	□ Masters degree (M.A. or M.S.)
□ Some college	□ Doctorate (M.D. or PhD)
□ College (B.A., B.S.)	□ Other, please specify: ___________

Who is your teacher at English Programs for Internationals (EPI), International Accelerator Program (IAP), or the University of South Carolina (USC)? Which teacher assigned you to complete this survey?

____

First language, native language, or primary language:
(If you have more than one native language, select the one that you think is your stronger language or the one that you use more often.)

□ Arabic	□ Hindi	□ Punjabi
□ Chinese	□ Indonesian	□ Russian
□ English	□ Japanese	□ Spanish
□ French	□ Korean	□ Turkish
□ German	□ Portuguese	□ Vietnamese

□ Other, please specify: _________________

Second languages that you know or have studied:
(Select as many as necessary.)

□ Arabic	□ Hindi	□ Punjabi
□ Chinese	□ Indonesian	□ Russian
□ English	□ Japanese	□ Spanish
□ French	□ Korean	□ Turkish
□ German	□ Portuguese	□ Vietnamese

□ Other, please specify: ________________
□ Other, please specify: ________________

C.3 LANGUAGE HISTORY QUESTIONS.

II. Language History

In this section, please answer some questions about your language history by typing an age or number of years into each box.

1. At what age did you **start learning** the following languages? (0=birth)
[First language]: _____
[Second language(s)]: _____

2. At what age did you **start to feel comfortable** using the following languages?
(0=as early as I can remember; 100=I still do not feel comfortable)
[First language]: _____
[Second language(s)]: _____

3. How many years of **classes (grammar, history, math, etc.)** have you had in the following languages (primary school through university)?
[First language]: _____
[Second language(s)]: _____

4. How many years have you **spent in a country/region** where the following languages are spoken?
[First language]: _____
[Second language(s)]: _____

5. How many years have you **spent in a family** where the following languages are spoken?
[First language]: _____
[Second language(s)]: _____

6. How many years have you **spent in a school/work environment** where the following languages are spoken?
[First language]: _____
[Second language(s)]: _____

C.4 LANGUAGE USE QUESTIONS.

III. Language Use

In this section, please answer some questions about your language use by typing a percentage into each box. Total use for all languages in a given question must equal 100%.

7. In an average week, what percentage of the time do you use the following languages **with friends**?
[First language]: _____%
[Second language(s)]: _____%

8. In an average week, what percentage of the time do you use the following languages **with family**?
[First language]: _____%
[Second language(s)]: _____%

9. In an average week, what percentage of the time do you use the following languages **at school/work**?
[First language]: _____%
[Second language(s)]: _____%

10. When you talk to yourself, how often do you **talk to yourself** in the following languages?
[First language]: _____%
[Second language(s)]: _____%

11. When you count, how often do you **count** in the following languages?
[First language]: _____%
[Second language(s)]: _____%

C.5 LANGUAGE PROFICIENCY QUESTIONS.

IV. Language Proficiency

In this section, please rate your language proficiency by giving marks from 1 to 6.

12. How well do you speak the following languages?

	Not well					Very well
	1	2	3	4	5	6
[First language]:						
[Second language(s)]:						

13. How well do you understand the following languages?

	Not well					Very well
	1	2	3	4	5	6
[First language]:						
[Second language(s)]:						

14. How well do you read the following languages?

	Not well					Very well
	1	2	3	4	5	6
[First language]:						
[Second language(s)]:						

15. How well do you write the following languages?

	Not well					Very well
	1	2	3	4	5	6
[First language]:						
[Second language(s)]:						

C.6 LANGUAGE ATTITUDE QUESTIONS.

V. Language Attitudes

In this section, please respond to the statements about language attitudes by giving marks from 1 to 6.

16. I feel like myself when I speak this language.

	Disagree					Agree
	1	2	3	4	5	6
[First language]:						
[Second language(s)]:						

17. I identify with a culture that speaks this language.

	Disagree					Agree
	1	2	3	4	5	6
[First language]:						
[Second language(s)]:						

18. It is important to me to use (or eventually use) this language like a native speaker.

	Disagree					Agree
	1	2	3	4	5	6
[First language]:						
[Second language(s)]:						

19. I want others to think that I am a native speaker of the following language.

	Disagree					Agree
	1	2	3	4	5	6
[First language]:						
[Second language(s)]:						

Appendix D

Self-paced reading task (SPRT)

D.1 SPRT introduction.

Purpose

The purpose of this study is to investigate how English language learners understand the meaning and grammar of English while reading.

Description

To find this out, I will be asking you to perform three tasks. You will perform the first task today and the other two tasks on another day.

Today, you will read short sentences and answer questions about them.

On some other day, you will answer some questions about yourself. The answers you give in this section will be kept secret. All answers are anonymous.

After this, you will also fill in the blanks of a short reading with 40 missing words.

Today's task should take about 45 to 60 minutes to complete.

Consent

Please write your name **and** check the box to agree to participate in this research.

Name: ________________________

□ I agree to participate in this research

D.2 SPRT INSTRUCTIONS.

Instructions

The first part of this study involves reading sentences and answering questions. There are 104 items in total: 8 practice items and 96 experimental items.

These sentences are presented **one word at a time**. **Tap the "Space Bar" to display the next word.** Please tap the "Space Bar" with the hand your write with. Please read each word as quickly as you can, but make sure you understand what you are reading.

Each sentence has a comprehension question after it. Each question is about the meaning of the sentence. **Answer each question by pressing "1" or "2"**. "1" selects the first answer and "2" selects the second answer. You have 10 seconds to answer each question.

When you are ready, continue to the practice items using the blue link below. There are 8 practice items. These items will help you to understand the task.

D.3 SPRT PRACTICE SENTENCES.

#	Type	Displayed text
1	Sentence	Press the Space Bar to tell the computer to show you the next word.
2	Sentence	This is a sentence to practice reading one word at a time like this.
3	Sentence	Reading one word at a time like this takes some practice, but you get used to it.
4	Sentence	This is another practice sentence, but this one has a practice question following it.
	Question	How would you like to answer this question?
	Answer 1	Press 1 to choose this answer.
	Answer 2	Press 2 to choose this answer.
5	Sentence	The sentences in the experiment have a correct answer based on something that's in the sentence.
	Question	What makes an answer correct?
	Answer 1	Some information in the sentence
	Answer 2	Something in the experiment
6	Sentence	While reading, try your best to read quickly and still get all of the questions correct.
	Question	How quickly should you read?
	Answer 1	Fast, but I should still be able to guess the correct answer
	Answer 2	So fast that I cannot guess the correct answer

7	Sentence	After each question, the computer will pause for one second before displaying the next sentence.
	Question	When does the computer pause the task?
	Answer 1	After each question
	Answer 2	After each sentence
8	Sentence	This is the last practice sentence before the experiment begins and your answers are recorded.
	Question	When you are ready to begin the experiment, select an answer from the below options.
	Answer 1	I'm ready.
	Answer 2	Let's go.

D.4 BOUNDEDNESS SPRT SENTENCES.

#	Cdn		Sentence displayed text Region of interest locations 1	2	3	4	5	
1	A	Carelessly, the landlord has addressed the	issues	with	his	tenant's	plumbing	that were causing tensions for months.
	B	Carelessly, the landlord addressed the	issues	with	his	tenant's	plumbing	that were causing tensions for months.
	C	Carelessly, the landlord has addressed	issues	with	his	tenant's	plumbing	that were causing tensions for months.
	D	Carelessly, the landlord addressed	issues	with	his	tenant's	plumbing	that were causing tensions for months.
2	A	Deliberately, the researcher has tested her	theory	on	the	circus	monkeys	who had to identify colors.
	B	Deliberately, the researcher tested her	theory	on	the	circus	monkeys	who had to identify colors.
	C	Deliberately, the researcher has tested	theories	on	the	circus	monkeys	who had to identify colors.
	D	Deliberately, the researcher tested	theories	on	the	circus	monkeys	who had to identify colors.
3	A	Fortunately, Barry has noticed two	changes	to	the	special	record	he was keeping about his friend Sarah's visit.
	B	Fortunately, Barry noticed two	changes	to	the	special	record	he was keeping about his friend Sarah's visit.
	C	Fortunately, Barry has noticed	changes	to	the	special	record	he was keeping about his friend Sarah's visit.
	D	Fortunately, Barry noticed	changes	to	the	special	record	he was keeping about his friend Sarah's visit.
4	A	Successfully, the professor has published a	paper	in	the	esteemed	journal	with the help of several colleagues who gave her feedback.
	B	Successfully, the professor published a	paper	in	the	esteemed	journal	with the help of several colleagues who gave her feedback.
	C	Successfully, the professor has published	papers	in	the	esteemed	journal	with the help of several colleagues who gave her feedback.
	D	Successfully, the professor published	papers	in	the	esteemed	journal	with the help of several colleagues who gave her feedback.
5	A	Reluctantly, Lauren has announced a	timeline	for	the	surprise	party	she was throwing for her best friend Jamie.
	B	Reluctantly, Lauren announced a	timeline	for	the	surprise	party	she was throwing for her best friend Jamie.
	C	Reluctantly, Lauren has announced	timelines	for	the	surprise	party	she was throwing for her best friend Jamie.
	D	Reluctantly, Lauren announced	timelines	for	the	surprise	party	she was throwing for her best friend Jamie.

6	A	Safely, the repairman has replaced the	window	on	the	seventh	story	of the building on the corner of Sumter and Main Streets.
	B	Safely, the repairman replaced the	window	on	the	seventh	story	of the building on the corner of Sumter and Main Streets.
	C	Safely, the repairman has replaced	windows	on	the	seventh	story	of the building on the corner of Sumter and Main Streets.
	D	Safely, the repairman replaced	windows	on	the	seventh	story	of the building on the corner of Sumter and Main Streets.
7	A	Warmly, the father has handled the	problem	with	his	cranky	children	by offering to take them all out for ice cream.
	B	Warmly, the father handled the	problem	with	his	cranky	children	by offering to take them all out for ice cream.
	C	Warmly, the father has handled	problems	with	his	cranky	children	by offering to take them all out for ice cream.
	D	Warmly, the father handled	problems	with	his	cranky	children	by offering to take them all out for ice cream.
8	A	Hurriedly, the official has counted the	ballots	from	the	voting	session	in which the incumbent was expected to win.
	B	Hurriedly, the official counted the	ballots	from	the	voting	session	in which the incumbent was expected to win.
	C	Hurriedly, the official has counted	ballots	from	the	voting	session	in which the incumbent was expected to win.
	D	Hurriedly, the official counted	ballots	from	the	voting	session	in which the incumbent was expected to win.
9	A	Eagerly, Samuel has ordered a	soda	for	each	friendly	person	at the table who wanted to listen to his story.
	B	Eagerly, Samuel ordered a	soda	for	each	friendly	person	at the table who wanted to listen to his story.
	C	Eagerly, Samuel has ordered	sodas	for	each	friendly	person	at the table who wanted to listen to his story.
	D	Eagerly, Samuel ordered	sodas	for	each	friendly	person	at the table who wanted to listen to his story.
10	A	Bravely, the soldier has survived four	attacks	by	the	vicious	raiders	who were trying to steal his precious supply of water.
	B	Bravely, the soldier survived four	attacks	by	the	vicious	raiders	who were trying to steal his precious supply of water.
	C	Bravely, the soldier has survived	attacks	by	the	vicious	raiders	who were trying to steal his precious supply of water.
	D	Bravely, the soldier survived	attacks	by	the	vicious	raiders	who were trying to steal his precious supply of water.

11	A	Foolishly, the swimmer has ignored the	warning	of	the	beach's	lifeguard	about the shark that was seen near the shore.
	B	Foolishly, the swimmer ignored the	warning	of	the	beach's	lifeguard	about the shark that was seen near the shore.
	C	Foolishly, the swimmer has ignored	warnings	of	the	beach's	lifeguard	about the shark that was seen near the shore.
	D	Foolishly, the swimmer ignored	warnings	of	the	beach's	lifeguard	about the shark that was seen near the shore.
12	A	Violently, the government has destroyed the	weapons	of	the	foreign	country	with whom they were fighting.
	B	Violently, the government destroyed the	weapons	of	the	foreign	country	with whom they were fighting.
	C	Violently, the government has destroyed	weapons	of	the	foreign	country	with whom they were fighting.
	D	Violently, the government destroyed	weapons	of	the	foreign	country	with whom they were fighting.
13	A	Repeatedly, the advertiser has promised the	rewards	to	the	lucky	winners	of the mail-in contest in South Carolina.
	B	Repeatedly, the advertiser promised the	rewards	to	the	lucky	winners	of the mail-in contest in South Carolina.
	C	Repeatedly, the advertiser has promised	rewards	to	the	lucky	winners	of the mail-in contest in South Carolina.
	D	Repeatedly, the advertiser promised	rewards	to	the	lucky	winners	of the mail-in contest in South Carolina.
14	A	Hungrily, the chef has combined two	onions	with	the	garlic	mixture	in order to make a sauce for the chicken.
	B	Hungrily, the chef combined two	onions	with	the	garlic	mixture	in order to make a sauce for the chicken.
	C	Hungrily, the chef has combined	onions	with	the	garlic	mixture	in order to make a sauce for the chicken.
	D	Hungrily, the chef combined	onions	with	the	garlic	mixture	in order to make a sauce for the chicken.
15	A	Cruelly, the monster has attacked the	city	on	a	peaceful	island	in the Pacific Ocean near Java.
	B	Cruelly, the monster attacked the	city	on	a	peaceful	island	in the Pacific Ocean near Java.
	C	Cruelly, the monster has attacked	cities	on	a	peaceful	island	in the Pacific Ocean near Java.
	D	Cruelly, the monster attacked	cities	on	a	peaceful	island	in the Pacific Ocean near Java.
16	A	Obediently, the scientist has obtained a	sample	of	the	icy	water	from the reservoir some thought was polluted.
	B	Obediently, the scientist obtained a	sample	of	the	icy	water	from the reservoir some thought was polluted.
	C	Obediently, the scientist has obtained	samples	of	the	icy	water	from the reservoir some thought was polluted.
	D	Obediently, the scientist obtained	samples	of	the	icy	water	from the reservoir some thought was polluted.
17	A	Stubbornly, the mayor has proposed four	taxes	on	the	local	workers	who make more than $100,000 per year.
	B	Stubbornly, the mayor proposed four	taxes	on	the	local	workers	who make more than $100,000 per year.
	C	Stubbornly, the mayor has proposed	taxes	on	the	local	workers	who make more than $100,000 per year.
	D	Stubbornly, the mayor proposed	taxes	on	the	local	workers	who make more than $100,000 per year.
18	A	Carefully, the journalist has confirmed his	belief	through	a	careful	study	of the facts from the initial reports.
	B	Carefully, the journalist confirmed his	belief	through	a	careful	study	of the facts from the initial reports.
	C	Carefully, the journalist has confirmed	beliefs	through	a	careful	study	of the facts from the initial reports.
	D	Carefully, the journalist confirmed	beliefs	through	a	careful	study	of the facts from the initial reports.

19	A	Promptly, the police have issued one	report	on	the	recent	events	surrounding the Johnsons' missing child.
	B	Promptly, the police issued one	report	on	the	recent	events	surrounding the Johnsons' missing child.
	C	Promptly, the police have issued	reports	on	the	recent	events	surrounding the Johnsons' missing child.
	D	Promptly, the police issued	reports	on	the	recent	events	surrounding the Johnsons' missing child.
20	A	Rightfully, the citizen has opposed the	taxes	on	all	rundown	houses	that the local government was suggesting.
	B	Rightfully, the citizen opposed the	taxes	on	all	rundown	houses	that the local government was suggesting.
	C	Rightfully, the citizen has opposed	taxes	on	all	rundown	houses	that the local government was suggesting.
	D	Rightfully, the citizen opposed	taxes	on	all	rundown	houses	that the local government was suggesting.
21	A	Sensibly, the activist has questioned the	judgment	of	the	local	police	who were keeping a man in jail without reason.
	B	Sensibly, the activist questioned the	judgment	of	the	local	police	who were keeping a man in jail without reason.
	C	Sensibly, the activist has questioned	judgments	of	the	local	police	who were keeping a man in jail without reason.
	D	Sensibly, the activist questioned	judgments	of	the	local	police	who were keeping a man in jail without reason.
22	A	Unexpectedly, the politician has convinced the	voter	in	her	civic	district	to agree with her stance on the issue.
	B	Unexpectedly, the politician convinced the	voter	in	her	civic	district	to agree with her stance on the issue.
	C	Unexpectedly, the politician has convinced	voters	in	her	civic	district	to agree with her stance on the issue.
	D	Unexpectedly, the politician convinced	voters	in	her	civic	district	to agree with her stance on the issue.
23	A	Generously, the boss has approved the	request	of	his	loyal	staffers	about increasing their role in the company.
	B	Generously, the boss approved the	request	of	his	loyal	staffers	about increasing their role in the company.
	C	Generously, the boss has approved	requests	of	his	loyal	staffers	about increasing their role in the company.
	D	Generously, the boss approved	requests	of	his	loyal	staffers	about increasing their role in the company.
24	A	Anxiously, the inventor has displayed an	image	of	her	newest	design	at the technology convention in California.
	B	Anxiously, the inventor displayed an	image	of	her	newest	design	at the technology convention in California.
	C	Anxiously, the inventor has displayed	images	of	her	newest	design	at the technology convention in California.
	D	Anxiously, the inventor displayed	images	of	her	newest	design	at the technology convention in California.

D.5 BOUNDEDNESS SPRT COMPREHENSION QUESTIONS AND ANSWERS.

#	Question	Answers	
		Correct	Incorrect
1	Who fixed the plumbing?	The landlord	The tenant
2	What do the monkeys need to identify?	Colors	Words
3	Does he have a schedule for her visit?	Yes	No
4	Did the professor's peers help her?	Yes	No
5	Who is planning the party?	Lauren	Jamie
6	What floor is/are the new window(s) on?	The seventh	The sixth
7	Who received ice cream?	The children	The father
8	Who counted the ballots?	The official	The incumbent
9	Did his friends buy him a soda(s)?	No	Yes
10	Does the soldier still have water?	Yes	No
11	Who was warned about the shark?	The swimmer	The lifeguard
12	Did the government make more weapons?	No	Yes
13	Will the winners receive a car?	No	Yes
14	Which is an ingredient in the recipe?	Onions	Peppers
15	Which ocean is the island in?	The Pacific	The Atlantic
16	Does the scientist study water?	Yes	No
17	Who is being taxed?	The workers	The mayor
18	Was the journalist's information correct?	Yes	No
19	Who released the information?	The police	The Johnsons
20	Does the citizen want to raise taxes on cars?	No	Yes
21	Does the activist think the man should be in jail?	No	Yes
22	Did the politician change people's minds?	Yes	No
23	Who approved the request(s)?	The boss	The staffers
24	What did the inventor show the audience?	A picture	A model

D.6 CURRENT RELEVANCE SPRT SENTENCES.

#	Cdn.	SubC	Sentence displayed text						
				Region of interest locations					
				1	2	3	4	5	
25	E	1	Since turning on, the transmitter has bounced the	signals	to	the	distant	station	where the message was received.
	F	1	Since turning on, the transmitter bounced the	signals	to	the	distant	station	where the message was received.
	G	1	After turning on, the transmitter has bounced the	signals	to	the	distant	station	where the message was received.
	H	1	After turning on, the transmitter bounced the	signals	to	the	distant	station	where the message was received.
26	E	1	Since filling the trash, the writer has crushed the	papers	in	the	packed	wastebin	to stop it from overflowing.
	F	1	Since filling the trash, the writer crushed the	papers	in	the	packed	wastebin	to stop it from overflowing.
	G	1	After filling the trash, the writer has crushed the	papers	in	the	packed	wastebin	to stop it from overflowing.
	H	1	After filling the trash, the writer crushed the	papers	in	the	packed	wastebin	to stop it from overflowing.
27	E	1	Since being called, Carol has danced with	Jason	to	the	wooden	platform	where the judges were seated.
	F	1	Since being called, Carol danced with	Jason	to	the	wooden	platform	where the judges were seated.
	G	1	After being called, Carol has danced with	Jason	to	the	wooden	platform	where the judges were seated.
	H	1	After being called, Carol danced with	Jason	to	the	wooden	platform	where the judges were seated.
28	E	1	Since dusting, the maid has pulled the	table	off	the	dirty	carpet	so that it could be cleaned too.
	F	1	Since dusting, the maid pulled the	table	off	the	dirty	carpet	so that it could be cleaned too.
	G	1	After dusting, the maid has pulled the	table	off	the	dirty	carpet	so that it could be cleaned too.
	H	1	After dusting, the maid pulled the	table	off	the	dirty	carpet	so that it could be cleaned too.
29	E	1	Since play began, Maria has rolled the	football	to	her	waiting	teammate	who had an open run to the goal.
	F	1	Since play began, Maria rolled the	football	to	her	waiting	teammate	who had an open run to the goal.
	G	1	After play began, Maria has rolled the	football	to	her	waiting	teammate	who had an open run to the goal.
	H	1	After play began, Maria rolled the	football	to	her	waiting	teammate	who had an open run to the goal.
30	E	1	Since arguing with Lila, David has shoved the	boxes	in	his	station	wagon	before slamming the trunk and driving off.
	F	1	Since arguing with Lila, David shoved the	boxes	in	his	station	wagon	before slamming the trunk and driving off.
	G	1	After arguing with Lila, David has shoved the	boxes	in	his	station	wagon	before slamming the trunk and driving off.
	H	1	After arguing with Lila, David shoved the	boxes	in	his	station	wagon	before slamming the trunk and driving off.
31	E	2	For twenty years, the cowboy has charged the	cattle	to	the	open	pasture	after releasing them from their pens.

	F	2	For twenty years, the cowboy charged the	cattle	to	the	open	pasture	after releasing them from their pens.
	G	2	In five minutes, the cowboy has charged the	cattle	to	the	open	pasture	after releasing them from their pens.
	H	2	In five minutes, the cowboy charged the	cattle	to	the	open	pasture	after releasing them from their pens.
32	E	2	For two minutes, the dog has chased the	foxes	to	their	nearby	burrow	while barking to alert the hunters.
	F	2	For two minutes, the dog chased the	foxes	to	their	nearby	burrow	while barking to alert the hunters.
	G	2	In two minutes, the dog has chased the	foxes	to	their	nearby	burrow	while barking to alert the hunters.
	H	2	In two minutes, the dog chased the	foxes	to	their	nearby	burrow	while barking to alert the hunters.
33	E	2	For thirty seconds, the fireman has climbed the	ladder	to	the	lofty	window	to save an elderly man.
	F	2	For thirty seconds, the fireman climbed the	ladder	to	the	lofty	window	to save an elderly man.
	G	2	In thirty seconds, the fireman has climbed the	ladder	to	the	lofty	window	to save an elderly man.
	H	2	In thirty seconds, the fireman climbed the	ladder	to	the	lofty	window	to save an elderly man.
34	E	2	For three months, the general has marched the	army	to	the	rebel	province	that wanted its freedom from rule.
	F	2	For three months, the general marched the	army	to	the	rebel	province	that wanted its freedom from rule.
	G	2	In three months, the general has marched the	army	to	the	rebel	province	that wanted its freedom from rule.
	H	2	In three months, the general marched the	army	to	the	rebel	province	that wanted its freedom from rule.
35	E	2	For two days, the diplomats have sailed to	London	on	the	Santa	Ana	with a message from the Queen of Spain.
	F	22	For two days, the diplomats sailed to	London	on	the	Santa	Ana	with a message from the Queen of Spain.
	G	2	In two days, the diplomats have sailed to	London	on	the	Santa	Ana	with a message from the Queen of Spain.
	H	2	In two days, the diplomats sailed to	London	on	the	Santa	Ana	with a message from the Queen of Spain.
36	E	2	For ten years, the dispatcher has steered the	buses	to	the	night-time	garage	where the drivers end their day.
	F	2	For ten years, the dispatcher steered the	buses	to	the	night-time	garage	where the drivers end their day.
	G	2	In one hour, the dispatcher has steered the	buses	to	the	night-time	garage	where the drivers end their day.

	H	2	In one hour, the dispatcher steered the	buses	to	the	night-time	garage	where the drivers end their day.
37	E	3	At present, the baby has crawled to	Momma	from	her	comfy	blanket	which Daddy caught on film.
	F	3	At present, the baby crawled to	Momma	from	her	comfy	blanket	which Daddy caught on film.
	G	3	At some point, the baby has crawled to	Momma	from	her	comfy	blanket	which Daddy caught on film.
	H	3	At some point, the baby crawled to	Momma	from	her	comfy	blanket	which Daddy caught on film.
38	E	3	At present, Jane has filed her	body	through	the	crowded	hallway	to get to her French class.
	F	3	At present, Jane filed her	body	through	the	crowded	hallway	to get to her French class.
	G	3	At some point, Jane has filed her	body	through	the	crowded	hallway	to get to her French class.
	H	3	At some point, Jane filed her	body	through	the	crowded	hallway	to get to her French class.
39	E	3	At present, Robert has hiked with	Emma	to	the	mountain	summit	in order to see the beautiful view.
	F	3	At present, Robert hiked with	Emma	to	the	mountain	summit	in order to see the beautiful view.
	G	3	At some point, Robert has hiked with	Emma	to	the	mountain	summit	in order to see the beautiful view.
	H	3	At some point, Robert hiked with	Emma	to	the	mountain	summit	in order to see the beautiful view.
40	E	3	At present, the trainer has rushed to	Noah	from	the	distant	sideline	to start treating Noah's injury.
	F	3	At present, the trainer rushed to	Noah	from	the	distant	sideline	to start treating Noah's injury.
	G	3	At some point, the trainer has rushed to	Noah	from	the	distant	sideline	to start treating Noah's injury.
	H	3	At some point, the trainer rushed to	Noah	from	the	distant	sideline	to start treating Noah's injury.
41	E	3	At present, the adventurer has tracked the	tigers	to	their	hiding	places	where he observed their habits.
	F	3	At present, the adventurer tracked the	tigers	to	their	hiding	places	where he observed their habits.
	G	3	At some point, the adventurer has tracked the	tigers	to	their	hiding	places	where he observed their habits.
	H	3	At some point, the adventurer tracked the	tigers	to	their	hiding	places	where he observed their habits.
42	E	3	At present, the detective has trailed the	taxi	to	its	final	target	while looking for a suspected criminal.
	F	3	At present, the detective trailed the	taxi	to	its	final	target	while looking for a suspected criminal.
	G	3	At some point, the detective has trailed the	taxi	to	its	final	target	while looking for a suspected criminal.
	H	3	At some point, the detective trailed the	taxi	to	its	final	target	while looking for a suspected criminal.
43	E	4	Often, the sluggish man has dropped his	mirror	on	the	fluffy	carpet	realizing he was late for an appointment.
	F	4	Often, the sluggish man dropped his	mirror	on	the	fluffy	carpet	realizing he was late for an appointment.
	G	4	Until drinking coffee, the sluggish man has dropped his	mirror	on	the	fluffy	carpet	realizing he was late for an appointment.

	H	4	Until drinking coffee, the sluggish man dropped his	mirror	on	the	fluffy	carpet	realizing he was late for an appointment.
44	E	4	Often, Jacob has skipped to	Jenny	at	the	local	playground	because he's so excited to see her.
	F	4	Often, Jacob skipped to	Jenny	at	the	local	playground	because he's so excited to see her.
	G	4	Until the fight, Jacob has skipped to	Jenny	at	the	local	playground	because he's so excited to see her.
	H	4	Until the fight, Jacob skipped to	Jenny	at	the	local	playground	because he's so excited to see her.
45	E	4	Often, Kyle has stepped to	Katie	from	the	treadmill	section	when seeing her enter the gym.
	F	4	Often, Kyle stepped to	Katie	from	the	treadmill	section	when seeing her enter the gym.
	G	4	Until this week, Kyle has stepped to	Katie	from	the	treadmill	section	when seeing her enter the gym.
	H	4	Until this week, Kyle stepped to	Katie	from	the	treadmill	section	when seeing her enter the gym.
46	E	4	Sometimes, the children have jumped the	fences	to	the	vacant	mansion	that people think is haunted.
	F	4	Sometimes, the children jumped the	fences	to	the	vacant	mansion	that people think is haunted.
	G	4	Until getting caught, the children have jumped the	fences	to	the	vacant	mansion	that people think is haunted.
	H	4	Until getting caught, the children jumped the	fences	to	the	vacant	mansion	that people think is haunted.
47	E	4	Sometimes, the busboy has slipped on	butter	to	the	wooden	flooring	while cleaning up the kitchen.
	F	4	Sometimes, the busboy slipped on	butter	to	the	wooden	flooring	while cleaning up the kitchen.
	G	4	Until mopping, the busboy has slipped on	butter	to	the	wooden	flooring	while cleaning up the kitchen.
	H	4	Until mopping, the busboy slipped on	butter	to	the	wooden	flooring	while cleaning up the kitchen.
48	E	4	Sometimes, the couple has walked to	Publix	for	some	frozen	dinners	to feed some unexpected guests.
	F	4	Sometimes, the couple walked to	Publix	for	some	frozen	dinners	to feed some unexpected guests.
	G	4	Until getting a car, the couple has walked to	Publix	for	some	frozen	dinners	to feed some unexpected guests.
	H	4	Until getting a car, the couple walked to	Publix	for	some	frozen	dinners	to feed some unexpected guests.

D.7 CURRENT RELEVANCE SPRT COMPREHENSION QUESTIONS AND ANSWERS.

#	Question	Answers	
		Correct	Incorrect
25	Did the station send the message?	No	Yes
26	Did the writer take the trash out?	No	Yes
27	Were they in a dance competition?	Yes	No
28	What was dirty?	The carpet	The table
29	Which person will probably score?	Her teammate	Maria
30	Who was moving out?	David	Lila
31	What were put out to pasture?	Cows	Horses
32	How many dogs are hunting?	One	Two
33	Was the elderly man on the first floor?	No	Yes
34	Was the province fighting the empire?	Yes	No
35	Why were they sailing?	Politics	Business
36	Was it early in the morning?	No	Yes
37	Did the baby want her blanket?	No	Yes
38	Is Jane probably in a school?	Yes	No
39	Who did Emma hike with?	Robert	Steven
40	Was Noah playing a game when he got hurt?	Yes	No
41	What was the adventurer doing?	Researching	Hunting
42	Who was the taxi carrying?	The criminal	The detective
43	Was he in a rush to leave?	Yes	No
44	Are Jacob and Jenny friends?	Yes	No
45	Who was in the gym first?	Kyle	Katie
46	Who snuck into the mansion?	Children	Teens
47	Who fell onto the kitchen floor?	The busboy	The chef
48	Was the couple buying dinner for two?	No	Yes

APPENDIX E

RATING TASK

E.1 RATING TASK INTRODUCTION.

Purpose

The purpose of this study is to investigate how English language learners understand the meaning and grammar of English.

Description

To find this out, I will be asking you to perform three tasks.

First, you will read short sentences and indicate how much you agree with statements about them.

Second, you will answer some questions about yourself. The answers you give in this section will be kept secret. All answers are anonymous.

Third, you will fill in the blanks of a short reading with 40 missing words.

All three tasks should take about 75 minutes to complete

Consent

Please write your name **and** check the box to agree to participate in this research.

Name: ________________________

□ I agree to participate in this research

E.2 RATING TASK INSTRUCTIONS.

Instructions (1/5)

In the first part of this study, you read sentences and indicate how much you agree with statements. There are 56 items in total: 8 practice items and 48 experimental items.

These sentences are presented <u>**one word at a time**</u>. **Tap the "Space Bar" to display the next word**. Tap the "Space Bar" using the hand your write with. Read each word as quickly as you can, but make sure you understand what you are reading.

Each sentence has a statement after it. These statements describe the meaning of a part of the sentence. Your response says <u>how much you **agree** with the statement</u> from “1” to “6”.

"1" means that you “Strongly **Disagree**” and "6" means you “Strongly **Agree**.”

The full scale looks like this:

1 Strongly Disagree
2 Disagree
3 Kind of Disagree
4 Kind of Agree
5 Agree
6 Strongly Agree

The next pages explain the different kinds of statements, and they give example answers.

Instructions (2/5)

The first kind of statement says something about the subject of the sentence. It asks you to guess whether or not **the subject** is <u>**new**</u> **or** <u>**old**</u> **information**.

Compare the following sentences:

1. **Stewart** saw Joan. ← Sentence

   *This is probably the first time that **Stewart** is mentioned* ← Statement

   Strongly Disagree 1 2 3 4 5 6 Strongly Agree ← Rating Scale

In English, we usually only use names the first time we introduce someone, so this is probably the first time we are seeing "Stewart" in the conversation.

This statement should be rated "**6**." You **strongly agree** that this is probably the first time that "Stewart" is mentioned.

2. **He** saw Joan.

   *This is probably the first time that **He** is mentioned*

   Strongly Disagree 1 2 3 4 5 6 Strongly Agree

The speaker does not say who "he" is. The speaker assumes that we already know who "he" is. This is probably **<u>not</u>** the first time that "he" is mentioned.

This statement should be rated "**1**." You **strongly <u>dis</u>agree** that this is probably the first time that "He" is mentioned.

Instructions (3/5)

The second kind of statement says something about the verb and the words that follow it. It asks you whether or not the **action of the verb** can **<u>continue</u>** or if it is **<u>finished</u>**.

Compare the following sentences:

3. Peter **walked around** with his new shoes on.

   *When the action **walked around** is finished, it can be continued.*

   Strongly Disagree 1 2 3 4 5 6 Strongly Agree

Once Peter finishes walking around, he **can continue** walking around more. He can walk around all day, stopping and starting again whenever he wants to.

This statement should be rated "**6**." You **strongly agree** that this action can continue.

4. Peter **walked to his house** with his new shoes on.

   *When the action **walked to his house** is finished, it can be continued.*

   Strongly Disagree 1 2 3 4 5 6 Strongly Agree

Once Peter finishes walking home, he **cannot** **continue** walking home. He is already there!

This statement should be rated "**1**." You **strongly disagree** that this action can continue.

Instructions (4/5)

The third kind of statement says something about the action of the verb or its consequences. It asks you whether or not the past **action** is **relevant** at the present time. An action is relevant when the action or its consequences are **more important to the present** than to the past.

Compare the following sentences:

5. Maria **broke her leg** last week, and she cannot play soccer while it heals.

   *The fact that **her leg** was **broken** is relevant at the present time.*

   Strongly Disagree 1 2 3 4 5 6 Strongly Agree

The consequence that her leg is currently broken is **important now** because her broken leg **stops her from playing soccer now**. The fact that she cannot play soccer now is more important than the action of breaking it.

This statement should be rated "**6**." You **strongly agree** that the action or its consequences are relevant at the present time.

6. Maria **broke her leg** last year, and it healed in only six weeks.

   *The fact that **her leg** was **broken** is relevant at the present time.*

   Strongly Disagree 1 2 3 4 5 6 Strongly Agree

The fact that her leg was broken last year is <u>**not**</u> **important anymore** because it is not still broken. **It is healed <u>now</u>**. The fact that her leg was broken is less important than the fact that it is healed.

This statement should be rated "**1**." You **strongly <u>dis</u>agree** that the action or its consequences are relevant at the present time.

Instructions (5/5)

The fourth kind of statement says something about the adverbs. It asks you whether or not the **adverb indicates** that the past action is <u>**relevant**</u> at the present time. An action is relevant when the action or its consequences are **more important to the present** than to the past.

Compare the following sentences:

7. **Up till now**, Megan lived in Columbia.

   *The phrase **up till now** indicates that the action or its consequences are still relevant.*

   Strongly Disagree 1 2 3 4 5 6 Strongly Agree

The phrase indicates that the action is very recent. In this sentence, **the action is relevant** to the present because "up till now" indicates that **it just ended** right now.

This statement should be rated "**6**." You **strongly agree** that the phrase helps to indicate relevance to the present time.

8. **In the past**, Megan lived in Columbia.

   *The phrase **in the past** indicates that the action or its consequences are still relevant.*

   Strongly Disagree 1 2 3 4 5 6 Strongly Agree

The phrase relates the action to past. In this sentence, **the action is <u>not</u> relevant** at present because *in the past* indicates that **the past is more important** than the present.

This statement should be rated “**1**.” You **strongly <u>dis</u>agree** that the phrase helps to indicate relevance to the present time.

The directions are now finished. Continue to the practice items using the blue link below.

There are 8 practice items. These items will help you to understand the study.

E.3 RATING TASK PRACTICE SENTENCES.

#	Type	Displayed text
1	Sentence	Press the Space Bar to tell the computer to show you the next word.
2	Sentence	This is a sentence to practice reading one word at a time like this.
3	Sentence	Reading one word at a time like this takes some practice, but you get used to it.
4	Sentence	This is another practice sentence, but this one has a practice statement following it.
	Statement	Say how much you agree by clicking a blue number or by typing that number's key. Pick any number you want.
	Rating scale	Strongly Disagree 1 2 3 4 5 6 Strongly Agree
5	Sentence	There is no right or wrong answer to the statements that follow each sentence.
	Statement	This study is looking for your opinion and **not** a correct answer.
	Rating scale	Strongly Disagree 1 2 3 4 5 6 Strongly Agree
6	Sentence	While reading, try your best to read quickly and still understand the entire sentence.
	Statement	I should read quickly as long as I understand the meaning of the sentence.
	Rating scale	Strongly Disagree 1 2 3 4 5 6 Strongly Agree
7	Sentence	After each question, the computer will pause for one second before displaying the next sentence.
	Statement	The computer pauses between sentences. If you need to rest, do it before hitting the Space Bar.
	Rating scale	Strongly Disagree 1 2 3 4 5 6 Strongly Agree
8	Sentence	This is the last practice sentence before the study begins and your answers are recorded.
	Statement	I am ready to **begin the study!** Pick any number to continue.
	Rating scale	Strongly Disagree 1 2 3 4 5 6 Strongly Agree

E.4 BOUNDEDNESS RATING STATEMENTS.

#*	Cdn	Statement displayed text
2	A	When the action **has tested her theory** is finished, it can be continued.
	B	When the action **tested her theory** is finished, it can be continued.
	C	When the action **has tested theories** is finished, it can be continued.
	D	When the action **tested theories** is finished, it can be continued.
4	A	When the action **has published a paper** is finished, it can be continued.
	B	When the action **published a paper** is finished, it can be continued.
	C	When the action **has published papers** is finished, it can be continued.
	D	When the action **published papers** is finished, it can be continued.
5	A	When the action **has announced a timeline** is finished, it can be continued.
	B	When the action **announced a timeline** is finished, it can be continued.
	C	When the action **has announced timelines** is finished, it can be continued.
	D	When the action **announced timelines** is finished, it can be continued.
6	A	When the action **has replaced the window** is finished, it can be continued.
	B	When the action **replaced the window** is finished, it can be continued.
	C	When the action **has replaced windows** is finished, it can be continued.
	D	When the action **replaced windows** is finished, it can be continued.
8	A	When the action **has counted the ballots** is finished, it can be continued.
	B	When the action **counted the ballots** is finished, it can be continued.
	C	When the action **has counted ballots** is finished, it can be continued.
	D	When the action **counted ballots** is finished, it can be continued.
9	A	When the action **has ordered a soda** is finished, it can be continued.
	B	When the action **ordered a soda** is finished, it can be continued.
	C	When the action **has ordered sodas** is finished, it can be continued.
	D	When the action **ordered sodas** is finished, it can be continued.
11	A	When the action **has ignored the warning** is finished, it can be continued.
	B	When the action **ignored the warning** is finished, it can be continued.
	C	When the action **has ignored warnings** is finished, it can be continued.
	D	When the action **ignored warnings** is finished, it can be continued.
12	A	When the action **has destroyed the weapons** is finished, it can be continued.
	B	When the action **destroyed the weapons** is finished, it can be continued.
	C	When the action **has destroyed weapons** is finished, it can be continued.
	D	When the action **destroyed weapons** is finished, it can be continued.
13	A	When the action **has promised the rewards** is finished, it can be continued.
	B	When the action **promised the rewards** is finished, it can be continued.
	C	When the action **has promised rewards** is finished, it can be continued.
	D	When the action **promised rewards** is finished, it can be continued.
15	A	When the action **has attacked the city** is finished, it can be continued.
	B	When the action **attacked the city** is finished, it can be continued.
	C	When the action **has attacked cities** is finished, it can be continued.
	D	When the action **attacked cities** is finished, it can be continued.

16	A	When the action **has obtained a sample** is finished, it can be continued.
	B	When the action **obtained a sample** is finished, it can be continued.
	C	When the action **has obtained samples** is finished, it can be continued.
	D	When the action **obtained samples** is finished, it can be continued.
17	A	When the action **has proposed four taxes** is finished, it can be continued.
	B	When the action **proposed four taxes** is finished, it can be continued.
	C	When the action **has proposed taxes** is finished, it can be continued.
	D	When the action **proposed taxes** is finished, it can be continued.
19	A	When the action **have issued one report** is finished, it can be continued.
	B	When the action **issued one report** is finished, it can be continued.
	C	When the action **have issued reports** is finished, it can be continued.
	D	When the action **issued reports** is finished, it can be continued.
20	A	When the action **have opposed the taxes** is finished, it can be continued.
	B	When the action **opposed the taxes** is finished, it can be continued.
	C	When the action **have opposed taxes** is finished, it can be continued.
	D	When the action **opposed taxes** is finished, it can be continued.
23	A	When the action **has approved the request** is finished, it can be continued.
	B	When the action **approved the request** is finished, it can be continued.
	C	When the action **has approved requests** is finished, it can be continued.
	D	When the action **approved requests** is finished, it can be continued.
24	A	When the action **has displayed an image** is finished, it can be continued.
	B	When the action **displayed an image** is finished, it can be continued.
	C	When the action **has displayed images** is finished, it can be continued.
	D	When the action **displayed images** is finished, it can be continued.

*Note: These numbers reference sentences in Appendix D.4 above.

E.5 CURRENT RELEVANCE RATING STATEMENTS.

#*	Cdn.	SubC.	Statement displayed text
28	E	1	The phrase **since dusting** indicates that the action or its consequences are still relevant.
	F	1	The phrase **after dusting** indicates that the action or its consequences are still relevant.
29	E	1	The phrase **since play began** indicates that the action or its consequences are still relevant.
	F	1	The phrase **after play began** indicates that the action or its consequences are still relevant.
32	E	2	The phrase **for two minutes** indicates that the action or its consequences are still relevant.
	F	2	The phrase **in two minutes** indicates that the action or its consequences are still relevant.

34	E	2	The phrase **for three months** indicates that the action or its consequences are still relevant.
	F	2	The phrase **in three months** indicates that the action or its consequences are still relevant.
39	E	3	The phrase **at present** indicates that the action or its consequences are still relevant.
	F	3	The phrase **at some point** indicates that the action or its consequences are still relevant.
41	E	3	The phrase **at present** indicates that the action or its consequences are still relevant.
	F	3	The phrase **at some point** indicates that the action or its consequences are still relevant.
43	E	4	The phrase **often** indicates that the action or its consequences are still relevant.
	F	4	The phrase **until drinking coffee** indicates that the action or its consequences are still relevant.
46	E	4	The phrase **sometimes** indicates that the action or its consequences are still relevant.
	F	4	The phrase **until getting caught** indicates that the action or its consequences are still relevant.

*Note: These numbers reference sentences in Appendix D.6 above. Sentences in condition E and F here are matched with sentences in conditions E and G, respectively.

#	Cdn.	Statement displayed text
49	G	The fact that **a picture** has been **painted** is relevant at the present time.
	H	The fact that **a picture** was **painted** is relevant at the present time.
50	G	The fact that **the issue** has been **resolved** is relevant at the present time.
	H	The fact that **the issue** was **resolved** is relevant at the present time.
51	G	The fact that **the cables** have been **attached** is relevant at the present time.
	H	The fact that **the cables** were **attached** is relevant at the present time.
52	G	The fact that **a flower** has been **planted** is relevant at the present time.
	H	The fact that **a flower** was **planted** is relevant at the present time.
53	G	The fact that **the winner** has been **declared** is relevant at the present time.
	H	The fact that **the winner** was **declared** is relevant at the present time.
54	G	The fact that **an album** has been **released** is relevant at the present time.
	H	The fact that **an album** was **released** is relevant at the present time.
55	G	The fact that **a request** has been **refused** is relevant at the present time.
	H	The fact that **a request** was **refused** is relevant at the present time.
56	G	The fact that **the demands** have been **denied** is relevant at the present time.
	H	The fact that **the demands** were **denied** is relevant at the present time.

Sentences in condition G and H here are matched with sentences in Appendix E.6 below.

E.6 CURRENT RELEVANCE SENTENCES.

#	Cdn	Sentence displayed text	Region of interest locations 1	2	3	4	5	
49	G	Cheerfully, the artist has painted a	picture	of	the	splendid	mountains	when the sun was rising because it was so beautiful.
	H	Cheerfully, the artist painted a	picture	of	the	splendid	mountains	when the sun was rising because it was so beautiful.
50	G	Easily, the expert has resolved the	issue	with	the	broken	printer	that was plaguing the office for six weeks.
	H	Easily, the expert resolved the	issue	with	the	broken	printer	that was plaguing the office for six weeks.
51	G	Loudly, the builder has attached the	cables	near	the	highest	tower	to the foundation below to strengthen the bridge.
	H	Loudly, the builder attached the	cables	near	the	highest	tower	to the foundation below to strengthen the bridge.
52	G	Shyly, Richard has planted a	flower	in	the	public	garden	because he wanted to make his city more beautiful.
	H	Shyly, Richard planted a	flower	in	the	public	garden	because he wanted to make his city more beautiful.
53	G	Truthfully, the organizers have declared the	winner	of	the	football	contest	that was held at the stadium.
	H	Truthfully, the organizers declared the	winner	of	the	football	contest	that was held at the stadium.
54	G	Boldly, the band has released an	album	with	their	newest	music	that diverges greatly from their previous work.
	H	Boldly, the band released an	album	with	their	newest	music	that diverges greatly from their previous work.
55	G	Recklessly, the student has refused a	request	by	his	English	teacher	to do his homework and improve his grade.
	H	Recklessly, the student refused a	request	by	his	English	teacher	to do his homework and improve his grade.
56	G	Nervously, the mayor has denied the	demands	of	the	worried	farmers	to plant less corn and more soybeans.
	H	Nervously, the mayor denied the	demands	of	the	worried	farmers	to plant less corn and more soybeans.

APPENDIX F

BLP RESULTS – SPRT, FIRST ADMINISTRATION

F.1 SUMMARY EXPLANATION OF RESULTS.

The following paragraphs summarize the results of the BLP for the first administration, making special note of any potentially confounding factors between the main groups that would need to be included in the statistical models. They are divided in three ways: (i) native & nonnative English users, (ii) L2 English proficiency, and (iii) first language of the L2 English users.

L2 ENGLISH USERS *IN TOTO*. This section addresses the potential confounding differences that may affect performance on the self-paced reading task between the L2 English User and English NS groups. As indicated in the appendix, there are no statistically significant differences between the groups for the following two groups of collected variables: nonlinguistic biographical information (age, handedness, and highest level of education achieved) and language exposure (formal education in the first language, years spent immersed in a country in which the first language is spoken, and years spent immersed in a work/school environment in which the first language is spoken).

There are several factors that differentiate the two groups, but it is not expected that they confound the results. First, the two groups differ by gender ($\chi^2_3 = 5.649$, p = .130). The L2 English users are much more evenly balanced than the English NS group, which is female-dominant; this difference is not expected to cause any meaningful differences in the results, but it is factored into the statistical model on the chance that it does. The next two significant differences concern linguistic biographical information (age of exposure to the first language [$\chi^2_1 = 2.819$, $p = .093$] and age of comfort using the first language [$\chi^2_1 = 3.102$, $p = .078$]). These differences are considered to be a result of the inherent differences between the bilingual L2 User group and the overwhelmingly monolingual English NS group. The fourth significant difference concerns language exposure (years spent within a family in which the first language is spoken [$\chi^2_1 = 3.088$, $p = .079$]), which should not significantly affect the L2 English abilities of the L2 User group. The next three significant differences concern the first language (overall use of the first language [$\chi^2_1 = 53.714$, $p < .001$], overall self-assessed L1 proficiency [$\chi^2_1 = 15.203$, $p < .001$], overall language attitudes concerning the first language [$\chi^2_1 = 11.339$, $p < .001$]), which are again considered to be a result of the inherent differences between bilinguals and monolinguals. The final significant difference between the two groups concerns English proficiency ($F_{1,81} = 109.88$, $p < .001$); this result is expected, and variation in proficiency is considered in greater detail below where the L2 User group is further subdivided by English proficiency level. Based on the collected data, the results of these two groups are directly comparable with no significant caveats.

L2 ENGLISH PROFICIENCY GROUPS. This section focuses on potential confounding difference that may affect performance on the self-paced reading task within the L2 English User group divided by English proficiency. When necessary, meaningful differences between the proficiency subgroups and the English NS control group are discussed.[1] There are no statistically significant differences between the groups for the following three groups of collected variables: nonlinguistic biographical information (age and handedness), linguistic biographical information (age of exposure to both the first and second language and age of comfort using the second language), and language exposure (formal education in the first and second language, years spent immersed in a country in which the first and second language is spoken, years spent within a family in which the first and second language is spoken, and years spent immersed in a work/school environment in which the first language is spoken). The values for overall use of the second language, for overall self-assessed L2 proficiency, and overall language attitudes concerning the second language are also not significantly different among the proficiency subgroups.

There are a few factors that differentiate the proficiency groups. First and foremost, the proficiency groups differ by their proficiency score on the independent measure of proficiency as reported above. Second, the Int.-Low and the English NS groups differ by gender (DSCF = 3.841, $p = .052$); the Int.-Low group has a higher proportion of males to females than does the English NS group. Although females have been shown to be more successful language learners than males, this difference is not expected to affect the more meaningful comparisons between the L2 user groups. Third, the Low group is significantly more formally educated than the Int.-Low and the English

---

[1] Linguistic and biographical differences between bilingual and largely monolingual groups are excluded from this and all further discussions. This decision is made to accommodate the fundamental differences between the experiences of monolingual and bilingual people. The decision to exclude these differences affects information concerning the first language, including the following: age of comfort using the first language, overall use of the first language, overall self-assessed L1 proficiency and overall language attitudes concerning the first language.

NS groups (DSCF = 3.941, $p = .043$). Fourth, the proportion of L1 Chinese speakers is significantly higher in the two lowest proficiency groups than in the Adv. group (Int.-Low: DSCF = 4.177, $p = .026$; Low: DSCF = 3.508, $p = .095$); this factor is considered in detail when the L2 users are divided by first language. Overall, the L2 user groups divided by proficiency are expected to be comparable, and potential confounds are considered in the statistical models and in below analyses.

FIRST LANGUAGE GROUPS. This section focuses on potential confounding difference that may affect performance on the rating task within the L2 English User group divided by first language. As discussed previously in section 3.3, the L2 English users are divided into three groups based on L1 features that may transfer into the interlanguage grammar and affect their performance on these tasks. There are no statistically significant differences between the L2 user groups for the following three groups of collected variables: nonlinguistic biographical information (age, gender, handedness, and highest level of education achieved), linguistic biographical information (age of exposure to the first and second language and age of comfort using the first and second languages), and language exposure (formal education in the first and second language, years spent immersed in a country in which the first and the second language is spoken, years spent within a family in which the first and second language is spoken, and years spent immersed in a work/school environment in which the first and second language is spoken). The groups are also not significantly different for overall language attitudes concerning the second language.

There are several factors that differentiate the L1 groups. First, as mentioned above in the discussion of the proficiency groups, the Arabic and Chinese groups have scores that are significantly lower than the Other group but that are not different from each other ($p < .05$). As the critical comparisons are between the Arabic and Chinese groups, this difference should not affect the results. Second, the values for overall use of the second language are significantly lower for the Chinese group than for the Other group (DSCF = 3.612, $p = .029$). Third, the values for overall self-assessed L2 proficiency are significantly lower for the Chinese group than for the Other group (DSCF = 3.832, $p = .019$). These latter two findings are not expected to significantly affect the below analysis because the meaningful comparison is between the Arabic and Chinese groups, which do not differ. Overall, the L2 user groups divided by first language are expected to be comparable, and potential confounds are considered in the statistical models and in below analyses.

F.2 BIOGRAPHICAL INFORMATION.

This section captures basic biographical information, including biological data and personal and language history information. The responses reported here concern the following six survey items: (i) age in years, (ii) gender, (iii) hand used for writing, (iv) highest level of formal education, (v) first/native language(s), and (vi) second/foreign language(s) studied.

TABLE F.1. Descriptive statistics for current age (years) for L2 and native English speakers.

	n	$\bar{x}$	sd
L2 Users	53	19.000	3.201
English NS	40	19.975	3.840

An ANOVA revealed that there is no statistically significant difference for age between the groups ($F_{1,81} = 0.36$, $p = .549$).

TABLE F.2. Descriptive statistics for current age (years) for groups by L2 English proficiency.

	n	$\bar{x}$	sd
Adv.	11	20.273	2.760
Int.-High	11	21.636	4.225
Int.-Low	20	19.25	1.803
Low	11	21.455	4.009
English NS	40	19.975	3.840

An ANOVA revealed that there are no statistically significant differences for age among the proficiency groups ($F_{4,88} = 1.26$, $p = .293$).

TABLE F.3. Descriptive statistics for current age (years) for groups by L1.

	n	$\bar{x}$	sd
Arabic	10	22.600	4.006
Chinese	29	19.586	2.885
Other	13	13.462	5.410
English NS	40	19.975	3.840

An ANOVA revealed that there are no statistically significant differences for age among the language groups ($F_{3,89} = 2.07$, $p = .110$).

TABLE F.4. Frequency table for gender for L2 and native English speakers.

	n	male	female
L2 Users	52	29	23
English NS	40	15	25

A Wilcoxon Rank Sum Test revealed that there is a statistically significant difference for gender between the groups ($\chi^2_1 = 2.991$, $p = .083$).

TABLE F.5. Frequency table for gender for groups by L2 English proficiency.

	n	male	female
Adv.	11	5	6
Int.-High	11	4	7
Int.-Low	20	15	5
Low	10	5	5
English NS	40	15	25

A Kruskal-Wallis Test revealed that there is a statistically significant difference for gender among the proficiency groups ($\chi^2_4 = 8.161$, p = .086). A pairwise two-sided multiple comparison analysis using the Dwass, Steel, Critchlow-Fligner method revealed that this significant finding emerges from the difference in the proportions of gender between the Int.-Low and English NS groups (DSCF = 3.841, p = .052).

TABLE F.6. Frequency table for gender for groups by L1.

	n	male	female
Arabic	9	7	2
Chinese	29	16	13
Other	14	6	8
English NS	40	15	25

A Kruskal-Wallis Test revealed that there are no statistically significant differences for gender among the language groups ($\chi^2_3 = 5.649$, p = .130).

TABLE F.7. Frequency table for handedness for L2 and native English speakers.

	n	right	left	both equally
L2 Users	53	47	4	2
English NS	40	35	4	1

A Wilcoxon Rank Sum Test revealed that there is no statistically significant difference for handedness between the groups ($\chi^2_1 = 0.021$, p = .884).

TABLE F.8. Frequency table for handedness for groups by L2 English proficiency.

	n	right	left	both equally
Adv.	11	9	2	0
Int.-High	11	11	0	0
Int.-Low	20	19	1	0
Low	11	8	1	2
English NS	40	35	4	1

A Kruskal-Wallis Test revealed that there are no statistically significant differences for handedness among the proficiency groups ($\chi^2_4 = 5.643$, p = .227).

TABLE F.9. Frequency table for handedness for groups by L1.

	n	right	left	both equally
Arabic	10	8	2	0
Chinese	29	27	0	2
Other	14	12	2	0
English NS	40	35	4	1

A Kruskal-Wallis Test revealed that there are no statistically significant differences for handedness among the language groups ($\chi^2_3 = 1.117$, $p = .773$).

TABLE F.10. Frequency table for highest level of education for L2 and native English speakers.

	n	High school	Some college	College	Some grad. school	Masters	Doctorate
L2 Users	53	24	11	13	3	1	1
English NS	40	9	29	2	0	0	0

A Wilcoxon Rank Sum Test revealed that there is no statistically significant difference for highest level of education achieved between the groups ($\chi^2_1 = 0.013$, $p = .911$).

TABLE F.11. Frequency table for highest level of education for groups by L2 English proficiency.

	n	High school	Some college	College	Some grad. school	Masters	Doctorate
Adv.	11	4	4	2	1	0	0
Int.-High	11	4	3	3	1	0	0
Int.-Low	20	13	4	3	0	0	0
Low	11	3	0	5	1	1	1
English NS	40	9	29	2	0	0	0

A Kruskal-Wallis Test revealed that there are statistically significant differences for highest level of education achieved among the proficiency groups (χ^2_4 = 12.294, p = .015). A pairwise two-sided multiple comparison analysis using the Dwass, Steel, Critchlow-Fligner method revealed that this significant finding emerges from the differences in the proportions of education level between the Int.-Low and Low groups (DSCF = 3.941, p = .043) and between the Low and English NS groups (DSCF = 3.927, p = .044).

TABLE F.12. Frequency table for highest level of education for groups by L1.

	n	High school	Some college	College	Some grad. school	Masters	Doctorate
Arabic	10	4	0	6	0	0	0
Chinese	29	17	6	2	2	1	1
Other	14	3	5	5	1	0	0
English NS	40	9	29	2	0	0	0

A Kruskal-Wallis Test revealed that there are statistically significant differences for highest level of education achieved among the language groups (χ^2_3 = 6.612, p = .085). A pairwise two-sided multiple comparison analysis using the Dwass, Steel, Critchlow-Fligner method revealed that this significant finding does not obtain under greater scrutiny.

TABLE F.13. Frequency table for first languages groups by L2 English proficiency.

	n	English	Arabic	Chinese	French	Hindi	Japanese	Korean	Spanish	Ukrainian
Adv.	11	0	1	3	2	2	1	1	1	0
Int.-High	11	0	1	6	0	0	1	0	2	1
Int.-Low	20	0	6	13	0	0	1	0	0	0
Low	26	0	3	20	0	0	2	1	0	0
English NS	41	41	0	0	0	0	0	0	0	0

A Kruskal-Wallis Test revealed that there are statistically significant differences for first language among the proficiency groups ($\chi^2_4 = 89.887$, $p < .001$). A pairwise two-sided multiple comparison analysis using the Dwass, Steel, Critchlow-Fligner method revealed that this significant finding mostly emerges from the differences in the proportions of first language between the English L2 and English NS groups ($p < .001$), but it also emerges from the differences between the Adv. and Int.-Low groups (DSCF = 4.177, $p = .026$) and between the Adv. and Low groups (DSCF = 3.508, $p = .095$).

TABLE F.14. Frequency table for second/foreign languages studied for groups by L2 English proficiency.

	n	English	Chinese	French	German	Italian	Japanese	Korean	Latin	Russian	Spanish	Tagalog	Turkish
Adv.	11	11	1	1	1	1	1	1	0	0	0	0	0
Int.-High	11	11	0	0	0	0	1	0	0	1	0	0	0
Int.-Low	20	20	0	0	0	0	1	0	0	0	1	0	0
Low	26	26	0	1	1	0	2	3	0	0	1	0	0
English NS	40	0	2	11	5	4	5	1	3	1	28	1	0

TABLE F.15. Frequency table for second/foreign languages studied for groups by L1.

	n	English	Chinese	French	German	Italian	Japanese	Korean	Latin	Russian	Spanish	Tagalog	Turkish
Arabic	11	11	0	0	0	0	0	0	0	0	0	0	0
Chinese	42	42	0	1	1	0	4	2	0	0	1	0	0
Other	15	15	1	1	1	1	1	1	0	1	1	0	1
English NS	40	0	2	11	5	4	5	1	3	1	28	1	0

F.3 LANGUAGE HISTORY.

This section captures personal language history information, including language exposure, language comfort, education in a language, and immersion. The responses reported here concern the following six survey items: (i) age of exposure, (ii) age of comfort, (iii) formal education in that language, (iv) immersion in a country where that language is spoken, (v) immersion in a family where that language is spoken, and (vi) immersion in a school/work environment where that language is spoken.

TABLE F.16. Descriptive statistics for age of exposure for L2 and native English speakers.

		L1		L2 English	
	n	$\bar{x}$	sd	$\bar{x}$	sd
L2 Users	53	0.415	1.151	9.596	3.610
English NS	40	0.250	1.056		

A Wilcoxon Rank Sum Test revealed that there is a statistically significant difference between the groups for age of exposure to the first language ($\chi^2_1 = 2.819$, p = .093).

TABLE F.17. Descriptive statistics for age of exposure for groups by L2 English proficiency.

		L1		L2 English	
	n	$\bar{x}$	sd	$\bar{x}$	sd
Adv.	11	0.727	2.102	9.091	2.982
Int.-High	11	0.182	0.405	9.273	3.952
Int.-Low	20	0.300	0.733	9.800	3.622
Low	11	0.545	1.036	11.909	7.231
English NS	40	0.250	1.056		

A Kruskal-Wallis Test revealed that there are no statistically significant differences among the proficiency groups for age of exposure to the first language ($\chi^2_4 = 3.397$, p = .494) or for the age of exposure to the second language ($\chi^2_3 = 0.194$, p = .979).

TABLE F.18. Descriptive statistics for age of exposure for groups by L1.

		L1		L2 English	
	n	x̄	sd	x̄	sd
Arabic	10	0.200	0.422	11.200	5.453
Chinese	29	0.379	0.862	9.071	2.478
Other	14	0.643	1.865	9.500	3.898
English NS	40	0.250	1.056		

A Kruskal-Wallis Test revealed that there are no statistically significant differences among the language groups for age of exposure to the first language ($\chi^2_3 = 2.855$, p = .415) or for the age of exposure to the second language ($\chi^2_2 = 0.747$, p = .688).

TABLE F.19. Descriptive statistics for age of comfort using L2 and native English speakers.

		L1		L2 English	
	n	x̄	sd	x̄	Not comfortable
L2 Users	52	1.942	3.589	16.029	18
English NS	40	1.250	3.557		

A Wilcoxon Rank Sum Test revealed that there is a statistically significant difference between the groups for age of comfort using the first language ($\chi^2_1 = 3.102$, p = .078).

TABLE F.20. Descriptive statistics for age of comfort using groups by L2 English proficiency.

		L1		L2 English	
	n	x̄	sd	x̄	Not comfortable
Adv.	11	2.091	3.015	14.125	3
Int.-High	11	2.364	4.433	13.250	3
Int.-Low	20	0.300	1.129	17.000	8
Low	11	4.600	4.926	19.714	4
English NS	40	1.250	3.557		

A Kruskal-Wallis Test revealed that there is a statistically significant difference among the proficiency groups for age of comfort using the first language ($\chi^2_4 = 14.277$, p = .007) but that there are no significant differences for the age of comfort using the second language ($\chi^2_3 = 3.488$, $p = .3223$). A pairwise two-sided multiple comparison analysis using the Dwass, Steel, Critchlow-Fligner method revealed that the former significant finding emerges from the differences in the proportions of age of comfort between the Int.-Low and Low groups (DSCF = 4.357, $p = .018$) and between the Low and English NS groups (DSCF = 4.076, $p = .032$).

TABLE F.21. Descriptive statistics for age of comfort using groups by L1.

		L1		L2 English	
	n	x̄	sd	x̄	Not comfortable
Arabic	10	2.500	4.062	19.750	2
Chinese	28	1.821	3.539	14.938	12
Other	14	1.786	3.577	13.500	4
English NS	40	1.250	3.557		

A Kruskal-Wallis Test revealed that there are no statistically significant differences among the language groups for age of comfort using the first language ($\chi^2_3 = 3.602$, $p = .308$) or for the age of comfort using the second language ($\chi^2_2 = 0.838$, $p = .658$).

TABLE F.22. Descriptive statistics for years of formal education in that language for L2 and native English speakers.

		L1		L2 English	
	n	x̄	sd	x̄	sd
L2 Users	52	13.192	4.771	8.404	3.991
English NS	38	13.842	4.699		

A Kruskal-Wallis Test revealed that there is no statistically significant difference between the groups for years of formal education in the first language ($\chi^2_1 = 1.837$, $p = .175$).

TABLE F.23. Descriptive statistics for years of formal education in that language for groups by L2 English proficiency.

		L1		L2 English	
	n	x̄	sd	x̄	sd
Adv.	11	13.364	4.760	9.273	3.069
Int.-High	11	11.545	3.671	7.636	4.904
Int.-Low	20	14.500	3.873	8.250	3.323
Low	11	11.556	5.548	8.600	4.613
English NS	38	13.842	4.699		

A Kruskal-Wallis Test revealed that there are no statistically significant differences among the proficiency groups for years of formal education in the first language (χ^2_4 = 6.650, p = .156) or for years of formal education in the second language (χ^2_3 = 1.032, p = .794).

TABLE F.24. Descriptive statistics for years of formal education in that language for groups by L1.

		L1		L2 English	
	n	x̄	sd	x̄	sd
Arabic	10	13.700	5.579	8.600	4.452
Chinese	28	13.179	4.738	8.000	4.028
Other	14	12.857	4.312	9.071	3.772
English NS	38	13.842	4.699		

A Kruskal-Wallis Test revealed that there are no statistically significant differences among the language groups for years of formal education in the first language (χ^2_3 = 2.471, p = .481) or for years of formal education in the second language (χ^2_2 = 0.228, p = .893).

TABLE F.25. Descriptive statistics for years immersed in a country where that language is spoken for L2 and native English speakers.

		L1		L2 English	
	n	x̄	sd	x̄	sd
L2 Users	52	17.481	7.368	3.273	5.778
English NS	40	17.775	7.181		

A Wilcoxon Rank Sum Test revealed that there is no statistically significant difference between the groups for years spent immersed in a country in which the first language is spoken ($\chi^2_1 = 0.179$, $p = .672$).

TABLE F.26. Descriptive statistics for years immersed in a country where that language is spoken for groups by L2 English proficiency.

		L1		L2 English	
	n	x̄	sd	x̄	sd
Adv.	11	16.818	8.195	6.736	8.949
Int.-High	11	17.273	8.945	4.409	6.917
Int.-Low	20	17.750	4.411	1.505	2.643
Low	10	17.900	10.104	1.750	2.552
English NS	40	17.775	7.181		

A Kruskal-Wallis Test revealed that there are no statistically significant differences among the proficiency groups for years spent immersed in a country in which the first language is spoken ($\chi^2_4 = 2.262$, $p = .688$) or for years spent immersed in a country in which the second language is spoken ($\chi^2_3 = .663$, $p = .882$).

TABLE F.27. Descriptive statistics for years immersed in a country where that language is spoken for groups by L1.

		L1		L2 English	
	n	x̄	sd	x̄	sd
Arabic	10	19.700	7.931	1.300	1.703
Chinese	28	16.750	7.286	2.614	4.744
Other	14	17.357	7.365	6.000	8.383
English NS	40	17.775	7.181		

A Kruskal-Wallis Test revealed that there are no statistically significant differences among the language groups for years spent immersed in a country in which the first language is spoken ($\chi^2_3 = 2.233$, $p = .526$) or for years spent immersed in a country in which the second language is spoken ($\chi^2_2 = .322$, $p = .851$).

TABLE F.28. Descriptive statistics for years immersed in a family where that language is spoken for L2 and native English speakers.

		L1		L2 English	
	n	$\bar{x}$	sd	$\bar{x}$	sd
L2 Users	52	19.615	5.003	1.901	4.999
English NS	40	18.175	6.602		

A Wilcoxon Rank Sum Test revealed that there is a statistically significant difference between the groups for years spent within a family in which the first language is spoken ($\chi^2_1 = 3.088$, $p = .079$).

TABLE F.29. Descriptive statistics for years immersed in a family where that language is spoken for groups by L2 English proficiency.

		L1		L2 English	
	n	$\bar{x}$	sd	$\bar{x}$	sd
Adv.	11	17.000	8.343	5.664	8.278
Int.-High	11	21.182	4.771	0.455	1.508
Int.-Low	20	19.200	1.824	1.303	4.317
Low	10	21.600	4.006	0.550	0.960
English NS	40	18.175	6.602		

A Kruskal-Wallis Test revealed that there are no statistically significant differences among the proficiency groups for years spent within a family in which the first language is spoken ($\chi^2_4 = 5.828$, $p = .212$) but that there is a significant difference for years spent within a family in which the second language is spoken ($\chi^2_3 = 6.282$, $p = .099$). A pairwise two-sided multiple comparison analysis using the Dwass, Steel, Critchlow-Fligner method revealed that the latter significant finding is probably a type I error and that there is, in fact, no significant difference among the groups.

TABLE F.30. Descriptive statistics for years immersed in a family where that language is spoken for groups by L1.

		L1		L2 English	
	n	$\bar{x}$	sd	$\bar{x}$	sd
Arabic	10	22.400	4.006	1.830	5.682
Chinese	28	17.750	4.672	1.288	3.740
Other	14	19.357	5.839	3.179	6.638
English NS	40	18.175	6.602		

A Kruskal-Wallis Test revealed that there are statistically significant differences among the language groups for years spent within a family in which the first language is spoken ($\chi^2_3 = 8.034$, $p = .045$) but that there are no significant differences for years spent within a family in which the second language is spoken ($\chi^2_2 = 0.352$, $p = .839$). A pairwise two-sided multiple comparison analysis using the Dwass, Steel, Critchlow-Fligner method revealed that the former significant finding does not obtain under greater scrutiny.

TABLE F.31. Descriptive statistics for years immersed in work/school where that language is spoken for L2 and native English speakers.

		L1		L2 English	
	n	$\bar{x}$	sd	$\bar{x}$	sd
L2 Users	52	10.102	9.795	2.419	4.733
English NS	40	9.400	9.697		

A Wilcoxon Rank Sum Test revealed that there is no statistically significant difference between the groups for years spent immersed in a work/school environment in which the first language is spoken ($\chi^2_1 = 0.086$, $p = .769$).

TABLE F.32. Descriptive statistics for years immersed in work/school where that language is spoken for groups by L2 English proficiency.

		L1		L2 English	
	n	$\bar{x}$	sd	$\bar{x}$	sd
Adv.	11	10.727	9.530	5.273	7.295
Int.-High	11	6.818	10.196	3.636	5.904
Int.-Low	20	9.815	9.200	0.715	1.870
Low	10	13.600	11.007	1.350	1.492
English NS	40	9.400	9.697		

A Kruskal-Wallis Test revealed that there are no statistically significant differences among the proficiency groups for years spent immersed in a work/school environment in which the first language is spoken ($\chi^2_4 = 3.297$, $p = .509$) or for years spent immersed in a work/school environment in which the second language is spoken ($\chi^2_3 = 6.106$, $p = .107$).

TABLE F.33. Descriptive statistics for years immersed in work/school where that language is spoken for groups by L1.

		L1		L2 English	
	n	$\bar{x}$	sd	$\bar{x}$	sd
Arabic	10	7.730	9.358	1.680	1.755
Chinese	28	11.785	10.115	1.679	4.110
Other	14	8.429	9.476	4.429	6.688
English NS	40	9.400	9.697		

A Kruskal-Wallis Test revealed that there are no statistically significant differences among the language groups for years spent immersed in a work/school environment in which the first language is spoken ($\chi^2_3 = 0.935$, $p = .817$) or for years spent immersed in a work/school environment in which the second language is spoken ($\chi^2_2 = 3.350$, $p = .187$).

F.4 LANGUAGE USE.

This section captures information concerning language use on an average day. The data presented are generated from the averages of five survey items that ask for percentage of time in which that language is used (i) with friends, (ii) with family, (iii) at school/work, (iv) while talking to oneself, and (v) while counting.

TABLE F.34. Descriptive statistics for overall language use for L2 and native English speakers.

		L1		L2 English	
	n	$\bar{x}$	sd	$\bar{x}$	sd
L2 Users	53	63.688	17.691	35.547	18.035
English NS	40	93.815	8.992		

A Wilcoxon Rank Sum Test revealed that there is a statistically significant difference between the groups for overall use of the first language ($\chi^2_1 = 53.714$, $p < .001$).

TABLE F.35. Descriptive statistics for overall language use for groups by L2 English proficiency.

		L1		L2 English	
	n	x̄	sd	x̄	sd
Adv.	11	56.350	19.842	40.745	21.991
Int.-High	11	69.891	20.093	29.655	20.246
Int.-Low	20	65.950	14.430	33.980	14.475
Low	11	60.709	17.687	39.091	17.693
English NS	40	93.815	8.992		

A Kruskal-Wallis Test revealed that there is a statistically significant difference among the proficiency groups for overall use of the first language ($\chi^2_4 = 55.402$, $p < .001$) but that there are no significant differences for overall use of the second language ($\chi^2_3 = 2.434$, $p = .487$). A pairwise two-sided multiple comparison analysis using the Dwass, Steel, Critchlow-Fligner method revealed that the former significant finding emerges from the differences in the proportions of overall use between the L2 users and the English NS control group (Adv.: DSCF = 6.691, $p < .001$; Int.-High: DSCF = 5.958, $p < .001$; Int.-Low: DSCF = 7.788, $p < .001$; Low: DSCF = 6.674, $p < .001$), which is to be expected considering the largely monolingual English NS group.

TABLE F.36. Descriptive statistics for overall language use for groups by L1.

		L1		L2 English	
	n	x̄	sd	x̄	sd
Arabic	10	60.220	15.725	39.780	15.725
Chinese	29	71.069	13.734	28.766	13.899
Other	14	50.875	19.153	46.571	21.598
English NS	40	93.815	8.992		

A Kruskal-Wallis Test revealed that there is a statistically significant difference among the language groups for overall use of the first language ($\chi^2_3 = 58.859$, $p < .001$) and for overall use of the second language ($\chi^2_2 = 8.287$, $p = .016$). A pairwise two-sided multiple comparison analysis using the Dwass, Steel, Critchlow-Fligner method revealed that the former significant finding emerges from the differences in the proportions of overall use between the L2 users and the English NS control group (Arabic: DSCF = 6.450, $p < .001$; Chinese: DSCF = 8.376, $p < .001$; Other: DSCF = 7.584, $p < .001$) and between the Chinese and Other groups (DSCF = 4.531, p = .007). The same analysis revealed that the latter significant finding emerges from the differences in the proportions of overall use between the Chinese and Other groups (DSCF = 3.612, p = .029).

(i) Use with friends:

TABLE F.37. Descriptive statistics for language use with friends for L2 and native English speakers.

		L1		L2 English	
	n	x̄	sd	x̄	sd
L2 Users	52	61.500	24.020	38.830	25.446
English NS	40	97.225	5.731		

A Wilcoxon Rank Sum Test revealed that there is a statistically significant difference between the groups for first language use with friends ($\chi^2_1 = 54.212$, $p < .001$).

TABLE F.38. Descriptive statistics for language use with friends for groups by L2 English proficiency.

		L1		L2 English	
	n	x̄	sd	x̄	sd
Adv.	10	64.900	20.496	39.182	29.006
Int.-High	11	55.455	33.276	43.182	33.187
Int.-Low	20	64.700	24.446	35.050	24.528
Low	11	58.636	15.507	41.000	15.047
English NS	40	97.225	5.731		

A Kruskal-Wallis Test revealed that there is a statistically significant difference among the groups for first language use with friends ($\chi^2_4 = 54.817$, $p < .001$) but that there are no significant differences for second language use with friends ($\chi^2_3 = .895$, $p = .827$). A pairwise two-sided multiple comparison analysis using the Dwass, Steel, Critchlow-Fligner method revealed that the former significant finding emerges from the differences in the proportions of use with friends between the L2 users and the English NS control group (Adv.: DSCF = 6.573, $p < .001$; Int.-High: DSCF = 6.334, $p < .001$; Int.-Low: DSCF = 7.872, $p < .001$; Low: DSCF = 7.420, $p < .001$).

TABLE F.39. Descriptive statistics for language use with friends for groups by L1.

		L1		L2 English	
	n	$\bar{x}$	sd	$\bar{x}$	sd
Arabic	10	56.00	23.310	44.000	23.310
Chinese	29	70.138	19.608	29.517	19.829
Other	13	46.462	26.400	54.429	29.848
English NS	40	97.225	5.731		

A Kruskal-Wallis Test revealed that there is a statistically significant difference among the groups for first language use with friends ($\chi^2_3 = 58.160$, $p < .001$) and for second language use with friends ($\chi^2_2 = 8.346$, $p = .015$). A pairwise two-sided multiple comparison analysis using the Dwass, Steel, Critchlow-Fligner method revealed that the former significant finding emerges from the differences in the proportions of use with friends between the L2 users and the English NS control group (Arabic: DSCF = 7.062, $p < .001$; Chinese: DSCF = 8.570, $p < .001$; Other: DSCF = 7.512, $p < .001$). The same analysis revealed that the latter significant finding emerges from the differences in the proportions of use with friends between the Chinese and Other groups (DSCF = 3.784, $p = .020$).

(ii) Use with family:

TABLE F.40. Descriptive statistics for language use with family for L2 and native English speakers.

		L1		L2 English	
	n	$\bar{x}$	sd	$\bar{x}$	sd
L2 Users	53	91.717	16.553	7.094	14.941
English NS	40	95.450	15.032		

A Wilcoxon Rank Sum Test revealed that there is no statistically significant difference between the groups for first language use with family ($\chi^2_1 = 2.572$, $p = .109$).

TABLE F.41. Descriptive statistics for language use with family for groups by L2 English proficiency.

		L1		L2 English	
	n	x̄	sd	x̄	sd
Adv.	11	83.182	24.318	11.364	20.011
Int.-High	11	95.455	7.891	4.545	7.891
Int.-Low	20	97.45	6.370	2.550	6.370
Low	11	86.091	22.318	13.636	22.452
English NS	40	95.450	15.032		

A Kruskal-Wallis Test revealed that there are no statistically significant differences among the groups for first language use with family ($\chi^2_4 = 7.157$, $p = .128$) or for second language use with family ($\chi^2_3 = 2.358$, $p = .502$).

TABLE F.42. Descriptive statistics for language use with family for groups by L1.

		L1		L2 English	
	n	x̄	sd	x̄	sd
Arabic	10	92.500	9.204	7.500	9.204
Chinese	29	96.862	8.484	3.138	8.484
Other	14	80.500	26.218	15.000	24.019
English NS	40	95.450	15.032		

A Kruskal-Wallis Test revealed that there are statistically significant differences among the groups for first language use with family ($\chi^2_3 = 12.406$, $p = .006$) and for second language use with family ($\chi^2_2 = 5.089$, $p = .079$). A pairwise two-sided multiple comparison analysis using the Dwass, Steel, Critchlow-Fligner method revealed that the former significant finding emerges from the differences in the proportions of use with family between the Chinese and Other groups (DSCF = 3.947, p = .027) and the English and Other groups (DSCF = 3.995, p = .025). The same analysis revealed that the latter significant finding does not obtain under greater scrutiny.

(iii) Use at school/work:

TABLE F.43. Descriptive statistics for language use at school/work for L2 and native English speakers.

		L1		L2 English	
	n	x̄	sd	x̄	sd
L2 Users	53	31.660	26.513	67.906	26.477
English NS	40	93.850	7.509		

A Wilcoxon Rank Sum Test revealed that there is a statistically significant difference between the groups for first language use at school/work ($\chi^2_1 = 64.200$, $p < .001$).

TABLE F.44. Descriptive statistics for language use at school/work for groups by L2 English proficiency.

		L1		L2 English	
	n	x̄	sd	x̄	sd
Adv.	11	30.000	22.361	68.636	20.747
Int.-High	11	38.182	31.327	61.364	32.023
Int.-Low	20	26.650	26.656	73.250	26.752
Low	11	35.909	26.535	11.640	26.608
English NS	40	93.850	7.509		

A Kruskal-Wallis Test revealed that there is a statistically significant difference among the groups for first language use at school/work ($\chi^2_4 = 64.592$, $p < .001$) but that there are no significant differences for second language use at school/work ($\chi^2_3 = 1.550$, $p = .671$). A pairwise two-sided multiple comparison analysis using the Dwass, Steel, Critchlow-Fligner method revealed that the former significant finding emerges from the differences in the proportions of use at school/work between the L2 users and the English NS control group (Adv.: DSCF = 7.639, $p < .001$; Int.-High: DSCF = 7.163, $p < .001$; Int.-Low: DSCF = 8.439, $p < .001$; Low: DSCF = 7.245, $p < .001$).

TABLE F.45. Descriptive statistics for language use at school/work for groups by L1.

		L1		L2 English	
	n	x̄	sd	x̄	sd
Arabic	10	25.100	26.227	74.900	26.227
Chinese	29	37.828	26.283	61.931	26.619
Other	14	23.571	25.678	75.286	25.015
English NS	40	93.850	7.509		

A Kruskal-Wallis Test revealed that there is a statistically significant difference among the groups for first language use at school/work ($\chi^2_3 = 65.918$, $p < .001$) but that there are no significant differences for second language use at school/work ($\chi^2_2 = 4.212$, $p = .122$). A pairwise two-sided multiple comparison analysis using the Dwass, Steel, Critchlow-Fligner method revealed that the former significant finding emerges from the differences in the proportions of use at school/work between the L2 users and the English NS control group (Arabic: DSCF = 6.979, $p < .001$; Chinese: DSCF = 9.596, $p < .001$; Other: DSCF = 7.933, $p < .001$).

(iv) Use with self:

TABLE F.46. Descriptive statistics for language use with self for L2 and native English speakers.

		L1		L2 English	
	n	x̄	sd	x̄	sd
L2 Users	53	65.208	32.001	33.377	31.634
English NS	40	90.300	17.985		

A Wilcoxon Rank Sum Test revealed that there is a statistically significant difference between the groups for first language use with themselves ($\chi^2_1 = 18.052$, $p < .001$).

TABLE F.47. Descriptive statistics for language use with self for groups by L2 English proficiency.

		L1		L2 English	
	n	x̄	sd	x̄	sd
Adv.	11	55.909	33.972	37.727	33.118
Int.-High	11	78.545	32.654	21.000	32.939
Int.-Low	20	66.400	30.120	33.600	30.120
Low	11	59.000	32.326	41.000	32.326
English NS	40	90.300	17.985		

A Kruskal-Wallis Test revealed that there is a statistically significant difference among the groups for first language use with themselves ($\chi^2_4 = 22.157$, $p < .001$) but that there are no significant differences for second language use with themselves ($\chi^2_3 = 3.682$, $p = .298$). A pairwise two-sided multiple comparison analysis using the Dwass, Steel, Critchlow-Fligner method revealed that the former significant finding emerges from the differences in the proportions of use with themselves between the L2 users and the English NS control group (Adv.: DSCF = 5.120, p = .003; Int.-Low: DSCF = 4.701, p = .008; Low: DSCF = 4.485, p = .013).

TABLE F.48. Descriptive statistics for language use with self for groups by L1.

		L1		L2 English	
	n	x̄	sd	x̄	sd
Arabic	10	66.100	30.985	33.900	30.985
Chinese	29	72.759	27.729	27.069	27.887
Other	14	48.929	36.908	46.071	37.428
English NS	40	90.300	17.985		

A Kruskal-Wallis Test revealed that there is a statistically significant difference among the groups for first language use with themselves ($\chi^2_3 = 22.019$, $p < .001$) but that there are no significant differences for second language use with themselves ($\chi^2_2 = 3.347$, $p = .188$). A pairwise two-sided multiple comparison analysis using the Dwass, Steel, Critchlow-Fligner method revealed that the former significant finding emerges from the differences in the proportions of use with themselves between the L2 users and the

English NS control group (Arabic: DSCF = 4.575, p = .007; Chinese: DSCF = 3.971, p = .026; Other: DSCF = 7.933, p < .001).

(v) Use to count:

TABLE F.49. Descriptive statistics for language use to count for L2 and native English speakers.

		L1		L2 English	
	n	x̄	sd	x̄	sd
L2 Users	53	69.038	29.768	30.528	29.957
English NS	40	92.250	16.470		

A Wilcoxon Rank Sum Test revealed that there is a statistically significant difference between the groups for first language use to count ($\chi^2_1 = 17.964$, $p < .001$).

TABLE F.50. Descriptive statistics for language use to count for groups by L2 English proficiency.

		L1		L2 English	
	n	x̄	sd	x̄	sd
Adv.	11	51.364	43.019	46.818	44.739
Int.-High	11	81.818	20.405	18.182	20.405
Int.-Low	20	74.550	23.314	25.450	23.314
Low	11	63.909	26.365	35.818	25.965
English NS	40	92.250	16.470		

A Kruskal-Wallis Test revealed that there is a statistically significant difference among the groups for first language use to count ($\chi^2_4 = 21.914$, $p < .001$) but that there are no significant differences for second language use with themselves ($\chi^2_3 = 3.741$, $p = .291$). A pairwise two-sided multiple comparison analysis using the Dwass, Steel, Critchlow-Fligner method revealed that the former significant finding emerges from the differences in the proportions of use to count between the L2 users and the English NS control group (Adv.: DSCF = 4.906, p = .005; Int.-Low: DSCF = 3.549, p = .088; Low: DSCF = 4.975, p = .004).

TABLE F.51. Descriptive statistics for language use to count for groups by L1.

		L1		L2 English	
	n	x̄	sd	x̄	sd
Arabic	10	61.400	26.896	38.600	26.895
Chinese	29	77.759	24.553	22.172	24.372
Other	14	56.429	36.871	4.071	38.213
English NS	40	92.250	16.470		

A Kruskal-Wallis Test revealed that there is a statistically significant difference among the groups for first language use to count ($\chi^2_3 = 25.965$, $p < .001$) but that there are no significant differences for second language use with themselves ($\chi^2_2 = 6.037$, $p = .049$). A pairwise two-sided multiple comparison analysis using the Dwass, Steel, Critchlow-Fligner method revealed that the former significant finding emerges from the differences in the proportions of use to count between the L2 users and the English NS control group (Arabic: DSCF = 5.411, $p < .001$; Other: DSCF = 6.172, $p < .001$). The same analysis revealed that the latter significant finding does not obtain under greater scrutiny.

F.5 LANGUAGE PROFICIENCY.

This section captures information concerning self-assessed language proficiency. The data presented are generated from the averages of four survey items that ask for a Likert-style rating from 0 (not well at all) to 6 (very well) on four language skills (i) speaking, (ii) understanding, (iii) reading, and (iv) writing.

TABLE F.52. Descriptive statistics for self-assessed proficiency for L2 and native English speakers.

		L1		L2 English	
	n	x̄	sd	x̄	sd
L2 Users	53	5.500	0.704	3.660	1.051
English NS	40	5.944	0.183		

A Wilcoxon Rank Sum Test revealed that there is a statistically significant difference between the groups for overall self-assessed L1 proficiency ($\chi^2_1 = 15.203$, $p < .001$).

TABLE F.53. Descriptive statistics for self-assessed proficiency for groups by L2 English proficiency.

		L1		L2 English	
	n	x̄	sd	x̄	sd
Adv.	11	5.250	0.844	4.023	0.847
Int.-High	11	5.318	0.902	3.705	1.106
Int.-Low	20	5.525	0.601	3.375	1.128
Low	11	5.886	0.303	3.773	1.034
English NS	40	5.944	0.183		

A Kruskal-Wallis Test revealed that there is a statistically significant difference among the proficiency groups for overall self-assessed L1 proficiency ($\chi^2_4 = 22.327$, $p < .001$) but that there are no significant differences for overall self-assessed L2 proficiency ($\chi^2_3 = 2.456$, $p = .483$). A pairwise two-sided multiple comparison analysis using the Dwass, Steel, Critchlow-Fligner method revealed that the former significant finding emerges from the differences in the proportions of overall self-assessed proficiency between the L2 users and the English NS control group (Adv.: DSCF = 5.319, p = .002; Int.-High: DSCF = 4.594, p = .010; Int.-Low: DSCF = 5.233, p = .002).

TABLE F.54. Descriptive statistics for self-assessed proficiency for groups by L1.

		L1		L2 English	
	n	x̄	sd	x̄	sd
Arabic	10	5.850	0.268	3.950	0.941
Chinese	29	5.491	0.614	3.293	0.975
Other	14	5.268	0.983	4.214	1.032
English NS	40	5.944	0.183		

A Kruskal-Wallis Test revealed that there is a statistically significant difference among the language groups for overall self-assessed L1 proficiency ($\chi^2_3 = 18.249$, $p < .001$) and for overall self-assessed L2 proficiency ($\chi^2_2 = 8.795$, $p = .012$). A pairwise two-sided multiple comparison analysis using the Dwass, Steel, Critchlow-Fligner method revealed that the former significant finding emerges from the differences in the proportions of overall self-assessed proficiency between the L2 users and the English NS control group

(Chinese: DSCF = 5.655, p = .004; Other: DSCF = 4.490, p = .008). The same analysis revealed that the latter significant finding emerges from the differences in the proportions of overall self-assessed proficiency between the Chinese and Other groups (DSCF = 3.832, p = .019)

(i) Speaking self-assessment:

TABLE F.55. Descriptive statistics for self-assessed speaking proficiency for L2 and native English speakers.

		L1		L2 English	
	n	x̄	sd	x̄	sd
L2 Users	53	5.528	0.749	3.558	1.036
English NS	40	5.925	0.267		

A Wilcoxon Rank Sum Test revealed that there is a statistically significant difference between the groups for overall self-assessed L1 speaking proficiency (χ^2_1 = 9.492, p = .002).

TABLE F.56. Descriptive statistics for self-assessed speaking proficiency for groups by L2 English proficiency.

		L1		L2 English	
	n	x̄	sd	x̄	sd
Adv.	11	5.273	0.786	3.727	0.786
Int.-High	11	5.455	0.820	3.727	1.191
Int.-Low	20	5.550	0.826	3.316	1.108
Low	11	5.818	0.405	3.636	1.027
English NS	40	5.925	0.267		

A Kruskal-Wallis Test revealed that there is a statistically significant difference among the groups for self-assessed L1 speaking proficiency (χ^2_4 = 14.182, p = .007) but that there are no significant differences for self-assessed L2 speaking proficiency (χ^2_3 = 1.221, p = .748). A pairwise two-sided multiple comparison analysis using the Dwass, Steel, Critchlow-Fligner method revealed that the former significant finding emerges from the differences in the proportions of self-assessed L1 speaking proficiency between the more advanced L2 users and the English NS control group (Adv.: DSCF = 5.206, p = .002; Int.-High: DSCF = 3.605, p = .080).

TABLE F.57. Descriptive statistics for self-assessed speaking proficiency for groups by L1.

		L1		L2 English	
	n	x̄	sd	x̄	sd
Arabic	10	5.900	0.316	3.900	0.994
Chinese	29	5.483	0.785	28.3250	1.005
Other	14	5.357	0.842	3.929	0.997
English NS	40	5.925	0.267		

A Kruskal-Wallis Test revealed that there is a statistically significant difference among the groups for self-assessed L1 speaking proficiency ($\chi^2_3 = 14.025$, p = .003) and for self-assessed L2 speaking proficiency ($\chi^2_2 = 6.708$, p = .035). A pairwise two-sided multiple comparison analysis using the Dwass, Steel, Critchlow-Fligner method revealed that the former significant finding emerges from the differences in the proportions of self-assessed L1 speaking proficiency between the L2 users and the English NS control group (Chinese: DSCF = 4.447, p = .009; Other: DSCF = 4.459, p = .009). The same analysis revealed that the latter significant finding emerges from the differences in the proportions of self-assessed L2 speaking proficiency between the Chinese and Other groups (DSCF = 3.193, p = .062)

(ii) Understanding self-assessment:

TABLE F.58. Descriptive statistics for self-assessed understanding proficiency for L2 and native English speakers.

		L1		L2 English	
	n	x̄	sd	x̄	sd
L2 Users	53	5.566	0.721	3.788	1.194
English NS	40	5.950	0.221		

A Wilcoxon Rank Sum Test revealed that there is a statistically significant difference between the groups for overall self-assessed L1 understanding proficiency ($\chi^2_1 = 10.419$, $p = .001$).

TABLE F.59. Descriptive statistics for self-assessed understanding proficiency for groups by L2 English proficiency.

		L1		L2 English	
	n	$\bar{x}$	sd	$\bar{x}$	sd
Adv.	11	5.182	0.982	4.091	1.044
Int.-High	11	5.455	0.688	3.909	1.136
Int.-Low	20	5.650	0.671	3.526	1.349
Low	11	5.909	0.302	3.818	1.168
English NS	40	5.950	0.221		

A Kruskal-Wallis Test revealed that there is a statistically significant difference among the groups for self-assessed L1 understanding proficiency ($\chi^2_4 = 19.316$, $p < .001$) but that there are no significant differences for self-assessed L2 understanding proficiency ($\chi^2_3 = 1.319$, $p = .725$). A pairwise two-sided multiple comparison analysis using the Dwass, Steel, Critchlow-Fligner method revealed that the former significant finding emerges from the differences in the proportions of self-assessed L1 understanding proficiency between the more advanced L2 users and the English NS control group (Adv.: DSCF = 5.690, $p < .001$; Int.-High: DSCF = 4.883, $p = .005$).

TABLE F.60. Descriptive statistics for self-assessed understanding proficiency for groups by L1.

		L1		L2 English	
	n	$\bar{x}$	sd	$\bar{x}$	sd
Arabic	10	6.000	0	4.100	1.101
Chinese	29	5.552	0.632	3.464	1.138
Other	14	5.286	0.994	4.214	1.251
English NS	40	5.950	0.221		

A Kruskal-Wallis Test revealed that that there is a statistically significant difference among the groups for self-assessed L1 understanding proficiency ($\chi^2_3 = 18.505$, $p < .001$)

but that there are no significant differences for self-assessed L2 understanding proficiency (χ^2_2 = 4.554, p = .103). A pairwise two-sided multiple comparison analysis using the Dwass, Steel, Critchlow-Fligner method revealed that the former significant finding emerges from the differences in the proportions of self-assessed L1 understanding proficiency between the L2 users and the English NS control group (Chinese: DSCF = 4.885, p = .003; Other: DSCF = 4.929, p = .003) and between the Arabic and Other groups (DSCF = 3.272, p = .095).

(iii) Reading self-assessment:

TABLE F.61. Descriptive statistics for self-assessed reading proficiency for L2 and native English speakers.

		L1		L2 English	
	n	x̄	sd	x̄	sd
L2 Users	53	5.566	0.747	3.868	1.144
English NS	40	5.975	0.158		

A Wilcoxon Rank Sum Test revealed that there is a statistically significant difference between the groups for overall self-assessed L1 reading proficiency (χ^2_1 = 11.699, p < .001).

TABLE F.62. Descriptive statistics for self-assessed reading proficiency for groups by L2 English proficiency.

		L1		L2 English	
	n	x̄	sd	x̄	sd
Adv.	11	5.273	0.905	4.273	0.905
Int.-High	11	5.273	1.009	3.818	1.328
Int.-Low	20	5.650	0.671	3.526	1.349
Low	11	5.909	0.302	3.818	1.168
English NS	40	5.975	0.158		

A Kruskal-Wallis Test revealed that there is a statistically significant difference among the groups for self-assessed L1 reading proficiency (χ^2_4 = 19.504, p < .001) but that there are no significant differences for self-assessed L2 reading proficiency (χ^2_3 = 2.326, p =

.508). A pairwise two-sided multiple comparison analysis using the Dwass, Steel, Critchlow-Fligner method revealed that the former significant finding emerges from the differences in the proportions of self-assessed reading proficiency between the L2 users and the English NS control group (Adv.: DSCF = 5.560, $p < .001$; Int.-High: DSCF = 5.531, $p < .001$; Int.-Low: DSCF = 3.859, $p = .050$).

TABLE F.63. Descriptive statistics for self-assessed reading proficiency for groups by L1.

		L1		L2 English	
	n	$\bar{x}$	sd	$\bar{x}$	sd
Arabic	10	5.900	0.316	4.200	0.919
Chinese	29	5.586	0.628	3.414	1.119
Other	14	5.286	1.069	4.571	0.938
English NS	40	5.975	0.158		

A Kruskal-Wallis Test revealed that there is a statistically significant difference among the groups for self-assessed L1 reading proficiency ($\chi^2_3 = 15.426$, $p = .002$) and for self-assessed L2 reading proficiency ($\chi^2_2 = 11.183$, $p = .004$). A pairwise two-sided multiple comparison analysis using the Dwass, Steel, Critchlow-Fligner method revealed that former significant finding emerges from the differences in the proportions of self-assessed reading proficiency between the L2 users and the English NS control group (Chinese: DSCF = 5.045, $p = .002$; Other: DSCF = 4.862, $p = .003$). The same analysis revealed that the latter significant finding emerges from the differences in the proportions of self-assessed reading proficiency between the Chinese and Other groups (DSCF = 4.374, $p = .006$).

(iv) Writing self-assessment:

TABLE F.64. Descriptive statistics for self-assessed writing proficiency for L2 and native English speakers.

		L1		L2 English	
	n	x̄	sd	x̄	sd
L2 Users	53	5.340	0.939	3.453	1.218
English NS	39	5.923	0.270		

A Wilcoxon Rank Sum Test revealed that there is a statistically significant difference between the groups for overall self-assessed L1 writing proficiency ($\chi^2_1 = 13.622$, $p < .001$).

TABLE F.65. Descriptive statistics for self-assessed writing proficiency for groups by L2 English proficiency.

		L1		L2 English	
	n	x̄	sd	x̄	sd
Adv.	11	5.273	0.905	4.000	1.000
Int.-High	11	5.455	0.688	3.909	1.136
Int.-Low	20	5.200	0.951	3.100	1.165
Low	11	5.909	0.302	3.636	1.286
English NS	39	5.923	0.270		

A Kruskal-Wallis Test revealed that there is a statistically significant difference among the groups for self-assessed L1 writing proficiency ($\chi^2_4 = 21.584$, $p < .001$) but that there are no significant differences for self-assessed L2 reading proficiency ($\chi^2_3 = 4.775$, $p = .189$). A pairwise two-sided multiple comparison analysis using the Dwass, Steel, Critchlow-Fligner method revealed that the former significant finding emerges from the differences in the proportions of self-assessed writing proficiency between the L2 users and the English NS control group (Adv.: DSCF = 4.443, p = .015; Int.-High: DSCF = 5.133, p = .003; Int.-Low: DSCF = 5.400, p = .001).

TABLE F.66. Descriptive statistics for self-assessed writing proficiency for groups by L1.

		L1		L2 English	
	n	$\bar{x}$	sd	$\bar{x}$	sd
Arabic	10	5.600	0.843	3.600	1.174
Chinese	29	5.345	0.814	3.069	1.010
Other	14	5.143	1.231	4.143	1.231
English NS	39	5.923	0.270		

A Kruskal-Wallis Test revealed that that there is a statistically significant difference among the groups for self-assessed L1 writing proficiency ($\chi^2_3 = 15.862$, $p = .001$) and for self-assessed L2 reading proficiency ($\chi^2_2 = 8.350$, $p = .015$). A pairwise two-sided multiple comparison analysis using the Dwass, Steel, Critchlow-Fligner method revealed that former significant finding emerges from the differences in the proportions of self-assessed writing proficiency between the L2 users and the English NS control group (Chinese: DSCF = 5.465, $p < .001$; Other: DSCF = 4.454, $p = .009$). The same analysis revealed that the latter significant finding emerges from the differences in the proportions of self-assessed writing proficiency between the Chinese and Other groups (DSCF = 4.014, $p = .013$).

F.6 LANGUAGE ATTITUDES.

This section captures personal beliefs about the languages, their speakers, and personal identity. The data presented are generated from the averages of four survey items that ask for a Likert-style rating from 0 (strongly disagree) to 6 (strongly agree) on four attitudes (i) feeling like oneself when speaking that language, (ii) identification with the culture that speaks that language, (iii) desire to use that language like a native speaker, and (iv) importance of passing as a native speaker.

TABLE F.67. Descriptive statistics for language attitudes for L2 and native English speakers.

		L1		L2 English	
	n	x̄	sd	x̄	sd
L2 Users	53	5.509	0.670	4.601	0.991
English NS	40	5.788	0.614		

A Wilcoxon Rank Sum Test revealed that there is a statistically significant difference between the groups for overall language attitudes concerning the first language (χ^2_1 = 11.339, p < .001).

TABLE F.68. Descriptive statistics for language attitudes for groups by L2 English proficiency.

		L1		L2 English	
	n	x̄	sd	x̄	sd
Adv.	11	5.682	0.372	4.659	1.014
Int.-High	11	5.523	0.647	4.886	0.728
Int.-Low	20	5.400	0.829	4.613	1.018
Low	11	5.523	0.647	4.235	1.159
English NS	40	5.788	0.614		

A Kruskal-Wallis Test revealed that there is a statistically significant difference among the proficiency groups for overall language attitudes concerning the first language (χ^2_4 = 11.445, p = .022) but that there are no significant differences for overall language attitudes concerning the second language (χ^2_3 = 2.251, p = .522). A pairwise two-sided multiple comparison analysis using the Dwass, Steel, Critchlow-Fligner method revealed that the former significant finding emerges from the differences in the proportions of overall language attitudes concerning the second language between the L2 users and the English NS control group (Adv.: DSCF = 4.000, p = .038; Int.-Low: DSCF = 3.877, p = .048; Low: DSCF = 3.485, p = .099).

TABLE F.69. Descriptive statistics for language attitudes for groups by L1.

		L1		L2 English	
	n	x̄	sd	x̄	sd
Arabic	10	5.575	0.657	5.050	0.599
Chinese	29	5.491	0.614	3.293	0.975
Other	14	5.607	0.457	4.339	1.277
English NS	40	5.788	0.614		

A Kruskal-Wallis Test revealed that there is a statistically significant difference among the language groups for overall language attitudes concerning the first language (χ^2_3 = 11.545, p = .009) but that there are no significant differences for overall language attitudes concerning the second language (χ^2_2 = 2.697, p = .260). A pairwise two-sided multiple comparison analysis using the Dwass, Steel, Critchlow-Fligner method revealed that the former significant finding emerges from the differences in the proportions of overall language attitudes concerning the second language between the L2 users and the English NS control group (Chinese: DSCF = 4.016, p = .023; Other: DSCF = 4.352, p = .011).

(i) Feel like myself when speaking…

TABLE F.70. Descriptive statistics for feeling like oneself speaking that language for L2 and native English speakers.

		L1		L2 English	
	n	x̄	sd	x̄	sd
L2 Users	53	5.566	0.636	4.057	1.365
English NS	40	5.750	0.954		

A Wilcoxon Rank Sum Test revealed that there is a statistically significant difference between the groups for feeling like oneself speaking the L1 (χ^2_1 = 8.209, p = .004).

TABLE F.71. Descriptive statistics for feeling like oneself speaking that language for groups by L2 English proficiency.

		L1		L2 English	
	n	x̄	sd	x̄	sd
Adv.	11	5.364	0.674	3.636	1.804
Int.-High	11	5.545	0.688	4.364	1.120
Int.-Low	20	5.600	0.598	4.150	1.348
Low	11	5.727	0.647	4.000	1.183
English NS	40	5.750	0.954		

A Kruskal-Wallis Test revealed that there is a statistically significant difference among the groups for feeling like oneself speaking the L1 ($\chi^2_4 = 11.591$, $p = .021$) but that there are no significant differences for feeling like oneself speaking the L2 ($\chi^2_3 = 1.076$, $p = .783$). A pairwise two-sided multiple comparison analysis using the Dwass, Steel, Critchlow-Fligner method revealed that the former significant finding emerges from the differences in the proportions of feeling like oneself speaking the L1 between the Adv. and English NS groups (DSCF = 4.635, $p = .009$).

TABLE F.72. Descriptive statistics for feeling like oneself speaking that language for groups by L1.

		L1		L2 English	
	n	x̄	sd	x̄	sd
Arabic	10	5.800	0.422	4.500	0.972
Chinese	29	5.586	0.628	4.034	1.267
Other	14	5.357	0.745	3.786	1.762
English NS	40	5.750	0.954		

A Kruskal-Wallis Test revealed that there is a statistically significant difference among the groups for feeling like oneself speaking the L1 ($\chi^2_3 = 11.262$, $p = .010$) but that there are no significant differences for feeling like oneself speaking the L2 ($\chi^2_2 = 1.387$, $p = .500$). A pairwise two-sided multiple comparison analysis using the Dwass, Steel, Critchlow-Fligner method revealed that the former significant finding emerges from the differences in the proportions of feeling like oneself speaking the L1 between the L2 users and the English NS control group (Chinese: DSCF = 3.597, p = .054; Other: DSCF = 4.516, p = .008).

(ii) Identify with culture that speaks…

TABLE F.73. Descriptive statistics for identifying with the culture that speaks that language for L2 and native English speakers.

		L1		L2 English	
	n	$\bar{x}$	sd	$\bar{x}$	sd
L2 Users	53	5.472	0.868	4.208	1.433
English NS	40	5.775	0.862		

A Wilcoxon Rank Sum Test revealed that there is a statistically significant difference between the groups for identifying with the culture that uses the L1 ($\chi^2_1 = 6.707$, p = .010).

TABLE F.74. Descriptive statistics for identifying with the culture that speaks that language for groups by L2 English proficiency.

		L1		L2 English	
	n	$\bar{x}$	sd	$\bar{x}$	sd
Adv.	11	5.636	0.674	4.182	1.537
Int.-High	11	5.091	1.221	4.273	1.272
Int.-Low	20	5.600	0.681	4.300	1.658
Low	11	5.455	0.934	4.000	1.183
English NS	40	5.775	0.862		

A Kruskal-Wallis Test revealed that there is a statistically significant difference among the groups for identifying with the culture that uses the L1 ($\chi^2_4 = 8.466$, p = .076) but that there are no significant differences for identifying with the culture that uses the L2 ($\chi^2_3 = .521$, p = .914). A pairwise two-sided multiple comparison analysis using the Dwass, Steel, Critchlow-Fligner method revealed that the former significant finding emerges from the differences in the proportions of identifying with the culture that uses the L1 between the Int.-High and English NS groups (DSCF = 3.828, p = .053).

TABLE F.75. Descriptive statistics for identifying with the culture that speaks that language for groups by L1.

		L1		L2 English	
	n	$\bar{x}$	sd	$\bar{x}$	sd
Arabic	10	5.300	0.949	4.700	1.059
Chinese	29	5.483	0.829	4.172	1.513
Other	14	5.571	0.938	3.929	1.492
English NS	40	5.775	0.862		

A Kruskal-Wallis Test revealed that there is a statistically significant difference among the groups for identifying with the culture that uses the L1 ($\chi^2_3 = 8.707$, p = .033) but that there are no significant differences for identifying with the culture that uses the L2 ($\chi^2_2 = 1.754$, p = .415). A pairwise two-sided multiple comparison analysis using the Dwass, Steel, Critchlow-Fligner method revealed that the former significant finding emerges from the differences in the proportions of identifying with the culture that uses the L1 between the L2 users and the English NS control group (Arabic: DSCF = 3.950, p = .027; Chinese: DSCF = 3.395, p = .077).

(iii) Want to use … like a NS:

TABLE F.76. Descriptive statistics for wanting to use that language like a native speaker for L2 and native English speakers.

		L1		L2 English	
	n	$\bar{x}$	sd	$\bar{x}$	sd
L2 Users	53	5.453	1.218	5.308	1.307
English NS	40	5.875	0.791		

A Wilcoxon Rank Sum Test revealed that there is a statistically significant difference between the groups for wanting to use the L1 like a native speaker ($\chi^2_1 = 7.2049$, p = .007).

TABLE F.77. Descriptive statistics for wanting to use that language like a native speaker for groups by L2 English proficiency.

		L1		L2 English	
	n	x̄	sd	x̄	sd
Adv.	11	5.909	0.302	5.636	0.674
Int.-High	11	5.818	0.603	5.727	0.647
Int.-Low	20	5.100	1.683	5.250	1.482
Low	11	5.273	1.104	4.600	1.776
English NS	40	5.875	0.791		

A Kruskal-Wallis Test revealed that there is a statistically significant difference among the groups for wanting to use the L1 like a native speaker ($\chi^2_4 = 13.273$, p = .010) but that there are no significant differences for wanting to use the L2 like a native speaker ($\chi^2_3 =$ 3.511, p = .319). A pairwise two-sided multiple comparison analysis using the Dwass, Steel, Critchlow-Fligner method revealed that the former significant finding emerges from the differences in the proportions of wanting to use the L1 like a native speaker between the less proficient L2 users and the English NS control group (Int.-Low: DSCF = 4.297, p = .020; Low: DSCF = 4.551, p = .011).

TABLE F.78. Descriptive statistics for wanting to use that language like a native speaker for groups by L1.

		L1		L2 English	
	n	x̄	sd	x̄	sd
Arabic	10	5.200	1.751	5.600	1.265
Chinese	29	5.345	1.261	5.286	1.243
Other	14	5.857	0.363	5.143	1.511
English NS	40	5.875	0.791		

A Kruskal-Wallis Test revealed that there is a statistically significant difference among the groups for wanting to use the L1 like a native speaker ($\chi^2_3 = 8.827$, p = .032) but that there are no significant differences for wanting to use the L2 like a native speaker ($\chi^2_2 =$ 1.916, p = .384). A pairwise two-sided multiple comparison analysis using the Dwass, Steel, Critchlow-Fligner method revealed that the former significant finding emerges from the differences in the proportions of wanting to use the L1 like a native speaker between the Chinese and English groups (DSCF = 4.172, p = .017).

(iv) Want to be seen as a NS:

TABLE F.79. Descriptive statistics for wanting to be seen as a native speaker of that language for L2 and native English speakers.

		L1		L2 English	
	n	$\bar{x}$	sd	$\bar{x}$	sd
L2 Users	53	5.547	0.992	4.849	1.460
English NS	40	5.750	1.104		

A Wilcoxon Rank Sum Test revealed that there is a statistically significant difference between the groups for wanting to be seen as a native speaker of the L1 ($\chi^2_1 = 5.591$, p = .018).

TABLE F.80. Descriptive statistics for wanting to be seen as a native speaker of that language for groups by L2 English proficiency.

		L1		L2 English	
	n	$\bar{x}$	sd	$\bar{x}$	sd
Adv.	11	5.818	0.404	5.182	1.079
Int.-High	11	5.636	0.924	5.182	0.982
Int.-Low	20	5.300	1.342	4.750	1.713
Low	11	5.636	0.674	4.364	1.690
English NS	40	5.750	1.104		

A Kruskal-Wallis Test revealed that there are no statistically significant differences among the groups for wanting to be seen as a native speaker of the L1 ($\chi^2_4 = 6.922$, p = .140) or for wanting to be seen as a native speaker of the L2 ($\chi^2_3 = 1.613$, p = .657).

TABLE F.81. Descriptive statistics for wanting to be seen as a native speaker of that language for groups by L1.

		L1		L2 English	
	n	$\bar{x}$	sd	$\bar{x}$	sd
Arabic	10	6.000	0	5.400	1.075
Chinese	29	5.345	1.233	4.828	1.560
Other	14	5.643	0.633	4.500	1.454
English NS	40	5.750	1.104		

A Kruskal-Wallis Test revealed that there is a statistically significant difference among the groups for wanting to be seen as a native speaker of the L1 ($\chi^2_3 = 10.867$, $p = .013$) but that there are no significant differences for wanting to be seen as a native speaker of the L2 ($\chi^2_2 = 3.106$, $p = .212$). A pairwise two-sided multiple comparison analysis using the Dwass, Steel, Critchlow-Fligner method revealed that the former significant finding emerges from the differences in the proportions of wanting to be seen as a native speaker of the L1 between the Chinese and English groups (DSCF = 3.859, $p = .032$).

APPENDIX G

STATISTICAL INFORMATION – SPRT, FIRST ADMINISTRATION

G.1 GENERALIZED LINEAR MODELS.

BOUNDEDNESS, *IN TOTO*.

log(RT) = Group Condition Location Gender

Group*Condition Group*Gender

Group*Condition*Location

BOUNDEDNESS, ENGLISH PROFICIENCY.

log(RT) = ProficiencyGroup Condition Location Education Gender

ProficiencyGroup*Condition ProficiencyGroup*Gender

ProficiencyGroup*Condition*Location

BOUNDEDNESS, FIRST LANGUAGE.

$RT^{0.25}$ = LanguageGroup Condition Location Proficiency L2Use

LanguageGroup *Condition LanguageGroup *Proficiency

LanguageGroup *Condition*Location

CURRENT RELEVANCE, *IN TOTO*.

$RT^{-0.25}$ = Group Condition Location Gender

Group*Condition Group*Gender

Group*Condition*Location

$RT^{-0.25}$ = Group Subcondition Location Gender

Group* Subcondition Group*Gender

Group* Subcondition *Location

CURRENT RELEVANCE, ENGLISH PROFICIENCY.

$RT^{-0.25}$ = ProficiencyGroup Condition Location Education Gender

ProficiencyGroup*Condition ProficiencyGroup*Gender

ProficiencyGroup*Condition*Location

$RT^{-0.25}$ = ProficiencyGroup Subcondition Location Education Gender

ProficiencyGroup*Subcondition ProficiencyGroup*Gender

ProficiencyGroup*Subcondition*Location

CURRENT RELEVANCE, FIRST LANGUAGE.

$RT^{-0.25}$ = LanguageGroup Condition Location Proficiency L2Use

LanguageGroup *Condition LanguageGroup *Proficiency

LanguageGroup *Condition*Location

$RT^{-0.25}$ = LanguageGroup Subcondition Location Proficiency L2Use

LanguageGroup *Subcondition LanguageGroup *Proficiency

LanguageGroup *Subcondition*Location

G.2 Boundedness contrasts, descriptive statistics.

Table G.1. Descriptive statistics for reading times (ms) for L2 and native English speakers.

			Location									
			1		2		3		4		5	
	Cdn.	n	$\bar{x}_1$	sd_1	$\bar{x}_2$	sd_2	$\bar{x}_3$	sd_3	$\bar{x}_4$	sd_4	$\bar{x}_5$	sd_5
L2 Users	A	225	818	568	939	697	853	559	894	564	940	668
	B	216	830	503	878	650	860	526	862	527	848	565
	C	223	858	568	941	665	890	621	973	699	905	623
	D	224	819	488	856	493	840	483	781	439	795	482
English NS	A	244	447	170	446	183	433	174	431	187	435	207
	B	244	414	178	441	183	434	193	428	197	452	193
	C	245	458	195	432	179	463	202	425	200	430	187
	D	245	466	201	445	197	462	198	469	208	434	170

Table G.2. Descriptive statistics for reading times (ms) for groups by L2 English proficiency.

			Location									
			1		2		3		4		5	
	Cdn.	n	$\bar{x}_1$	sd_1	$\bar{x}_2$	sd_2	$\bar{x}_3$	sd_3	$\bar{x}_4$	sd_4	$\bar{x}_5$	sd_5
Adv.	A	38	552	279	636	473	681	433	635	400	583	359
	B	43	566	302	630	369	699	379	658	390	684	383
	C	36	717	439	776	477	660	558	927	769	795	548
	D	42	668	259	703	375	673	343	573	259	637	248
Int.-High	A	41	663	457	669	546	579	408	593	378	591	347
	B	47	675	310	604	332	719	359	659	413	664	390
	C	48	606	342	578	412	648	399	629	410	593	335
	D	42	639	322	627	343	606	340	693	345	633	286
Int.-Low	A	67	924	576	931	586	880	511	886	513	941	591
	B	62	827	450	913	533	896	478	943	518	893	436
	C	68	944	498	978	587	865	567	1019	749	951	602
	D	66	872	444	845	450	908	455	756	433	810	530
Low	A	79	936	656	1232	821	1054	639	1182	612	1292	781
	B	64	1123	619	1211	885	1038	683	1069	595	1051	781
	C	71	1019	723	1236	806	1196	700	1185	688	1127	729
	D	74	961	633	1081	568	1006	563	970	502	963	567
English NS	A	244	447	170	446	183	433	174	431	187	435	207
	B	244	414	178	441	183	434	193	428	197	452	193
	C	245	458	195	432	179	463	202	425	200	430	187
	D	245	466	201	445	197	462	198	469	208	434	170

TABLE G.3. Descriptive statistics for reading times (ms)for groups by L1.

			Location									
			1		2		3		4		5	
	Cdn.	n	$\bar{x}_1$	sd_1	$\bar{x}_2$	sd_2	$\bar{x}_3$	sd_3	$\bar{x}_4$	sd_4	$\bar{x}_5$	sd_5
Arabic	A	35	983	610	964	582	913	539	1048	511	1030	649
	B	32	860	426	868	469	882	497	1078	552	1006	542
	C	30	901	509	967	695	997	632	1202	841	1062	619
	D	31	887	493	886	307	749	422	928	419	827	557
Chinese	A	133	853	593	996	734	894	591	925	589	984	693
	B	119	903	540	894	668	881	560	915	570	888	628
	C	141	858	588	999	678	901	616	976	666	931	613
	D	134	846	543	922	561	930	514	817	466	828	513
Other	A	57	634	425	791	659	720	479	727	500	782	601
	B	65	681	436	853	699	811	476	659	342	698	408
	C	52	836	553	771	588	799	627	833	674	743	629
	D	59	724	317	688	353	681	386	621	336	703	344
English NS	A	244	447	170	446	183	433	174	431	187	435	207
	B	244	414	178	441	183	434	193	428	197	452	193
	C	245	458	195	432	179	463	202	425	200	430	187
	D	245	466	201	445	197	462	198	469	208	434	170

G.3 CURRENT RELEVANCE CONTRASTS, DESCRIPTIVE STATISTICS.

TABLE G.4. Descriptive statistics for reading times (ms) for L2 and native English speakers.

			Location									
			1		2		3		4		5	
	Cdn.	n	$\bar{x}_1$	sd_1	$\bar{x}_2$	sd_2	$\bar{x}_3$	sd_3	$\bar{x}_4$	sd_4	$\bar{x}_5$	sd_5
L2 Users	E	232	625	311	498	213	487	201	629	371	736	452
	F	225	654	347	503	201	468	190	675	408	743	451
	G	216	636	357	478	170	444	157	625	351	821	555
	H	226	673	360	509	210	466	167	667	368	746	401
English NS	E	246	400	171	390	157	351	115	379	146	409	188
	F	244	393	163	388	135	352	119	381	164	416	194
	G	243	387	170	377	139	350	114	376	153	405	180
	H	246	391	152	414	295	353	111	392	241	396	167

TABLE G.5. Descriptive statistics for reading times (ms) for L2 and native English speakers.

	SubC.	n	x̄	sd	min	Q1	med	Q3	max
L2 Users	E1	235	565	383	113	316	464	657	2013
	F1	226	615	390	118	376	464	736	1969
	G1	265	612	416	112	348	488	702	2222
	H1	250	615	387	125	376	481	736	1751
	E2	300	557	298	120	368	473	656	2013
	F2	340	562	307	117	368	484	659	1969
	G2	285	603	390	143	376	480	656	2222
	H2	305	616	321	104	408	536	745	1751
	E3	280	632	322	160	408	580	788	2048
	F3	265	676	393	160	416	553	816	1969
	G3	260	581	317	120	368	508	688	2222
	H3	270	601	316	160	377	539	736	1751
	E4	346	619	342	216	399	512	712	2013
	F4	290	596	323	124	376	500	744	1969
	G4	270	606	367	167	376	504	721	2222
	H4	305	616	311	176	409	520	760	1751
English NS	E1	325	367	146	136	270	349	419	1000
	F1	315	398	172	160	281	360	454	994
	G1	285	368	147	144	269	333	424	895
	H1	295	389	225	168	267	347	420	2186
	E2	290	373	162	135	248	345	448	1000
	F2	295	365	131	144	270	357	432	800
	G2	325	374	141	140	280	352	420	895
	H2	320	404	241	137	273	343	464	2506
	E3	325	402	152	151	288	381	465	1000
	F3	315	391	166	151	268	363	464	994
	G3	280	394	175	101	257	354	488	895
	H3	295	373	179	144	263	336	416	1792
	E4	290	401	171	136	272	369	486	1000
	F4	295	390	157	108	281	358	437	994
	G4	325	381	153	159	269	360	456	895
	H4	320	390	163	152	290	356	463	1261

TABLE G.6. Descriptive statistics for reading times (ms) for L2 and native English speakers.

			Location									
			1		2		3		4		5	
	SubC.	n	$\bar{x}_1$	sd_1	$\bar{x}_2$	sd_2	$\bar{x}_3$	sd_3	$\bar{x}_4$	sd_4	$\bar{x}_5$	sd_5
L2 Users	E1	47	570	323	446	213	453	212	584	397	773	567
	F1	46	663	368	463	158	429	148	708	458	812	524
	G1	53	621	362	452	160	423	157	657	372	906	636
	H1	50	679	402	467	170	443	155	674	447	813	498
	E2	60	607	287	468	198	488	225	559	293	662	408
	F2	68	616	313	447	144	447	188	574	285	726	430
	G2	57	700	426	457	147	440	134	605	342	813	574
	H2	61	706	383	519	214	466	166	681	313	709	387
	E3	56	657	306	546	239	509	200	690	393	760	370
	F3	53	694	374	559	234	525	211	802	490	802	485
	G3	52	588	313	518	192	452	165	610	319	736	448
	H3	54	671	353	522	242	457	173	632	343	725	355
	E4	69	650	328	521	196	494	172	670	389	757	463
	F4	58	654	348	548	235	472	193	650	384	657	373
	G4	54	630	311	486	178	460	172	629	377	825	545
	H4	61	638	310	521	206	494	173	679	376	746	365
English NS	E1	65	371	150	373	148	337	99	349	125	405	189
	F1	63	403	172	405	148	357	130	390	177	433	219
	G1	57	374	141	350	115	328	101	380	167	411	184
	H1	59	389	154	417	340	348	119	393	260	399	181
	E2	58	387	174	372	153	340	117	365	143	404	205
	F2	59	369	147	363	118	348	110	371	152	371	128
	G2	65	378	154	370	121	349	106	373	134	402	177
	H2	64	399	144	453	398	358	113	398	254	412	183
	E3	65	433	188	398	133	377	124	397	139	407	167
	F3	63	399	167	391	142	367	129	373	167	422	212
	G3	56	401	199	406	167	370	134	382	161	411	206
	H3	59	378	149	405	256	344	104	368	199	371	152
	E4	58	408	167	420	189	350	118	406	172	422	194
	F4	59	400	164	392	131	336	104	388	159	435	198
	G4	65	396	183	383	148	352	115	373	152	400	159
	H4	64	398	162	380	101	361	108	407	249	401	153

TABLE G.7. Descriptive statistics for reading times (ms) for groups by L2 English proficiency.

			Location									
			1		2		3		4		5	
	Cdn.	n	$\bar{x}_1$	sd_1	$\bar{x}_2$	sd_2	$\bar{x}_3$	sd_3	$\bar{x}_4$	sd_4	$\bar{x}_5$	sd_5
Adv.	E	41	484	187	472	190	431	165	530	282	692	437
	F	43	509	232	453	189	422	163	574	329	616	320
	G	39	528	348	395	137	383	140	475	233	570	339
	H	39	575	303	443	203	409	129	509	228	585	345
Int.-High	E	44	553	281	481	220	476	216	560	340	618	365
	F	50	550	297	490	218	403	151	530	302	565	332
	G	42	551	292	497	199	422	163	534	255	646	458
	H	43	569	323	414	127	403	116	495	223	561	344
Int.-Low	E	80	654	273	532	223	514	196	672	382	754	421
	F	61	807	383	514	177	491	174	820	453	864	479
	G	65	650	302	508	163	452	116	711	392	962	602
	H	77	747	342	590	198	515	157	766	367	907	383
Low	E	67	722	389	484	210	498	214	684	411	819	530
	F	71	682	352	532	213	523	223	714	434	844	512
	G	70	735	419	484	164	482	183	683	381	934	590
	H	67	713	411	515	234	484	203	755	438	774	408
English NS	E	246	400	171	390	157	351	115	379	146	409	188
	F	244	393	163	388	135	352	119	381	164	416	194
	G	243	387	170	377	139	350	114	376	153	405	180
	H	246	391	152	414	295	353	111	392	241	396	167

TABLE G.8. Descriptive statistics for reading times (ms) for groups by L2 English proficiency.

	SubC.	n	$\bar{x}$	sd	min	Q1	med	Q3	max
Adv.	E1	60	478	349	113	256	420	523	2013
	F1	50	515	292	118	332	432	623	1439
	G1	50	493	267	176	356	424	560	1520
	H1	45	370	156	125	296	324	422	912
	E2	40	527	272	244	345	482	568	1560
	F2	65	474	213	117	347	448	560	1072
	G2	50	485	299	143	341	386	549	1668
	H2	50	550	262	131	407	481	704	1333
	E3	50	573	290	247	359	509	676	1605
	F3	40	531	270	216	368	443	656	1755
	G3	30	499	143	290	396	462	584	838
	H3	40	478	271	184	241	420	610	1401
	E4	55	519	193	216	400	492	624	1144
	F4	60	549	284	124	399	500	684	1818
	G4	65	429	276	167	274	352	491	1853
	H4	60	584	277	257	423	536	631	1646

Int.-High	E1	60	574	338	185	332	460	704	1850
	F1	35	578	317	299	376	448	693	1560
	G1	55	445	299	152	313	376	520	2222
	H1	40	532	340	160	310	445	646	1704
	E2	55	464	251	257	313	379	560	1768
	F2	85	489	262	192	311	413	589	1560
	G2	45	526	263	232	376	460	536	1668
	H2	60	511	234	176	384	476	646	1232
	E3	50	583	314	168	354	550	709	1580
	F3	60	515	287	160	311	506	572	1818
	G3	55	490	241	192	320	440	581	1232
	H3	60	513	271	200	335	489	560	1711
	E4	55	531	248	222	352	452	623	1265
	F4	70	488	247	198	300	404	618	1208
	G4	55	658	338	168	456	593	776	2222
	H4	55	405	146	184	281	400	480	848
Int.-Low	E1	35	561	238	296	438	496	640	1256
	F1	61	757	473	192	416	624	1016	1969
	G1	85	651	452	112	400	520	696	2222
	H1	80	648	346	280	424	505	748	1734
	E2	125	600	296	120	422	504	733	1632
	F2	95	651	314	256	423	544	784	1576
	G2	85	658	439	291	400	504	688	2222
	H2	110	738	354	296	488	612	921	1734
	E3	75	701	330	328	472	600	824	2048
	F3	80	795	428	232	490	652	1017	1969
	G3	90	620	300	272	440	541	698	2222
	H3	75	657	267	328	472	625	753	1751
	E4	165	623	350	224	400	496	703	2013
	F4	65	600	331	248	392	486	696	1969
	G4	65	712	400	280	448	560	808	2222
	H4	120	743	335	280	480	668	925	1751
Low	E1	80	625	473	136	252	464	864	2013
	F1	80	586	380	208	392	460	684	1969
	G1	75	770	463	168	416	640	1048	2222
	H1	85	754	458	200	424	584	1015	1751
	E2	80	568	332	184	360	444	672	2013
	F2	95	598	360	200	352	488	712	1969
	G2	105	648	417	168	400	507	728	2222
	H2	85	572	316	104	384	504	720	1751
	E3	105	635	329	160	384	600	880	1576
	F3	85	746	418	184	472	625	944	1969
	G3	85	626	399	120	344	512	833	2222
	H3	95	666	367	160	368	632	969	1688
	E4	71	754	425	248	431	608	1030	2013
	F4	95	703	358	232	424	592	935	1848
	G4	85	627	378	208	400	507	728	2222
	H4	70	589	290	176	392	500	720	1344

TABLE G.9. Descriptive statistics for reading times (ms) for groups by L2 English proficiency.

			Location									
			1		2		3		4		5	
	SubC.	n	$\bar{x}_1$	sd_1	$\bar{x}_2$	sd_2	$\bar{x}_3$	sd_3	$\bar{x}_4$	sd_4	$\bar{x}_5$	sd_5
Adv.	E1	12	449	235	391	170	373	135	457	261	722	630
	F1	10	555	301	407	113	365	115	567	368	682	372
	G1	10	553	381	400	138	387	108	511	285	613	299
	H1	9	402	216	334	102	364	94	326	129	423	206
	E2	8	479	177	510	276	468	205	479	200	701	422
	F2	13	447	169	402	137	446	231	470	212	603	266
	G2	10	683	519	388	151	380	116	395	145	578	272
	H2	10	595	299	437	125	460	183	657	243	599	369
	E3	10	515	144	471	171	439	145	690	404	749	371
	F3	8	502	214	512	219	454	173	577	169	609	486
	G3	6	476	92	463	144	426	78	585	174	546	182
	H3	8	524	224	463	305	352	107	455	256	597	387
	E4	11	498	190	533	142	460	184	504	182	600	255
	F4	12	542	254	507	252	424	97	691	456	581	212
	G4	13	415	178	367	127	364	198	458	265	540	475
	H4	12	721	358	517	217	439	97	558	159	686	375
Int.-High	E1	12	571	219	463	241	489	236	691	482	653	421
	F1	7	674	301	438	102	383	42	655	422	738	419
	G1	11	436	183	391	132	358	147	438	185	602	582
	H1	8	529	272	406	117	401	130	505	246	819	587
	E2	11	447	191	432	155	430	232	423	163	587	422
	F2	17	580	379	452	200	378	110	493	194	543	315
	G2	9	637	407	479	155	403	68	559	298	550	252
	H2	12	650	327	440	168	390	138	529	217	547	221
	E3	10	660	429	603	315	550	256	496	220	606	351
	F3	12	478	211	535	258	438	148	585	435	541	324
	G3	11	482	263	506	201	415	190	486	222	560	326
	H3	12	619	417	431	108	426	115	569	268	517	305
	E4	11	541	244	440	102	442	132	610	349	621	296
	F4	14	514	248	523	249	413	221	465	197	524	321
	G4	11	663	273	609	242	509	183	656	288	853	545
	H4	11	455	225	372	108	392	91	370	103	435	164
Int.-Low	E1	7	618	301	548	280	500	169	467	100	674	282
	F1	13	887	496	520	228	516	226	841	449	1011	650
	G1	17	586	264	461	144	416	149	766	428	1026	714
	H1	16	669	313	505	161	489	183	657	329	918	487
	E2	25	704	286	469	190	508	230	639	297	680	385
	F2	19	767	307	455	84	476	159	716	341	838	380
	G2	17	640	335	492	178	449	106	668	393	1041	686
	H2	22	898	414	649	224	497	109	729	312	915	435

	E3	15	650	240	615	231	560	216	863	498	816	309
	F3	16	896	405	562	178	558	190	1015	568	945	451
	G3	18	595	216	569	184	490	112	645	323	803	463
	H3	15	719	300	581	152	495	132	689	326	801	280
	E4	33	626	278	538	226	500	167	652	396	800	510
	F4	13	676	321	534	219	409	77	714	404	665	441
	G4	13	824	365	506	125	451	77	786	451	992	520
	H4	24	678	280	599	209	561	187	922	427	957	314
Low	E1	16	638	436	430	189	464	256	651	474	943	696
	F1	16	545	227	467	134	424	92	720	535	776	543
	G1	15	843	453	521	191	503	179	793	375	1190	618
	H1	17	907	486	531	187	462	154	952	557	917	507
	E2	16	631	333	469	207	505	234	568	363	667	458
	F2	19	611	279	463	142	480	231	575	295	860	577
	G2	21	784	468	451	109	477	171	673	358	853	618
	H2	17	561	344	454	201	482	222	740	388	622	331
	E3	21	729	332	505	228	485	179	658	337	797	420
	F3	17	745	393	595	280	590	264	860	473	942	536
	G3	17	688	441	491	211	444	216	660	397	847	531
	H3	19	728	393	559	316	490	233	701	412	851	370
	E4	14	913	439	532	213	547	201	891	453	879	547
	F4	19	814	425	601	233	588	228	717	409	797	412
	G4	17	624	284	483	148	508	177	623	412	898	575
	H4	14	641	348	510	196	507	204	609	302	679	357

TABLE G.10. Descriptive statistics for reading times (ms) for groups by L1.

			Location									
			1		2		3		4		5	
	Cdn.	n	$\bar{x}_1$	sd_1	$\bar{x}_2$	sd_2	$\bar{x}_3$	sd_3	$\bar{x}_4$	sd_4	$\bar{x}_5$	sd_5
Arabic	E	34	787	314	600	232	587	238	1015	415	1095	493
	F	33	891	297	555	182	544	175	879	398	1186	440
	G	33	874	372	502	110	473	154	908	451	1305	615
	H	40	868	335	582	209	556	182	952	413	1051	397
Chinese	E	140	584	281	464	191	468	187	551	313	667	413
	F	128	609	344	496	210	465	198	651	421	676	442
	G	127	597	341	466	169	436	155	568	296	713	473
	H	130	625	341	496	205	446	160	627	344	682	364
Other	E	58	627	351	520	234	475	197	590	337	696	426
	F	64	621	333	489	192	435	173	616	359	649	333
	G	56	586	336	489	202	443	163	587	327	779	549
	H	56	646	382	488	214	449	153	558	284	677	389
English NS	E	246	400	171	390	157	351	115	379	146	409	188
	F	244	393	163	388	135	352	119	381	164	416	194
	G	243	387	170	377	139	350	114	376	153	405	180
	H	246	391	152	414	295	353	111	392	241	396	167

TABLE G.11. Descriptive statistics for reading times (ms) for groups by L2 English proficiency.

	SubC.	n	x̄	sd	min	Q1	med	Q3	max
Arabic	E1	10	872	466	352	536	688	1200	1848
	F1	35	800	392	336	456	739	1104	1719
	G1	50	850	523	240	472	588	1176	2222
	H1	45	824	429	280	504	664	1184	1751
	E2	60	801	368	248	528	764	1000	2013
	F2	50	762	372	336	456	641	960	1969
	G2	35	777	468	305	424	584	1000	2222
	H2	65	798	349	328	512	703	1032	1734
	E3	40	916	410	328	596	864	1178	2048
	F3	50	867	422	376	528	724	1144	1969
	G3	35	706	425	304	408	561	752	2222
	H3	25	776	401	328	504	656	912	1751
	E4	60	757	428	240	468	580	1000	2013
	F4	30	810	399	368	464	776	968	1969
	G4	45	881	520	280	480	720	1184	2222
	H4	65	800	361	360	519	647	992	1751
Chinese	E1	135	564	388	136	312	464	660	2013
	F1	131	569	375	192	368	448	632	1969
	G1	150	553	383	112	320	456	648	2222
	H1	140	605	365	160	386	484	752	1751
	E2	180	479	216	120	336	437	552	1560
	F2	195	540	299	192	343	464	650	1840
	G2	175	568	333	168	372	471	624	2222
	H2	165	568	304	104	398	501	712	1751
	E3	185	575	270	160	376	557	711	1576
	F3	140	646	404	160	376	545	799	1969
	G3	165	552	310	120	344	486	664	2222
	H3	170	567	283	160	364	539	728	1711
	E4	201	571	311	216	368	472	680	2013
	F4	170	577	324	198	352	487	700	1848
	G4	145	550	259	168	376	488	664	1600
	H4	175	566	283	176	376	488	715	1751
Other	E1	90	533	354	113	326	440	608	2013
	F1	60	609	397	118	366	452	733	1969
	G1	65	566	326	176	360	480	672	1668
	H1	65	494	350	125	309	407	507	1751
	E2	60	546	309	208	324	471	620	1580
	F2	95	500	239	117	352	456	568	1552
	G2	75	603	453	143	360	434	670	2222
	H2	75	564	275	131	400	512	648	1648
	E3	55	620	307	247	408	512	727	1580
	F3	75	605	308	216	415	536	656	1818

G3	60	588	242	290	406	525	688	1344
H3	75	622	338	184	377	512	808	1711
E4	85	634	318	244	408	560	704	1799
F4	90	561	265	124	392	516	704	1592
G4	80	553	366	167	304	463	653	2222
H4	65	565	258	257	416	519	632	1719

TABLE G.12. Descriptive statistics for reading times (ms) for groups by L1.

			Location									
			1		2		3		4		5	
	SubC.	n	$\bar{x}_1$	sd_1	$\bar{x}_2$	sd_2	$\bar{x}_3$	sd_3	$\bar{x}_4$	sd_4	$\bar{x}_5$	sd_5
Arabic	E1	2	544	272	527	13	610	41	1086	161	1592	362
	F1	7	917	362	499	137	464	109	913	370	1209	332
	G1	10	859	371	486	72	470	156	964	409	1468	637
	H1	9	818	323	516	74	608	224	988	496	1188	516
	E2	12	779	237	596	235	653	291	839	267	1140	511
	F2	10	922	236	493	83	522	126	656	218	1218	459
	G2	7	1062	420	433	65	490	189	843	519	1057	575
	H2	13	955	391	615	244	580	177	902	318	938	398
	E3	8	996	311	679	288	566	258	1288	467	1053	285
	F3	10	906	363	561	111	645	225	1020	430	1204	529
	G3	7	682	283	524	128	458	137	830	513	1037	624
	H3	5	643	99	618	314	397	59	990	441	1231	331
	E4	12	698	346	564	214	532	186	998	465	995	586
	F4	6	782	232	716	336	504	170	974	530	1076	453
	G4	9	893	376	556	137	475	160	956	461	1523	559
	H4	13	901	327	579	204	556	169	963	470	1000	309
Chinese	E1	27	583	341	442	219	455	219	599	444	742	551
	F1	27	575	332	465	177	439	162	631	456	735	537
	G1	30	560	332	432	161	401	155	553	317	818	619
	H1	28	709	407	472	170	408	108	651	402	783	460
	E2	36	550	245	428	169	451	190	442	170	522	269
	F2	39	584	339	456	165	438	179	585	316	640	388
	G2	35	660	407	449	134	435	123	606	310	689	462
	H2	33	604	347	503	218	432	150	642	320	661	376
	E3	37	573	223	512	215	490	179	588	291	709	360
	F3	28	618	336	530	251	499	225	827	541	757	489
	G3	33	565	331	497	204	432	152	571	296	692	439
	H3	34	633	318	488	194	462	188	590	322	659	315
	E4	40	625	318	466	160	473	174	583	314	705	440
	F4	34	657	375	537	236	490	220	597	360	604	380
	G4	29	595	275	488	169	479	187	534	265	656	336
	H4	35	571	296	516	232	475	174	628	348	641	309
Other	E1	18	553	316	444	219	431	213	508	295	727	561
	F1	12	714	398	439	128	388	138	756	498	746	514

G1	13	581	375	470	205	438	163	663	355	677	425
H1	13	519	414	424	212	404	117	504	428	617	458
E2	12	607	391	457	205	432	184	631	401	604	290
F2	19	519	175	403	113	425	227	509	244	642	326
G2	15	626	416	485	200	429	134	489	277	987	745
H2	15	714	374	470	153	442	154	577	199	616	346
E3	11	693	387	563	266	529	230	598	238	715	387
F3	15	693	419	610	264	495	153	610	364	618	279
G3	12	596	291	572	191	502	213	586	193	683	302
H3	15	766	466	568	311	465	167	607	307	706	329
E4	17	677	352	618	224	517	161	644	391	714	393
F4	18	608	330	512	180	427	141	642	342	617	236
G4	16	545	270	445	208	416	150	618	420	740	561
H4	13	555	175	480	106	484	172	532	155	773	444

G.4 BOUNDEDNESS CONTRASTS, SELECTED INFERENTIAL STATISTICS.

TABLE G.13. Tukey posthoc results for Condition by Group for L2 English users and English NS.

		English NS				L2 Users			
		A	B	C	D	A	B	C	D
Eng. NS	A		.996	1.00	.655	<.001	<.001	<.001	<.001
	B	.996		.991	.207	<.001	<.001	<.001	<.001
	C	1.00	.991		.728	<.001	<.001	<.001	<.001
	D	.655	.207	.728		<.001	<.001	<.001	<.001
L2 Users	A	<.001	<.001	<.001	<.001		1.00	.946	.991
	B	<.001	<.001	<.001	<.001	1.00		.799	1.00
	C	<.001	<.001	<.001	<.001	.946	.799		.484
	D	<.001	<.001	<.001	<.001	.991	1.00	.484	

TABLE G.14. Tukey posthoc results for Condition by Proficiency Group for L2 English users and English NS.

		Low				Int.-L				Int.-H				Adv.				Eng. NS			
		A	B	C	D	A	B	C	D	A	B	C	D	A	B	C	D	A	B	C	D
Low	A		.790	1.00	.362	.822	.952	.992	.142	<.001	<.001	<.001	<.001	<.001	<.001	.003	<.001	<.001	<.001	<.001	<.001
	B	.790		1.00	1.00	1.00	1.00	1.00	1.00	<.001	.002	<.001	.000	<.001	.001	.873	.029	<.001	<.001	<.001	<.001
	C	1.00	1.00		.994	1.00	1.00	1.00	.974	<.001	<.001	<.001	<.001	<.001	<.001	.268	.001	<.001	<.001	<.001	<.001
	D	.362	1.00	.994		1.00	1.00	.996	1.00	<.001	.025	<.001	.007	.001	.009	.995	.187	<.001	<.001	<.001	<.001
Int.-L	A	.822	1.00	1.00	1.00		1.00	1.00	.999	<.001	<.001	<.001	<.001	<.001	<.001	.387	.000	<.001	<.001	<.001	<.001
	B	.952	1.00	1.00	1.00	1.00		1.00	.978	<.001	<.001	<.001	<.001	<.001	<.001	.209	<.001	<.001	<.001	<.001	<.001
	C	.992	1.00	1.00	.996	1.00	1.00		.832	<.001	<.001	<.001	<.001	<.001	<.001	.078	<.001	<.001	<.001	<.001	<.001
	D	.142	1.00	.974	1.00	.999	.978	.832		<.001	.002	<.001	.000	<.001	.000	.985	.043	<.001	<.001	<.001	<.001
Int.-H	A	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001		.890	1.00	.995	1.00	.988	.006	.492	.000	<.001	.000	.008
	B	<.001	.002	<.001	.025	<.001	<.001	<.001	.002	.890		.818	1.00	1.00	1.00	.720	1.00	<.001	<.001	<.001	<.001
	C	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	1.00	.818		.987	1.00	.974	.003	.368	<.001	<.001	<.001	.004
	D	<.001	.000	<.001	.007	<.001	<.001	<.001	.000	.995	1.00	.987		1.00	1.00	.418	1.00	<.001	<.001	<.001	<.001
Adv.	A	<.001	<.001	<.001	.001	<.001	<.001	<.001	<.001	1.00	1.00	1.00	1.00		1.00	.140	.984	<.001	<.001	<.001	<.001
	B	<.001	.001	<.001	.009	<.001	<.001	<.001	.000	.988	1.00	.974	1.00	1.00		.498	1.00	<.001	<.001	<.001	<.001
	C	.003	.873	.268	.995	.387	.209	.078	.985	.006	.720	.003	.418	.140	.498		.986	<.001	<.001	<.001	<.001
	D	<.001	.029	.001	.187	.000	<.001	<.001	.043	.492	1.00	.368	1.00	.984	1.00	.986		<.001	<.001	<.001	<.001
Eng. NS	A	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	.000	<.001	<.001	<.001	<.001	<.001	<.001	<.001		1.00	1.00	.964
	B	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	1.00		1.00	.555
	C	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	.000	<.001	<.001	<.001	<.001	<.001	<.001	<.001	1.00	1.00		.982
	D	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	.008	<.001	.004	<.001	<.001	<.001	<.001	<.001	.964	.555	.982	

TABLE G.15. Tukey posthoc results for Condition by L1 Group for L2 English users and English NS.

		English				Arabic				Chinese				Other L1			
		A	B	C	D	A	B	C	D	A	B	C	D	A	B	C	D
Eng.	A		1.00	1.00	.859	<.001	<.001	<.001	<.001	.011	<.001	.001	.001	<.001	<.001	<.001	<.001
	B	1.00		1.00	.673	<.001	<.001	<.001	<.001	.006	<.001	.000	.001	<.001	<.001	<.001	<.001
	C	1.00	1.00		.989	<.001	<.001	<.001	<.001	.027	.000	.002	.004	<.001	<.001	<.001	<.001
	D	.859	.673	.989		<.001	<.001	<.001	<.001	.231	.010	.044	.062	<.001	<.001	<.001	<.001
Arab.	A	<.001	<.001	<.001	<.001		1.00	1.00	1.00	<.001	.000	<.001	<.001	.002	.022	.116	.016
	B	<.001	<.001	<.001	<.001	1.00		1.00	.989	<.001	<.001	<.001	<.001	.000	.003	.024	.002
	C	<.001	<.001	<.001	<.001	1.00	1.00		.999	<.001	.000	<.001	<.001	.002	.017	.087	.013
	D	<.001	<.001	<.001	<.001	1.00	.989	.999		.001	.007	.002	.003	.061	.279	.635	.231
Chin.	A	.011	.006	.027	.231	<.001	<.001	<.001	.001		1.00	1.00	1.00	.948	.384	.116	.514
	B	<.001	<.001	.000	.010	.000	<.001	.000	.007	1.00		1.00	1.00	1.00	.843	.448	.919
	C	.001	.000	.002	.044	<.001	<.001	<.001	.002	1.00	1.00		1.00	.991	.576	.210	.710
	D	.001	.001	.004	.062	<.001	<.001	<.001	.003	1.00	1.00	1.00		.992	.601	.229	.730
Other	A	<.001	<.001	<.001	<.001	.002	.000	.002	.061	.948	1.00	.991	.992		1.00	.964	1.00
	B	<.001	<.001	<.001	<.001	.022	.003	.017	.279	.384	.843	.576	.601	1.00		1.00	1.00
	C	<.001	<.001	<.001	<.001	.116	.024	.087	.635	.116	.448	.210	.229	.964	1.00		1.00
	D	<.001	<.001	<.001	<.001	.016	.002	.013	.231	.514	.919	.710	.730	1.00	1.00	1.00	

G.5 CURRENT RELEVANCE CONTRASTS, SELECTED INFERENTIAL STATISTICS.

TABLE G.16. Tukey posthoc results for Condition by Group for L2 English users and English NS.

		English NS				L2 Users			
		E	F	G	H	E	F	G	H
Eng. NS	E		1.00	.998	1.00	<.001	<.001	<.001	<.001
	F	1.00		.993	1.00	<.001	<.001	<.001	<.001
	G	.998	.993		.998	<.001	<.001	<.001	<.001
	H	1.00	1.00	.998		<.001	<.001	<.001	<.001
L2 Users	E	<.001	<.001	<.001	<.001		1.00	.998	.762
	F	<.001	<.001	<.001	<.001	1.00		.956	.955
	G	<.001	<.001	<.001	<.001	.998	.956		.351
	H	<.001	<.001	<.001	<.001	.762	.955	.351	

TABLE G.17. Tukey posthoc results for Condition by Proficiency Group for L2 English users and English NS.

		Low				Int.-L				Int.-H				Adv.				Eng. NS			
		E	F	G	H	E	F	G	H	E	F	G	H	E	F	G	H	E	F	G	H
Low	E		.704	.993	1.00	.982	.067	.777	.000	.104	.002	.042	.000	.153	.403	<.001	.053	<.001	<.001	<.001	<.001
	F	.704		1.00	.688	1.00	1.00	1.00	.881	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001
	G	.993	1.00		.991	1.00	.991	1.00	.366	.000	<.001	.000	<.001	.001	.005	<.001	.000	<.001	<.001	<.001	<.001
	H	1.00	.688	.991		.988	.106	.828	.001	.183	.005	.083	.001	.249	.542	<.001	.102	<.001	<.001	<.001	<.001
Int.-L	E	.982	1.00	1.00	.988		.626	1.00	.003	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001
	F	.067	1.00	.991	.106	.626		.994	.991	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001
	G	.777	1.00	1.00	.828	1.00	.994		.124	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001
	H	.000	.881	.366	.001	.003	.991	.124		<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001
Int.-H	E	.104	<.001	.000	.183	<.001	<.001	<.001	<.001		1.00	1.00	.954	1.00	1.00	.581	1.00	<.001	<.001	<.001	<.001
	F	.002	<.001	<.001	.005	<.001	<.001	<.001	<.001	1.00		1.00	1.00	1.00	.962	.995	1.00	<.001	<.001	<.001	<.001
	G	.042	<.001	.000	.083	<.001	<.001	<.001	<.001	1.00	1.00		.994	1.00	1.00	.798	1.00	<.001	<.001	<.001	<.001
	H	.000	<.001	<.001	.001	<.001	<.001	<.001	<.001	.954	1.00	.994		.935	.607	1.00	.996	<.001	<.001	<.001	<.001
Adv.	E	.153	<.001	.001	.249	<.001	<.001	<.001	<.001	1.00	1.00	1.00	.935		1.00	.497	1.00	<.001	<.001	<.001	<.001
	F	.403	<.001	.005	.542	<.001	<.001	<.001	<.001	1.00	.962	1.00	.607	1.00		.151	1.00	<.001	<.001	<.001	<.001
	G	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	.581	.995	.798	1.00	.497	.151		.820	<.001	<.001	<.001	<.001
	H	.053	<.001	.000	.102	<.001	<.001	<.001	<.001	1.00	1.00	1.00	.996	1.00	1.00	.820		<.001	<.001	<.001	<.001
Eng. NS	E	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001		1.00	1.00	1.00
	F	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	1.00		1.00	1.00
	G	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	1.00	1.00		1.00
	H	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	1.00	1.00	1.00	

TABLE G.18. Abbr. Tukey posthoc results for Subcondition within Proficiency Groups.

	Comparison Condition	p-value
Low	E3-F3	.007
Int-L	Ø	
Int-H	G4-H4	.048
Adv	F4-G4	.051
	G4-H4	.002
NS	Ø	

TABLE G.19. Tukey posthoc results for Condition by L1 Group for L2 English users and English NS.

		English				Arabic				Chinese				Other L1			
		E	F	G	H	E	F	G	H	E	F	G	H	E	F	G	H
Eng.	E		1.00	1.00	1.00	<.001	<.001	<.001	<.001	.000	.000	<.001	<.001	<.001	<.001	.019	.000
	F	1.00		1.00	1.00	<.001	<.001	<.001	<.001	.000	.000	<.001	<.001	<.001	<.001	.027	.000
	G	1.00	1.00		1.00	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	.006	<.001
	H	1.00	1.00	1.00		<.001	<.001	<.001	<.001	.000	.000	<.001	<.001	<.001	<.001	.018	.000
Arab.	E	<.001	<.001	<.001	<.001		1.00	.969	1.00	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001
	F	<.001	<.001	<.001	<.001	1.00		1.00	1.00	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001
	G	<.001	<.001	<.001	<.001	.969	1.00		1.00	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001
	H	<.001	<.001	<.001	<.001	1.00	1.00	1.00		<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001
Chin.	E	.000	.000	<.001	.000	<.001	<.001	<.001	<.001		1.00	1.00	1.00	.996	1.00	1.00	1.00
	F	.000	.000	<.001	.000	<.001	<.001	<.001	<.001	1.00		1.00	1.00	.998	1.00	1.00	1.00
	G	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	1.00	1.00		1.00	1.00	1.00	.999	1.00
	H	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	1.00	1.00	1.00		1.00	1.00	.981	1.00
Other	E	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	.996	.998	1.00	1.00		1.00	.692	.991
	F	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	1.00	1.00	1.00	1.00	1.00		.996	1.00
	G	.019	.027	.006	.018	<.001	<.001	<.001	<.001	1.00	1.00	.999	.981	.692	.996		1.00
	H	.000	.000	<.001	.000	<.001	<.001	<.001	<.001	1.00	1.00	1.00	1.00	.991	1.00	1.00	

APPENDIX H

BLP RESULTS – RATING TASK, SECOND ADMINISTRATION

H.1 SUMMARY EXPLANATION OF RESULTS.

The following paragraphs summarize the results of the BLP for the second administration, making special note of any potentially confounding factors between the main groups that would need to be included in the statistical models. They are divided in three ways: (i) native & nonnative English users, (ii) L2 English proficiency, and (iii) first language of the L2 English users.

L2 ENGLISH USERS *IN TOTO*. This section addresses potential confounding group differences that may affect performance on the rating task between the L2 English User and English NS groups. There are no statistically significant differences between the groups for the following three groups of collected variables: nonlinguistic biographical information (gender, handedness, and highest level of education achieved), linguistic biographical information (age of exposure to the first language and age of comfort using the first language), and language exposure (formal education in the first language, years spent immersed in a country in which the first language is spoken, years spent within a family in which the first language is spoken, and years spent immersed in a work/school environment in which the first language is spoken).

There are several factors that differentiate the two groups, but it is not expected that they confound the results. First, the two groups differ by age ($F_{1,94} = 3.76$, $p = .056$). The L2 English users are on average slightly older than the native speakers, but this difference is not expected to cause any meaningful differences in the results because reaction time is not considered. The next three significant differences concern the first language (overall use of the first language [$\chi^2_1 = 54.307$, $p < .001$], overall self-assessed L1 proficiency [$\chi^2_1 = 12.687$, $p < .001$], overall language attitudes concerning the first language [$\chi^2_1 = 5.426$, $p = .020$]), which is considered more a result of the inherent differences between the bilingual L2 User group and the overwhelmingly monolingual English NS group. The final significant difference between the two groups concerns English proficiency ($F_{1,95} = 131.81$, $p < .001$); this result is expected, and variation in proficiency is considered in greater detail below where the L2 User group is further subdivided by English proficiency level. Based on the collected data, the results of these two groups are directly comparable with no significant caveats.

L2 ENGLISH PROFICIENCY GROUPS. This section focuses on potential confounding difference that may affect performance on the rating task within the L2 English User group divided by English proficiency. When necessary, meaningful differences between the proficiency subgroups and the English NS control group are discussed. There are no statistically significant differences between the groups for the following three groups of collected variables: nonlinguistic biographical information (age, gender, handedness, and highest level of education achieved), linguistic biographical information (age of exposure to both the first and second language and age of comfort using the second language), and language exposure (formal education in the first and second language, years spent immersed in a country in which the first and second language is spoken, years spent within a family in which the first and second language is spoken, and years spent immersed in a work/school environment in which the first language is spoken). The values for overall use of the second language, for overall self-assessed L2 proficiency, and overall language attitudes concerning the second language are also not significantly different among the proficiency subgroups.

There are a few factors that differentiate the proficiency groups. First and foremost, the groups all differ by their score on the independent measure of proficiency; this score is the basis on which group membership is assigned. Second, the groups differ according to the number of years spent immersed in a school/work environment in which English is spoken; the pairwise analysis reveals that these differences emerge from the Int.-Low group, whose members on average spent significantly less time than their peers in such an environment. If exposure to English specifically at school/work affects the ratings in a way dissimilar to the other measures of immersion, then it is expected that the Int.-Low group will perform more poorly than another proficiency-matched sample of participants. Due to the fact that other measures of immersion were also captured for which the groups do not significantly differ, this difference is not expected to affect the results.

FIRST LANGUAGE GROUPS. This section focuses on potential confounding difference that may affect performance on the rating task within the L2 English User group divided by first language. As discussed previously in section 3.3, the L2 English users are divided into three groups based on L1 features that may transfer into the interlanguage grammar and affect their performance on these tasks. There are no statistically significant differences between the L2 user groups for the following three groups of collected variables: nonlinguistic biographical information (gender, handedness, and highest level of education achieved), linguistic biographical information (age of exposure to the first language and age of comfort using the first and second languages), and language exposure (formal education in the first language, years spent immersed in a country in which the first and the second language is spoken, years spent within a family in which the first language is spoken, and years spent immersed in a work/school environment in which the first language is spoken). The English proficiency scores and the values for overall language attitudes concerning the second language are also not significantly different among the L1 subgroups.

There are several factors that differentiate the L1 groups. First, the Arabic group differs from the Chinese and Other groups by age ($F_{3,92} = 5.550$, $p = .002$); the members of this group are generally older than the other two groups, but age is not expected to meaningfully affect Likert ratings. Second, the members of the Arabic group were exposed to English at a later age than the Chinese (DSCF = 4.982, p = .001) and Other groups (DSCF = 3.765, p = .021). Third, and related to the second difference, the Arabic group has also received less formal education in English than the other two groups (Chinese: DSCF = 4.593, p = .003; Other: DSCF = 4.126, p = .010). Third, the members of the Other group have significantly higher values for overall use of English than the Arabic (DSCF = 3.024, p = .082) and Chinese groups (DSCF = 3.548, p = .033); if the Other group performs in a more nativelike manner than the Arabic or Chinese groups, this difference may indicate that a higher proportion of L2 use leads to improved accuracy. Finally, the Arabic and Chinese groups differ by their self-assessed English proficiency (DSCF = 3.378, p = .045); this result obtains despite the fact that this assessment is not justified by a difference in proficiency scores, and it may be due to cultural differences in self-reporting. Based on the collected data, the results of these L1 groups are directly comparable with the caveats that the Arabic group was exposed to English comparatively late and has received less English education; if the Arabic group performs worse than the other groups, these differences may be a cause and not just grammatical differences in the L1.

H.2 BIOGRAPHICAL INFORMATION.

This section captures the same data as the first administration. The collected results are presented here.

TABLE H.1. Descriptive statistics for current age (years) for L2 and native English speakers.

	n	$\bar{x}$	sd
L2 Users	66	20.409	3.263
English NS	30	19.200	1.472

An ANOVA revealed that there is a statistically significant difference for age between the groups ($F_{1,94} = 3.76$, $p = .056$).

TABLE H.2. Descriptive statistics for current age (years) for groups by L2 English proficiency.

	n	$\bar{x}$	sd
Adv.	13	20.538	2.332
Int.-High	24	20.250	2.863
Int.-Low	14	19.286	1.201
Low	15	21.600	5.207
English NS	30	19.200	1.472

An ANOVA revealed that there are statistically significant differences for age among the proficiency groups ($F_{4,90} = 2.22$, $p = .072$). A Tukey-Kramer posthoc test revealed no significant differences pairwise between each group at an alpha of .05.

TABLE H.3. Descriptive statistics for current age (years) for groups by L1.

	n	$\bar{x}$	sd
Arabic	7	23.571	3.552
Chinese	44	19.773	3.176
Other	15	20.800	2.569
English NS	30	19.200	1.472

An ANOVA revealed that there is a statistically significant difference for age among the language groups ($F_{3,92} = 5.550$, $p = .002$). A Tukey-Kramer posthoc test revealed significant differences between the Arabic group and the three additional groups at an alpha of .05.

TABLE H.4. Frequency table for gender for L2 and native English speakers.

	n	male	female
L2 Users	66	32	34
English NS	30	10	20

A Wilcoxon Rank Sum Test revealed that there is no statistically significant difference for gender between the groups ($\chi^2_1 = 1.904$, $p = .168$).

TABLE H.5. Frequency table for gender for groups by L2 English proficiency

	n	male	female
Adv.	13	5	8
Int.-High	24	11	13
Int.-Low	14	8	6
Low	15	8	7
English NS	30	10	20

A Kruskal-Wallis Test revealed that there no statistically significant differences for gender among the proficiency groups ($\chi^2_4 = 3.061$, $p = .548$).

TABLE H.6. Frequency table for gender for groups by L1.

	n	male	female
Arabic	7	5	2
Chinese	44	21	23
Other	15	6	9
English NS	30	10	20

A Kruskal-Wallis Test revealed that there are no statistically significant differences for gender among the language groups ($\chi^2_3 = 3.830$, $p = .280$).

TABLE H.7. Frequency table for handedness for L2 and native English speakers.

	n	right	left	both equally
L2 Users	66	59	4	3
English NS	30	26	4	0

A Wilcoxon Rank Sum Test revealed that there is no statistically significant difference for handedness between the groups ($\chi^2_1 = 0.090$, $p = .764$).

TABLE H.8. Frequency table for handedness for groups by L2 English proficiency.

	n	right	left	both equally
Adv.	13	11	2	0
Int.-High	24	23	1	0
Int.-Low	14	12	0	2
Low	15	13	1	1
English NS	30	26	4	0

A Kruskal-Wallis Test revealed that there are no statistically significant differences for handedness among the proficiency groups ($\chi^2_4 = 1.790$, $p = .774$).

TABLE H.9. Frequency table for handedness for groups by L1.

	n	right	left	both equally
Arabic	7	7	0	0
Chinese	44	40	1	3
Other	15	12	3	0
English NS	30	26	4	0

A Kruskal-Wallis Test revealed that there are no statistically significant differences for handedness among the language groups ($\chi^2_3 = 2.020$, $p = .5683$).

TABLE H.10. Frequency table for highest level of education for L2 and native English speakers.

	n	High school	Some college	College	Some grad. school	Masters	Doctorate
L2 Users	66	25	13	20	6	1	1
English NS	30	4	26	0	0	0	0

A Wilcoxon Rank Sum Test revealed that there is no statistically significant difference for highest level of education achieved between the groups ($\chi^2_1 = 1.021$, p = .312).

TABLE H.11. Frequency table for highest level of education for groups by L2 English proficiency.

	n	High school	Some college	College	Some grad. school	Masters	Doctorate
Adv.	13	5	4	4	0	0	0
Int.-High	24	8	4	10	1	1	0
Int.-Low	14	4	2	4	4	0	0
Low	15	8	3	2	1	0	1
English NS	30	4	26	0	0	0	0

A Kruskal-Wallis Test revealed that there are no statistically significant differences for highest level of education achieved among the proficiency groups ($\chi^2_4 = 5.731$, p = .220).

TABLE H.12. Frequency table for highest level of education for groups by L1.

	n	High school	Some college	College	Some grad. school	Masters	Doctorate
Arabic	7	2	1	4	0	0	0
Chinese	44	18	7	12	6	1	0
Other	15	5	5	4	0	0	1
English NS	30	4	26	0	0	0	0

A Kruskal-Wallis Test revealed that there are no statistically significant differences for highest level of education achieved among the language groups ($\chi^2_3 = 1.428$, $p = .699$).

TABLE H.13. Frequency table for first languages groups by L2 English proficiency.

	n	English	Arabic	Chinese	French	Hindi	Japanese	Korean	Spanish	Thai	Vietnamese
Adv.	15	0	3	5	2	1	1	1	2	0	0
Int.-High	29	0	5	20	0	1	1	1	0	0	1
Int.-Low	16	0	3	11	0	0	2	0	0	0	0
Low	27	0	0	17	0	0	6	2	1	1	0
English NS	31	31	0	0	0	0	0	0	0	0	0

A Kruskal-Wallis Test revealed that there are statistically significant differences for first language among the proficiency groups ($\chi^2_4 = 80.513$, $p < .001$). A pairwise two-sided multiple comparison analysis using the Dwass, Steel, Critchlow-Fligner method revealed that this significant finding emerges from the differences in the proportions of first language between the English L2 and English NS groups ($p < .001$).

TABLE H.14. Frequency table for second/foreign languages studied for groups by L2 English proficiency.

	n	English	Arabic	Chinese	French	German	Gujarati	Italian	Japanese	Korean	Latin	Russian	Spanish	Tagalog	Taiwanese	Thai	Vietnamese
Adv.	15	15	1	0	1	0	1	0	0	0	0	0	0	0	0	1	1
Int.-H	29	29	0	0	1	0	0	0	0	1	0	0	0	0	1	0	0
Int.-L	16	16	0	0	1	1	0	0	0	3	0	0	0	0	0	0	0
Low	27	27	0	2	0	1	0	0	1	2	0	0	1	0	0	0	0
Eng NS	31	0	1	2	13	4	0	1	3	2	2	2	20	1	0	0	0

TABLE H.15. Frequency table for second/foreign languages studied for groups by L1.

	n	English	Arabic	Chinese	French	German	Gujarati	Italian	Japanese	Korean	Latin	Russian	Spanish	Tagalog	Taiwanese	Thai	Vietnamese
Arabic	11	11	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
Chinese	53	53	0	0	1	1	0	0	1	5	0	0	0	0	1	0	0
Other	23	23	0	1	2	1	1	0	0	1	0	0	1	0	0	1	1
Eng NS	31	0	1	2	13	4	0	1	3	2	2	2	20	1	0	0	0

H.3 LANGUAGE HISTORY.

This section captures the same data as the first administration. The collected results are presented here.

TABLE H.16. Descriptive statistics for age of exposure for L2 and native English speakers.

		L1		L2 English	
	n	x̄	sd	x̄	sd
L2 Users	67	0.604	1.223	9.881	4.287
English NS	30	0.067	0.365		

A Wilcoxon Rank Sum Test revealed that there is a statistically significant difference between the groups for age of exposure to the first language (χ^2_1 = 7.660, p = .005).

TABLE H.17. Descriptive statistics for age of exposure for groups by L2 English proficiency.

		L1		L2 English	
	n	x̄	sd	x̄	sd
Adv.	13	0.500	0.913	10.461	5.651
Int.-High	24	0.750	1.595	8.833	3.460
Int.-Low	14	0.429	0.756	10.929	5.370
Low	16	0.625	1.204	10.063	2.977
English NS	30	0.067	0.365		

A Kruskal-Wallis Test revealed that there are no statistically significant differences among the proficiency groups for age of exposure to the first language (χ^2_4 = 7.735, p = .102) or for the age of exposure to the second language (χ^2_3 = 2.443, p = .486).

TABLE H.18. Descriptive statistics for age of exposure for groups by L1.

		L1		L2 English	
	n	x̄	sd	x̄	sd
Arabic	7	0.571	0.976	15.857	3.671
Chinese	44	0.523	0.976	8.750	3.667
Other	16	0.844	1.841	10.375	4.031
English NS	30	0.067	0.365		

A Kruskal-Wallis Test revealed that there are statistically significant differences among the language groups for age of exposure to the first language (χ^2_3 = 7.786, p = .051) and for the age of exposure to the second language (χ^2_2 = 13.298, p = .001). A pairwise two-sided multiple comparison analysis using the Dwass, Steel, Critchlow-Fligner method revealed that the former significant finding emerges from the differences in the proportions of age of exposure between the L2 users and the English NS control group (Chinese: DSCF = 3.700, p = .044; Other: DSCF = 3.706, p = .044). The same analysis revealed that the latter significant finding emerges from the differences in the proportions of age of exposure between the Arabic and Chinese groups (DSCF = 4.982, p = .001) and between the Arabic and Other groups (DSCF = 3.765, p = .021).

TABLE H.19. Descriptive statistics for age of comfort using L2 and native English speakers.

		L1		L2 English	
	n	x̄	sd	x̄	Not comfortable
L2 Users	67	1.896	2.887	14.818	23
English NS	30	0.067	0.365		

A Wilcoxon Rank Sum Test revealed that there is a statistically significant difference between the groups for age of comfort using the first language (χ^2_1 = 13.863, p < .001).

TABLE H.20. Descriptive statistics for age of comfort using groups by L2 English proficiency.

		L1		L2 English	
	n	x̄	sd	x̄	Not comfortable
Adv.	13	1.615	2.844	16.778	4
Int.-High	24	1.958	3.141	14.058	7
Int.-Low	14	1.643	2.790	11.667	5
Low	16	2.250	2.840	17.444	7
English NS	30	0.067	0.365		

A Kruskal-Wallis Test revealed that there is a statistically significant difference among the proficiency groups for age of comfort using the first language (χ^2_4 = 14.809, p = .005) but that there are no significant differences for the age of comfort using the second language (χ^2_3 = 2.390, p = .496). A pairwise two-sided multiple comparison analysis using the Dwass, Steel, Critchlow-Fligner method revealed that the former significant finding emerges from the differences in the proportions of age of comfort between the L2 users and the English NS control group (Adv.: DSCF = 4.378, p = .017; Int.-High: DSCF = 4.579, p = .011; Int.-Low: DSCF = 4.122, p = .029; Low: DSCF = 5.466, p = .001).

TABLE H.21. Descriptive statistics for age of comfort using groups by L1.

		L1		L2 English	
	n	$\bar{x}$	sd	$\bar{x}$	Not comfortable
Arabic	7	4.143	4.845	22.200	2
Chinese	44	1.636	2.553	12.759	15
Other	16	1.625	2.446	15.667	7
English NS	30	0.067	0.365		

A Kruskal-Wallis Test revealed that there is a statistically significant difference among the language groups for age of comfort using the first language ($\chi^2_3 = 15.897$, p = .001) but that there are no significant differences for the age of comfort using the second language ($\chi^2_2 = 3.540$, p = .170). A pairwise two-sided multiple comparison analysis using the Dwass, Steel, Critchlow-Fligner method revealed that the former significant finding emerges from the differences in the proportions of age of comfort between the L2 users and the English NS control group (Arabic: DSCF = 5.353, p < .001; Chinese: DSCF = 4.745, p = .004; Other: DSCF = 4.890, p = .003).

TABLE H.22. Descriptive statistics for years of formal education in that language for L2 and native English speakers.

		L1		L2 English	
	n	$\bar{x}$	sd	$\bar{x}$	sd
L2 Users	67	12.776	5.949	8.470	4.446
English NS	30	15.067	8.262		

A Wilcoxon Rank Sum Test revealed that there is no statistically significant difference between the groups for years of formal education in the first language ($\chi^2_1 = 2.201$, p = .138).

TABLE H.23. Descriptive statistics for years of formal education in that language for groups by L2 English proficiency.

		L1		L2 English	
	n	$\bar{x}$	sd	$\bar{x}$	sd
Adv.	13	13.769	4.567	8.346	4.404
Int.-High	24	13.208	4.303	8.833	4.669
Int.-Low	14	9.643	7.344	6.857	5.489
Low	16	14.063	7.197	9.438	2.874
English NS	30	15.067	8.262		

A Kruskal-Wallis Test revealed that there are no statistically significant differences among the proficiency groups for years of formal education in the first language (χ^2_4 = 6.372, p = .173) or for years of formal education in the second language (χ^2_3 = 2.016, p = .569).

TABLE H.24. Descriptive statistics for years of formal education in that language for groups by L1.

		L1		L2 English	
	n	$\bar{x}$	sd	$\bar{x}$	sd
Arabic	7	11.429	8.162	3.000	2.000
Chinese	44	13.386	5.068	9.250	4.368
Other	16	11.688	7.227	8.719	3.856
English NS	30	15.067	8.262		

A Kruskal-Wallis Test revealed that there are no statistically significant differences among the language groups for years of formal education in the first language (χ^2_3 = 2.387, p = .496) but that there is a significant difference for years of formal education in the second language (χ^2_2 = 11.529, p = .003). A pairwise two-sided multiple comparison analysis using the Dwass, Steel, Critchlow-Fligner method revealed that the latter significant finding emerges from the differences in the proportions of age of comfort between the Arabic and Chinese groups (DSCF = 4.593, p = .003) and between the Arabic and Other groups (DSCF = 4.126, p = .010).

TABLE H.25. Descriptive statistics for years immersed in a country where that language is spoken for L2 and native English speakers.

		L1		L2 English	
	n	x̄	sd	x̄	sd
L2 Users	66	17.778	6..428	2.981	4.579
English NS	30	18.500	3.803		

A Wilcoxon Rank Sum Test revealed that there is no statistically significant difference between the groups for years spent immersed in a country in which the first language is spoken ($\chi^2_1 = 0.168$, $p = .682$).

TABLE H.26. Descriptive statistics for years immersed in a country where that language is spoken for groups by L2 English proficiency.

		L1		L2 English	
	n	x̄	sd	x̄	sd
Adv.	13	18.231	5.747	4.092	7.302
Int.-High	24	18.138	5.058	2.200	3.301
Int.-Low	14	16.643	5.995	1.125	1.655
Low	15	17.867	9.296	4.875	4.628
English NS	30	18.500	3.803		

A Kruskal-Wallis Test revealed that there are no statistically significant differences among the proficiency groups for years spent immersed in a country in which the first language is spoken ($\chi^2_4 = 1.866$, $p = .760$) or for years spent immersed in a country in which the second language is spoken ($\chi^2_3 = 5.938$, $p = .115$).

TABLE H.27. Descriptive statistics for years immersed in a country where that language is spoken for groups by L1.

		L1		L2 English	
	n	x̄	sd	x̄	sd
Arabic	7	18.857	8.915	1.800	0.902
Chinese	44	18.023	5.415	2.631	3.672
Other	15	16.553	8.082	4.463	7.065
English NS	30	18.500	3.803		

A Kruskal-Wallis Test revealed that there are no statistically significant differences among the language groups for years spent immersed in a country in which the first language is spoken (χ^2_3 = 2.623, p = .454) or for years spent immersed in a country in which the second language is spoken (χ^2_2 = 1.135, p = .567).

TABLE H.28. Descriptive statistics for years immersed in a family where that language is spoken for L2 and native English speakers.

		L1		L2 English	
	n	x̄	sd	x̄	sd
L2 Users	66	18.591	5.997	1.119	3.136
English NS	30	19.167	1.510		

A Wilcoxon Rank Sum Test revealed that there is no statistically significant difference between the groups for years spent within a family in which the first language is spoken (χ^2_1 = 0.010, p = .920).

TABLE H.29. Descriptive statistics for years immersed in a family where that language is spoken for groups by L2 English proficiency.

		L1		L2 English	
	n	x̄	sd	x̄	sd
Adv.	13	20.308	2.016	1.462	4.977
Int.-High	24	18.917	5.064	0.833	2.615
Int.-Low	14	17.357	5.168	0.500	0.855
Low	15	17.733	9.558	1.813	3.351
English NS	30	19.167	1.510		

A Kruskal-Wallis Test revealed that there are no statistically significant differences among the proficiency groups for years spent within a family in which the first language is spoken (χ^2_4 =5.297, p = .258) or for years spent within a family in which the second language is spoken (χ^2_3 = 4.735, p = .192).

TABLE H.30. Descriptive statistics for years immersed in a family where that language is spoken for groups by L1.

		L1		L2 English	
	n	x̄	sd	x̄	sd
Arabic	7	22.571	3.735	0.286	0.488
Chinese	44	17.977	5.773	0.864	2.258
Other	15	18.533	7.0139	2.188	5.180
English NS	30	19.167	1.510		

A Kruskal-Wallis Test revealed that there are statistically significant differences among the language groups for years spent within a family in which the first language is spoken ($\chi^2_3 = 9.384$, $p = .025$) but that there are no significant differences for years spent within a family in which the second language is spoken ($\chi^2_2 = 0.326$, $p = .850$). A pairwise two-sided multiple comparison analysis using the Dwass, Steel, Critchlow-Fligner method revealed that the former significant finding emerges from the differences in the proportions of years spent within a family in which the first language is spoken between the Arabic and Chinese groups (DSCF = 3.323, p = .087)

TABLE H.31. Descriptive statistics for years immersed in work/school where that language is spoken for L2 and native English speakers.

		L1		L2 English	
	n	$\bar{x}$	sd	$\bar{x}$	sd
L2 Users	65	12.466	9.240	1.510	3.132
English NS	30	9.767	7.881		

A Wilcoxon Rank Sum Test revealed that there is no statistically significant difference between the groups for years spent immersed in a work/school environment in which the first language is spoken ($\chi^2_1 = 0.9586$, $p = .328$).

TABLE H.32. Descriptive statistics for years immersed in work/school where that language is spoken for groups by L2 English proficiency.

		L1		L2 English	
	n	$\bar{x}$	sd	$\bar{x}$	sd
Adv.	13	14.667	8.206	1.731	3.844
Int.-High	24	14.138	8.616	1.738	3.756
Int.-Low	14	8.643	9.001	0.071	0.267
Low	15	11.600	10.762	2.250	2.646
English NS	30	9.767	7.881		

A Kruskal-Wallis Test revealed that there are no statistically significant differences among the proficiency groups for years spent immersed in a work/school environment in which the first language is spoken ($\chi^2_4 = 6.724$, $p = .151$) but there are significant differences for years spent immersed in a work/school environment in which the second language is spoken (χ^2_3 = 11.635, p = .009). A pairwise two-sided multiple comparison analysis using the Dwass, Steel, Critchlow-Fligner method revealed that the latter significant finding emerges from the differences in the proportions of years spent immersed in a work/school environment between the Int.-Low and Int.-High (DSCF = 4.285, p = .0131) and between the Int.-Low and Low groups (DSCF = 4.499, p = .008).

TABLE H.33. Descriptive statistics for years immersed in work/school where that language is spoken for groups by L1.

		L1		L2 English	
	n	$\bar{x}$	sd	$\bar{x}$	sd
Arabic	7	13.714	11.398	0.429	0.787
Chinese	44	13.841	8.915	1.318	2.041
Other	16	7.521	7.992	2.513	5.408
English NS	30	9.767	7.881		

A Kruskal-Wallis Test revealed that there are no statistically significant differences among the language groups for years spent immersed in a work/school environment in which the first language is spoken (χ^2_3 = 5.665, p = .129) or for years spent immersed in a work/school environment in which the second language is spoken (χ^2_2 = 1.276, p = .528).

H.4 LANGUAGE USE.

This section captures the same data as the first administration. The collected results are presented here.

TABLE H.34. Descriptive statistics for overall language use for L2 and native English speakers.

		L1		L2 English	
	n	$\bar{x}$	sd	$\bar{x}$	sd
L2 Users	66	68.597	15.509	32.309	18.118
English NS	30	95.566	5.939		

A Wilcoxon Rank Sum Test revealed that there is a statistically significant difference between the groups for overall use of the first language (χ^2_1 = 54.307, p < .001).

TABLE H.35. Descriptive statistics for overall language use for groups by L2 English proficiency.

		L1		L2 English	
	n	x̄	sd	x̄	sd
Adv.	13	68.862	19.691	29.523	19.853
Int.-High	24	71.550	13.671	28.583	15.108
Int.-Low	14	66.629	16.566	34.136	17.178
Low	15	65.480	13.897	38.561	21.203
English NS	30	95.566	5.939		

A Kruskal-Wallis Test revealed that there is a statistically significant difference among the proficiency groups for overall use of the first language ($\chi^2_4 = 55.525$, $p < .001$) but that there are no significant differences for overall use of the second language ($\chi^2_3 = 4.236$, $p = .236$). A pairwise two-sided multiple comparison analysis using the Dwass, Steel, Critchlow-Fligner method revealed that the former significant finding emerges from the differences in the proportions of overall use between the L2 users and the English NS control group (Adv.: DSCF = 6.714, $p < .001$; Int.-High: DSCF = 8.430, $p < .001$; Int.-Low: DSCF = 7.258, $p < .001$; Low: DSCF = 7.343, $p < .001$), which is to be expected considering the largely monolingual English NS group.

TABLE H.36. Descriptive statistics for overall language use for groups by L1.

		L1		L2 English	
	n	x̄	sd	x̄	sd
Arabic	7	72.543	8.756	27.457	8.756
Chinese	44	71.691	14.134	28.625	15.191
Other	15	57.680	17.455	44.561	23.265
English NS	30	95.566	5.939		

A Kruskal-Wallis Test revealed that there is a statistically significant difference among the language groups for overall use of the first language ($\chi^2_3 = 58.655$, $p < .001$) and for overall use of the second language ($\chi^2_2 = 7.170$, $p = .028$). A pairwise two-sided multiple comparison analysis using the Dwass, Steel, Critchlow-Fligner method revealed that the former significant finding emerges from the differences in the proportions of overall use between the L2 users and the English NS control group (Arabic: DSCF = 5.665, $p < .001$; Chinese: DSCF = 9.603, $p < .001$; Other: DSCF = 7.412, $p < .001$) and between the L2 user groups and the Other group (Arabic: DSCF = 3.440, $p = .071$; Chinese: DSCF = 3.929, $p = .028$). The same analysis revealed that the latter significant finding emerges from the differences in the proportions of overall use between the L2 user groups and the Other group (Arabic: DSCF = 3.024, $p = .082$; Chinese: DSCF = 3.548, $p = .033$).

(i) Use with friends:

TABLE H.37. Descriptive statistics for language use with friends for L2 and native English speakers.

		L1		L2 English	
	n	x̄	sd	x̄	sd
L2 Users	66	65.258	28.684	35.381	29.644
English NS	30	97.667	3.960		

A Wilcoxon Rank Sum Test revealed that there is a statistically significant difference between the groups for first language use with friends ($\chi^2_1 = 46.502$, $p < .001$).

TABLE H.38. Descriptive statistics for language use with friends for groups by L2 English proficiency.

		L1		L2 English	
	n	x̄	sd	x̄	sd
Adv.	13	57.615	35.293	41.231	35.970
Int.-High	24	75.750	20.096	24.042	20.060
Int.-Low	14	62.643	34.191	37.321	34.132
Low	15	57.533	26.262	45.938	29.105
English NS	30	97.667	3.960		

A Kruskal-Wallis Test revealed that there is a statistically significant difference among the groups for first language use with friends ($\chi^2_4 = 48.900$, $p < .001$) but that there are no significant differences for second language use with friends ($\chi^2_3 = 5.694$, $p = .128$). A pairwise two-sided multiple comparison analysis using the Dwass, Steel, Critchlow-Fligner method revealed that the former significant finding emerges from the differences in the proportions of use with friends between the L2 users and the English NS control group (Adv.: DSCF = 6.414, $p < .001$; Int.-High: DSCF = 7.746, $p < .001$; Int.-Low: DSCF = 7.275, $p < .001$; Low: DSCF = 7.011, $p < .001$).

TABLE H.39. Descriptive statistics for language use with friends for groups by L1.

		L1		L2 English	
	n	$\bar{x}$	sd	$\bar{x}$	sd
Arabic	7	62.143	28.557	37.857	28.557
Chinese	44	72.273	25.827	27.602	25.800
Other	15	46.133	29.503	55.688	31.813
English NS	30	97.667	3.960		

A Kruskal-Wallis Test revealed that there is a statistically significant difference among the groups for first language use with friends ($\chi^2_3 = 52.012$, $p < .001$) and for second language use with friends ($\chi^2_2 = 9.543$, $p = .009$). A pairwise two-sided multiple comparison analysis using the Dwass, Steel, Critchlow-Fligner method revealed that the former significant finding emerges from the differences in the proportions of use with friends between the L2 users and the English NS control group (Arabic: DSCF = 6.167, $p < .001$; Chinese: DSCF = 8.537, $p < .001$; Other: DSCF = 7.602, $p < .001$) and between the Chinese and Other groups (DSCF = 4.142, $p = .018$). The same analysis revealed that the latter significant finding emerges from the differences in the proportions of use with friends between the Chinese and Other groups (DSCF = 4.271, $p = .007$).

(ii) Use with family:

TABLE H.40. Descriptive statistics for language use with family for L2 and native English speakers.

		L1		L2 English	
	n	$\bar{x}$	sd	$\bar{x}$	sd
L2 Users	66	93.258	16.706	6.650	18.536
English NS	30	99.367	2.059		

A Wilcoxon Rank Sum Test revealed that there is a statistically significant difference between the groups for first language use with family ($\chi^2_1 = 4.568$, $p = .033$).

TABLE H.41. Descriptive statistics for language use with family for groups by L2 English proficiency.

		L1		L2 English	
	n	x̄	sd	x̄	sd
Adv.	13	91.538	20.350	3.077	8.549
Int.-High	24	90.625	20.763	8.348	20.749
Int.-Low	14	99.429	2.138	0.357	1.336
Low	15	93.200	13.251	12.619	26.588
English NS	30	99.367	2.059		

A Kruskal-Wallis Test revealed that there are statistically significant differences among the groups for first language use with family ($\chi^2_4 = 12.725$, $p = .013$) and for second language use with family ($\chi^2_3 = 9.404$, $p = .024$). A pairwise two-sided multiple comparison analysis using the Dwass, Steel, Critchlow-Fligner method revealed that the former significant finding emerges from the differences in the proportions of use with family between the Low and English NS groups (DSCF = 4.159, p = .027) and between the Low and Int.-Low groups (DSCF = 3.653, p = .074). Another pairwise analysis using this method revealed that the latter finding emerges from the differences in the proportions of use with family between the Low and Int.-Low groups (DSCF = 3.962, p = .026).

TABLE H.42. Descriptive statistics for language use with family for groups by L1.

		L1		L2 English	
	n	x̄	sd	x̄	sd
Arabic	7	96.429	4.756	3.571	4.756
Chinese	44	94.227	16.011	5.070	15.732
Other	15	88.833	21.694	12.243	27.236
English NS	30	99.367	2.059		

A Kruskal-Wallis Test revealed that there are no statistically significant differences among the groups for first language use with family ($\chi^2_3 = 5.888$, p = .117) or for second language use with family ($\chi^2_2 = 1.314$, p = .519).

(iii) Use at school/work:

TABLE H.43. Descriptive statistics for language use at school/work for L2 and native English speakers.

		L1		L2 English	
	n	$\bar{x}$	sd	$\bar{x}$	sd
L2 Users	66	35.848	30.115	64.657	30.199
English NS	30	94.667	7.649		

A Wilcoxon Rank Sum Test revealed that there is a statistically significant difference between the groups for first language use at school/work ($\chi^2_1 = 53.323$, p < .001).

TABLE H.44. Descriptive statistics for language use at school/work for groups by L2 English proficiency.

		L1		L2 English	
	n	$\bar{x}$	sd	$\bar{x}$	sd
Adv.	13	36.538	32.364	63.462	32.364
Int.-High	24	42.083	29.961	57.917	29.961
Int.-Low	14	37.000	33.013	63.000	33.013
Low	15	24.200	24.748	77.188	24.696
English NS	30	94.667	7.649		

A Kruskal-Wallis Test revealed that there is a statistically significant difference among the groups for first language use at school/work ($\chi^2_4 = 54.958$, p < .001) but that there are no significant differences for second language use at school/work ($\chi^2_3 = 3.609$, p = .307). A pairwise two-sided multiple comparison analysis using the Dwass, Steel, Critchlow-Fligner method revealed that the former significant finding emerges from the differences in the proportions of use at school/work between the L2 users and the English NS control group (Adv.: DSCF = 7.063, p < .001; Int.-High: DSCF = 8.359, p < .001; Int.-Low: DSCF = 6.827, p < .001; Low: DSCF = 7.796, p < .001).

TABLE H.45. Descriptive statistics for language use at school/work for groups by L1.

		L1		L2 English	
	n	$\bar{x}$	sd	$\bar{x}$	sd
Arabic	7	42.857	34.864	57.143	34.864
Chinese	44	38.182	30.137	61.818	30.137
Other	15	25.733	27.366	75.750	27.234
English NS	30	94.667	7.649		

A Kruskal-Wallis Test revealed that there is a statistically significant difference among the groups for first language use at school/work ($\chi^2_3 = 54.553$, $p < .001$) but that there are no significant differences for second language use at school/work ($\chi^2_2 = 2.920$, $p = .232$). A pairwise two-sided multiple comparison analysis using the Dwass, Steel, Critchlow-Fligner method revealed that the former significant finding emerges from the differences in the proportions of use at school/work between the L2 users and the English NS control group (Arabic: DSCF = 5.493, $p < .001$; Chinese: DSCF = 9.558, $p < .001$; Other: DSCF = 7.645, $p < .001$).

(iv) Use with self:

TABLE H.46. Descriptive statistics for language use with self for L2 and native English speakers.

		L1		L2 English	
	n	$\bar{x}$	sd	$\bar{x}$	sd
L2 Users	66	74.841	25.023	25.925	26.599
English NS	30	93.397	17.360		

A Wilcoxon Rank Sum Test revealed that there is a statistically significant difference between the groups for first language use with themselves ($\chi^2_1 = 19.169$, $p < .001$).

TABLE H.47. Descriptive statistics for language use with self for groups by L2 English proficiency.

		L1		L2 English	
	n	x̄	sd	x̄	sd
Adv.	13	79.462	30.201	19.385	30.508
Int.-High	24	78.083	17.874	24.792	17.971
Int.-Low	14	63.179	28.682	36.429	28.788
Low	15	76.533	25.685	28.250	31.334
English NS	30	93.397	17.360		

A Kruskal-Wallis Test revealed that there is a statistically significant difference among the groups for first language use with themselves ($\chi^2_4 = 22.690$, $p < .001$) but that there are no significant differences for second language use with themselves ($\chi^2_3 = 4.265$, $p = .234$). A pairwise two-sided multiple comparison analysis using the Dwass, Steel, Critchlow-Fligner method revealed that the former significant finding emerges from the differences in the proportions of use with themselves between the L2 users and the English NS control group (Int.-High.: DSCF = 5.615, $p < .001$; Int.-Low: DSCF = 5.136, $p = .003$; Low: DSCF = 4.023, $p = .036$).

TABLE H.48. Descriptive statistics for language use with self for groups by L1.

		L1		L2 English	
	n	x̄	sd	x̄	sd
Arabic	7	77.143	14.100	22.857	14.100
Chinese	44	77.034	24.865	22.773	24.872
Other	15	67.333	29.066	35.938	33.461
English NS	30	93.397	17.360		

A Kruskal-Wallis Test revealed that there is a statistically significant difference among the groups for first language use with themselves ($\chi^2_3 = 21.599$, $p < .001$) but that there are no significant differences for second language use with themselves ($\chi^2_2 = 2.846$, $p = .241$). A pairwise two-sided multiple comparison analysis using the Dwass, Steel, Critchlow-Fligner method revealed that the former significant finding emerges from the differences in the proportions of use with themselves between the L2 users and the

English NS control group (Arabic: DSCF = 3.849, $p = .033$; Chinese: DSCF = 5.012, $p = .022$; Other: DSCF = 5.792, $p < .001$).

(v) Use to count:

TABLE H.49. Descriptive statistics for language use to count for L2 and native English speakers.

		L1		L2 English	
	n	$\bar{x}$	sd	$\bar{x}$	sd
L2 Users	65	73.954	23.438	28.134	26.327
English NS	30	97.733	4.785		

A Wilcoxon Rank Sum Test revealed that there is a statistically significant difference between the groups for first language use to count ($\chi^2_1 = 28.834$, $p < .001$).

TABLE H.50. Descriptive statistics for language use to count for groups by L2 English proficiency.

		L1		L2 English	
	n	$\bar{x}$	sd	$\bar{x}$	sd
Adv.	13	79.154	27.691	20.462	27.361
Int.-High	24	71.208	20.805	28.667	20.932
Int.-Low	14	71.538	23.397	33.571	29.511
Low	15	75.933	25.018	28.813	30.734
English NS	30	97.733	4.785		

A Kruskal-Wallis Test revealed that there is a statistically significant difference among the groups for first language use to count ($\chi^2_4 = 31.718$, $p < .001$) but that there are no significant differences for second language use with themselves ($\chi^2_3 = 2.486$, $p = .478$). A pairwise two-sided multiple comparison analysis using the Dwass, Steel, Critchlow-Fligner method revealed that the former significant finding emerges from the differences in the proportions of use to count between the L2 users and the English NS control group (Adv.: DSCF = 4.387, $p = .017$; Int.-High: DSCF = 7.797, $p < .001$; Int.-Low: DSCF = 5.19, $p = .002$; Low: DSCF = 4.125, $p = .029$).

TABLE H.51. Descriptive statistics for language use to count for groups by L1.

		L1		L2 English	
	n	x̄	sd	x̄	sd
Arabic	7	84.143	17.034	15.857	17.034
Chinese	44	77.070	20.683	24.614	23.560
Other	15	60.267	28.654	43.188	31.526
English NS	30	97.733	4.785		

A Kruskal-Wallis Test revealed that there is a statistically significant difference among the groups for first language use to count ($\chi^2_3 = 32.853$, $p < .001$) and for second language use with themselves ($\chi^2_2 = 6.044$, $p = .049$). A pairwise two-sided multiple comparison analysis using the Dwass, Steel, Critchlow-Fligner method revealed that the former significant finding emerges from the differences in the proportions of use to count between the L2 users and the English NS control group (Arabic: DSCF = 4.06911, p = .021; Chinese: DSCF = 6.586, p < .001; Other: DSCF = 6.707, p < .001). The same analysis revealed that the latter significant finding emerges from the differences in the proportions of use to count between the Chinese and Other groups (DSCF = 3.145, p = .067).

H.5 LANGUAGE PROFICIENCY.

This section captures the same data as the first administration. The collected results are presented here.

TABLE H.52. Descriptive statistics for self-assessed proficiency for L2 and native English speakers.

		L1		L2 English	
	n	x̄	sd	x̄	sd
L2 Users	66	5.638	0.578	3.882	0.917
English NS	30	5.992	0.046		

A Wilcoxon Rank Sum Test revealed that there is a statistically significant difference between the groups for overall self-assessed L1 proficiency ($\chi^2_1 = 12.687$, $p < .001$).

TABLE H.53. Descriptive statistics for self-assessed proficiency for groups by L2 English proficiency.

		L1		L2 English	
	n	x̄	sd	x̄	sd
Adv.	13	5.654	0.650	4.385	1.069
Int.-High	24	5.615	0.590	3.667	0.789
Int.-Low	14	5.482	0.690	3.929	0.988
Low	15	5.806	0.347	3.755	0.818
English NS	30	5.992	0.046		

A Kruskal-Wallis Test revealed that there is a statistically significant difference among the proficiency groups for overall self-assessed L1 proficiency ($\chi^2_4 = 15.371$, $p = .004$) but that there are no significant differences for overall self-assessed L2 proficiency ($\chi^2_3 = 3.881$, $p = .275$). A pairwise two-sided multiple comparison analysis using the Dwass, Steel, Critchlow-Fligner method revealed that the former significant finding emerges from the differences in the proportions of overall self-assessed proficiency between the L2 users and the English NS control group (Adv.: DSCF = 3.695, p = .068; Int.-High: DSCF = 4.969, p = .004; Int.-Low: DSCF = 5.352, p = .002).

TABLE H.54. Descriptive statistics for self-assessed proficiency for groups by L1.

		L1		L2 English	
	n	x̄	sd	x̄	sd
Arabic	7	5.536	0.728	4.571	0.800
Chinese	44	5.716	0.527	3.722	0.829
Other	15	5.456	0.641	4.021	1.074
English NS	30	5.992	0.046		

A Kruskal-Wallis Test revealed that there is a statistically significant difference among the language groups for overall self-assessed L1 proficiency ($\chi^2_3 = 15.934$, $p = .001$) and for overall self-assessed L2 proficiency ($\chi^2_2 = 5.473$, $p = .065$). A pairwise two-sided multiple comparison analysis using the Dwass, Steel, Critchlow-Fligner method revealed that the former significant finding emerges from the differences in the proportions of overall self-assessed proficiency between the L2 users and the English NS control group

(Arabic: DSCF = 4.324, p = .012; Chinese: DSCF = 4.302, p = .013; Other: DSCF = 5.655, p < .001). The same analysis revealed that the latter significant finding emerges from the differences in the proportions of overall self-assessed proficiency between the Arabic and Chinese groups (DSCF = 3.378, p = .045)

(i) Speaking self-assessment:

TABLE H.55. Descriptive statistics for self-assessed speaking proficiency for L2 and native English speakers.

		L1		L2 English	
	n	$\bar{x}$	sd	$\bar{x}$	sd
L2 Users	66	5.742	0.535	3.746	0.910
English NS	30	5.967	0.183		

A Wilcoxon Rank Sum Test revealed that there is a statistically significant difference between the groups for overall self-assessed L1 speaking proficiency ($\chi^2_1 = 5.010$, p = .025).

TABLE H.56. Descriptive statistics for self-assessed speaking proficiency for groups by L2 English proficiency.

		L1		L2 English	
	n	$\bar{x}$	sd	$\bar{x}$	sd
Adv.	13	5.692	0.630	4.154	0.987
Int.-High	24	5.667	0.637	3.583	0.830
Int.-Low	14	5.786	0.426	3.714	0.994
Low	15	5.867	0.352	3.688	0.873
English NS	30	5.967	0.183		

A Kruskal-Wallis Test revealed that there are not statistically significant differences among the groups for self-assessed L1 speaking proficiency ($\chi^2_4 = 6.169$, p = .187) or for self-assessed L2 speaking proficiency ($\chi^2_3 = 2.590$, p = .459).

TABLE H.57. Descriptive statistics for self-assessed speaking proficiency for groups by L1.

		L1		L2 English	
	n	$\bar{x}$	sd	$\bar{x}$	sd
Arabic	7	5.857	0.380	4.429	0.976
Chinese	44	5.795	0.462	3.614	0.841
Other	15	5.33	0.743	3.813	0.981
English NS	30	5.967	0.183		

A Kruskal-Wallis Test revealed that there is a statistically significant difference among the groups for self-assessed L1 speaking proficiency (χ^2_3 = 7.582, p = .056) but there are no significant differences for self-assessed L2 speaking proficiency (χ^2_2 = 4.460, p = .108). A pairwise two-sided multiple comparison analysis using the Dwass, Steel, Critchlow-Fligner method revealed that the former significant finding emerges from the differences in the proportions of self-assessed L1 speaking proficiency between the Other and the English NS group (DSCF = 3.952, p = .027).

(ii) Understanding self-assessment:

TABLE H.58. Descriptive statistics for self-assessed understanding proficiency for L2 and native English speakers.

		L1		L2 English	
	n	$\bar{x}$	sd	$\bar{x}$	sd
L2 Users	65	5.785	0.450	3.924	1.071
English NS	30	6.000	0		

A Wilcoxon Rank Sum Test revealed that there is a statistically significant difference between the groups for overall self-assessed L1 understanding proficiency (χ^2_1 = 6.868, p = .009).

TABLE H.59. Descriptive statistics for self-assessed understanding proficiency for groups by L2 English proficiency.

		L1		L2 English	
	n	x̄	sd	x̄	sd
Adv.	13	5.846	0.376	4.462	1.127
Int.-High	23	5.739	0.541	3.783	1.043
Int.-Low	14	5.786	0.426	3.929	1.141
Low	16	5.800	0.414	3.688	0.946
English NS	30	6.000	0		

A Kruskal-Wallis Test revealed that there are no statistically significant differences among the groups for self-assessed L1 understanding proficiency ($\chi^2_4 = 7.218$, $p = .125$) or for self-assessed L2 understanding proficiency ($\chi^2_3 = 3.563$, $p = .313$).

TABLE H.60. Descriptive statistics for self-assessed understanding proficiency for groups by L1.

		L1		L2 English	
	n	x̄	sd	x̄	sd
Arabic	7	6.000	0	5.000	0.817
Chinese	43	5.791	0.466	3.674	0.919
Other	15	5.667	0.488	4.125	1.258
English NS	30	6.000	0		

A Kruskal-Wallis Test revealed that that there is a statistically significant difference among the groups for self-assessed L1 understanding proficiency ($\chi^2_3 = 11.382$, $p = .010$) and for self-assessed L2 understanding proficiency ($\chi^2_2 = 9.332$, $p = .009$). A pairwise two-sided multiple comparison analysis using the Dwass, Steel, Critchlow-Fligner method revealed that the former significant finding emerges from the differences in the proportions of self-assessed L1 understanding proficiency between the L2 users and the English NS control group (Chinese: DSCF = 3.514, p = .062; Other: DSCF = 4.690, p = .005). The same analysis revealed that the latter significant finding emerges from the differences in the proportions of self-assessed L1 understanding proficiency between the Arabic and Chinese groups (DSCF = 4.378, p = .006).

(iii) Reading self-assessment:

TABLE H.61. Descriptive statistics for self-assessed reading proficiency for L2 and native English speakers.

		L1		L2 English	
	n	x̄	sd	x̄	sd
L2 Users	65	5.631	0.741	3.985	1.045
English NS	30	6.000	0		

A Wilcoxon Rank Sum Test revealed that there is a statistically significant difference between the groups for overall self-assessed L1 reading proficiency (χ^2_1 = 8.071, p = .005).

TABLE H.62. Descriptive statistics for self-assessed reading proficiency for groups by L2 English proficiency.

		L1		L2 English	
	n	x̄	sd	x̄	sd
Adv.	13	5.538	0.967	4.538	1.198
Int.-High	24	5.708	0.624	3.750	0.989
Int.-Low	14	5.357	0.929	4.000	1.109
Low	15	5.857	0.363	3.867	0.834
English NS	30	6.000	0		

A Kruskal-Wallis Test revealed that there is a statistically significant difference among the groups for self-assessed L1 reading proficiency (χ^2_4 = 11.229, p = .024) but that there are no significant differences for self-assessed L2 reading proficiency (χ^2_3 = 3.349, p = .341). A pairwise two-sided multiple comparison analysis using the Dwass, Steel, Critchlow-Fligner method revealed that the former significant finding emerges from the differences in the proportions of self-assessed reading proficiency between the L2 users and the English NS control group (Adv.: DSCF = 3.810, p = .055; Int.-High: DSCF = 3.673, p = .071; Int.-Low: DSCF = 4.855, p = .005).

TABLE H.63. Descriptive statistics for self-assessed reading proficiency for groups by L1.

		L1		L2 English	
	n	$\bar{x}$	sd	$\bar{x}$	sd
Arabic	7	5.429	0.976	4.571	0.787
Chinese	44	5.750	0.576	3.818	0.971
Other	15	5.357	1.008	4.200	1.265
English NS	30	6.000	0		

A Kruskal-Wallis Test revealed that there is a statistically significant difference among the groups for self-assessed L1 reading proficiency ($\chi^2_3 = 11.155$, $p = .011$) but there are no significant differences for self-assessed L2 reading proficiency ($\chi^2_2 = 3.645$, $p = .162$). A pairwise two-sided multiple comparison analysis using the Dwass, Steel, Critchlow-Fligner method revealed that former significant finding emerges from the differences in the proportions of self-assessed reading proficiency between the L2 users and the English NS control group (Arabic: DSCF = 4.199, p = .016; Chinese: DSCF = 3.468, p = .068; Other: DSCF = 4.851, p = .003).

(iv) Writing self-assessment:

TABLE H.64. Descriptive statistics for self-assessed writing proficiency for L2 and native English speakers.

		L1		L2 English	
	n	$\bar{x}$	sd	$\bar{x}$	sd
L2 Users	66	5.394	1.006	3.896	1.116
English NS	60	6.000	0		

A Wilcoxon Rank Sum Test revealed that there is a statistically significant difference between the groups for overall self-assessed L1 writing proficiency ($\chi^2_1 = 14.101$, $p < .001$).

TABLE H.65. Descriptive statistics for self-assessed writing proficiency for groups by L2 English proficiency.

		L1		L2 English	
	n	x̄	sd	x̄	sd
Adv.	13	5.538	0.967	4.385	1.261
Int.-High	24	5.333	0.963	3.583	0.929
Int.-Low	14	5.000	1.414	4.071	1.269
Low	15	5.733	0.458	3.813	1.047
English NS	60	6.000	0		

A Kruskal-Wallis Test revealed that there is a statistically significant difference among the groups for self-assessed L1 writing proficiency ($\chi^2_4 = 18.170$, $p = .001$) but that there are no significant differences for self-assessed L2 reading proficiency ($\chi^2_3 = 4.236$, $p = .237$). A pairwise two-sided multiple comparison analysis using the Dwass, Steel, Critchlow-Fligner method revealed that the former significant finding emerges from the differences in the proportions of self-assessed writing proficiency between the L2 users and the English NS control group (Adv.: DSCF = 3.810, p = .055; Int.-High: DSCF = 5.458, p = .001; Int.-Low: DSCF = 5.882, p < .001; Low: DSCF = 4.144, p = .028).

TABLE H.66. Descriptive statistics for self-assessed writing proficiency for groups by L1.

		L1		L2 English	
	n	x̄	sd	x̄	sd
Arabic	7	4.857	1.676	4.286	1.113
Chinese	44	5.523	0.876	3.800	1.025
Other	15	5.267	0.961	4.000	1.367
English NS	60	6.000	0		

A Kruskal-Wallis Test revealed that that there is a statistically significant difference among the groups for self-assessed L1 writing proficiency ($\chi^2_3 = 16.004$, $p = .001$) but there are no significant differences for self-assessed L2 reading proficiency ($\chi^2_2 = 1.423$, $p = .491$). A pairwise two-sided multiple comparison analysis using the Dwass, Steel, Critchlow-Fligner method revealed that former significant finding emerges from the differences in the proportions of self-assessed writing proficiency between the L2 users and the English NS control group (Arabic: DSCF = 5.212, p = .001; Chinese: DSCF = 4.797, p = .004; Other: DSCF = 5.673, p < .001).

H.6 LANGUAGE ATTITUDES.

This section captures the same data as the first administration. The collected results are presented here.

TABLE H.67. Descriptive statistics for language attitudes for L2 and native English speakers.

		L1		L2 English	
	n	x̄	sd	x̄	sd
L2 Users	66	5.813	0.404	4.596	0.892
English NS	30	5.950	0.201		

A Wilcoxon Rank Sum Test revealed that there is a statistically significant difference between the groups for overall language attitudes concerning the first language (χ^2_1 = 5.426, p = .020).

TABLE H.68. Descriptive statistics for language attitudes for groups by L2 English proficiency.

		L1		L2 English	
	n	x̄	sd	x̄	sd
Adv.	13	5.865	0.300	4.538	0.713
Int.-High	24	5.736	0.584	4.611	1.037
Int.-Low	14	5.875	0.214	4.893	0.764
Low	15	5.833	0.244	4.359	0.890
English NS	30	5.950	0.201		

A Kruskal-Wallis Test revealed that there are no statistically significant differences among the proficiency groups for overall language attitudes concerning the first language (χ^2_4 = 6.422, p = .170) or for overall language attitudes concerning the second language (χ^2_3 = 3.610, p = .307).

TABLE H.69. Descriptive statistics for language attitudes for groups by L1.

		L1		L2 English	
	n	x̄	sd	x̄	sd
Arabic	7	5.786	0.393	4.750	0.722
Chinese	44	5.881	0.272	4.574	0.977
Other	15	5.628	0.642	4.589	0.734
English NS	30	5.950	0.201		

A Kruskal-Wallis Test revealed that there is a statistically significant difference among the language groups for overall language attitudes concerning the first language ($\chi^2_3 = 9.376$, $p = .025$) but that there are no significant differences for overall language attitudes concerning the second language ($\chi^2_2 = 0.164$, $p = .921$). A pairwise two-sided multiple comparison analysis using the Dwass, Steel, Critchlow-Fligner method revealed that the former significant finding emerges from the differences in the proportions of overall language attitudes concerning the second language between the Other and the English NS group (DSCF = 4.315, $p = .012$).

(i) Feel like myself when speaking…

TABLE H.70. Descriptive statistics for feeling like oneself speaking that language for L2 and native English speakers.

		L1		L2 English	
	n	x̄	sd	x̄	sd
L2 Users	566	5.833	0.514	4.091	1.262
English NS	30	6.000	0		

A Wilcoxon Rank Sum Test revealed that there is a statistically significant difference between the groups for feeling like oneself speaking the L1 ($\chi^2_1 = 3.920$, $p = .048$).

TABLE H.71. Descriptive statistics for feeling like oneself speaking that language for groups by L2 English proficiency.

		L1		L2 English	
	n	x̄	sd	x̄	sd
Adv.	13	5.769	0.599	4.154	1.144
Int.-High	24	5.875	0.612	3.957	1.364
Int.-Low	14	5.857	0.363	4.357	1.393
Low	15	5.800	0.414	4.000	1.155
English NS	30	6.000	0		

A Kruskal-Wallis Test revealed that there are no statistically significant differences among the groups for feeling like oneself speaking the L1 (χ^2_4 = 7.872, p = .131) or for feeling like oneself speaking the L2 (χ^2_3 = 1.434, p = .698).

TABLE H.72. Descriptive statistics for feeling like oneself speaking that language for groups by L1.

		L1		L2 English	
	n	x̄	sd	x̄	sd
Arabic	7	6.000	0	4.143	1.069
Chinese	44	5.909	0.291	4.047	1.290
Other	15	5.533	0.915	4.188	1.328
English NS	30	6.000	0		

A Kruskal-Wallis Test revealed that there is a statistically significant difference among the groups for feeling like oneself speaking the L1 (χ^2_3 = 10.265, p = .016) but that there are no significant differences for feeling like oneself speaking the L2 (χ^2_2 = 0.051, p = .975). A pairwise two-sided multiple comparison analysis using the Dwass, Steel, Critchlow-Fligner method revealed that the former significant finding emerges from the differences in the proportions of feeling like oneself speaking the L1 between the Other and the English NS group (DSCF = 4.139, p = .018).

(ii) Identify with culture that speaks…

TABLE H.73. Descriptive statistics for identifying with the culture that speaks that language for L2 and native English speakers.

		L1		L2 English	
	n	$\bar{x}$	sd	$\bar{x}$	sd
L2 Users	66	5.727	0.692	4.030	1.324
English NS	30	5.800	0.805		

A Wilcoxon Rank Sum Test revealed that there is no statistically significant difference between the groups for identifying with the culture that uses the L1 ($\chi^2_1 = 1.831$, p = .176).

TABLE H.74. Descriptive statistics for identifying with the culture that speaks that language for groups by L2 English proficiency.

		L1		L2 English	
	n	$\bar{x}$	sd	$\bar{x}$	sd
Adv.	13	5.846	0.376	3.538	0.967
Int.-High	24	5.542	0.977	4.348	1.335
Int.-Low	14	5.929	0.267	4.500	1.225
Low	15	5.733	0.594	3.563	1.459
English NS	30	5.800	0.805		

A Kruskal-Wallis Test revealed that there are no statistically significant differences among the groups for identifying with the culture that uses the L1 ($\chi^2_4 = 4.402$, p = .354) or for identifying with the culture that uses the L2 ($\chi^2_3 = 5.659$, p = .129).

TABLE H.75. Descriptive statistics for identifying with the culture that speaks that language for groups by L1.

		L1		L2 English	
	n	$\bar{x}$	sd	$\bar{x}$	sd
Arabic	7	5.429	1.512	3.714	1.799
Chinese	44	5.841	0.428	4.279	1.221
Other	15	5.533	0.743	3.500	1.265
English NS	30	5.800	0.805		

A Kruskal-Wallis Test revealed that there is no statistically significant difference among the groups for identifying with the culture that uses the L1 ($\chi^2_3 = 5.330$, p = .149) or for identifying with the culture that uses the L2 ($\chi^2_2 = 3.508$, p = .173).

(iii) Want to use … like a NS:

TABLE H.76. Descriptive statistics for wanting to use that language like a native speaker for L2 and native English speakers.

		L1		L2 English	
	n	$\bar{x}$	sd	$\bar{x}$	sd
L2 Users	65	5.923	0.322	5.227	1.174
English NS	29	6.000	0		

A Wilcoxon Rank Sum Test revealed that there is no statistically significant difference between the groups for wanting to use the L1 like a native speaker ($\chi^2_1 = 1.843$, p = .175).

TABLE H.77. Descriptive statistics for wanting to use that language like a native speaker for groups by L2 English proficiency.

		L1		L2 English	
	n	$\bar{x}$	sd	$\bar{x}$	sd
Adv.	13	5.923	0.277	5.308	1.109
Int.-High	23	5.870	0.458	5.130	1.424
Int.-Low	14	6.000	0	5.571	0.756
Low	15	5.933	0.258	5.000	1.155
English NS	29	6.000	0		

A Kruskal-Wallis Test revealed that there are no statistically significant differences among the groups for wanting to use the L1 like a native speaker ($\chi^2_4 = 3.592$, p = .464) or for wanting to use the L2 like a native speaker ($\chi^2_3 = 2.367$, p = .500).

TABLE H.78. Descriptive statistics for wanting to use that language like a native speaker for groups by L1.

		L1		L2 English	
	n	$\bar{x}$	sd	$\bar{x}$	sd
Arabic	7	6.000	0	5.286	1.113
Chinese	44	5.932	0.255	5.091	1.309
Other	15	5.857	0.535	5.600	0.632
English NS	29	6.000	0		

A Kruskal-Wallis Test revealed that there is no statistically significant difference among the groups for wanting to use the L1 like a native speaker ($\chi^2_3 = 2.575$, $p = .462$) or for wanting to use the L2 like a native speaker ($\chi^2_2 = 0.991$, $p = .609$).

(iv) Want to be seen as a NS:

TABLE H.79. Descriptive statistics for wanting to be seen as a native speaker of that language for L2 and native English speakers.

		L1		L2 English	
	n	x̄	sd	x̄	sd
L2 Users	66	5.788	0.621	5.075	1.306
English NS	29	6.000	0		

A Wilcoxon Rank Sum Test revealed that there is a statistically significant difference between the groups for wanting to be seen as a native speaker of the L1 ($\chi^2_1 = 3.790$, $p = .052$).

TABLE H.80. Descriptive statistics for wanting to be seen as a native speaker of that language for groups by L2 English proficiency.

		L1		L2 English	
	n	x̄	sd	x̄	sd
Adv.	13	5.923	0.277	5.154	1.345
Int.-High	24	5.708	0.806	5.125	1.191
Int.-Low	14	5.714	0726	5.143	1.460
Low	15	5.867	0.352	4.875	1.408
English NS	29	6.000	0		

A Kruskal-Wallis Test revealed that there are no statistically significant differences among the groups for wanting to be seen as a native speaker of the L1 ($\chi^2_4 = 4.316$, $p = .365$) or for wanting to be seen as a native speaker of the L2 ($\chi^2_3 = 0.681$, $p = .878$).

TABLE H.81. Descriptive statistics for wanting to be seen as a native speaker of that language for groups by L1.

		L1		L2 English	
	n	$\bar{x}$	sd	$\bar{x}$	sd
Arabic	7	5.714	0.756	5.857	0.378
Chinese	44	5.841	0.479	4.932	1.388
Other	15	5.667	0.900	5.125	1.258
English NS	29	6.000	0		

A Kruskal-Wallis Test revealed that there is no statistically significant difference among the groups for wanting to be seen as a native speaker of the L1 ($\chi^2_3 = 3.968$, $p = .265$) or for wanting to be seen as a native speaker of the L2 ($\chi^2_2 = 3.239$, $p = .198$).

APPENDIX I

STATISTICAL INFORMATION – RATING TASK, SECOND ADMINISTRATION

I.1 CURRENT RELEVANCE CONTRASTS, DESCRIPTIVE STATISTICS.

TABLE I.1. Descriptive statistics for CR rating for L2 and native English speakers.

	SubC.	n	x̄	sd	min.	Q1	med.	Q3	max.
L2 Users	E1	84	4.607	1.552	1	3.5	5	6	6
	E2	82	4.110	1.812	1	3	5	6	6
	E3	86	4.733	1.514	1	4	5	6	6
	E4	84	4.393	1.708	1	3	5	6	6
	F1	85	4.447	1.562	1	4	5	6	6
	F2	83	4.217	1.732	1	3	5	6	6
	F3	84	4.226	1.710	1	3	5	6	6
	F4	86	4.384	1.702	1	3	5	6	6
English NS	E1	31	4.258	2.016	1	2	5	6	6
	E2	31	4.516	1.930	1	3	5	6	6
	E3	31	5.645	0.661	4	5	6	6	6
	E4	31	4.839	1.695	1	3	6	6	6
	F1	30	4.867	1.613	1	4	6	6	6
	F2	30	4.100	2.057	1	2	5	6	6
	F3	31	2.968	1.816	1	1	3	4	6
	F4	31	4.290	2.209	1	1	6	6	6

TABLE I.2. Descriptive statistics for CR rating for groups by L2 English proficiency.

	SubC.	n	x̄	sd	min.	Q1	med.	Q3	max.
Adv.	E1	15	4.467	1.885	1	3	5	6	6
	E2	15	3.133	2.031	1	1	3	5	6
	E3	15	4.533	1.846	1	4	5	6	6
	E4	15	4.000	1.890	1	2	4	6	6
	F1	15	4.533	1.685	1	4	5	6	6
	F2	15	3.400	1.957	1	1	4	5	6
	F3	15	2.800	1.612	1	1	2	4	6
	F4	15	3.933	1.831	1	3	5	5	6
Int.-High	E1	28	4.500	1.732	1	3	5	6	6
	E2	28	4.357	1.890	1	3	5	6	6
	E3	29	5.034	1.210	2	4	5	6	6

	E4	27	4.519	1.909	1	3	5	6	6
	F1	29	4.379	1.741	1	4	5	6	6
	F2	28	4.357	1.789	1	3	5	6	6
	F3	28	4.750	1.531	1	4	5	6	6
	F4	28	4.929	1.538	1	4	6	6	6
Int.-Low	E1	16	4.750	1.390	1	4.5	5	6	6
	E2	16	4.250	1.770	1	4	5	5.5	6
	E3	17	4.059	1.952	1	3	5	6	6
	E4	16	4.625	1.746	1	4	5.5	6	6
	F1	16	4.500	1.592	1	4	5	6	6
	F2	16	4.813	1.377	1	4	5	6	6
	F3	16	4.438	1.632	1	4	5	6	6
	F4	17	4.294	1.572	1	3	5	5	6
Low	E1	25	4.720	1.275	2	4	5	6	6
	E2	23	4.348	1.465	1	3	4	6	6
	E3	25	4.960	1.172	2	4	5	6	6
	E4	26	4.346	1.384	2	4	4	6	6
	F1	25	4.440	1.325	1	4	4	6	6
	F2	24	4.167	1.633	1	3	4.5	5.5	6
	F3	25	4.360	1.630	1	3	5	6	6
	F4	26	4.115	1.818	1	3	4.5	6	6
English NS	E1	31	4.258	2.016	1	2	5	6	6
	E2	31	4.516	1.930	1	3	5	6	6
	E3	31	5.645	0.661	4	5	6	6	6
	E4	31	4.839	1.695	1	3	6	6	6
	F1	30	4.867	1.613	1	4	6	6	6
	F2	30	4.100	2.057	1	2	5	6	6
	F3	31	2.968	1.816	1	1	3	4	6
	F4	31	4.290	2.209	1	1	6	6	6

TABLE I.3. Descriptive statistics for CR rating for groups by first language.

	SubC.	n	$\bar{x}$	sd	min.	Q1	med.	Q3	max.
Arabic	E1	E1	11	5.364	0.809	4	5	6	6
	E2	E2	11	3.909	2.386	1	1	5	6
	E3	E3	11	4.909	1.578	1	4	6	6
	E4	E4	11	5.091	1.814	1	5	6	6
	F1	F1	11	4.273	2.240	1	1	6	6
	F2	F2	11	4.182	2.136	1	1	5	6
	F3	F3	11	3.545	1.864	1	1	4	5
	F4	F4	11	5.455	0.688	4	5	6	6
Chinese	E1	E1	51	4.490	1.580	1	3	5	6
	E2	E2	48	4.292	1.675	1	3	5	6
	E3	E3	52	4.596	1.563	1	4	5	6
	E4	E4	52	4.288	1.753	1	3	5	6

	F1	F1	52	4.308	1.553	1	4	4.5	6
	F2	F2	50	4.260	1.614	1	3	5	6
	F3	F3	51	4.529	1.654	1	4	5	6
	F4	F4	53	4.491	1.577	1	3	5	6
Other	E1	E1	22	4.500	1.711	1	4	5	6
	E2	E2	23	3.826	1.825	1	2	4	5
	E3	E3	23	4.957	1.397	1	4	6	6
	E4	E4	21	4.286	1.521	1	4	4	5
	F1	F1	22	4.864	1.125	2	4	5	6
	F2	F2	22	4.136	1.859	1	3	4.5	6
	F3	F3	22	3.864	1.670	1	3	4	5
	F4	F4	22	3.591	2.016	1	2	3.5	6
English NS	E1	31	4.258	2.016	1	2	5	6	6
	E2	31	4.516	1.930	1	3	5	6	6
	E3	31	5.645	0.661	4	5	6	6	6
	E4	31	4.839	1.695	1	3	6	6	6
	F1	30	4.867	1.613	1	4	6	6	6
	F2	30	4.100	2.057	1	2	5	6	6
	F3	31	2.968	1.816	1	1	3	4	6
	F4	31	4.290	2.209	1	1	6	6	6

I.2 Current relevance contrasts, selected inferential statistics.

TABLE I.4. Abbr. Kruskal-Wallis results for CR rating between L2 Users & Eng. NS.

Group	Comparison Group	SubC.	df	χ^2	p
L2 Users	English NS	E3	1	10.198	.001
		E4	1	2.284	.131
		F1	1	2.871	.090
		F3	1	10.218	.001

TABLE I.5. Abbr. DSCF pairwise comparison results for CR rating between L2 proficiency groups.

Group	Comparison Group	SubC.	Wilcoxon Z	DSCF	p
Adv.	Eng. NS	E2	2.339	3.307	.133
	Int.-High	F3	-3.376	4.775	.007
	Int.-Low	F3	-2.538	3.589	.082
	Low	F3	-2.698	3.816	.054
Int.-High	Eng. NS	F3	-3.593	5.082	.003
Int.-Low	Eng. NS	E3	3.378	4.777	.007
		F3	-2.506	3.545	.089
Low	Eng. NS	E3	2.386	3.374	.119
		F3	-2.742	3.878	.048

TABLE I.6. Abbr. DSCF pairwise comparison results for CR rating between L1 groups.

Group	Comparison Group	SubC.	Wilcoxon Z	DSCF	p
Arabic	Other	F4	2.397	3.389	.078
Chinese	Eng. NS	E3	-3.424	4.842	.004
		F3	3.634	5.139	.002

TABLE I.7. Abbr. DSCF pairwise comparison results for CR rating within groups.

Group	SubC.	Wilcoxon Z	DSCF	p
English NS	E1 vs. E3	-2.880	4.073	.077
	E2 vs. F3	3.119	4.411	.038
	E3 vs. F2	3.241	4.584	.026
	E3 vs. F3	5.571	7.879	<.001
	E4 vs. F3	3.803	5.378	.004
	F1 vs. F3	3.878	5.485	.003

TABLE I.8. Abbr. DSCF pairwise comparison results for CR rating within L1 groups.

Group	SubC.	Wilcoxon Z	DSCF	p
Arabic	F3 vs. F4	-2.693	3.808	.125