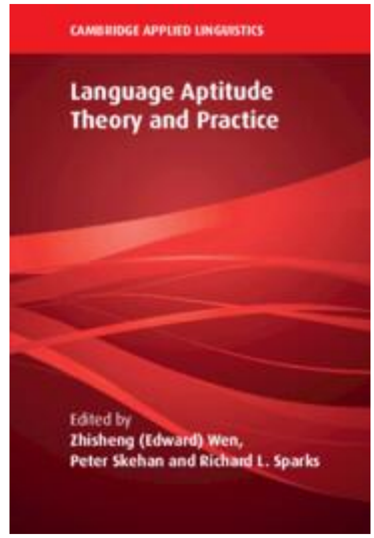
Please cite as:

Suzuki, Y., DeKeyser, R. M., & Huang, H. Y. (2023). Implicit (not explicit) learning aptitude predicts the acquisition of difficult (not easy) structure: A visual-world eye-tracking study. In E. Wen, P. Skehan, & R. Sparks (Eds.), *Language Aptitude Theory and Practice* (pp. 381-413). New York, NY: Cambridge University Press.



Chapter 15

Implicit (not explicit) learning aptitude predicts the acquisition of difficult (not easy) structure: A visual-world eye-tracking study

Yuichi Suzuki

Kanagawa University

Robert DeKeyser and Yi Ting Huang

University of Maryland College Park

Abstract

The current study investigated to what extent two aptitude components, one for explicit and the other for implicit learning, could predict the acquisition of English grammatical structures by late L2 English learners in a naturalistic acquisition context. Sixty-five L2 English learners of Chinese Mandarin, as well as a group of English native speakers performed a visual-world eye-tracking task. In this task, real-time processing of two grammatical properties of the English nominal phrase that differ in terms of L2 psycholinguistic difficulty: (1) definiteness (a difficult structure) and (2) mass-count (an easy structure). Predictors were implicit learning aptitude, measured by the serial-reaction time (SRT) task and explicit learning aptitude measured by subtests of the Modern Language Aptitude Test (MLAT Part 4) and the LLAMA_F test. A majority of L2 learners were unaware of target grammatical structures tested in the visual-world task, which is construed as evidence for the use of linguistic knowledge without awareness (i.e., implicit knowledge). While explicit learning aptitude was not related to either grammatical feature, implicit learning aptitude was related to definiteness, but not to mass-count. These findings suggest that explicit and implicit learning mechanisms are recruited differentially for learning different grammatical properties.

Key words: Cognitive aptitudes, visual-world task, grammatical difficulty, implicit knowledge, explicit and implicit learning, awareness naturalistic adult L2 acquisition

Acknowledgement

This study was supported by IGERT: Biological and Computational Foundations of Language Diversity (NSF DGE-0801465). We would like to express our gratitude to Craig Chambers for sharing experimental stimuli and advice on study design. We also acknowledge Colin Phillips, Robert Slevc, and Jared Novick for their input during the early stages of this project. We are also grateful to Scott Kaufman for allowing us to use the serial-reaction time task and Liu Shuihui for sharing the Chinese version of the MLAT. Finally, we thank Mary Bounds, Brandi Lynch, Alyssa Cameron, Maggie Kinser, Ariel Borten, Lani Alden, and Aliya Candeloro for their assistance with data collection and coding.

Introduction

Researching individual differences in cognitive aptitude has proven to be a useful approach for understanding the relationship between L2 *learning* and resultant *knowledge*. SLA researchers have strived to identify systematic relationships between cognitive aptitude and L2 knowledge to elucidate underlying learning processes (DeKeyser, 2012). Of particular interest to us as SLA researchers is that associations between aptitude and grammatical knowledge have been examined from explicit and implicit perspectives (e.g., Bolibaugh & Foster, 2021; DeKeyser, 2000; Granena, 2013a; Suzuki & DeKeyser, 2015, 2017). Explicit learning refers to conscious learning processes, whereas implicit learning refers to the learning process without intention or awareness (Andringa & Rebuschat, 2015; Hulstijn, 2005).

Recent innovations in two related lines of L2 research—development of tests for cognitive aptitude and for grammar knowledge—have pushed the boundary of our understanding of the interface of explicit and implicit learning and knowledge, which has had a profound impact on L2 theories and practical issues. First, in the last decade, a proposal was made for distinguishing two aptitude components, one for explicit and one for implicit learning. Explicit learning aptitude is defined as cognitive abilities that are important for intentional and rote learning and deliberate hypothesis testing, whereas aptitude for implicit learning refers to cognitive capacity for learning transitional/distributional probabilities of linguistic input without awareness as well as absence of conscious attribution of the resulting knowledge (Granena, 2019; Li & DeKeyser, 2021; Linck et al., 2013).

Second, the validation research of explicit-implicit knowledge tests has offered a set of research tools to measure explicit and implicit knowledge (e.g., R. Ellis, 2005; Suzuki, 2017; see Isbell & Rogers, 2021 for a recent review). Grammatical knowledge is distinguished as explicit or implicit based on whether or not awareness is involved (DeKeyser, 2003; Williams, 2009). Although it is challenging to isolate implicit knowledge from explicit knowledge (access to the latter can also be speeded up), accumulating evidence suggests that using finely-tuned reaction-time (RT) and eye-tracking measures allows for assessing implicit knowledge, at least for L2 learners with sufficient immersion experience (Suzuki, 2017; Suzuki & DeKeyser, 2015; Suzuki et al., 2022; Vafaee, Suzuki, & Kachinske, 2017).

Based on these two emerging lines of research, the current study aims to achieve two goals pertaining to the interface of explicit and implicit learning and knowledge. First, we investigate the roles of explicit and implicit aptitude on the acquisition of grammatical knowledge measured by a visual-world eye-tracking task. Because grammatical knowledge presumably results from a combination of explicit and implicit learning, this cross-sectional study attempts to elucidate underlying learning processes that may be facilitated by aptitudes for explicit and implicit learning. Second, because the roles of aptitude in learning of different types of structures vary considerably (e.g., Robinson, 1997; Yalcin & Spada, 2016), the current study investigates to what extent explicit and implicit learning aptitude would predict two grammatical structures that differ in learning difficulty.

In the remainder of this section, we first provide a focused review on the roles of explicit and implicit learning aptitudes in acquisition of morphosyntax in naturalistic settings—target-L2 speaking countries where extensive L2 exposure is putatively sufficient for (certain levels of) implicit learning. Given the current study's scope, our review concerns late or adult L2 learners with arrival in target-L2 speaking countries after age 12 (Abrahamsson & Hyltenstam, 2008).

The last part of this section discusses the complex interactions between aptitude and types of grammatical structures both in intervention and cross-sectional studies.

Explicit Learning Aptitudes and Explicit Knowledge

Many SLA researchers have investigated the relationship between explicit learning aptitudes and L2 grammar acquisition in naturalistic acquisition settings (Abrahamsson & Hyltenstam, 2008; DeKeyser, 2000; DeKeyser, Alfi-Shabtay, & Ravid, 2010; Granena, 2013b; Granena & Long, 2013). In these studies, explicit learning aptitudes have been measured with different tests across studies (e.g., Modern Language Aptitude Test [MLAT], LLAMA tests), but all the tests require the ability to consciously reflect on linguistic aspects of languages. For the assessment of grammatical knowledge, grammaticality judgement tests (GJTs) have been commonly used across the studies with different task parameters (e.g., presence of time pressure).

The earlier three studies (Abrahamsson & Hyltenstam, 2008; DeKeyser, 2000; DeKeyser et al., 2010) have consistently found a weak to moderate positive correlation for post-puberty learners between explicit aptitudes and the grammatical knowledge, which was measured by *untimed* GJTs ($.33 < r < .53$). GJTs in DeKeyser's and Abrahamsson and Hyltenstam's studies were *untimed* or "off-line" in which participants were given enough time to allow for accessing explicit knowledge. These findings lend support for the reliance on explicit learning processes for the acquisition of explicit knowledge. A subsequent study by Granena and Long (2013a) failed to find any effects of explicit aptitudes on the acquisition of L2 morphosyntactic knowledge, which was measured with a *time-pressured* GJT. The lack of relationship appears to be due to the differences in how the GJTs were administered. Granena and Long's study told participants to make a grammatical judgment as quickly as they could, and no pause was inserted between the test items, which made the task more conducive to draw on automatic processes. Although the role of (explicit) aptitude may be more important for explicit, not implicit knowledge, which is arguably influenced by conditions under which GJTs were administered (Granena, 2013b), it remains inconclusive whether linguistic knowledge, measured by the time-pressured GJTs, was explicit or implicit (e.g., Godfroid et al., 2015; Vafaee et al., 2017). It is thus still an open question to what extent explicit aptitude influences the acquisition of implicit knowledge. In order to scrutinize the role of explicit aptitude, more fine-grained measures that tap into implicit knowledge should be employed, which will be discussed next.

Implicit Learning Aptitude and Implicit Knowledge

There is a growing interest in implicit learning aptitude in SLA (Granena, 2019; Li & DeKeyser, 2021; Linck et al., 2013). Despite the incipient nature of our understanding of this new construct of aptitude, researchers have started to find some systematic relationship between implicit learning aptitude and L2 grammar acquisition in naturalistic settings. Most relevant to the present study are two studies (Granena, 2013a; Suzuki & DeKeyser, 2015) that used the same set of implicit aptitude and language tests. In these studies, implicit learning aptitude was measured with a serial-reaction time (SRT) task (Kaufman et al., 2010), which measures the ability of probabilistic sequence learning without awareness. Grammatical knowledge was measured by a reaction-time (RT) task called the word-monitoring task. The word-monitoring task is claimed to be a measure of implicit knowledge because it can capture real-time grammar processing while learners' attention is directed to meaning (Godfroid, 2016; Suzuki, 2017; Suzuki et al., 2022; Vafaee et al., 2017). As illustrated in Figure 1, participants listen to a sentence which includes a monitoring word, to which participants need to respond by pressing a button as soon as they hear it. When they can detect the grammatical error (i.e., a past-tense error "*purchase"* in Sentence (b)), which occurs immediately before the monitoring word, their RT to the monitored word is expected to slow down in relative to the grammatical sentence (a). The RT difference between the ungrammatical and grammatical items thus indexes the online sensitivity to grammatical errors during sentence comprehension. Importantly, participants are told to answer a yes/no comprehension question following the stimulus sentence presentation, in order to focus participants' attention on meaning, i.e., minimize the conscious attention directed to search for grammatical errors.

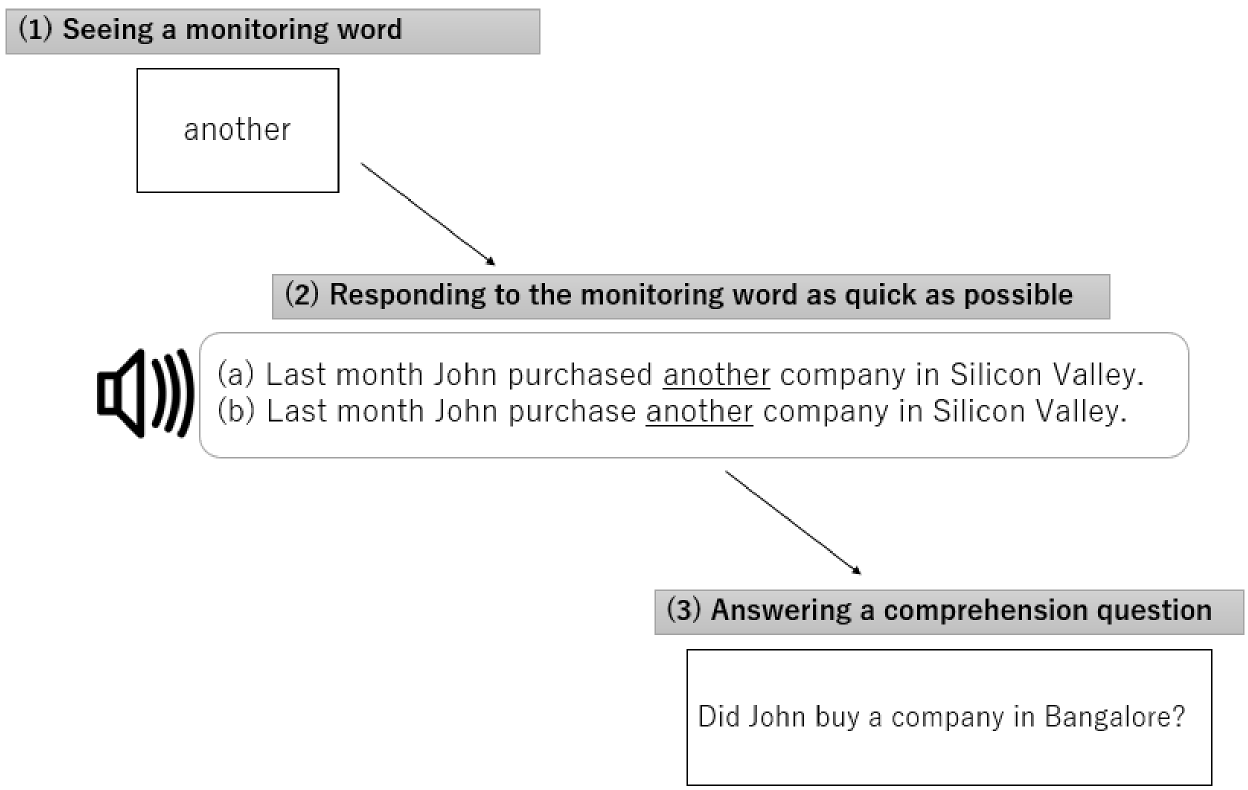


Figure 1. Word-monitoring task used in Granena (2013a) and Suzuki and DeKeyser (2015)

In the aforementioned two studies, significant positive correlations were found between the SRT score (implicit aptitude) and the word-monitoring task performance (implicit knowledge) among L1 Chinese speakers who had lived in Japan at least for 2 years and a half (Suzuki & DeKeyser, 2015) and L1 Chinese speakers who lived in Spain for a minimum of 5 years (Granena, 2013a). These findings provide initial converging evidence regarding the role of implicit learning aptitude on L2 grammar acquisition of adult learners (cf. Suzuki & DeKeyser, 2017, who found more limited effects of implicit aptitude when a link between explicit knowledge and aptitude has been established in the statistical model). However, these previous studies used RT data from both aptitude and grammar tests. The positive correlation found between the aptitude and grammar test scores could at least in part be an artifact of the shared metrics (i.e., RT) between the two tasks.

In order to overcome this potential limitation, a visual-world eye-tracking task is employed in the current study to assess L2 grammatical knowledge. In this task, participants' eye-movement toward a display with several possible referents during aural comprehension is analyzed to examine real-time grammar processing. The sensitivity to grammatical manipulations (index of grammatical knowledge) is captured using eye-movement data, and an underlying association between knowledge and aptitude can be explored more rigorously than the relation between the RT measures. More importantly, the visual-world task is very unlikely to raise awareness of target structures. As a case in point, the post-task debriefing results have shown that participants remained unaware of the real purpose or target linguistic structure of a visual-world task (Andringa & Curcic, 2015; Dussias, Valdés Kroff, Guzzardo Tamargo, & Gerfen, 2013). The finely-tuned eye-tracking technique is arguably a promising test of L2 implicit knowledge because it is a more direct measure of real-time grammar processing (also characterized as fast and ballistic) than RT tasks (Suzuki, 2017). Despite its task constraints, a variety of grammatical structures such as gender agreement, case-markers, tense-aspect, and pronouns have been tested in this visual-world paradigm (Godfroid, 2019). In the current study, a visual-world task was designed to test two grammatical properties of the English nominal phrase—definiteness mass-count distinctions.

Explicit and Implicit Learning Aptitudes and Grammatical Difficulty

To what extent a certain grammatical structure is amenable to explicit and implicit learning is presumably influenced by characteristics of that structure and the specific problems they pose for a given learner. According to Krashen (1982), explicit learning is conducive to easy structures, whereas implicit learning favors difficult structures. The definition of easy and difficult structures—grammatical difficulty—is elusive and most likely multi-faceted (DeKeyser, 2005). Grammatical difficulty should be conjointly determined by taking into account not only linguistic factors (e.g., abstractness of form-meaning mapping), but also contextual (e.g., predisposition of intentional or incidental learning) and learner-related (e.g., aptitude) factors (DeKeyser, 2016; Housen & Simoens, 2016).

Findings from instructed SLA research indicate that the role of explicit learning aptitudes is moderated by grammatical difficulty of target structures (e.g., Yalcin & Spada, 2016; Robinson, 1997). Robinson (1997) examined the role of aptitudes on the acquisition of L2 English syntactic rules under different types of training conditions (implicit, incidental, rule-search, and inductive). He trained intermediate ESL learners in the university language programs on an easy (a subject-verb inversion rule, e.g., Into the house John ran.) and a difficult English syntactic rule (a pseudo-clefts of location, e.g., Where Mary and John live is in Chicago not in New York.). The difficulty of the structures was determined based on expert judgements by ESL teachers. The explicit learning aptitude score (MLAT Part 4 [Words in Sentences]) was related to the learning of easy structures, not to that of difficult structures, measured by the GJT, in the rule-search (explicit-inductive) condition. In contrast, a recent EFL classroom research by Yalçın and Spada (2016) showed the opposite pattern of findings. Turkish L2 English learners received explicit, form-focused instruction on an easy (past progressive) and a difficult (passive) structure. The difficulty of the structures was determined according to various linguistic criteria such as transparency of form-meaning mapping, frequency, saliency and perceived difficulty by learners. Explicit learning aptitude (the LLAMA_F test) turned out to be a significant predictor of the acquisition of the difficult structure, not for the easy structure, as assessed by the GJT.

The findings of Robinson (1997) and Yalçın and Spada (2016) may not be as contradictory as they seem, because the effect of aptitudes could have become significant only for the linguistic structures that imposed a *moderate* level of difficulty for their targeted participants (DeKeyser, 2016). In Robinson (1997), the pseudo-clefts of location were *too difficult* for the intermediate ESL learners, whereas in Yalçın and Spada (2016), the past progressive structure was *too easy* for the Turkish learners (hence, lack of significant aptitude effects for either cases). In contrast, higher aptitude could have facilitated the learning of the subject-verb inversion rule (easier than the "too difficult" structure) and the passive structure (harder than the "too easy" structure), respectively.

While these short-term intervention studies examined the effects of explicit learning aptitude in instructed settings, the role of implicit learning aptitude for different types of grammatical structures was examined in the aforementioned naturalistic acquisition study by Granena (2013a). The word-monitoring task measured the real-time sensitivity to six grammatical structures categorized into two types: the agreement structures (noun–adjective gender agreement, subject–verb agreement, and noun–adjective number agreement) and the non-agreement structures (subjunctive mood, perfective/imperfective aspect, and passives with ser/estar). These two categories were based on a developmental stage in L1 acquisition. The first type of structures is mastered by age 3 in the L1 acquisition of Spanish, whereas the latter is not fully acquired until age 7 or later in L1 Spanish. The results showed that implicit learning aptitude, measured by the SRT task, significantly predicted the acquisition of the agreement structures only, not for that of the non-agreement structures.

According to Granena's interpretation, the implicit learning aptitude might have compensated for the lack of inflectional morphology in L1 Chinese in order to acquire the Spanish agreement structures involving a rich inflectional paradigm. Consistent with prior research on explicit learning aptitude in instructed settings (Robinson and Yalçın & Spada), the effects of implicit learning aptitude should also be examined by taking the linguistic difficulty for a given group of L2 learners into account. In order to understand explicit and implicit learning processes, it is of high importance to study the complex interaction between explicit-implicit aptitudes and linguistic difficulty.

The Current Study

In order to explore the relationship between explicit-implicit learning and grammatical knowledge, the current study aimed at clarifying the contribution of explicit and implicit learning aptitude for the acquisition of L2 implicit knowledge by adult L2 English speakers with L1 Chinese who lived in the United States. The visual-world task was employed to assess L2 real-time grammatical processing as an indicator of implicit knowledge of definiteness and the mass-count distinction. Two distinctions within the English NP (definiteness and the mass-count distinction) are chosen because they are interesting cases to examine whether the contribution of aptitudes is moderated by grammatical difficulty. Definiteness is distinguished by *the* and *a*, and pragmatically guided by discourse information. Definite descriptions like "*the can*" are used when referents are uniquely identifiable, while indefinite descriptions like "*a can*" are used when multiple referents are possible (Lyons, 1999). English uses a numerical counter for a count noun like "*two candles*" and a quantifier for a mass noun like "*two pieces of bacon*" to distinguish between quantification of referents with rigid boundaries from ones without them. It is ungrammatical to use a numeral counter with a mass noun (e.g., two bacon).

From a psycholinguistic perspective, definiteness is essentially much more challenging for acquisition than the mass-count distinction for the current L2 learners. First, the two systems are different in terms of the degree of transfer effects from L1. An analogous distinction of definiteness does not exist in L1 Mandarin Chinese, whereas the mass-count distinction bears resemblance to the Chinese classifier systems (Cheng & Sybesma, 1998, 1999). Based on similar structures in L1, L2 learners can selectively attend to relevant features within the L2 input (N. C. Ellis et al., 2012), which may facilitate the acquisition of the mass-count distinction. Second, the two systems differ in complexity of meaning. Definiteness is a more abstract concept than the mass-count distinction; learners have to learn how to map the definite and indefinite articles to variable discourse semantics (e.g., identifiability of referents in context). On the other hand, the mass vs. count status is based on rigid boundaries of referents, although the distinction can sometimes be arbitrary for L2 learners. These factors are related to each other and all contribute to the different level of difficulty in their acquisition (Graus & Coppen, 2015; Housen, 2014; Yalçın & Spada, 2016).

Empirical evidence supports the higher difficulty of definiteness over mass-count. Hua and Lee (2005) showed that L2 English learners with L1 Chinese (i.e., third-year Chinese college students) often correctly rejected descriptions like “ten beef” and “three rice.” Their GJT performance was indistinguishable from native English speakers. On the other hand, definiteness is notoriously difficult, especially for L2 learners whose L1 does not make the distinction (Ionin, Zubizarreta, & Philippov, 2009; Robertson, 2000; Snape, 2008; Trenkic, 2008). Although some types of article errors seem to decrease in frequency as proficiency increases (Lu, 2001; Trenkic, 2002; Young, 1996), non-target like use appears to persist for an extended period of time under some circumstances (Lardiere, 2007; White, 2003). Critically, the current sample of L2 English learners with L1 Chinese residing in the United States were able to make the mass-count distinction, whereas they had substantial difficulty in distinguishing the definiteness (see the Preliminary Analysis of Visual-World Task section).

Explicit and implicit learning aptitudes were used as predictors for the acquisition of these two grammatical structures. Explicit learning aptitude, operationalized as language-analytic ability, was measured with two aptitude tests that are characteristically involved in explicit learning: the “Words in Sentences” of MLAT, or MLAT_4 (Carroll & Sapon, 1959) for assessing grammatical sensitivity, and the LLAMA_F (Meara, 2005) for assessing inductive language learning ability. Implicit learning aptitude was measured by the probabilistic SRT task (Kaufman et al., 2010). The current study addresses the following research questions (RQs):

1. To what extent do explicit and implicit learning aptitude contribute to the attainment of real-time grammar processing?
2. Do the contributions of explicit and implicit aptitude vary by the L2 structures (definiteness and count-mass distinction)?

With regard to RQ1, we hypothesize that implicit, not explicit, learning aptitude will be a significant predictor for real-time grammar processing ability, measured by the visual-world task. The second RQ further explores whether the effects of aptitudes change depending on linguistic structures. We hypothesize that implicit learning aptitude may play a facilitative role in the acquisition of definiteness because of the complexity of mastering the article system. Due to the lack of similar L1 system to definiteness for Chinese speakers in the current study, learners may need to rely on implicit learning mechanisms to successfully acquire the definiteness distinction from scratch, leading to a systematic relationship between aptitude for implicit learning and acquisition of this distinction. For the mass-count distinction, we predict that there will be no systematic contributions of implicit learning aptitude because the learning difficulty is less burdensome. With regard to the role of explicit learning aptitude, we leave the question open for the context of this study—naturalistic L2 acquisition—given the scarce research findings even in instructed settings (Robinson, 1997; Yalçın & Spada, 2016).

Methods

Participants

Sixty-five English L2 learners with L1 Chinese participated in the current study. Most of them were university students (15 undergraduate, 26 master, and 21 doctoral students), and three participants were employees (two participants with MA degree and one with PhD degree). Since all participants had received or were receiving college education entirely in English, their proficiency is deemed advanced. They all arrived in an English-speaking country after the age of 15; the mean age of arrival was 21.29 (*SD* = 3.41, range 15-28 years). They had received EFL/ESL instruction for an average of 11.66 years (SD = 3.43, range 5-20 years). The mean length of residence in English-speaking countries was 46.22 months (*SD* = 29.67, range: 1-121 months). Twenty-eight native speakers (NSs) were also recruited to ensure that the visual-world task would work as expected.

Instruments

Visual-world Task: Definiteness. Figure 2 (Panel A) illustrates that displays for the definiteness trials consisted of two possible locations, a distractor location, and a theme (cf., Chambers, Tanenhaus, Eberhard, Filip, & Carlson, 2002; Trenkic, Mirkovic, & Altmann, 2014). Possible locations involved two same-category locations (e.g., big and small cans), while the distractor location involved a different kind of location (e.g., a bowl). Sentences involved pairs of instructions: "*Pick up the pig. Now put it inside* ____." Two critical trial types were created by manipulating the definiteness of the article as a within-subject factor: definite (e.g., "*the can*") and indefinite trials (e.g., "*a can*"). The indefinite description (inside *a* can) matched the display where multiple goal locations (two cans) were available, whereas the definite description (inside *the* can) was pragmatically inappropriate because the display does not include a single uniquely identifiable goal. If participants were sensitive to the definiteness distinction, the reference resolution would be facilitated (resulting in faster convergence of eye-movement to a goal location) when hearing an indefinite description compared to a definite description. Sixteen additional filler trials were also included so that participants could not predict the goal location prior to the instruction (see Appendix A for a list of stimuli).

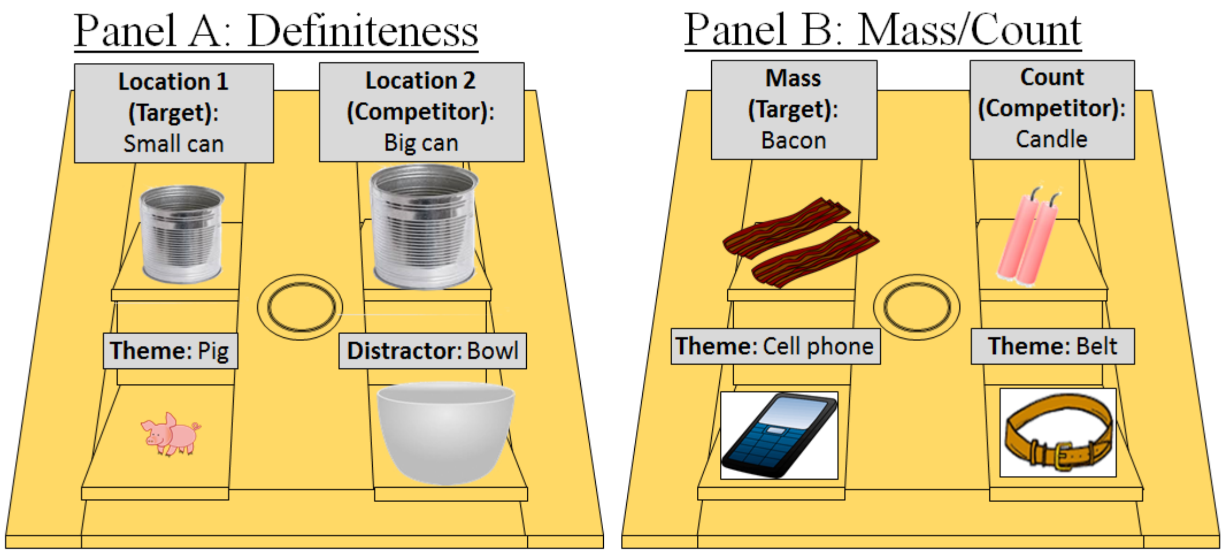


Figure 2. Sample displays of the visual-world task for definiteness condition (panel A) and mass-count condition (panel B).

Visual-world Task: Mass-count. Figure 2 (Panel B) illustrates that displays for the mass-count trials consisted of two possible locations and two themes. Themes involved pairs of printed pictures of count nouns (e.g., belt and cell phone), while the two possible locations involved pairs of three-dimensional miniatures of count nouns (e.g., candles) and mass nouns (e.g., bacon). Sentences involved pairs of instructions: "*Pick up the belt. Now put it on top of* ______." Two critical trial types were created by manipulating the referring expressions as a within-subject factor: quantifier (e.g., "*some bacon*") and numeral trials (e.g., "*the two bacon*"). Since the numeral description was ungrammatical (e.g., two bacon), the numeral "*two*" was expected to mislead participants to look for the count noun (e.g., two candles). On the other hand, the quantifier "*some*" can follow either a count or a mass noun. The referent resolution (choosing the mass noun) was thus expected to be facilitated in the quantifier trials than in the numeral trials. If participants were sensitive to the mass-count distinction, eye-movements should converge to the mass noun faster after hearing a quantified expression than a numerical expression. Eight additional fillers trials asked participants to put the theme on top of the count (competitor) noun (e.g.., "*Now put it on top of some/the two candles*") so that participants could not predict the goal location prior to the instruction (see Appendix B for a list of stimuli). While the modifier again varied across trials, it was always grammatically correct. The count and mass nouns were chosen for this study based on a norming study (see Appendix C).

Visual-world Task: Procedure. Participants were seated in front of a podium consisting of four shelves. An object was placed on each quadrant, and a camera was placed at the center of the display to record participants' eye movements on these objects. A second camera was located behind the participants and recorded their actions following the sentences. Each trial began with the experimenter taking four objects out of a bag and labeling them using a bare singular form (e.g., "*bacon*," "*candle*," "*cell phone*," '*belt*"). Each object was placed on a dedicated quadrant on the podium. The experimenter then played a pair of audio-recorded sentences that instructed participants to move an object in/on the container/object in one of the quadrants to another quadrant. All sentences were pre-recorded by a female native speaker of English.

A total of 40 trials were presented, consisting of 16 critical trials (four definite trials, four indefinite trials, four quantifier trials, and four numeral trials) and 24 filler trials. All trials were presented in semi-randomized order such that the same trial type never occurred more than twice in a row. Multiple versions of items in the critical trials were distributed across four counterbalanced lists, constructed to control for the number of items in each condition within subjects and vary the condition for each item between subjects. Within each counterbalancing list, the placement of each object type was rotated through the four quadrants on the podium, such that goal locations/objects could not be predicted based on their location.

Visual-world Task: Debriefing. The experimenter asked the participants two questions immediately after the visual-world task was completed. The first question asked about their general impression of performing the task ("How was the task?"). The second question specifically aimed at the participants' awareness of the target grammatical structures embedded in the task ("Did you notice anything weird in the sentences you heard?"). When a participant reported that they noticed infelicitous/ungrammatical features in the sentences, we interpreted this as indicating they became aware of the target structures either during or after the visual-world task. This debriefing procedure was only conducted for L2 learners.

SRT task. The SRT task was administered to measure the domain-general ability of learning sequences without awareness (i.e., implicit learning aptitude). In the current study, the probabilistic SRT task was adopted from Kaufman et al. (2010). In the SRT task, participants saw a dot appearing at one of four locations on the computer screen and responded to it as quickly and accurately as possible by pressing the corresponding key. Their reaction times (RTs) were recorded for each response. Unbeknownst to the participants, the sequence of stimuli was generated by a probabilistic rule: 85% of the sequences followed the rule (probable, training condition), whereas the other 15% of the sequences were generated by another rule (improbable, the control condition). This probabilistic nature of the SRT task made it difficult to learn the sequence explicitly. There were eight blocks, and each block consisted of 120 trials, 960 trials in total. The scoring method devised in Kaufman et al. (2010) was used in the current study, resulting in the score ranging from 0 to a maximum of 6 (see Appendix D for details). The internal consistency, indexed by Cronbach alpha, was .46. Although this reliability index is considered very low for more traditional cognitive tests such as working memory task, it was deemed acceptable for the implicit learning task in both psychology and SLA (Granena, 2013a; Kaufman et al., 2010; Reber, Walkenfeld, & Hernstadt, 1991; Suzuki & DeKeyser, 2015).

In order to provide some evidence that the knowledge developed and assessed on the task was implicit, a surprise recognition test was also administered immediately after the SRT task was completed. This test assessed whether participants became aware of the sequence patterns in the SRT task, i.e., whether they developed explicit knowledge about the sequence (Suzuki, 2015 for details). The results from the recognition task confirmed that knowledge the participants developed through the SRT task was implicit (see Appendix E).

LLAMA F. LLAMA_F (Meara, 2005) was administered to measure inductive language learning ability. Inductive language learning ability refers to the capacity to induce rules governing a given foreign language with conscious effort (Carroll, 1991). This test consisted of a learning phase and a test phase. In the learning phase, participants were given five minutes to learn a new language by seeing sentences matched with pictures. In the testing phase, the program displayed a picture and two sentences, one grammatical and the other ungrammatical, and the learners' task was to choose the grammatical sentence. In order to increase reliability, ten test items were added to the original 20 test items; the total number of items was 30 (see Suzuki & DeKeyser, 2017 for details). To reduce the skewness, the LLAMA F score was square-root-transformed for statistical analyses. The internal consistency, indexed by Cronbach's alpha, was .66.

Chinese MLAT Part 4. The Chinese version of MLAT_4, "Words in Sentences," was administered to measure grammatical sensitivity. Grammatical sensitivity is the awareness of the syntactic patterns and grammatical functions of sentences in the test-takers' L1. This Chinese-version MLAT was validated and found to be a significant predictor for English achievement test scores (Xia, 2011). The participants were asked to identify the parts of speech in the sentences in L1. They were first presented with the key sentence in which one word is underlined and bolded, and their task was to select the word in the second sentence that had the same grammatical function as the underlined word in the key sentence. The total number of items was 20. To reduce the skewness, the MLAT_4 score was log-transformed for statistical analyses. The internal consistency, indexed by Cronbach's alpha, was .59.

Procedure

All the L2 learners participated in two individual sessions in a quiet laboratory. They took the visual-world task in the first session. In the second session, which took place on a separate day (usually after one week), they were administered the following tasks in fixed order: SRT task, LLAMA F, and MLAT_4.

Data Coding of Visual-World Task

In the visual-world task, participants' actions for the critical trials were coded by the first author and five trained research assistants. Action accuracy (whether participants put correct themes in correct locations) was near ceiling across all conditions and groups (> 99%). Eye-movement data from the incorrect trials in which participants misidentified themes or locations were excluded, which only accounted for 0.4% (NS) and 0.6% (L2) of all trials in the definiteness condition and 0% (NS) and 0.2% (L2) of those in the mass-count condition.

The same individuals also coded participants' eye-movements using the frame-by-frame annotation software Vcode (Hagedorn, Hailpern, & Karahalios, 2008). All coders were blind to the condition of the trial. For every frame, eye-movements were coded as fixations on one of the quadrants (upper-left, upper-right, lower-left, lower-right), center of the display, or missing due to blinks, looks outside of the podium, or track loss. Missing frames accounted for 2.5% of the data in the NS group and 9.4% in the L2 learners. For all remaining frames, fixation locations were then recoded based on the displayed object. We were primarily interested in looks to two possible locations, coded as Targets and Competitors. In the critical definiteness trials where a little theme was included, a Target was defined as a location that matched the size of the theme (e.g., small can), while a Competitor was defined as a non-size-matching location (e.g., big can). This is because when coders scored which target location participants selected, results revealed a strong tendency for participants to match the size of the theme with the size of the location (e.g., small pig placed in the small can). Matching responses exceeded chance for both NS (M = 77.2%, SD = 19.6%, $t(27) = 1331.79$, $p < .001$) and L2 learners (M = 68.6%, SD = 20.6%, $t(64)$

= 1933.86, $p < .001$).[1] In the mass-count trials, Targets were defined as mentioned locations (i.e., mass nouns) while Competitors were defined as non-mentioned locations (i.e., count nouns) because mass nouns were the goal locations in both quantifier and numeral trials.

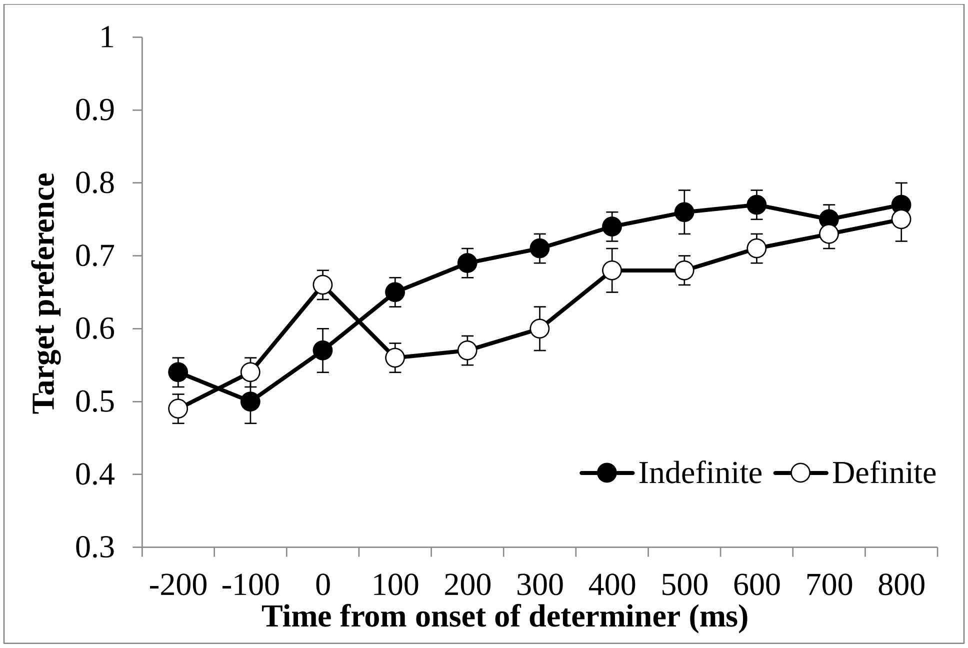
Preliminary Analysis of Visual-World Task

Since we did not have specific hypotheses about when the effects of linguistic processing would emerge in eye-movements, we focused on eye-movements over an extended period, beginning from 200 ms before the onset of the determiner until 1000 ms later. For each linguistic cue (i.e., a/the and two/some), the window was shifted 200 ms after the linguistic cue in the speech stream to account for the time it takes to generate a saccadic eye-movement (Matin, Shao, & Boff, 1993). Within each window, our dependent measure was the preference for the Target over the Competitor. This was calculated as the number of samples (for a given trial and window) in which participants looked at the Target minus the number of samples in which they looked at the Competitor. If this number was positive, Target preference was 1. If it was negative, then Target preference was 0. If participants looked at neither object, or at both objects equally, this sample was excluded from the analysis.

***Definiteness*.** Figure 3 illustrates that as expected, NS participants were quicker to look at the Target following indefinite descriptions compared to definite descriptions (Panel A). This led to a Target advantage that began approximately 100 ms after determiner onset and continued until the 600 ms window. In contrast, L2 learners showed the opposite pattern. Target preference was unexpectedly greater in the definite trial compared to the indefinite trial (Panel B). This pattern, opposite from what was seen in the NS group, began approximately 200 ms before determiner onset and continued until the 400 ms window, possibly due to the co-articulation in the target phrase, "*inside a/the*".

[1] We also conducted the analyses in which the target location was coded based on the actual goal location that participants placed the theme, which is the same procedure used in the previous studies (Chambers et al., 2002). Their coding resulted in the same pattern as the current coding, but the effect was smaller for the NS group.

(Panel A) Native Speakers



(Panel B) L2 learners

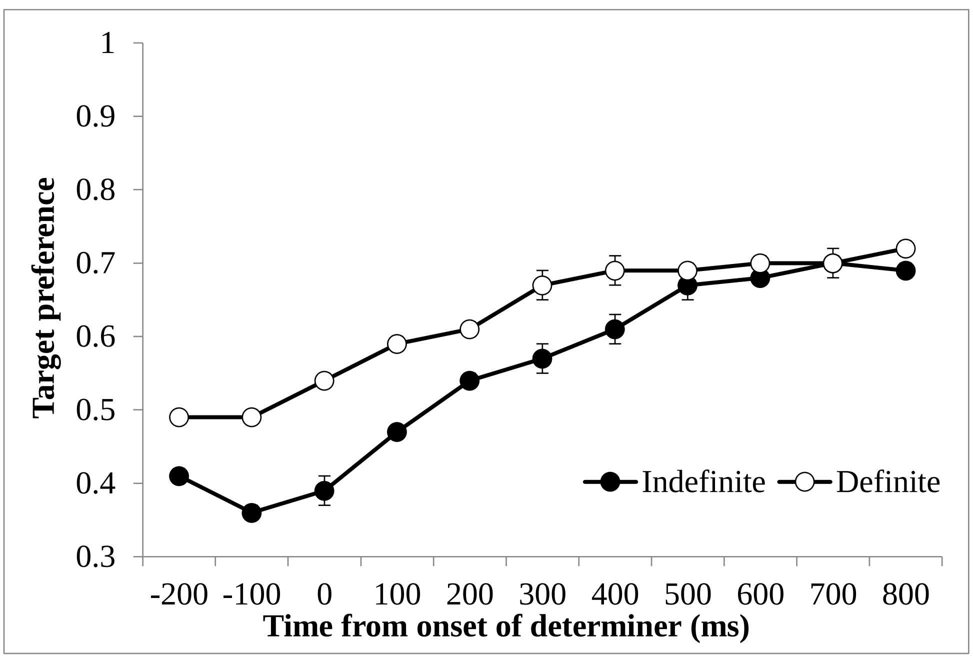


Figure 3. Time-course of target preference for the definiteness condition

Note. The figures illustrate the time windows that are shifted 200 ms after the linguistic cues. The mean length of the determiner and noun was 140 ms and 573 ms for definite trials and 171 ms and 575 ms for indefinite trials.

As Target preference successfully captured the sensitivity of definiteness by the native English speakers, we used Target preference to compute a “sensitivity index” for each L2 learner by subtracting Target preference in the definite trials from that in the indefinite trials. The greater index indicates higher sensitivity to the distinction. The index was calculated based on the critical time region in which NSs consistently showed the sensitivity (i.e., 100-600 ms). We used NSs’ critical region as a reference for computing L2 sensitivity index because the study aimed to assess to what extent L2 learners’ linguistic processing ability is qualitatively similar to NSs’ ability (implicit knowledge).

***Mass-Count*.** Figure 4 illustrates that as expected, the NS group was quicker to look at the Target following the quantifier compared to the numeral (Panel A). This led to a Target advantage that began immediately after quantifier onset and continued until the 600 ms window. Similar patterns were also observed for the L2 learners. The same Target advantage for quantified expressions was observed from the -200 ms window to the 800 ms window (Panel B).

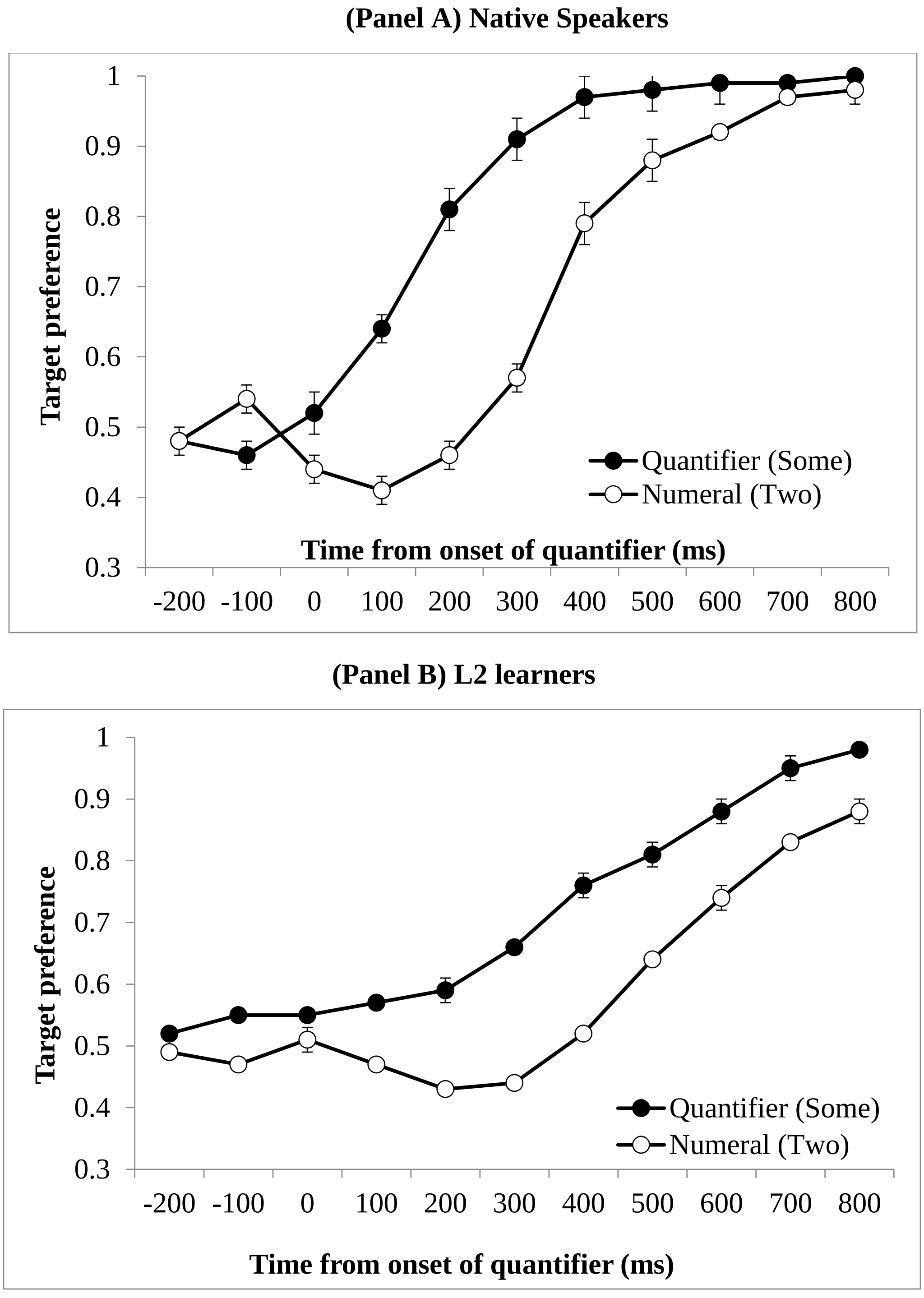


Figure 4. Time-course of target preference for the mass-count condition

Note. The figures illustrate the time windows that are shifted 200 ms after the linguistic cues. The mean length of the modifier and noun was 281 ms and 642 ms for quantifier trials and 435 ms and 679 ms for numeral trials.

Similar to the definiteness condition, we also computed a "sensitivity index" for each L2 learner by subtracting the Target preference in the numeral trials from that in the quantifier trials. The index was calculated across a critical region that began 100 ms from the onset of quantifier to the 600 ms window so that we could fairly compare the sensitivity to definiteness.[2]

Results

Debriefing of Visual-world Task

In answer to the first debriefing question (How was the task?), most participants reported "interesting", "fun", "easy", and "simple." No participants mentioned anything about ungrammatical/infelicitous aspects of stimulus sentences. In response to the more specific second question (Did you notice anything weird in the sentences you heard?), three participants reported that they found the determiner of the noun was wrong (e.g., "the sentence said the cup, but there were two cups), and eight participants reported that they noticed the ungrammaticality of "numeral + mass noun" (e.g., "two meat should be two pieces of meat"). These "aware" participants were excluded from further analysis in order to ensure that the visual-world task never triggered noticing of the target structures. Additionally, we excluded two participants due to the experimenter's errors in procedure. Note that excluding these participants did not change the pattern of the current findings. Results including all 65 participants before excluding the aforementioned participants are presented in Appendix F.

Descriptive Statistics for Dependent and Independent Variables

Descriptive statistics for the eye-tracking measures and the aptitude tests are presented in Table 1. The eye-tracking measure for both structures was computed by summing up the *z* scores of "sensitivity index" (see Preliminary Analysis of Visual-World Task) for definiteness and mass-count structures in order to equally weight the indices for the two structures (see Granena, 2013 and Suzuki and DeKeyser, 2015 for a similar approach). This composite score indicates online sensitivity for the two systems altogether. Although we predicted that the two structures would involve different learning problems, we were also interested in whether a more broadly defined ability for real-time grammar processing can index implicit knowledge. As L2 learners were sensitive to the mass-count distinction, the mean eye-tracking score for mass-count is positive (i.e., a higher target preference in the numeral trials than that in the quantifier trials). The score for the definiteness is negative as the participants showed sensitivity in the opposite direction to NS participants' (i.e., a higher target preference in the indefinite trials than that in the definite trials). Despite the difference in the means, distributions were almost identical between definiteness and mass-count conditions (*SD* = 0.29 and 0.30, respectively). According to the Kolmogorov-Smirnov test, the sensitivity indices for definiteness and mass-count conditions were normally distributed ($ps > .1$), but the sensitivity index for the two structures combined was not ($p = .01$). Square-root transformation was applied to reduce the skewness for that sensitivity index.

[2] The critical region that was more consistent with the empirical data for mass-count distinction was from 0 ms to 600 ms. However, the region from 100 ms to 600 ms was chosen in order to directly compare with the sensitivity to definiteness. We also confirmed that the same pattern of results was borne out in the analysis based on the critical region derived from the empirical data (i.e., 0 -600 ms).

Table 1.

Descriptive Statistics for Language and Aptitude Measures for L2 learners

		N	*M*	*SD*	Min	Max	Possible Max
Dependent V.	Def. + Count/Mass	55	0.14	1.46	-4.20	3.34	-
	Definiteness	60	-0.06	0.30	-0.73	0.73	1
	Count/Mass	55	0.20	0.31	-0.47	0.82	1
Independent V.	SRT	65	2.14	1.46	0	6	6
	MLAT_4	65	14.77	2.66	7	19	20
	LLAMAF	65	24.88	3.24	17	30	30

Relationship of Eye-Movement Data with Explicit and Implicit Learning Aptitudes

In order to examine to what extent explicit and implicit aptitudes predict real-time grammar processing, correlations and multiple regression analyses were conducted. The dependent variables were eye-tracking scores for definiteness and the mass-count distinction separately and their combined score (square-root transformed). The independent variables were SRT scores, LLAMA F scores (square-root transformed), and MLAT_4 scores (log transformed).

Table 2 presents correlations between the three eye-tracking scores and the three aptitude scores. Significant correlations with the eye-tracking scores for both structures (definiteness + mass-count) and definiteness only (r = .30 and .35, p < .05) were found only for the SRT scores. The eye-tracking score for the mass-count distinction was not significantly correlated with any of the aptitude test scores; the largest magnitude was .19 with MLAT_4.

Table 2.

Correlation Coefficients (p values) Between Eye-tracking Scores and Aptitude Scores

	SRT	LLAMA F	MLAT_4
Combined	.30*	.10	.10
	(.03)	(.48)	(.48)
Definiteness	.35**	.03	-.04
	(.01)	(.85)	(.79)
Mass-count	.07	.04	.19
	(.63)	(.77)	(.16)

* p < .05, ** p < .01.

Note. The SRT score was not correlated with either of the LLAMA F (r = - .12, p = 34) or MLAT_4 scores (r = .01, p = 92). The scores from the two explicit learning aptitude tests were significantly correlated (r = .27, p = .03), suggesting that grammatical sensitivity and inductive language learning ability are related constructs.

Three multiple regression analyses were conducted separately on the composite score of definiteness and mass-count (RQ1) and the separate scores for each (RQ 2). Inspection of the data showed no violation of multicollinearity (which would be indexed by VIF less than 10 and tolerance above .02 (Field, 2009). First, a multiple regression analysis was conducted on the combined eye-tracking scores. The omnibus test revealed that the model was not significant, F (3, 51) = 1.93, p = .14, R^2 = .10. Regression coefficients in the multiple regression model are presented in Table 3. Results showed that the SRT score was the only significant predictor (β = .29, p = .04). None of the other predictors were statistically significant (p > .1).

Next, a multiple regression analysis was conducted on the eye-tracking scores for definiteness. The omnibus test revealed that the model was significant, F (3, 56) = 2.90, p = .04, R^2 = .13. The SRT score was the only significant predictor (β = .36, p = .01). Again, none of the other predictors were statistically significant (p > .1).

Finally, a multiple regression analysis was conducted on the eye-tracking scores for the mass-count distinction. The omnibus test revealed that the model was not significant, F (3, 51) = 0.81, p = .50, R^2 = .05. None of the predictors were statistically significant (p > .1). The largest standardized coefficient was for MLAT_4 (β = .21, p = .16).

Table 3.

Multiple Regression Results

[Composite of Definiteness and Mass-Count (*n* = 55)]

	B	SE	*β*	*t*	*p*	Partial-*r*
SRT	.07	.03	.28	2.13	.04	.29
LLAMA F	.01	.02	.09	.64	.52	.09
MLAT_4	.01	.02	.07	.48	.63	.07

[Definiteness (*n* = 60)]

	B	SE	*β*	*t*	*p*	Partial-*r*
SRT	.08	.03	.36	2.91	.01	.36
LLAMA F	.01	.01	.08	.61	.54	.08
MLAT_4	-.01	.02	-.09	-.67	.51	-.09

[Mass-count (*n* = 55)]

	B	SE	*β*	*t*	*p*	Partial-*r*
SRT	.01	.03	.05	.34	.74	.05
LLAMA F	.00	.01	-.02	-.14	.89	-.02
MLAT_4	.03	.0	.21	1.43	.16	.20

Note. B and β indicate the unstandardized and standardized regression coefficients, respectively.

Discussion

The Role of Explicit and Implicit Learning Aptitudes in Adult SLA

The first RQ addressed to what extent explicit and implicit learning aptitude contributed to the attainment of real-time grammar processing. The findings indicate that aptitude for implicit learning (SRT task), not explicit learning aptitudes (LLAMA_F and MLAT_4), significantly predicted the overall sensitivity to the distinctions in the English NP.

As hypothesized at the outset of this study, a positive relationship was detected between implicit learning aptitude and real-time comprehension of grammatical structures, which is argued to index implicit knowledge (Suzuki, 2017). With virtually no overlap of the measurement metrics between eye-movements (the visual-world task) and RT (the SRT task), their systematic correlation can be construed as empirical evidence—more convincing than what was found in prior research using the RT-based word-monitoring task (Granena, 2013b; Suzuki & DeKeyser, 2015; see also Godfroid & Kim, 2021)—that aptitude and language tests tap into a common underlying ability pertaining to implicit learning. In contrast, there was no significant association between explicit learning aptitude and the visual-world task performance. Unlike the lingering ambiguous interpretations of GJT scores as explicit, implicit, or a mixture of both (Godfroid et al., 2015; Vafaee et al., 2017), the current finely-tuned eye-tracking measure was useful to demonstrate that explicit learning aptitudes played no significant role in the acquisition of implicit knowledge.

The sensitivity index from the visual-world task, however, may need to be interpreted with caution as indicators of implicit knowledge. Unlike regular visual-world task design, which does not include any ungrammatical sentences (e.g., see Suzuki, 2017, in which predictive sentence processing based on Japanese case-markers and classifiers were examined without using ungrammatical sentences in the visual-world task), the current task included ungrammatical noun phrases (e.g., two bacon) and infelicitous use of the definite article. This might have raised learners' awareness of the target structures. Yet, according to the debriefing results, only a few participants noticed infelicitous/ungrammatical target(s) in the auditory sentences. The majority of the participants remained unaware of the linguistic target as they were performing the visual-world task. The aptitude-outcome correlation essentially remained the same, regardless of learners' post-task self-reported awareness (see Appendix F). This finding, combined with the systematic correlation with implicit learning, rather than explicit learning, is promising in the sense that the visual-world task can tap into automatic linguistic processing with little or no awareness, i.e., implicit knowledge.

A broader picture of the present findings illustrates the significant effect of implicit learning aptitude AND no effects of explicit learning aptitude on the acquisition of implicit knowledge. In our view, this overall pattern is consistent with the major claim in the SLA field that there are independent routes of explicit and implicit learning (e.g., Hulstijn, 2002; Krashen, 1985; Paradis, 2009). A systematic correlation between implicit aptitude and grammatical knowledge suggests that adult L2 learners can still use implicit learning mechanisms to acquire L2 grammatical properties to the extent they can be used rapidly—and possibly without awareness at some stage in the learning process.

Having said that, interpreting a systematic association of eye-tracking measures with implicit (but not explicit) aptitude may not always be straightforward (DeKeyser & Li, 2021; cf., Godfroid & Kim, 2021). First, the grammatical knowledge that was retrieved for the current visual-world task may not be necessarily the same as the knowledge that is initially acquired by recruiting explicit or/and implicit aptitudes (DeKeyser & Li, 2021). In order to shed light on the developmental processes, a longitudinal study is urgently needed to examine the differential effects of explicit and implicit aptitudes in early and late stages of L2 learning (e.g., Kim, 2020; Li & DeKeyser, 2021). Second, even if the two learning mechanisms operate independently, explicit learning aptitude (and mechanism) could have "indirectly" influenced the acquisition of implicit knowledge. For instance, Li and DeKeyser (2021) argue that "explicit aptitude may contribute to implicit knowledge indirectly by providing fodder (declarative knowledge) for implicit learning." (p. 479) Hence, the lack of explicit aptitude effect does NOT lead to the conclusion that the role explicit learning plays in adult SLA is marginal. The empirical question still remains to what extent explicit knowledge (i.e., the product of explicit learning) influences the acquisition of implicit knowledge, i.e., the interface issue. It is conceivable that explicit learning aptitude is related to explicit knowledge, and explicit knowledge further plays facilitative or essential roles for the acquisition of implicit knowledge, as demonstrated by the structural equation modelling in Suzuki and DeKeyser (2017) study.

The SLA field has just started rigorous empirical investigations into how different individuals recruit explicit and implicit learning mechanisms—through a set of cognitive aptitudes—for attaining explicit and implicit knowledge in naturalistic settings. These complex interfaces among explicit-implicit learning and knowledge, presumably mediated by a set of cognitive aptitudes, need to be further understood by a more rigorous study from multiple perspectives.

Implicit Learning Aptitude Predicted the Learning of "Difficult" Grammatical Structure

Regarding the second RQ, the current findings indicated that implicit learning aptitude was a significant predictor of the acquisition of definiteness, but not the mass-count distinction.[3] The relatively easier mass-count distinction was acquired so easily that it might have led to the diminished effects of explicit aptitude, while the acquisition of definiteness imposed much more burdensome and might have necessitated the compensatory role of aptitude (cf., DeKeyser, 2016). Although the implicit aptitude test (SRT) primarily requires domain-general sequence learning of "dots" that appear on the computer screen and thus requires no semantic processing component, it is surprising that the SRT task consistently predicted the acquisition of L2 implicit knowledge (weak to moderate correlations ranging from r = .36 to .43) across different studies targeting different structures—agreement structures (Granena, 2013a), case-markers (Suzuki & DeKeyser, 2015), and definiteness in this study (see Table 4).

[3] The lack of association of the index for mass-count with the SRT cannot be attributed to the range restriction in the dependent variable because standard deviations for definiteness (.29) and mass-count (.30) were almost comparable.

Table 4.

Summary of Previous Research on Implicit Learning Aptitude

	Granena (2013b)	Suzuki & DeKeyser (2015)	Current Study
Participants	50 L2 Spanish learners	63 L2 Japanese learners	65 L2 English learners
Age of Arrival	> 16 years old	> 18 years old	> 15 years old
Length of Residence	101 months	55 months	46 months
L1	Chinese	Chinese	Chinese
Language Test	Word-monitoring task	Word-monitoring task	Visual-world task
Target Structures	3 agreement structures 3 morpho-semantic structures	Five Japanese particles	Definiteness Count/Mass
Aptitude test	SRT task	SRT task	SRT task LLAMA F MLAT Part 4
Results	SRT was only related to agreement structures ($r = .36$, $p < .05$)	SRT was related to five Japanese particles ($r = .43$, $p < .05$)	SRT was only related to definiteness ($r = .36$, $p < .05$)

All three studies recruited Chinese speakers whose L1 has no overt morphological marking; all L2 grammatical structures that were significantly facilitated by the implicit learning aptitude have no equivalent grammatical system to their learner's L1.[4] These incongruent L2 structures are notoriously difficult for integration into learners' L2 linguistic system and automatization for real-time comprehension (Jiang, Hu, Chrabaszcz, & Ye, 2015; Jiang, Novokshanova, Masuda, & Wang, 2011; Roberts & Liszka, 2013). The role of implicit learning aptitude may become more important when L2 learners are left with no congruent L1 feature to fall back on. More specifically, implicit learning aptitude becomes more important when new semantic distinctions are to be learned (cf., Murakami & Alexopoulou, 2016). In the case of learning definiteness distinction, it is speculated that more efficient sequence learning may facilitate accumulating determiner-noun sequences from the input in the first place (N. C. Ellis, 2015). The *initial* statistical tallying of co-occurrence of determiner and noun pairs essentially corresponds to the nature of implicit learning assessed in the probabilistic SRT task. The higher sequence learning ability might have promoted rapid association of the determiner-noun pair across different pieces of discourse. Subsequently, these determiner-noun sequences committed to memory need to be

[4] Suzuki (2021) re-analyzed the data of Suzuki and DeKeyser (2015) and revealed that the SRT was particularly related to transitive-intransitive verb pairs and the locative particles ($r = .42$ in both cases), which presumably rely on item-based learning.

abstracted through more complex mappings of the determiners with the semantics of definiteness, which may necessitate different cognitive abilities beyond what is tapped by the SRT task.

With regard to the contribution of explicit aptitude, no systematic relation was found for either definiteness or mass-count distinction. Definiteness involves a very abstract pattern of form-meaning mapping requiring the learning of novel semantic distinctions for Chinese speakers, which might have made explicit learning very difficult (DeKeyser, 2005; Robinson, 2005). Possibly, the absence of a role for explicit aptitude might have left more room for other predictors (e.g., implicit aptitude). While definiteness might have been too hard to learn explicitly, let alone to automatize, the mass-count distinction might have been easy enough to neutralize the effects of aptitude (DeKeyser, 2016). Since the mass-count distinction is salient for Chinese learners (e.g., by virtue of a similar classifier system), learners were automatically able to tune into the relevant features in input (Leung & Williams, 2014).

Suggestions for Future Research: Aptitude and Language Test Development

The current study opens up new lines of investigation into explicit and implicit learning processes in adult SLA. First and foremost, we call for further research addressing the validity of the explicit and implicit measures for both aptitudes and linguistic knowledge. The current study did not include any tests for explicit knowledge, and it is important to more comprehensively examine the relationships among explicit and implicit knowledge and aptitudes (see Suzuki & DeKeyser, 2017).

For measures for implicit knowledge, a wider variety of target structures should be tested in the visual-world task. While the current study only targeted two easy and difficult structures, a wider variety of linguistic structures can be examined through a useful operationalization of grammatical difficulty such as saliency (Gass, Spinner, & Behney, 2017). Because explicit learning processes are more influenced by the saliency of linguistic structures than implicit learning processes, it is highly relevant to investigate the roles of explicit and implicit aptitudes for salient and non-salient linguistic structures (see, e.g., DeKeyser, Alfi-Shabtay, Ravid, & Shi, 2017, who demonstrated that adult learners, presumably relying more on explicit learning mechanisms, found it extremely difficult to acquire the low salient grammatical structures).

Although a majority of L2 learners remained unaware of target grammatical structures tested in this study, it is possible that some participants simply could not articulate what they noticed. One way to solve that potential problem might be to have participants agree or disagree with a list of comments including simple descriptions of the infelicity that they could have noticed. Nonetheless, the visual-world task can be devised without any ungrammatical sentences, which can help further reduce the risk of raising awareness (see Suzuki, 2017). Although it requires more laborious work to develop a visual-world task than reaction-time tasks, it is applicable to a wider L2 populations including children.

For measures of aptitudes, the field can avail itself of a new aptitude test battery such as Hi-Level Language Aptitude Battery (e.g., Linck et al., 2013). The aptitude measure for implicit learning in the current study was the SRT task, which focuses on probabilistic/statistical sequence learning; other aspects of implicit learning should be examined. Future research could thus scrutinize the role of more specific forms of implicit learning (Granena, 2019, 2020). Interestingly, in the current study, MLAT_4 (grammatical sensitivity in L1) may seem to be related to the acquisition of the mass-count distinction, the learning of which is assumed to rely on a similar L1 structure, more strongly than LLAMA_F (inductive language learning ability). While grammatical sensitivity and inductive language learning ability are subsumed under the larger construct "language analytic ability" (Skehan, 2002), it may be useful to further examine how different subcomponents of explicit aptitude are more or less related to different grammatical structures. The reliability of the aptitude tests was acceptable (e.g., Granena, 2013; Suzuki & DeKeyser, 2017) but not very high in this study, which points to the need for improvement in the test instruments' reliability and validity. We hope research on both language and aptitude tests can advance hand in hand to enable better understanding of L2 learning processes.

Conclusions

The aim of the current study was to investigate the extent to which explicit and implicit learning aptitudes influence the attainment of real-time processing by adult L2 English learners with L1 Chinese. By employing the visual-world task, we assessed real-time processing of definiteness and the mass-count distinction with little contamination from explicit knowledge. The post-task debriefing results showed that a majority of L2 learners were unaware of target grammatical structures tested, which lends support to the use of linguistic knowledge without awareness (i.e., implicit knowledge). Implicit learning aptitude was particularly important for the acquisition of definiteness, but not for the acquisition of the mass-count distinction. These findings underscore the importance of understanding L2 learning processes underlying different structures through the lens of explicit and implicit learning aptitudes.

References

Abrahamsson, N., & Hyltenstam, K. (2008). The robustness of aptitude effects in near-native second language acquisition. *Studies in Second Language Acquisition, 30*(4), 481-509. doi:10.1017/S027226310808073X

Andringa, S., & Curcic, M. (2015). How explicit knowledge affects online L2 processing. *Studies in Second Language Acquisition, 37*(s2), 237-268. doi:10.1017/S0272263115000017

Andringa, S., & Rebuschat, P. (2015). New directions in the study of implicit and explicit learning. *Studies in Second Language Acquisition, 37*(s2), 185-196. doi:10.1017/S027226311500008X

Bolibaugh, C., & Foster, P. (2021). Implicit statistical learning in naturalistic and instructed morphosyntactic attainment: An aptitude-treatment interaction design. *Language Learning, Early View*. doi:10.1111/lang.12465

Carroll, J. B. (1991). Cognitive abilities in foreign language aptitude: Then and now. In T. Parry & C. Stansfield (Eds.), *Language aptitude reconsidered* (pp. 11-29). Englewood Cliffs, N. J.: Prentice Hall.

Carroll, J. B., & Sapon, S. M. (1959). *Modern language aptitude test: MLAT*. New York: Psychological Corporation.

Chambers, C. G., Tanenhaus, M. K., Eberhard, K. M., Filip, H., & Carlson, G. N. (2002). Circumscribing referential domains during real-time language comprehension. *Journal of memory & language, 47*(1), 30-49.

Cheng, L., & Sybesma, R. (1998). Yi-wan tang, yi-ge tang: Classifiers and massifiers. *Tsing Hua Journal of Chinese Studies, 28*(3), 385-412.

Cheng, L., & Sybesma, R. (1999). Bare and not-so-bare nouns and the structure of np. *Linguistic inquiry, 30*(4), 509-542.

DeKeyser, R. M. (2000). The robustness of critical period effects in second language acquisition. *Studies in Second Language Acquisition*, *22*(4), 499-534. doi:10.1017/S0272263100004022

DeKeyser, R. M. (2003). Implicit and explicit learning. In C. J. Doughty & H. M. Long (Eds.), *The handbook of second language acquisition* (pp. 312-348). Oxford: Blackwell Publishers.

DeKeyser, R. M. (2005). What makes learning second language grammar difficult? A review of issues. *Language Learning, 55*(S1), 1-25. doi:10.1111/j.0023-8333.2005.00294.x

DeKeyser, R. M. (2012). Interactions between individual differences, treatments, and structures in SLA. *Language Learning, 62*(s2), 189-200. doi:10.1111/j.1467-9922.2012.00712.x

DeKeyser, R. M. (2016). Of moving targets and chameleons: Why the concept of difficulty is so hard to pin down. *Studies in Second Language Acquisition, 38*(2), 353-363. doi:10.1017/S0272263116000024

DeKeyser, R. M., Alfi-Shabtay, I., & Ravid, D. (2010). Cross-linguistic evidence for the nature of age effects in second language acquisition. *Applied Psycholinguistics, 31*(3), 413-438. doi:10.1017/S0142716410000056

DeKeyser, R. M., Alfi-Shabtay, I., Ravid, D., & Shi, M. (2017). The role of salience in the acquisition of hebrew as a second language: Interaction with age of acquisition. In S. Gass, P. Spinner, & J. Behney (Eds.), *Salience in second language acquisition* (pp. 131-146). New York: Routledge.

DeKeyser, R. M., & Li, S. (2021). Epilogue: Evidence for the validity of implicit aptitude and the need for construct validation. *Studies in Second Language Acquisition, 43*(3), 692-697. doi:10.1017/S0272263121000140

Dussias, P. E., Valdés Kroff, J. R., Guzzardo Tamargo, R. E., & Gerfen, C. (2013). When gender and looking go hand in hand. *Studies in second language acquisition, 35*(2), 353-387. doi:10.1017/S0272263112000915

Ellis, N. C. (2015). Implicit and explicit language learning: Their dynamic interface and complexity. In P. Rebuschat (Ed.), *Implicit and explicit learning of languages* (pp. 3-24). Amsterdam, the Netherlands: John Benjamins Publishing Company.

Ellis, N. C., Hafeez, K., Martin, K. I., Chen, L., Boland, J., & Sagarra, N. (2012). An eye-tracking study of learned attention in second language acquisition. *Applied Psycholinguistics, 35*(3), 547-579.

Ellis, R. (2005). Measuring implicit and explicit knowledge of a second language. *Studies in Second Language Acquisition, 27*(2), 141-172. doi:10.1017/S0272263105050096

Field, A. (2009). *Discovering statistics using spss*: Sage Publications Limited.

Gass, S. M., Spinner, P., & Behney, J. (2017). *Salience in second language acquisition*. New York: Routledge.

Godfroid, A. (2016). The effects of implicit instruction on implicit and explicit knowledge development. *Studies in Second Language Acquisition, 38*(2), 177-215. doi:10.1017/S0272263115000388

Godfroid, A. (2019). *Eye tracking in second language acquisition and bilingualism: A research synthesis and methodological guide*. New York: Routledge.

Godfroid, A., & Kim, K. (2021). The contributions of implicit-statistical learning aptitude to implicit second-language knowledge. *Studies in Second Language Acquisition, 43*(3), 606-634. doi:10.1017/S0272263121000085

Godfroid, A., Loewen, S., Jung, S., Park, J., Gass, S., & Ellis, R. (2015). Timed and untimed grammaticality judgements measure distinct types of knowledge. *Studies in Second Language Acquisition, 37*(s2), 269-297. doi:10.1017/S0272263114000850

Granena, G. (2013a). Individual differences in sequence learning ability and second language acquisition in early childhood and adulthood. *Language Learning, 63*(4), 665–703. doi:10.1111/lang.12018

Granena, G. (2013b). Reexamining the robustness of aptitude in second language acquisition. In G. Granena, & Long, M. H. (Ed.), *Sensitive periods, language aptitude, and ultimate L2 attainment.* (pp. 179-204). Philadelphia: PA: John Benjamins.

Granena, G. (2019). Cognitive aptitudes and L2 speaking proficiency: Links between LLAMA and Hi-LAB. *Studies in Second Language Acquisition, 41*(2), 313-336. doi:10.1017/S0272263118000256

Granena, G., & Long, M. H. (2013). Age of onset, length of residence, language aptitude, and ultimate L2 attainment in three linguistic domains. *Second Language Research, 29*(3), 311-343. doi:10.1177/0267658312461497

Graus, J., & Coppen, P.-A. (2015). Defining grammatical difficulty: A student teacher perspective. *Language Awareness, 24*(2), 101-122. doi:10.1080/09658416.2014.994639

Hagedorn, J., Hailpern, J., & Karahalios, K. G. (2008). *Vcode and vdata: Illustrating a new framework for supporting the video annotation workflow*. Paper presented at the Advanced Video Interface. http://social.cs.uiuc.edu/papers/pdfs/vcode.pdf

Housen, A. (2014). Difficulty and complexity of language features and second language instruction. In C. Chapelle (Ed.), *The encyclopedia of applied linguistics* (pp. 1-7): John Wiley & Sons, Inc.

Housen, A., & Simoens, H. (2016). Introduction: Cognitive perspectives on difficulty and complexity in L2 acquisition. *Studies in Second Language Acquisition, 38*(2), 163-175. doi:10.1017/S0272263116000176

Hua, D., & Lee, H. (2005). *Chinese ESL learners' understanding of the English count-mass distinction.* Paper presented at the Proceedings of the 7th Generative Approaches to Second Language Acquisition Conference (GASLA 2004) pp, Somerville, MA.

Hulstijn, J. H. (2002). Towards a unified account of the representation, processing and acquisition of second language knowledge. *Second Language Research, 18*(3), 193-223. doi:10.1191/0267658302sr207oa

Hulstijn, J. H. (2005). Theoretical and empirical issues in the study of implicit and explicit second-language learning. *Studies in Second Language Acquisition, 27*(2), 129-140. doi:10.1017/S0272263105050084

Ionin, T., Zubizarreta, M. L., & Philippov, V. (2009). Acquisition of article semantics by child and adult L2-English learners. *Bilingualism: Language and Cognition, 12*(3), 337-361. doi:10.1017/S1366728909990149

Jiang, N., Hu, G., Chrabaszcz, A., & Ye, L. (2015). The activation of grammaticalized meaning in L2 processing: Toward an explanation of the morphological congruency effect. *International Journal of Bilingualism, 21*(1), 81-98. doi:10.1177/1367006915603823

Jiang, N., Novokshanova, E., Masuda, K., & Wang, X. (2011). Morphological congruency and the acquisition of L2 morphemes. *Language Learning, 61*(3), 940–967. doi:10.1111/j.1467-9922.2010.00627.x

Kaufman, S. B., DeYoung, C. G., Gray, J. R., Jiménez, L., Brown, J., & Mackintosh, N. (2010). Implicit learning as an ability. *Cognition, 116*(3), 321-340. doi:10.1016/j.cognition.2010.05.011

Kim, M. (2020). *Exploring the interface of explicit and implicit second-language knowledge: A longitudinal perspective.* (Ph.D.), Michigan State University, Ann Arbor.

Krashen, S. D. (1982). *Principles and practice in second language acquisition*. Oxford, UK: Pergamon Press.

Krashen, S. D. (1985). *The input hypothesis: Issues and implications*. New York, NY: Longman.

Lardiere, D. (2007). *Ultimate attainment in second language acquisition: A case study*. Mahwah, NJ: Lawrence Erlbaum Associates.

Leung, J. H. C., & Williams, J. N. (2014). Crosslinguistic differences in implicit language learning. *Studies in Second Language Acquisition, 36*(4), 733-755. doi:10.1017/S0272263114000333

Li, S., & DeKeyser, R. M. (2021). Implicit language aptitude: Conceptualizing the construct, validating the measures, and examining the evidence: Introduction to the special issue. *Studies in Second Language Acquisition, 43*(3), 473-497. doi:10.1017/S0272263121000024

Linck, J. A., Hughes, M. M., Campbell, S. G., Silbert, N. H., Tare, M., Jackson, S. R., . . . Doughty, C. J. (2013). Hi-LAB: A new measure of aptitude for high-level language proficiency. *Language Learning, 63*(3), 530-566. doi:10.1111/lang.12011

Lu, C. F.-C. (2001). The acquisition of English articles by Chinese learners. *Second Language Studies, 20*(1), 43-78.

Lyons, C. (1999). *Definiteness*. Cambridge, UK: Cambridge University Press.
Matin, E., Shao, K., & Boff, K. R. (1993). Saccadic overhead: Information-processing time with and without saccades. *Perception and Psychophysics, 53*(4), 372-380. doi:10.3758/BF03206780
Meara, P. M. (2005). LLAMA language aptitude tests: The manual. Swansea, UK: Lognostics. Retrieved from http://www.lognostics.co.uk/tools/llama/llama_manual.pdf
Murakami, A., & Alexopoulou, T. (2016). L1 influency on the acquisition order of English grammartical morphemes: A learner corpus study. *Studies in Second Language Acquisition, 38*(3), 365-401. doi:10.1017/S0272263115000352
Paradis, M. (2009). *Declarative and procedural determinants of second languages*. Philadelphia, PA: John Benjamins Publishing Company.
Reber, A. S., Walkenfeld, F. F., & Hernstadt, R. (1991). Implicit and explicit learning: Individual differences and IQ. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 17*(5), 888-896. doi:10.1037/0278-7393.17.5.888
Roberts, L., & Liszka, S. A. (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. doi:10.1177/0267658313503171
Robertson, D. (2000). Variability in the use of the English article system by Chinese learners of English. *Second Language Research, 16*(2), 135-172.
Robinson, P. (1997). Individual differences and the fundamental similarity of implicit and explicit adult SLA. *Language Learning, 47*(1), 45-99. doi:10.1111/0023-8333.21997002
Robinson, P. (2005). Cognitive abilities, chunk-strength and frequency effects during implicit artificial grammar, and incidental second language learning: Replications of Reber, Walkenfeld and Hernstadt (1991) and Knowlton and Squire (1996) and their relevance to SLA. *Studies in Second Language Acquisition, 27*(2), 235-268. doi:10.1017/S0272263105050126
Skehan, P. (2002). Theorising and updating aptitude. In P. Robinson (Ed.), *Individual differences and instructed language learning* (pp. 69-94). Philadelphia, PA: John Benjamins.
Snape, N. (2008). Resetting the nominal mapping parameter in L2 English: Definite article use and the count-mass distinction. *Bilingualism Language and Cognition, 11*(1), 63-79. doi:10.1017/S1366728907003215
Suzuki, Y. (2015). *Using new measures of implicit L2 knowledge to study the interface of explicit and implicit knowledge.* (Doctoral dissertation), University of Maryland, College Park, MD, USA.
Suzuki, Y. (2017). Validity of new measures of implicit knowledge: Distinguishing implicit knowledge from automatized explicit knowledge. *Applied Psycholinguistics, 38*(5), 1229-1261. doi:10.1017/S014271641700011X
Suzuki, Y. (2021). How do Japanese grammatical knowledge develop in naturalistic setting: Roles of grammatical features and individual difference factors. [dainigengokankyou de nihongo no bunpou chishiki wa donoyouni hattatsushiteikuka: Bunpoukoumoku no tokuchou to gakushusha no kojinsa no eikyou]. In W. Suzuki, Y. Sakuma, & T. Terasawa (Eds.), *What are the roles of explicit and implicit knowledge in foreign language learning* (pp. 33-48). Tokyo, Japan: Taishukan.
Suzuki, Y., & DeKeyser, R. M. (2015). Comparing elicited imitation and word monitoring as measures of implicit knowledge. *Language Learning, 65*(4), 860-895. doi:10.1111/lang.12138

Suzuki, Y., & DeKeyser, R. M. (2017). The interface of explicit and implicit knowledge in a second language: Insights from individual differences in cognitive aptitudes. *Language Learning, 67*(4), 747-790. doi:10.1111/lang.12241

Suzuki, Y., Jeong, H., Cui, H., Okamoto, K., Kawashima, R., & Sugiura, M. (2022). An fMRI validation study of the word-monitoring task as a measure of implicit knowledge: Exploring the role of explicit and implicit aptitudes in behavioral and neural processing. *Studies in Second Language Acquisition, Early View*. doi:10.1017/S0272263122000043

Trenkic, D. (2002). Form meaning connections in the acquisition of English articles. *EUROSLA Yearbook, 2*(1), 115-133.

Trenkic, D. (2008). The representation of English articles in second language grammars: Determiners or adjectives? *Bilingualism: Language and Cognition, 11*(1), 1-18. doi:10.1017/S1366728907003185

Trenkic, D., Mirkovic, J., & Altmann, G. T. M. (2014). Real-time grammar processing by native and non-native speakers: Constructions unique to the second language. *Bilingualism: Language and Cognition, 17*(2), 237-257. doi:10.1017/S1366728913000321

Vafaee, P., Suzuki, Y., & Kachinske, I. (2017). Validating grammaticality judgment tests: Evidence from two new psycholinguistic measures. *Studies in Second Language Acquisition, 39*(1), 59-95. doi:10.1017/S0272263115000455

White, L. (2003). Fossilization in steady state L2 grammars: Persistent problems with inflectional morphology. *Bilingualism: Language and Cognition, 6*(02), 129-141. doi:10.1017/S1366728903001081

Williams, J. N. (2009). Implicit learning in second language acquisition. In W. C. Ritchie & T. K. Bhatia (Eds.), *The new handbook of second language acquisition* (pp. 319-353). Bingly, UK: Emerald Group Publishing Limited.

Xia, H.-Y. (2011). A language aptitude test for native Chinese: Development and validation. [汉语语境下的外语学能测试研究及效度验证]. *Shandong Foreign Language Teaching Journal, 5*, 52-55.

Yalçın, Ş., & Spada, N. (2016). Language aptitude and grammatical difficulty. *Studies in Second Language Acquisition, 38*(2), 239-263. doi:10.1017/S0272263115000509

Young, R. (1996). Form-function relations in articles in English interlanguage. In R. Bayley & D. R. Preston (Eds.), *Second language acquisition and linguistic variation* (pp. 135-175). Amsterdam: John Benjamins.