L1-Korean speakers' definiteness processing in L2 English: A visual world paradigm eye-tracking study

Journal:	*Bilingualism: Language and Cognition*
Manuscript ID	BLC-20-RA--0191.R5
Manuscript Type:	Research Article
Date Submitted by the Author:	04-Nov-2022
Complete List of Authors:	Ahn, Hyunah; Seoul National University, English Language and Literature Song, Mi-jeong; Seoul National University, English Language and Literature
Content Areas:	Grammatical Processing
Linguistic Areas:	Articles/Determiners, Definiteness, Semantics
Methods:	Behavioural Measurements, Eye-movement Monitoring, Statistics: Mixed Effects Modelling
Populations:	Adult Second Language Development
Languages:	English
Abstract:	This study investigated whether L1-Korean L2-English speakers could use (in)definiteness to predict unmentioned (non-)unique referents. In previous studies on the topic, methodological considerations made it difficult to separate the effects of grammar from the effects of non-linguistic knowledge. We used a visual world paradigm eye-tracking task to resolve such issues and designed stimuli that focused only on (in)definiteness. Participants' eye movements were recorded as they heard "Click on the/a blue circle." L1 and advanced L2 speakers used definiteness information to predict unique referents before the critical noun (circle) was heard while indefinite articles were not utilized to the same extent. Intermediate L2 speakers relied heavily on color words, not articles, to locate a referent. The results imply that predicting a referent solely based on definiteness (without real-world knowledge) requires substantial advancement in L2 development, and indefinite articles do not predict non-unique referents as clearly as definite articles predict unique referents.

Short title: Running-Head: L2 definiteness processing via VWP

Full title: L1-Korean speakers' definiteness processing in L2 English: A visual world paradigm eye-tracking study

Author names: Hyunah Ahn[1], Mi-Jeong Song[2]

Affiliations: 1 Kunsan National University, Gunsan, Korea
2 Seoul National University, Seoul, Korea

Acknowledgments:
This study was supported by the Ministry of Education of the Republic of Korea and the National Research Foundation of Korea (NRF-2018S1A5B5A07071643). We would like to express our gratitude to Dr. Sung-Eun Lee in the Department of German Language and Literature at Seoul National University for allowing access to an eye-tracker in the Brain and Humanities Lab and Mr. Jonghyun Lee for helping us with running participants in the lab. We are also grateful to the editors Harald Clahsen, Annie Tremblay, and Kira Gor and three anonymous reviewers for their insightful feedback on earlier versions of our paper. All errors that remain are ours.

Address for correspondence:
Department of English Language and Literature
College of Humanities, Kunsan National University
558 Daehakro, Gunsan, Korea 54150
hyunah1112@gmail.com

Keywords:
L2 definiteness processing, visual world paradigm eye-tracking, definiteness as uniqueness, real-world knowledge, L2 predictive processing

Abstract

This study investigated whether L1-Korean L2-English speakers could use (in)definiteness to predict unmentioned (non-)unique referents. In previous studies on the topic, methodological considerations made it difficult to separate the effects of grammar from the effects of non-linguistic knowledge. We used a visual world paradigm eye-tracking task to resolve such issues and designed stimuli that focused only on (in)definiteness. Participants' eye movements were recorded as they heard "Click on *the*/*a* blue circle." L1 and advanced L2 speakers used definiteness information to predict unique referents before the critical noun (*circle*) was heard while indefinite articles were not utilized to the same extent. Intermediate L2 speakers relied heavily on color words, not articles, to locate a referent. The results imply that predicting a referent solely based on definiteness (without real-world knowledge) requires substantial advancement in L2 development, and indefinite articles do not predict non-unique referents as clearly as definite articles predict unique referents.

Keywords

L2 definiteness processing, visual world paradigm eye-tracking, definiteness as uniqueness, real-world knowledge, L2 predictive processing

1 Introduction

Since English articles are often produced in a phonologically reduced form, investigating the processing of *the* and *a(n)* is a complicated enterprise. English articles are also semantically complex in that, in addition to definiteness, number (a car vs. cars) and countability (a chair vs. furniture) constraints should be factored in to use articles correctly. Many L2 studies documented L2 English article use regarding a wide range from singular count nouns to generic and customary uses (Butler, 2002; Huebner, 1983; Master, 1987, 1997, among others), but such comprehensive documentation has made it challenging to identify the source of non-target-like behavior in L2 English article use.

Systematic investigations with a narrower focus on singular count nouns have been conducted by Ionin and colleagues (Ionin et al., 2004; Ionin et al., 2008; Ionin et al., 2009; Ko et al., 2010), who put forward the Fluctuation Hypothesis, inspiring several subsequent studies in the tradition of semantic universals (Snape, 2009; Snape et al., 2006). These studies mainly used offline methods that could not probe real-time sentence processing (c.f., Ionin et al., 2021), and their stimuli differed systematically between conditions, which could have provided participants with metalinguistic cues (see Ahn, 2021 for a detailed review of this issue). Trenkic *et al.* (2014) was one of a few studies designed to investigate the real-time processing of definiteness in L2, but their study design had definiteness conflated with real-world knowledge, which we will review more thoroughly in the next section.

Recently, many researchers have focused on making predictions or generating expectations in L2 sentence processing (Grüter et al, 2017) and reported different results regarding the ability to use ‘articles’ in predictive L2 processing (Dussias et al., 2013; Henry et al., 2020; Hopp, 2015; Hopp & Lemmerth, 2018). These studies on L2 article processing focused on gender and case in Spanish and German articles rather than definiteness in English articles; however, their findings diverged depending on the involvement of another source of information than articles. This observation necessitates that one conducts a study in which L2 speakers are tested only on articles in the absence of any other information that can conflate the grammatical feature at issue.

It is challenging to isolate the effect of definiteness from other factors that can influence L2 processing. L2 processing has been reported to be susceptible to different information types (Clahsen & Felser, 2006; 2018; Cunnings, 2017; Sorace, 2011; Sorace & Filliaci, 2006), non-linguistic information can play a role alongside linguistic information (Ahn, 2022), and metalinguistic knowledge could play a role when offline methods are used (Ionin et al., 2004; Ko et al., 2010). In this study, we present a carefully designed study to minimize the interference of real-world knowledge and the use of metalinguistic knowledge. Investigating the real-time processing of definiteness using eye-tracking in the visual world paradigm can shed light on such issues from previous findings. The method can be used to observe participants' eye-gaze movements time-locked to auditory stimuli (Conklin & Pellier-Sanchez, 2016; Dussias, 2010; Roberts & Siyanova-Chanturia, 2013). Using a method that probes online processing can help prevent the potential problem of participants' relying on metalinguistic knowledge. The present study was also designed to separate definiteness from other factors such as real-world knowledge. We designed auditory and visual stimuli such that real-world knowledge could not interfere with definiteness processing. Below we report how definite articles can play a role in predictive sentence processing by L1-Korean L2-English speakers.

2 Background & Motivation

2.1 Definiteness in L1 research

Results and implications from previous studies on the processing of English articles by native speakers vary greatly. Much of the variability in the results arises from testing different aspects of the construct 'definiteness.' Definiteness has been studied and defined in many different frameworks involving constructs such as unique identifiability (Russell, 1905; Hawkins, 1991), givenness (Heim, 1982), the old vs. new distinction (Prince, 1992), the givenness hierarchy (Gundel et al., 1993), and the accessibility hierarchy (Ariel, 1991) among others. Discussing which is the optimal way of defining definiteness is beyond the scope of this study. However, distinguishing definite NPs from indefinite NPs via previously mentioned and unmentioned entities can cause difficulty in result interpretation. When definiteness is operationalized by the distinction of discourse-old vs. discourse-new (Prince, 1992), a previously mentioned referent is referred to by a definite NP (discourse-old), and a referent newly introduced into the discourse is referred to by an indefinite NP (discourse-new). When definiteness is operationalized in terms of unique identifiability, on the other hand, a referent identified to be unique by interlocutors is marked definite. If mutual knowledge of the unique identifiability of an entity cannot be established, the entity is accompanied by an indefinite NP regardless of whether the entity has been previously mentioned in the discourse. The two definitions make an essential difference in carrying out a controlled study. The givenness-based research design will result in the repetition of an NP and lack thereof in definite and indefinite conditions, respectively, which could provide participants with a metalinguistic clue in distinguishing when to use which article.

The sentences in (1) are German experimental stimuli used in an ERP study (Schumacher, 2009). As one can see, the flow from the contextual sentences to the critical sentence sounds natural only when the given or inferred contextual sentence leads to the definite critical sentence, as in (1a) and (1b), respectively.

(1) a. Peter besuchte neulich *einen Redner* in München. — DEFINITE GIVEN
Er erzählte, dass der Redner sehr net war.
"Peter has recently visited *a speaker* in München.
He said that the speaker was very nice."

b. Peter besuchte neulich *einen Vortrag* in München. — DEFINITE INFERRED
Er erzählte, dass der Redner sehr net war.
"Peter has recently visited *a lecture* in München.
He said that the speaker was very nice."

c. Peter traf neulich *Hannah* in München. — DEFINITE NEW
Er erzählte, dass der Redner sehr net war.
"Peter has recently met *Hannah* in München.
He said that the speaker was very nice."

d. Peter besuchte neulich *einen Redner* in München. — INDEFINITE GIVEN
Er erzählte, dass ein Redner sehr net war.
"Peter has recently visited *a speaker* in München.
He said that a speaker was very nice."

e. Peter besuchte neulich *einen Vortrag* in München. — INDEFINITE INFERRED
Er erzählte, dass ein Redner sehr net war.
"Peter has recently visited *a lecture* in München.
He said that a speaker was very nice."

f. Peter traf neulich *Hannah* in München. — Indefinite New
Er erzählte, dass ein Redner sehr net war.
"Peter has recently met *Hannah* in München.
He said that a speaker was very nice."

The experiment results showed that a significantly smaller late positivity was observed in the 550-700*ms* time window only in the definite-given combination, which means that participants perceived only the definite given condition (1a) as natural. In this experimental setting, however, it is hard to tease apart the effect of the givenness of a discourse entity and the effect of the repetition of an NP. For the two effects to be separated, experiments should be set up so that definite and indefinite NPs can be grammatical and ungrammatical in all contexts. In this experiment, the combination of the contextual and critical sentences does not allow such counterbalancing, making it difficult to conclude that the effect of givenness induced the late positivity.

Murphy (1984) distinguished the effect of givenness from that of repetition. Using sentences as in (2), he showed that a definite NP is processed faster than an indefinite NP in an experimental setting where both definite and indefinite NPs can be grammatical, representing given and new conditions, respectively. In this case, the critical noun *truck* is repeated in both definite and indefinite conditions; therefore, the effect of givenness is separated from the effect of repetition.

(2) a. Though driving 55, Steve was passed by a truck. *context*

b. Later, George was passed by *the/a* truck, too. *critical sentence*

However, the effect of (un)grammaticality could not be tested because the materials did not include ungrammatical counterparts of the definite and indefinite conditions. Also, the experiment measured the reading time of the entire sentence rather than the time it took to read the critical NPs. Thus, the resulting reading times might not necessarily reflect the processing loads of definite and indefinite NPs.

Unlike Murphy (1984), where the definite and indefinite conditions were equalized by having both conditions repeat the NPs, Clifton (2013) removed the potential confounding of givenness and NP repetition by manipulating both definite and indefinite NPs to be first mentions. To show that the processing of definite and indefinite NPs can rely on the presupposition made regarding the unique identifiability of a referent, he contextualized critical NPs to refer to either a unique or a non-unique referent (3).

(3) a. In the kitchen, Jason checked out *the/a* stove very carefully.

b. In the appliance store, Jason checked out *the/a* stove very carefully.

(3a) and (3b) contextualize the sentence to presuppose a unique or a non-unique referent and create a grammatical sentence for definite and indefinite NPs, respectively. The study revealed that definite and indefinite NPs do not differ in processing costs when presuppositions for unique and non-unique referents are accommodated by the (in)definiteness of a referring NP.

In brief, the processing of definiteness has been investigated via different experimental settings and resulted in different findings and implications based on the definition of definiteness adopted. This variability issue is not much different in L2 studies. Below is a brief overview of L2 studies that have probed the use and processing of English articles by L2 speakers.

2.2 Definiteness as unique identifiability in L2 sentence processing

Among very few L2 studies investigating definiteness defined as unique identifiability are Trenkic et al. (2014) and Ahn (2021; 2022). These studies report distinct findings regarding L2 definiteness processing. Ahn (2021), via a self-paced reading task with materials adapted from Clifton (2013), showed that definite and indefinite NPs do not display the same effect in accommodating presupposed unique and non-unique referents, respectively, in L1-Korean L2-English speakers. The phenomenon was observed not only in L2 speakers; L1 speakers also read definite NPs faster when unique referents were presupposed than indefinite NPs when non-unique referents were presupposed. Ahn (2021) also reported that only advanced L2 speakers, not their intermediate peers, showed sensitivity to definiteness with regard to unique identifiability the same way L1 speakers did. Her findings show how definite and indefinite NPs can be accommodated depending on the presupposition of a unique or a non-unique referent; however, the study was not intended to investigate whether definite and indefinite NPs can lead participants to predict unique and non-unique referents, respectively. It was Trenkic *et al*. (2014) that investigated whether the (in)definiteness of an NP could be utilized in identifying (non-)unique referents.

Trenkic et al. (2014) reported that L1 Chinese speakers of L2 English at an intermediate level could use the definiteness of an NP to predict a unique referent the same way L1 speakers did. However, their findings and interpretations need reconsidering due to the way their experiment was set up.

Chambers et al. (2002), the original study from which Trenkic *et al.* (2014) borrowed the experimental setup, intended to probe the effect of non-linguistic information in a situational context in L1 sentence processing. There were two visual conditions (one-target and two-target conditions), and the main experimental task was to click on the target referred to as ‘can’ at the end of the auditory stimuli in (4).[1]

(4) a. The man will put the cube inside the can.

b. The man will put the cube inside a can.

There were two cans in both visual conditions. In the one-target condition, only one of the two cans was open and could be a viable target, while, in the two-target condition, both cans were open, thus, two potential targets. Since the preposition used in the auditory stimuli was ‘inside,’ a bowl in the visual scene played the role of a competitor. A rubber duck and a hammer were distractors. This experimental setting allows participants to predict a container-type goal based on the preposition “inside.”

Instead of manipulating only the number of targets (cans) in the visual scene, the original study (Chambers et al*.,* 2002) was designed to measure the effect of real-world knowledge, e.g., that a closed can is not a good candidate as a goal to place an object in. Indeed, Chambers et al. (2002) interpreted their results as supporting evidence for the use of situational contexts (real-world knowledge) in L1 sentence processing, while Trenkic et al. (2014) interpreted their results to indicate that L2 speakers have the same grammatical knowledge as L1 speakers. In both experiments, grammatical knowledge might have played a role in sentence processing to a certain degree; however, the experimental design may not allow the role of grammatical knowledge to be separated from that of real-world knowledge.

The pattern of L2 behavior observed in Trenkic et al. (2014) provides support for the concerns raised above. L2 participants in their study displayed the opposite pattern of behavior from that of L1 speakers in the grammatical conditions. L2 participants took longer to fixate on a target in the indefinite/non-unique condition than in the definite/unique condition, while L1 speakers were quickest to locate a target in the indefinite/non-unique condition (see Figure S1 in *Supplementary Material A*). One should also note that the L2 speakers showed a larger proportion of looks to unique distractors in the indefinite condition than in the definite condition

(Trenkic et al., 2014, p. 247, Figure 6). This discrepancy between L1 and L2 speakers indicates that, to predict a referent, L2 speakers might have relied on both the real-world knowledge that a cube cannot be put into a closed can and the non-target-like interlanguage grammar that the indefinite article *a(n)* is a singular marker. This speculation is further supported by Robertson's (2000) observation that L1-Chinese L2-English speakers tend to use *one* in place of an indefinite article.

Ahn (2022) reported that L2 behavior regarding definiteness changes depending on the availability of other usable non-linguistic information. In the absence of real-world knowledge, advanced L2 speakers displayed target-like behavior and predicted a unique referent at the cue of a definite article significantly more than an indefinite article. On the other hand, in the presence of real-world knowledge, they prioritized it over (in)definiteness information to identify a referent. L1 participants, on the other hand, still used (in)definiteness to identify a referent even when real-world knowledge alone could do so. Ahn's (2022) findings were obtained using a referent prediction task and a referent identification task. The former is a type of forced-choice task, while the latter measures participants' reaction time to given stimuli and instructions. Although both tasks were designed to tap into participants' reaction to stimuli without resorting to metalinguistic knowledge, the tasks were not what could observe online responses time-locked to the exact location of a critical word within auditory stimuli.

In sum, Trenkic et al.'s (2014) and Ahn's (2022) findings need to be revisited to investigate the source of the discrepancies. A study designed to exclude real-world knowledge that can observe participants' real-time sentence processing will elucidate *bona fide* L2 grammatical knowledge.

2.3 To predict or not in L2 sentence processing

Another important question that motivates this study is whether L2 speakers can use a nuanced and subtle grammatical feature (e.g, articles) in predictive processing. Recently, many researchers have paid attention to the role of prediction in language processing and acquisition. Some have argued that predictive processing is a necessary mechanism for acquiring language (Kuperberg & Jaeger, 2016; Phillips & Ehrenhofer, 2015), while others have been more skeptical about its role in understanding and learning language (Huettig & Mani, 2016; Kaan & Grüter, 2021).

Among numerous studies on the role of prediction in language processing and acquisition, studies that used articles to probe predictive processing are worth mentioning. Delong et al.'s (2005) seminal work showed that L1 English speakers could predict an unmentioned lexical item based on top-down information (e.g., event structure) made available. A sentence like (5) is likely to end with *a kite* rather than *an airplane*. Delong et al.'s study reported correlations between ERP components and the cloze probabilities of both nouns and articles. N400 decreased not only when the cloze probability of the noun was high (e.g., *kite* over *airplane*) but also when the cloze probability of the article was high (e.g., *a* over *an*). Martin et al. (2015) used a similar scheme to test whether L2 speakers will show sensitivity to the cloze probability of both the noun and the article. They report that L2 speakers' ERP components responded to the nouns but not to the articles.

(5) The day was breezy so the boy went out to fly…

Another study that used articles to investigate predictive processing is Kamide et al. (2003), who used the case-marking system of the German language to test how syntactic and semantic features are integrated. Since German allows variable word orders, the cases marked on the noun phrases can provide cues regarding the thematic role of the referent leading to the prediction of different referents depending on the verb. The sentence below (6) is an example from Kamide et al (2003).

(6) a. Der Hase frißt gleich den Kohl.

The hare-NOM eats shortly the cabbage-ACC.

"The hare will shortly eat the cabbage."

b. Den Hasen frißt gleich der Fuchs.

The hare-ACC eats shortly the fox-NOM.

"The fox will shortly eat the hare."

(6a) has the nominative case on the first NP (NP1) and the accusative on the second NP (NP2) while (6b) has the cases reversed on NP1 and NP2. Such linguistic stimuli were accompanied by visual stimuli in which a hare, cabbage, a fox, and a distractor were presented. German native speakers used both the syntactic and semantic information to predict NP2 before it is uttered. Here, the syntactic information is the cases marked on the article and the noun in NP1, and the semantic information is such real-world knowledge that a hare will more likely eat cabbage than a fox and it is more likely to be eaten by a fox than to eat one. Hopp (2015) and Henry et al. (2020) used a similar scheme and showed that L2 German speakers could use the semantic information but not the morphosyntactic information.

Not all L2 studies reported that L2 speakers could not predict upcoming signals the way L1 speakers do. Dussias et al (2013) and Hopp and Lemmerth (2018) show that L2 speakers can use gender marking in Spanish and case marking in German, respectively, to predict an upcoming referent in the visual world paradigm. What separates the studies in which L2 speakers could not use articles to predict upcoming signals (Henry et al., 2020; Hopp, 2015; Martin et al., 2015) from those in which they could (Dussias et al., 2013; Hopp & Lemmerth, 2018)? The answer lies in the involvement of multiple sources of information. Hopp (2015) and Henrty et al. (2020) both included real-world knowledge that a hare is more likely to eat cabbage than a fox, and the cloze probability used in Martin et al (2015) is also related to real-world knowledge on the relationship between winds and kites.

What should be noted here is that the cloze probability of the nouns used in Martin et al. (2015) and the thematic relation between nouns and verbs in Hopp (2015) and Henry et al (2020), along with real-world knowledge in Ahn (2022), can be labeled top-down information in the umbrella term. At the other end of top-down information such as one's understanding of the world, an event, or a situation is bottom-up information such as a phonetic feature, a sound segment, a syllable, a lexical item or a phrase. Such categorization of different types of information leads to the question whether L2 speakers experience difficulty using top-down information to make a prediction for bottom-up information. These observations, along with the discrepancies observed between Ahn (2021, 2022) and Trenkic et al (2014) in the previous section, motivate this research to separate definiteness from any potential interference of real-world knowledge.

2.4 Research Questions and Predictions

The previous studies described above called for research designed to operationalize the construct of definiteness separated from real-world knowledge and to probe the topography of real-time definiteness processing. The current study, therefore, used visual world paradigm eye-tracking in which L1-Korean L2-English speakers were tested using stimuli that did not involve real-world knowledge. With such a design, the current study aimed to answer the following two questions.

RQ 1. Do L1 and L2 speakers of English use definite and indefinite articles in the same manner to predict uniquely identifiable referents and non-unique referents, respectively, when no real-world knowledge is involved?

RQ 2. Do L2 English speakers of different proficiency levels process (in)definiteness in different manners?

Suppose the language behavior observed in both L1 and intermediate L2 speakers in Trenkic et al. (2014) was attributable to grammatical knowledge, as they claim. In that case, L1 and L2 speakers of the current study should also be able to predictively look at unique and non-unique referents when definite and indefinite articles are given, respectively. However, if the findings of Trenkic et al. (2014) could be, to a certain degree, accounted for by the use of real-world knowledge, L2 speakers, especially at the intermediate level, might not be able to use definiteness information to predict (non-)unique referents.

L1 speakers might also behave differently from those in Trenkic et al. (2014). According to Ahn's recent findings (2021; 2022), indefinite NPs do not play clear roles in accommodating

(Ahn, 2021) or predicting (Ahn, 2022) non-unique referents not only in L2 but also in L1 processing. Indeed, L1 speakers in Ahn (2021; 2022) did not display great sensitivity to the grammaticality of indefinite NPs. If the difference between Trenkic et al. (2014) and Ahn (2021; 2022) comes from the presence and absence of real-world knowledge, L1 speakers might not be able to use indefinite articles to predictively locate a non-unique referent in the current study, which is designed to exclude real-world knowledge.

3 Methods

3.1 Participants

A total of 134 L1 and L2 speakers of English participated in the present study. Forty-six L1-English speakers were recruited in the metropolitan Seoul area via advertisements, and 88 L1-Korean speakers who learned English as a second language were recruited from the Seoul National University community. L1 participants were born and raised in English-speaking countries but were residing in Seoul, Korea, at the time of research. To ensure that they did not have extensive exposure to a language other than English in childhood, we used a questionnaire to screen out those who used another language with their parents or caregivers while growing up and had lived for more than six consecutive months before age 7 in a community where a language other than English was spoken.[2] The same screening criterion was applied to L1-Korean L2-English speakers as well. We invited those who spoke only Korean with parents and caregivers and had not lived outside Korea for more than six consecutive months before age 7. All participants were compensated for their participation monetarily. The L2 population was divided into advanced and intermediate groups using a C-test[3] adapted from Schultz (2006). The demographics and the C-test statistics are summarized in Table 1. L1 speakers' C-test scores ranged from 36 to 40. Using this as a criterion, L2 speakers whose C-test scores were 36 or higher were categorized into the advanced group,[4] and those who scored 34 or lower were placed into the intermediate group.

<Insert Table 1 here>

3.2 Stimuli

Visual stimuli (Figure 1) were used with the auditory stimuli (7) (see *Supplementary Material B* for a complete list of critical stimuli). The auditory stimuli included a color word between the article and the disambiguating noun to allow time to observe whether participants would predictively fixate on the target before the final disambiguating noun was provided. Visual stimuli included competitors of the same color as that of the target to prevent the infelicitousness of the auditory stimuli.

<Insert Figure 1 here>

(7) a. Click on the blue circle.

b. Click on a blue circle.

There were a total of 24 critical stimuli and 48 fillers. Each critical stimulus had two visual conditions (one vs. two target conditions) and two auditory conditions (*the* vs. *a*). The four conditions (2 visual x 2 auditory) for each item was Latin-squared such that each participant saw each item in only one of the four conditions. The location of the target referent was rotated around the five positions on the screen, as in Figure 1. The color and shape combinations used for the critical stimuli were not used for the filler stimuli. The order of items was pseudo-randomized such that every critical item would be followed by two fillers. The mean length of the definite article in the auditory stimuli was 149*ms,* and that of the indefinite article was 115*ms.* The mean length of the color word was 555*ms,* and the shape word was around 716*ms*.

3.3 Procedure

After participants arrived at the lab, they read the guidelines for the experiment procedure and filled out a consent form. The eye-tracking portion of the experiment was conducted before the C-test. In the main experiment, participants' eye gazes were calibrated using nine points on the screen before written instructions were given, after which participants practiced clicking on colors (all given in the same shape) and shapes (all given in the same color). The practice trials were administered to ensure that all participants agreed on the names of the colors and shapes. Another round of calibration was conducted between practice trials and experimental trials. An SMI RED eye-tracking system was used with participants' heads placed on a chin rest, so the location of their eyes was constant throughout the experimental session. The computer screen was approximately 60 *cm* away from the participant's eyes, and their eye gazes were recorded at 120Hz. In every trial, a visual stimulus was given first, and an auditory stimulus followed 2.5 seconds after the onset of the visual stimulus. Between trials, a cross was shown in the center of the screen for participants to click on, which was to ensure that the location of the cursor or participants' eye gazes were not biased to the target area of interest (AOI) of the previous trial (Barr et al., 2011).

In the C-test adapted from Schultz (2006), participants were given 15 minutes to read two short paragraphs and fill 40 blanks in total such that the filled words would fit the contexts. A blank was placed in every other content word. L1 speakers also took the C-test for a comparison purpose. The entire experiment, including the eye-tracking experiment and the C-test, took around 45-50 minutes.

3.4 Analysis

AOIs were labeled as follows in the two conditions. There were two each of competitors and distractors in the one-target condition. Therefore, they were numbered for distinction: competitor1, competitor2, distractor1, and distractor2. In the two-target condition, there were two targets, and the two were distinguished as a selected target (targetS) and an unselected target (targetU). The selected target is the target that was clicked on by the participant between the two target AOIs. Although there was only one competitor in the two-target condition, it was still labeled competitor1. Also, the label targetS was used for the only target in the one-target condition.

Separate models were fit for targets and competitors because we were interested in the effect of (in)definite articles in predicting (non-)unique referents, and also because fixations to different AOIs in the same trial were not independent. Growth curve analysis was used to analyze looks to AOIs over the course of time, following the guidelines of Mirman, Dixon, and Magnuson (2008) and Mirman (2014). Raw proportions of fixations were transformed into empirical logits[5]. We first conducted overall analyses for the 1500*ms* time window from the article onset, separately for each AOI (targets and competitors) in the two visual conditions. The overall models had two versions each. All participants' data were analyzed first with participants' first language (English vs. Korean) as a fixed factor. To see if proficiency of L2 participants has an effect in fixating on targets (and competitors), we also analyzed L2 data separately with L2 participants' proficiency (C-test scores) as a continuous variable. For the overall models, fixation rates were modeled with third-order orthogonal polynomial terms of time and fixed effects of L1[6] (or proficiency in L2 only models[7]) and definiteness on all time terms.

Then, within the 1500*ms* time span, two 500-*ms* periods of interest were selected and labeled Phase 1 and Phase 2. Phase 1 began 200*ms* after the article onset and lasted until the 700*ms* time point. Phase 2 began 800*ms* after the article onset and lasted until the 1300*ms* time point (Please see Figures 2 and 3 for the visualizations of the two phases). The two phases were set this way because it usually takes around 150*ms* to 200*ms* for auditory input to influence eye gaze movements (Matin et al., 1993), and we were interested in finding out whether participants could predict a target referent without hearing the shape word. The 200*ms*-to-700*ms* time slot (Phase 1) is where participants processed information only from the article and the color. For example, when "the blue…" is given in Phase 1, the article information hints at a uniquely identifiable referent while the color information hints at a shape that is blue. In Phase 2, which begins around 100*ms* after the shape word begins, participants will be able to either confirm or revise their prediction made in Phase 1.[8]

For each phase, fixation rates were also modeled with third-order orthogonal polynomial terms of time and fixed effects of definiteness and speaker group on all time terms. The L1 participants in the definite condition was treated as the baseline. All factors were deviation-coded with the baseline as -1 and the comparisons as 1. Due to convergence failures, maximal models with random intercepts and slopes were reduced to include random intercepts only (Barr et al., 2013; Matuschek et al., 2017). All analyses were carried out in R version 3.6.0 (R Core Team, 2019) using packages lme4 version 1.1-23 (Bates et al., 2015) and lmerTest version 3.1.0 (Kuznetsova et al., 2017). Based on these overall models, pairwise comparisons were conducted for AOIs, definiteness conditions, and groups using the lsmeans package version 2.30.0 (Lenth, 2016).

4 Results

The one-target models with L1 as a factor showed significant interaction effects of orthogonal polynomials of time, definiteness, and language (*ps* <.0001). Table S1 in *Supplementary Material C* provide the Type III ANOVA summaries of the lmer models (Kuznetsova et al., 2017), which is useful when interactions are of interest (Fox, 2015; p. 201). The significant interactions mean that participants' looks to targets and competitors differed by time, articles given, and first languages they spoke. However, the-two target models with L1 as a factor showed a significant interaction of time, definiteness, and L1 only in targetU and competitor1. Significant interactions of definiteness and L1 or of time, definiteness, and L1 was not observed for targetS (see Table S2 in *Supplementary Material C*). For the L2-only models with proficiency as a factor, a significant interaction of definiteness and proficiency was observed in the one-target models of the target and the competitors, but a significant interaction of time, definiteness, and proficiency was observed only for competitor1 (see Table S3 in *Supplementary*

Material C). For the two-target L2-only models, significant interactions of definiteness and proficiency and also of time, definiteness, and proficiency were observed in all three AOIs (see Table S4 in *Supplementary Material C*).

Confirming that both L1 and proficiency had significant effects in interactions with either definiteness or with both time and definiteness in the overall analyses, we used groups L1, advanced L2, and intermediate L2 to analyze eye movements in separate speaker groups phase by phase. The type III ANOVA summaries of the models for each AOI in Phases 1 and 2 in the two visual conditions can be found in *Supplementary Material C* (Tables S5 to S8). All models showed significant interactions of definiteness and speaker group (*p*s < .05). Since we were interested in finding out how each group reacted to different articles for different AOIs, we conducted pairwise comparisons to probe the effect of definiteness to fixations on targets and competitors by language group. Our discussion mainly relies on the pairwise analyses comparing eye gaze fixations on different AOIs between the two definiteness conditions per language group (see Tables 2 and 3 for the summary of the pairwise analyses).

Visualizations of all AOIs in different auditory and visual conditions can be found in Figures S2 and S3 in Supplementary Material A. Since our main interests were whether participants' eye gaze patterns differed between definite and indefinite conditions for targets and competitors, Figures 2 and 3 display the empirical logits of fixations on three AOIs in the one-target and two-target conditions, respectively. The leftmost column in each figure reports fixation rates from L1 speakers, the center column is from advanced L2 speakers, and the rightmost column is from the intermediate L2 group. Each row shows fixation rates in targetS, competitor1, and competitor2 from top to bottom in Figure 2 and targetS, targetU, and competitor1 in Figure 3. Solid lines indicate eye movement patterns in the definite condition, while dotted lines represent those in the indefinite condition.

4.1 One Target Condition

In the one-target condition (Figure 1a), a definite article was expected to lead to fixations on the unique target whereas an indefinite article was expected to mislead participants to fixate on one of the two competitors. L1 and advanced L2 speakers showed the effect of the definite article and fixated more on the target in the definite condition than in the indefinite condition, but their behavior regarding the indefinite article and competitors was delayed compared to that towards the target. The intermediate L2 group did not seem to use articles to predict upcoming referents.

<Insert Figure 2 here>

<Insert Table 2 here>

Target

Phase 1. The first row of Figure 2 shows the three speaker groups' eye movement patterns for targetS in the one-target condition. L1 and advanced L2 speakers both displayed more fixations on the unique target in the definite condition than in the indefinite condition during Phase 1 before the disambiguating shape word was heard (L1: $\beta = 0.162$; $SE = 0.049$; $z = 3.337$; $p = .0008$; Adv L2: $\beta = 0.321$; $SE = 0.056$; $z = 5.752$; $p < 0.0001$; see Table 2 for complete statistics). The right most column on the first row shows that intermediate L2 speakers also momentarily fixated more on targetS in the definite condition than in the indefinite condition (β = 0.155; *SE* = 0.051; *z* = 3.018; *p* = .0025). However, the momentary increase in the intermediate L2 group did not show gradual increase unlike the L1 and advanced L2 groups who also showed a statistically significant difference in the linear time term estimate with the definite condition showing a faster increase of fixations than the indefinite condition (L1: β = 36.89; *SE* = 8.78; *z* = 4.202; *p* < .0001; Adv L2: β = 28.91; *SE* = 10.04; *z* = 2.880; *p* = 0.004).

Phase 2. The different patterns of behavior between the intermediate group and the rest became apparent in Phase 2. While L1 and advanced L2 speakers fixated significantly more on the target in the definite condition (L1: β = 0.431; *SE* = 0.045; *z* = 9.641; *p* < .0001; Adv L2: β = 0.269; *SE* = 0.047; *z* = 5.700; *p* < 0.0001), the intermediate L2 group did not show a clear difference between the two conditions (β = 0.0356; *SE* = 0.047; *z* = 0.758; *p* = .4485).

Competitors

Phase 1. Regarding competitors during Phase 1, all three groups behaved differently. L1 speakers did not show a significant difference between the definite and indefinite conditions in fixating on comeptitor1 or competitor2 (*ps* < 1; see Table 2 for complete statistics). Advanced L2 speakers showed an inconsistent pattern of behavior regarding competitors. They did show a higher fixation rates to competitor1 in the indefinite condition (β = -0.329; *SE* = 0.051; *z* = -6.505; *p* < .0001), but no significant difference was observed between the definite and indefinite conditions regarding competitor2 (β = -0.011; *SE* = 0.048; *z* = -0.227; *p* = 0.8200). The intermediate L2 group showed a marginally significant difference in competitor1 (β = - 0.084; *SE* = 0.047; *z* = - 1.802; *p* = .0715) and a statistically significant difference in competitor2 (β = -

0.167; *SE* = 0.044; *z* = -3.793; *p* = .0001), showing higher fixation rates in the indefinite condition.

Phase 2. In Phase 2, L1 and advanced L2 speakers behaved alike regarding the two competitors while the intermediate L2 group's behavior was inconsistent. The indefinite condition showed higher fixation rates in both competitors in both groups (competitor1- L1: β = -0.148; *SE* = 0.044; *z* = -3.385; *p* = .0007; adv: β = -0.125; *SE* = 0.046; *z* = -2.712; *p* = .0067; competitor2 – L1: β = -0.248; *SE* = 0.040; *z* = -6.154; *p* < .0001; adv L2: β = -0.096; *SE* = 0.043; *z* = -2.267; *p* = .0234). The intermediate L2 group also showed a significantly higher rate of fixation on competitor1 in the indefinite condition (β = -0.125; *SE* = 0.046; *z* = -2.712; *p* = .0067), but no difference was observed for competitor2 (β = 0.052; *SE* = 0.042; *z* = 1.219; *p* = .2228).

4.2 Two-target Condition

In the two-target condition (Figure 1b), none of the three speaker groups showed the expected pattern of behavior: More eye gaze fixations were expected on one of the two non-unique targets in the indefinite condition than in the definite condition and the reverse for the unique competitor. The effect of the indefinite article was not clearly observed in any of the three groups, and a very subtle effect of the definite article (to predict the unique competitor) was observed only in the L1 group.

<Insert Figure 3 here>

<Insert Table 3 here>

Targets

Phase 1. L1 and intermediate L2 participants showed an inconsistent pattern of eye movements.[9] As the top left panel of Figure 3 shows, L1 speakers' fixation rates were statistically significantly higher on targetS in the definite condition in Phase 1 ($\beta = 0.216$, $SE = 0.045$, $z = 4.760$, $p < .0001$). For targetU, more fixations, if statistically insignificant, were observed in the indefinite condition for L1 speakers ($\beta = -0.065$, $SE = 0.040$, $z = -1.643$, $p = .1004$). The intermediate group also showed an inconsistent pattern of behavior for targetS and targetU. Their fixation rates for targetS was higher in the definite condition ($\beta = 0.186$, $SE = 0.050$, $z = 3.763$, $p = .0002$) while that for targetU was statistically significantly higher in the indefinite condition ($\beta = -0.158$, $SE = 0.043$, $z = -3.667$, $p = .0002$). Advanced L2 speakers showed the opposite behavior from what was predicted, and their fixation rates were statistically significantly higher in the definite condition for both targetS ($\beta = 0.199$, $SE = 0.051$, $z = 3.919$, $p = .0001$) and targetU ($\beta = 0.173$, $SE = 0.044$, $z = 3.899$, $p = .0001$).

Phase 2. The incoherent behavioral pattern of L1 and intermediate L2 speakers continued in Phase 2. L1 speakers did not reach statistical significance regarding the difference between definite and indefinite conditions for targets during Phase 2 and continued to show the inconsistent behavior. This time, only the conditions for higher fixations were reversed from Phase 1. Their fixation rate on targetS was numerically higher in the indefinite condition ($\beta = -0.077$, $SE = 0.048$, $z = -1.626$, $p = .1039$), but it was numerically higher in the definite condition for targetU ($\beta = 0.051$, $SE = 0.034$, $z = 1.534$, $p = .1250$). The intermediate group maintained their behavioral pattern from Phase 1 and fixated on targetS significantly more in the definite condition (β = 0.2878, *SE* = 0.0555, z = 5.180, p < .0001) while fixating on targetU significantly more in the indefinite condition (β = -0.2293, *SE* = 0.0391, z = -5.867, p < .0001). Advanced L2 speakers showed a statistically significantly higher fixation rate for targetS in the indefinite condition (β = -0.151, *SE* = 0.052, z = -2.912, p = .0036), but such statistical significance was not observed for targetU (β = -0.027, *SE* = 0.037, z = -0.725, p = .4682).

Competitor

Phase 1. As for L1 speakers, a temporary increase of fixation rate was observed in the definite condition (β = 0.089, *SE* = 0.044, z = 2.013, p = .0441). On the other hand, both advanced and intermediate L2 speakers showed a statistically significantly higher fixation rate on competitor1 in the indefinite condition (Advanced: β = -0.257, *SE* = 0.050, z = -5.163, p < .0001; Intermediate: β = -0.261, *SE* = 0.049, z = -5.378, p < .0001).

Phase 2. The three speaker groups all behaved differently as for the competitor during Phase 2. L1 speakers fixated more on competitor1 in the definite condition (β = 0.101, *SE* = 0.034, z = 2.338, p = .0194) while advanced L2 speakers did not display any statistical difference (β = 0.0602, *SE* = 0.0471, z = 1.277, p = .2015) and intermediate L2 speakers fixated significantly more on competitor1 in the indefinite condition (β = -0.252, *SE* = 0.050, z = -5.000, p < .0001).

5 Discussion

The aims of this study were two-fold: (1) to investigate whether (in)definiteness without the interference of real-world knowledge could lead to the prediction of (non-)unique referents in a visual scene in L1 and L2 sentence processing and (2) to examine whether L2 speakers' eye movements in response to definite and indefinite articles would pattern differently depending on their proficiency. The results of the experiment showed that L1 speakers and advanced L2 speakers used the definite article in the one-target condition to anticipatorily fixate on targetS (the unique target) during Phase 1 before the disambiguating noun was heard. Also, the higher fixation rate on targetS in the definite condition continued until the disambiguating noun was fully processed towards the end of Phase 2. Such a pattern of eye movements, however, was not observed in the intermediate L2 group. In the two-target condition, none of the three speaker groups formed a coherent pattern of eye movements that could be interpreted as signs of using (in)definite articles to predict (non-)unique referents except the L1 group's subtle sign of using the definite article to predict the unique competitor.

Our results point back to the issue raised at the beginning of the paper: Can intermediate L2 speakers predict a referent solely based on definiteness without resorting to real-world knowledge? The intermediate L2 group in the present study did not use English articles in predicting upcoming linguistic material, unlike the participants in Trenkic et al. (2014), who were intermediate-level Chinese L2 speakers of English. The proficiency levels of the intermediate learner groups in the two studies cannot be directly compared since different proficiency measures were employed. In any case, the findings from the intermediate L2 group of the current study suggest that the results of Trenkic et al. (2014) be carefully re-interpreted in discussing L2 speakers' ability to use grammatical knowledge (definiteness) in sentence processing. If the experimental design of Trenkic et al. (2014) had not involved real-world knowledge, the intermediate-level L1-Chinese L2-English speakers might not have been able to locate a target as quickly as they did.

Another difference between our findings and those of the previous studies is that an indefinite article did not induce a clear predictive processing effect even in L1 or advanced L2 speakers of the current study. In Chambers et al. (2002) and Trenkic et al. (2014), L1 participants could locate a target the fastest with the combination of an indefinite article and a non-unique target. This discrepancy should be attributable to the different research purposes and methodologies between the current study and the two previous studies. We aimed to investigate the role of (in)definiteness only in predicting (non-)unique referents without the influence of real-world knowledge. On the contrary, Chambers et al. (2002) focused on the function of situational contexts or real-world knowledge involved in definiteness processing. Therefore, Trenkic et al.'s (2014) argument that their findings indicate L2 speakers' using grammatical knowledge of definiteness needs reconsidering.

A possible account for the delayed effect of an indefinite article could be that indefinite articles might not necessarily predict a non-unique referent. Clifton's (2013) study on L1 definiteness processing suggested that when a uniquely identifiable referent or a non-unique referent is presupposed in a given context, definite and indefinite NPs, respectively, will take an equal amount of processing loads and time to accommodate the presupposition. This was not replicated in Ahn (2021), where for both L1 and advanced L2 groups, indefinite NPs did not lead to a clear contrast in accommodating the presupposition of unique and non-unique referents. That is, indefinite NPs did not elicit different processing times in response to the (non-)uniqueness of a referent to the extent that definite NPs did.

Before we discuss the theoretical implications of our findings, a nagging question remains: Is it possible that L2 speakers could not aurally distinguish *the* from *a*? We did not conduct a separate test to see whether that was a case. However, assuming the perceptual indistinction between the two articles in the L2 groups will make it difficult to produce the most parsimonious account for our findings. Advanced L2 speakers clearly showed the effect of definiteness in predicting unique referents. Even for intermediate speakers alone, separate accounts will be required to explain their behaviors in different AOIs, visual settings, and auditory conditions since their behavior, if not targetlike, differed between one-target and two-target conditions, amongst different AOIs, and between the definite and indefinite conditions. If they could not distinguish definite and indefinite articles, they would not have shown different responses to the same visual scene when the only difference was the articles (see Figures S2 and S3 in *Supplementary Material A* for the overall view of their eye gaze patterns).

Comparing the findings of the current study and those of previous studies (Ahn, 2021; 2022; Henry et al., 2020; Hopp, 2015; Martin et al., 2015; Trenkic et al*.,* 2014) has implications for recent theories in predictive L2 processing. Hopp (2021) lists constraints that make cues in predictive L2 processing unreliable such as "lacking and inconsistent lexical and grammatical knowledge, L1 influence, shallow parsing, and cognitive resource limitations (p. 180)," and Kaan and Grüter (2021) argues that unreliable cues decrease the utility of prediction in L2 sentence processing. We believe that, based on our findings, the presence of top-down information (e.g., real-world knowledge) should be considered an important factor in calculating the utility of prediction. At lower proficiency levels, L2 speakers might lack lexical and/or grammatical knowledge in L2 to base their predictions on. If top-down information, such as knowledge that a rabbit is more likely to eat cabbage than a fox (Henry et al., 2020; Hopp, 2015) or that a closed can cannot be a viable goal for a cube to be placed in (Trenkic et al., 2014), is available, L2 speakers, even when not very advanced, will be able to use such knowledge to make a prediction for how the sentence will unfold, which was the case in Trenkic et al. (2014).

At more advanced levels, if top-down information provides cues clear enough to generate expectations, the grammatical redundancy of nuanced and subtle features such as articles might lack utility for L2 speakers to predict what is to be heard/read next. When L2 speakers do not use articles to make predictions, it is not just because unreliable cues will not contribute to learning, but also because it is suboptimal to process additional information when more explicit and less subtle information is available to make predictions necessary for communication.

6 Conclusion

To sum up, a definite article clearly has an effect in predicting a uniquely identifiable referent yet to be mentioned in L1 and advanced L2 sentence processing. On the other hand, the effect of an indefinite article was observed only in limited circumstances in a delayed manner. These behavioral patterns were observed in both L1 and advanced L2 speakers, but intermediate L2 speakers did not show any meaningful differences in their responses to definite and indefinite articles in either of the unique and non-unique target conditions.

A future study is called for to address the following issues. Our findings indicate that advanced L2 speakers' eye gazes were moving more quickly and to a greater extent in response to indefinite articles and color words (observed in the one-target condition) than L1 speakers. Investigating potential causes of such L2 behavior will shed light on the nature of L2 processing that sets it apart from L1 processing. Also, although we believe that the effect of definiteness was not confounded with that of color words' lexical information, an experimental design that can keep a distance between an article and a disambiguating noun without involving additional information (i.e., color in the current study) will lead to a more straightforward interpretation of participants' eye movements.

Supplementary Materials

A: Figures S1 to S3

B: List of critical stimuli

C: Tables S1 to S8

Competing Interests Declaration

The authors declare no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

References

Ahn, H. (2021). From Interlanguage grammar to target grammar in L2 processing of definiteness as uniqueness. *Second Language Research, 37*(1), 91-119. https://doi.org/10.1177/0267658319868003

Ahn, H. (2022). L2 processing of linguistic and non-linguistic information: L2 speakers use definiteness if real-world knowledge is unusable. *Studies in Second Language Acquisition, 44*(2), 507-535. https://doi.org/10.1017/S0272263121000322

Ariel, M. (1991). The function of accessibility in a theory of grammar. *Journal of Pragmatics, 16*(5), 443-463. https://doi.org/https://doi.org/10.1016/0378-2166(91)90136-L

Barr, D. J., Gann, T. M., & Pierce, R. S. (2011, Jun). Anticipatory baseline effects and information integration in visual world studies. *Acta Psychologica (Amst), 137*(2), 201-207. https://doi.org/10.1016/j.actpsy.2010.09.011

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language, 68*(3), 255–278. https://dx.doi.org/10.1016/j.jml.2012.11.001

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software, 67*(1), 1–48.

Butler, Y. G. (2002). Second language learners' theories on the use of English articles. *Studies in Second Language Acquisition, 24*(3), 451-480.

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 and Language, 47*(1), 30-49. https://doi.org/10.1006/jmla.2001.2832

Clahsen, H., & Felser, C. (2006). Grammatical processing in language learners. *Applied Psycholinguistics, 27*(1), 3-42.

Clahsen, H., & Felser, C. (2018). Some Notes on the Shallow Structure Hypothesis. *Studies in Second Language Acquisition, 40*(3), 693-706. https://doi.org/10.1017/s0272263117000250

Clifton Jr., C. (2013). Situational context affects definiteness preferences: Accommodation of presuppositions. *Journal of Experimental Psychology: Learning, Memory and Cognition, 39*(2), 487-501. https://doi.org/10.1037/a0028975

Conklin, K., & Pellicer-Sanchez, A. (2016). Using eye-tracking in applied linguistics and second language research. *Second Language Research, 32*(3), 453-467. https://doi.org/10.1177/0267658316637401

Cunnings, I. (2017). Parsing and Working Memory in Bilingual Sentence Processing. *Bilingualism: Language and Cognition, 20*(4), 659-678. https://doi.org/10.1017/s1366728916000675

DeLong, K. A., Urbach, T. P., & Kutas, M. (2005). Probabilistic word pre-activation during language comprehension inferred from electrical brain activity. *Nature Neuroscience, 8*(8), 1117–1121.

Dink, J., & Ferguson, B. (2018). eyetrackingR: An R Library for Eye-tracking Data Analysis. Retrieved from http://www.eyetrackingr.com

Dussias, P. E. (2010). Uses of Eye-Tracking Data in Second Language Sentence Processing Research. *Annual Review of Applied Linguistics, 30*, 149-166. https://doi.org/10.1017/s026719051000005x

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. https://doi.org/10.1017/s0272263112000915

Fox, J. (2015). *Applied regression analysis and generalized linear models*. Sage Publications.

Grüter, T., Rohde, H., & Schafer, A. J. (2017). Coreference and discourse coherence in L2. *Linguistic Approaches to Bilingualism*, *7*(2), 199-229. https://doi.org/10.1075/lab.15011.gru

Gundel, J. K., Hedberg, N., & Zacharski, R. (1993). Cognitive status and the form of referring expressions in discourse. *Language, 69*(2), 274-307.

Hawkins, J. A. (1991). On (In)definite Articles: Implicatures and (un)grammaticality prediction. *Journal of Linguistics, 27*(2), 405-442. https://doi.org/10.2307/4176125

Heim, I. (1982). *The semantics of definite and indefinite noun phrases*. (Doctor of Philosophy). University of Massachusetts, Amherst, MA.

Henry, N., Jackson, C. N., & Hopp, H. (2020). Cue coalitions and additivity in predictive processing: The interaction between case and prosody in L2 German. *Second Language Research*, *38*(3), 397-422. https://doi.org/10.1177/0267658320963151

Hopp, H. (2015). Semantics and morphosyntax in predictive L2 processing. *International Review of Applied Linguistics in Language Teaching, 23*(3), 277-306. https://doi.org/10.1515/iral-2015-0014

Hopp, H. (2021). Prediction and grammatical learning in second language sentence processing. In E. Kaan & T. Gruter (Eds.), *Prediction in Second Language Processing and Learning* (Vol. 12, pp. 167). John Benjamins.

Hopp, H., & Lemmerth, N. (2018). Lexical and Syntactic Congruency in L2 Predictive Gender Processing. *Studies in Second Language Acquisition*, *40*(1), 171-199.

https://doi.org/10.1017/s0272263116000437

Huebner, T. (1983). *A longitudinal analysis of the acquisition of English*. Karoma Publishers.

Huettig, F., & Mani, N. (2015). Is prediction necessary to understand language? Probably not. *Language, Cognition and Neuroscience*, *31*(1), 19-31. https://doi.org/10.1080/23273798.2015.1072223

Ionin, T., Choi, S. H., & Liu, Q. (2021). Knowledge of indefinite articles in L2-English: Online vs. offline performance. *Second Language Research, 37*(1), 121-160. https://doi.org/10.1177/0267658319857466

Ionin, T., Ko, H., & Wexler, K. (2004). Article Semantics in L2 Acquisition: *The Role of Specificity. Language Acquisition, 12*(1), 3-69. https://doi.org/10.1207/s15327817la1201_2

Ionin, T., Zubizarreta, M. L., & Maldonado, S. B. (2008). Sources of linguistic knowledge in the second language acquisition of English articles. *Lingua, 118*(4), 554-576. https://doi.org/10.1016/j.lingua.2006.11.012

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. https://doi.org/10.1017/s1366728909990149

Kaan, E., & Grüter, T. (2021). Prediction in second language processing and learning: Advances and directions. In E. Kaan & T. Grüter (Eds.), *Prediction in second language processing and learning* (Vol. 12). John Benjamins.

Kamide, Y., Altmann, G. T. M., & Haywood, S. L. (2003). The time-course of prediction in incremental sentence processing: Evidence from anticipatory eye movements. *Journal of Memory and Language, 49*(1), 133-156. https://doi.org/10.1016/s0749-596x(03)00023-8

Kamide, Y., Scheepers, C., & Altmann, G. T. (2003). Integration of syntactic and semantic information in predictive processing: cross-linguistic evidence from German and English. *Journal of Psycholinguistic Research, 32*(1), 37–55.

Ko, H., Ionin, T., & Wexler, K. (2010). The role of presuppositionality in the second language acquisition of English articles. *Linguistic Inquiry, 41*(2), 213-254. https://doi.org/10.1162/ling.2010.41.2.213

Kuperberg, G. R., & Jaeger, T. F. (2016). What do we mean by prediction in language comprehension? *Language, Cognition and Neuroscience*, *31*(1), 32-59. https://doi.org/10.1080/23273798.2015.1102299

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest Package: Tests in Linear Mixed Effects Models. *Journal of Statistical Software, 82*(13). https://doi.org/10.18637/jss.v082.i13

Lenth, R. V. (2016). Least-Squares Means: The R Package lsmeans. *Journal of Statistical Software, 69*(1). https://doi.org/10.18637/jss.v069.i01

Martin, C. D., Garcia, X., Breton, A., Thierry, G., & Costa, A. (2015). World knowledge integration during second language comprehension. *Language, Cognition and Neuroscience, 31*(2), 206-216. https://doi.org/10.1080/23273798.2015.1084012

Master, P. A. (1987). *A cross-linguistic interlanguage analysis of the acquisition of the English article system,* University of California. Los Angeles.

Master, P. (1997). The English article system: acquisition, function and pedagogy. *System, 25*(2), 215-232.

Matin, E., Shao, K. C., & Boff, K. (1993). Saccadic overhead: Information-processing time with and without saccades. *Perception & Psychophysics, 53*(4), 372-380.

Matuschek, H., Kliegl, R., Vasishth, S., Baayen, H., & Bates, D. (2017). Balancing Type I error and power in linear mixed models. *Journal of Memory and Language, 94*, 305-315. https://doi.org/10.1016/j.jml.2017.01.001

Mirman, D. (2014). *Growth Curve Analysis and Visualization Using R*. Chapman and Hall / CRC Press.

Mirman, D., Dixon, J. A., & Magnuson, J. S. (2008). Statistical and computational models of the visual world paradigm: Growth curves and individual differences. *Journal of Memory and Language, 59*(4), 475-494. https://doi.org/10.1016/j.jml.2007.11.006

Murphy, G. L. (1984). Establishing and accessing referents in discourse. Memory & *Cognition, 12*(5), 489-497. https://doi.org/10.3758/BF03198311

Phillips, C., & Ehrenhofer, L. (2015). The role of language processing in language acquisition. *Linguistic Approaches to Bilingualism, 5*(4), 409–453.

Prince, E. F. (1992). The ZPG Letter: Subjects, definiteness, and information-status. In W. C. Mann & S. A. Thompson (Eds.), *Discourse Description: Diverse linguistic analyses of a fund-raising text* (pp. 295-325). John Benjamins.

R Core Team. (2019). R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from https://www.R-project.org.

Roberts, L., & Siyanova-Chanturia, A. (2013). Using Eye-Tracking to Investigate Topics in L2 Acquisition and L2 Processing. *Studies in Second Language Acquisition, 35*(2), 213-235. https://doi.org/10.1017/s0272263112000861

Robertson, D. (2000). Variability in the use of the English article system by Chinese learners of English. *Second Language Research, 16*(2), 135-172. https://doi.org/10.1191/026765800672262975

Russell, B. (1905). On denoting. *Mind, 14*(56), 479-493.

Schultz, B. (2006). *Wh-scope marking in English interlanguage grammars: Transfer and processing effects on the second language acquisition of complex wh-questions*. (Doctoral dissertation). University of Hawaii at Manoa, Honolulu, HI.

Schumacher, P. B. (2009). Definiteness marking shows late effects during discourse processing: Evidence from ERPs. In S. L. Devi, A. Branco, & R. Mitkov (Eds.), *Anaphora processing and applications, Lecture notes in artificial intelligence* (Vol. 5847, pp. 91-106). Heidelberg: Springer.

Sorace, A. (2011). Pinning down the concept of "interface" in bilingualism. *Linguistic Approaches to Bilingualism, 1*(1), 1-33. https://doi.org/10.1075/lab.1.1.01sor

Sorace, A., & Filliaci, F. (2006). Anaphora resolution in near-native speakers of Italian. *Second Language Research, 22*(3), 339-368. https://doi.org/10.1191/0267658306sr271oa

Snape, N. (2009). Exploring Mandarin Chinese speakers' L2 article use. In N. Snape, Y.-k. I. Leung, & M. Sharwood Smith (Eds.), *Representational Deficits in SLA: Studies in honor of Roger Hawkins* (pp. 27-52): John Benjamins Publishing.

Snape, N., Leung, Y.-k. I., & Ting, H.-C. (2006). Comparing Chinese, Japanese, and Spanish speakers in L2 English article acquisition: Evidence against the Fluctuation Hypothesis? In M. G. O'Brien, C. Shea, & J. Archibald (Eds.), *Proceedings of the 8th Generative Approaches to Second Language Acquisition Conference* (pp. 132-139). Somerville, MA: Cascadilla Proceedings Project.

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. https://doi.org/10.1017/s1366728913000321

Notes

1. Please note that the examples in (4) are from Trenkic et al. (2013), where they presented visual stimuli on a computer screen. In Chambers et al. (2002), experimental props were presented as realia, and participants were asked to move the objects as they listened to instructions such as "Put the cube inside the can."
2. Comparing L1 participants' mean age and their mean length of residence in an English-speaking country in Table 1 clarifies that L1 participants lived most of their lives in English-speaking countries.
3. The C-test we used was not a standardized test; however, we used a standardized test called TEPS (Test of English Proficiency developed by Seoul National University) in recruiting participants when they had a test score. Participants' C-test scores and TEPS scores showed a strong correlation (r = .75; t = 11.27; p < .0001). We used the C-test scores for data analysis to compare L1 and L2 participants. Since all L1 participants and some L2 participants did not have TEPS scores, we had all participants take the C-test after the main experiment. When the available TEPS scores were calculated, the mean of the advanced L2 group was 910 out of 990 maximum possible points, and that of the intermediate L2 group was 714.
4. A reviewer pointed out that the advanced group who showed indistinguishable C-test scores from L1 speakers should be labeled as near-native speakers, but C-test results alone could not identify near-native speakers. We would like to remain conservative about the matter and simply label them 'advanced.' A concern was also raised that Korean children start their English education earlier than those in Europe or North America, who usually begin their foreign language instruction at age 8. Most Korean children begin their

compulsory English education with 2-3 hours of exposure to classroom English at around age 9. Affluent families send their children to English immersion kindergartens and extra-curricular classes; therefore, as in many EFL settings, language learning experience in Korea dramatically varies depending on one's socio-economic status. However, such high-quality English education is not available to all Korean children. Especially those currently in their twenties (our participants) are not such a homogenous group who had exposure to a larger quantity and better quality of English education than Europeans.

5. Because logistic GCA must be used to model binary data of fixations, calculating logits often results in undefined functions (Mirman, 2014, p. 110). An empirical logit transformation was applied to the raw data to avoid such cases, and the calculation was performed using the eyetrackingR package (Dink & Ferguson, 2015).

6. elog ~ poly(Time, 3) * def * L1 + (1|participant) + (1|item)

7. elog ~ poly(Time, 3) *def*proficiency + (1|participant) + (1|item)

8. This resulted in a gap of 100*ms* between 700 and 800*ms* after the onset of the article. Since the 700*ms* time point was where the color word ended and the shape word began, we thought excluding the shape word from Phase 1 was necessary to observe the effect of definiteness without giving a clue of the critical word. Also, one might wonder whether the shorter length of an indefinite article (mean length = 115*ms*) than that of the definite article (mean length = 149*ms*) led to earlier disambiguation in the indefinite condition. The combined length of article and color word in the indefinite condition was 670*ms,* while that in the definite condition was 704*ms*. However, Phase 1 ended at 700*ms* after the onset of the article, and the last 30*ms* at the end of Phase 1 window could not have contributed to the patterns observed in Phase 1, thus, its statistical output.

9. One might wonder, in the two-target condition, looks to targetU in the indefinite condition should also be considered as evidence of using an indefinite article to predict a referent. We also considered combining looks to both targetS and targetU to calculate looks to target; however, we did not do so because it would create an imbalance across conditions. In the definite condition where going back and forth between the two targets cannot be the effect of a definite article, counting looks to both targetS and targetU would not have made sense, and only targetS should have been counted. Combining looks to both targets in the indefinite condition would have made 40% of all shapes on the visual scene candidates for target referents, while only 20% of all referents would have been targets in the definite condition.

List of Figures

Table 1. Participant demographics

		L1 (*n* = 46; f = 26)	L2 advanced (n = 44; f = 32)	L2 intermediate (n = 44; f = 22)
Age	Mean (*SD*)	28.9 (8.7)	21.6 (2.2)	22.6 (2.8)
	Range	19 - 54	18 - 26	18 - 33
LOR (years)	Mean (*SD*)	25.4 (7)	0.7 (0.9)	0.1 (0.3)
	Range	17 - 50	0 - 3.8	0 - 1
C-test	Mean (*SD*)	38.5 (1.6)	37.8 (1.3)	29.8 (2.9)
Max=40	Range	36 - 40	36 - 40	21 - 34

Note. LOR = Length of residence in an English-speaking country

Table 2. Pairwise comparisons between definite and indefinite conditions on the target and the competitors in the one-target condition

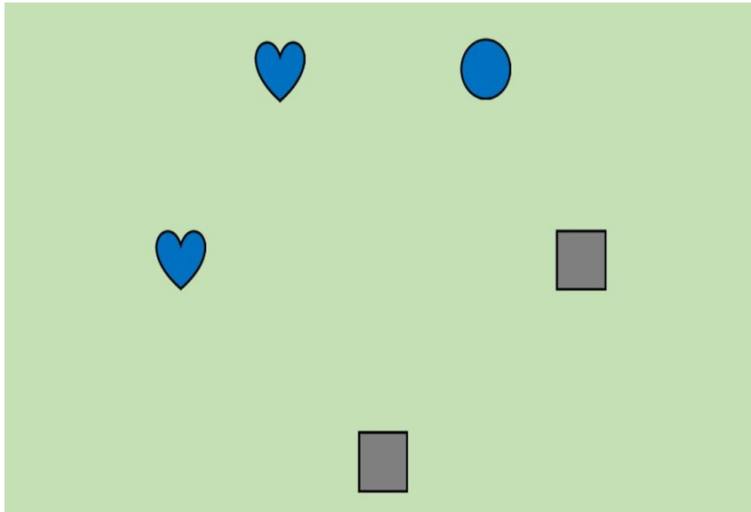
	AOI	Group		estimate	SE	z.ratio	p.value
Phase 1	targetS	L1	intercept	0.162	0.049	3.337	0.0008
			linear	36.890	8.780	4.202	<.0001
		Advanced	intercept	0.321	0.056	5.752	<.0001
			linear	28.910	10.040	2.880	0.0040
		Intermediate	intercept	0.155	0.051	3.018	0.0025
	competitor1	L1	intercept	0.048	0.044	1.088	0.2767
			linear	-21.195	7.950	-2.665	0.0077
		Advanced	intercept	-0.329	0.051	-6.505	<.0001
			linear	30.451	9.100	3.348	0.0008
		Intermediate	intercept	-0.084	0.047	-1.802	0.0715
	competitor2	L1	intercept	0.006	0.042	0.155	0.8766
		Advanced	intercept	-0.011	0.048	-0.227	0.8200
		Intermediate	intercept	-0.167	0.044	-3.793	0.0001
			linear	16.750	7.940	2.110	0.0348
Phase 2	targetS	L1	intercept	0.431	0.045	9.641	<.0001
			linear	30.293	8.190	3.698	0.0002
			quadratic	-251.411	65.190	-3.856	0.0001
		Advanced	intercept	0.269	0.047	5.700	<.0001
			linear	-31.466	8.870	-3.547	0.0004
			quadratic	-153.657	70.030	-2.194	0.0282
		Intermediate	intercept	0.036	0.047	0.758	0.4485
	competitor1	L1	intercept	-0.148	0.044	-3.385	0.0007
			cubic	855.870	412.320	2.076	0.0379
		Advanced	intercept	-0.125	0.046	-2.712	0.0067
			quadratic	256.410	68.690	3.733	0.0002
		Intermediate	intercept	-0.125	0.046	-2.712	0.0067
			linear	39.240	8.670	4.525	<.0001
			quadratic	150.450	68.440	2.198	0.0279
	competitor2	L1	intercept	-0.248	0.040	-6.154	<.0001
			quadratic	224.139	58.770	3.814	0.0001
		Advanced	intercept	-0.096	0.043	-2.267	0.0234
			linear	25.859	8.000	3.233	0.0012
		Intermediate	intercept	0.052	0.042	1.219	0.2228
			linear	-34.012	7.970	-4.267	<.0001

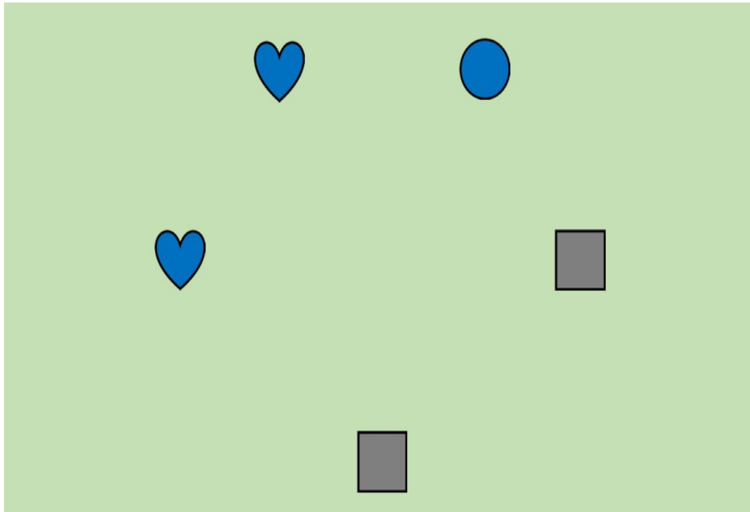
Note. Positive values indicate larger estimates in the definite condition. Only statistically significant time terms were listed.

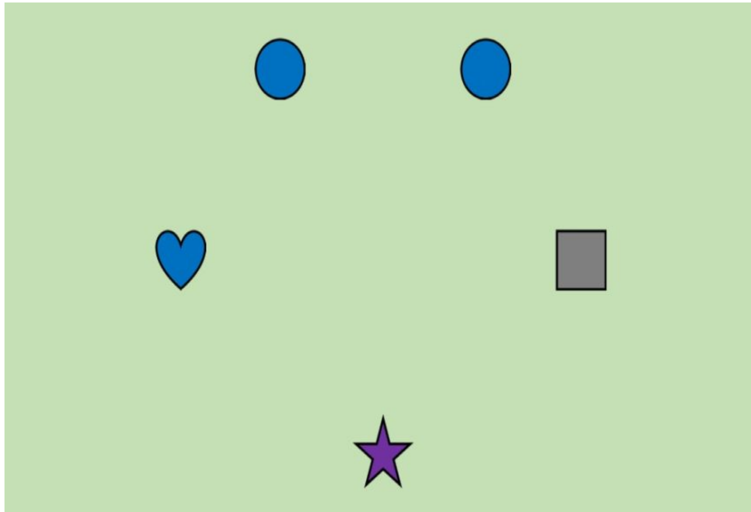
Table 3. Pairwise comparisons between definite and indefinite conditions on the targets and the competitor in the two-target condition

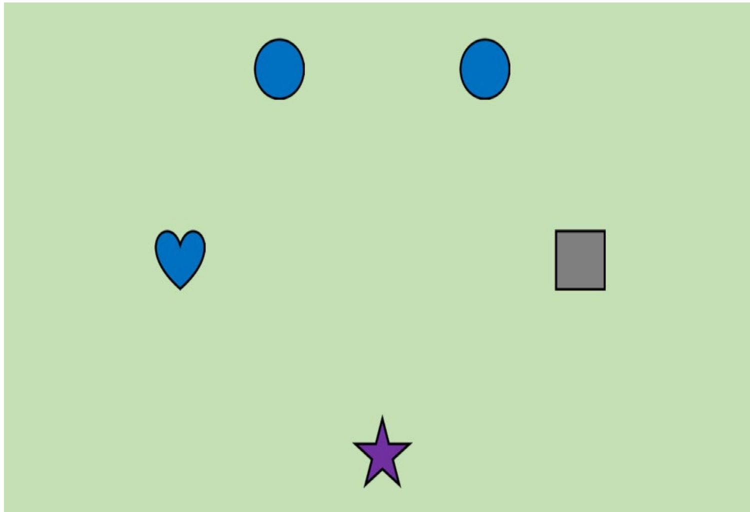
				estimate	SE	z.ratio	p.value
Phase 1	targetS	L1	intercept	0.216	0.045	4.760	<.0001
		Advanced	intercept	0.199	0.051	3.919	0.0001
			quadratic	-219.880	72.790	-3.021	0.0025
		Intermediate	intercept	0.186	0.050	3.763	0.0002
	targetU	L1	intercept	-0.065	0.040	-1.643	0.1004
		Advanced	intercept	0.173	0.044	3.899	0.0001
		Intermediate	intercept	-0.158	0.043	-3.667	0.0002
			linear	42.100	7.930	5.312	<.0001
			cubic	-1480.410	406.720	-3.640	0.0003
	competitor1	L1	intercept	0.089	0.044	2.013	0.0441
		Advanced	intercept	-0.257	0.050	-5.163	<.0001
			linear	24.180	9.000	2.687	0.0072
			quadratic	203.040	71.330	2.847	0.0044
		Intermediate	intercept	-0.261	0.049	-5.378	<.0001
			linear	-71.420	8.920	-8.008	<.0001
Phase 2	targetS	L1	intercept	-0.077	0.048	-1.626	0.1039
		Advanced	intercept	-0.151	0.052	-2.912	0.0036
			linear	-25.150	9.670	-2.602	0.0093
		Intermediate	intercept	0.288	0.056	5.180	<.0001
			quadratic	-195.030	79.810	-2.444	0.0145
	targetU	L1	intercept	0.051	0.034	1.534	0.1250
		Advanced	intercept	-0.027	0.037	-0.725	0.4682
			linear	-33.790	6.800	-4.966	<.0001
		Intermediate	intercept	-0.229	0.039	-5.867	<.0001
	competitor1	L1	intercept	0.101	0.043	2.338	0.0194
			linear	14.950	8.000	1.868	0.0618
		Advanced	intercept	0.060	0.047	1.277	0.2015
			linear	52.820	8.770	6.025	<.0001
		Intermediate	intercept	-0.252	0.050	-5.000	<.0001
			linear	24.230	9.120	2.656	0.0079
			quadratic	182.580	72.390	2.522	0.0117

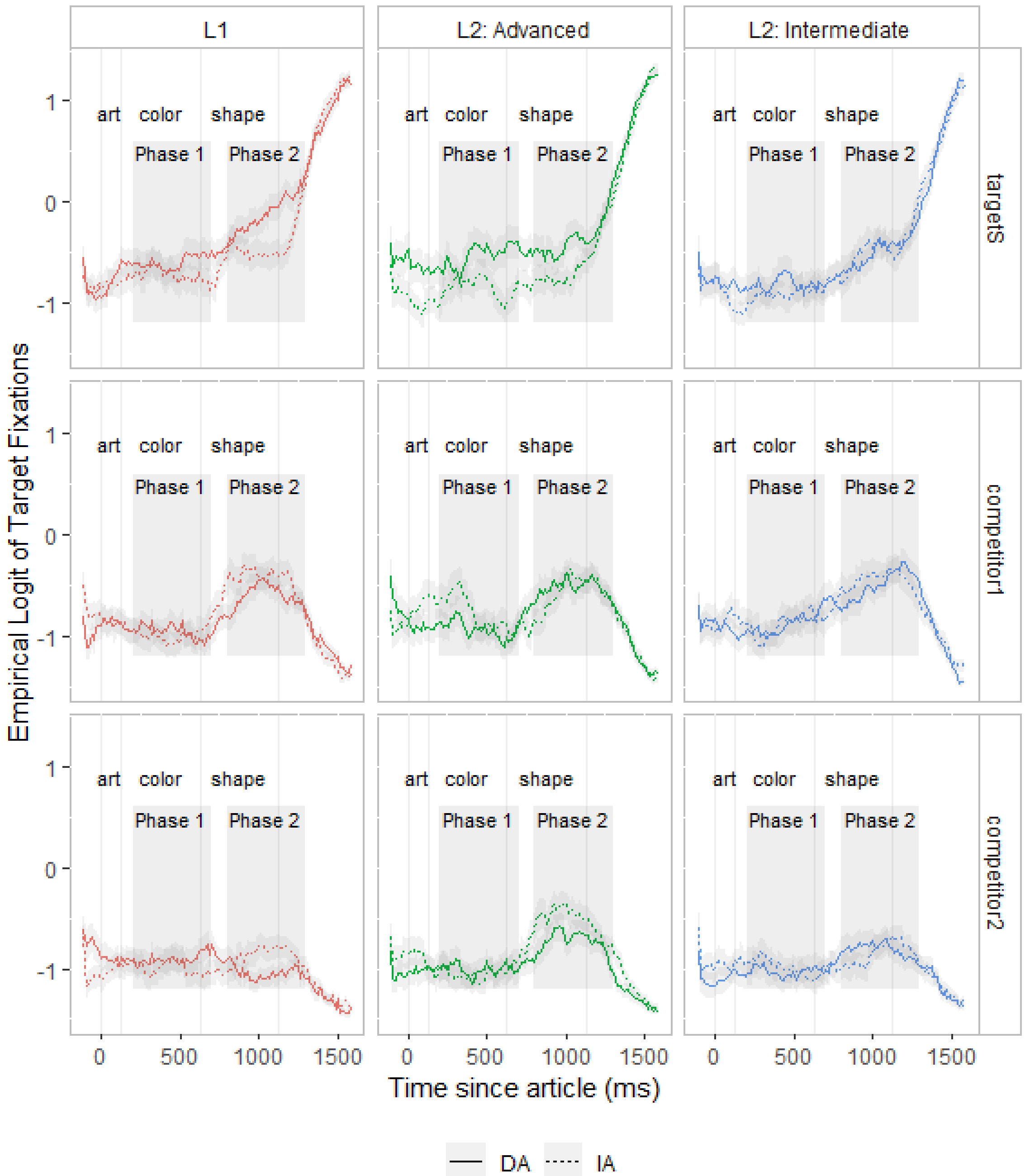
Note. Positive values indicate larger estimates in the definite condition. Only statistically significant time terms were listed.

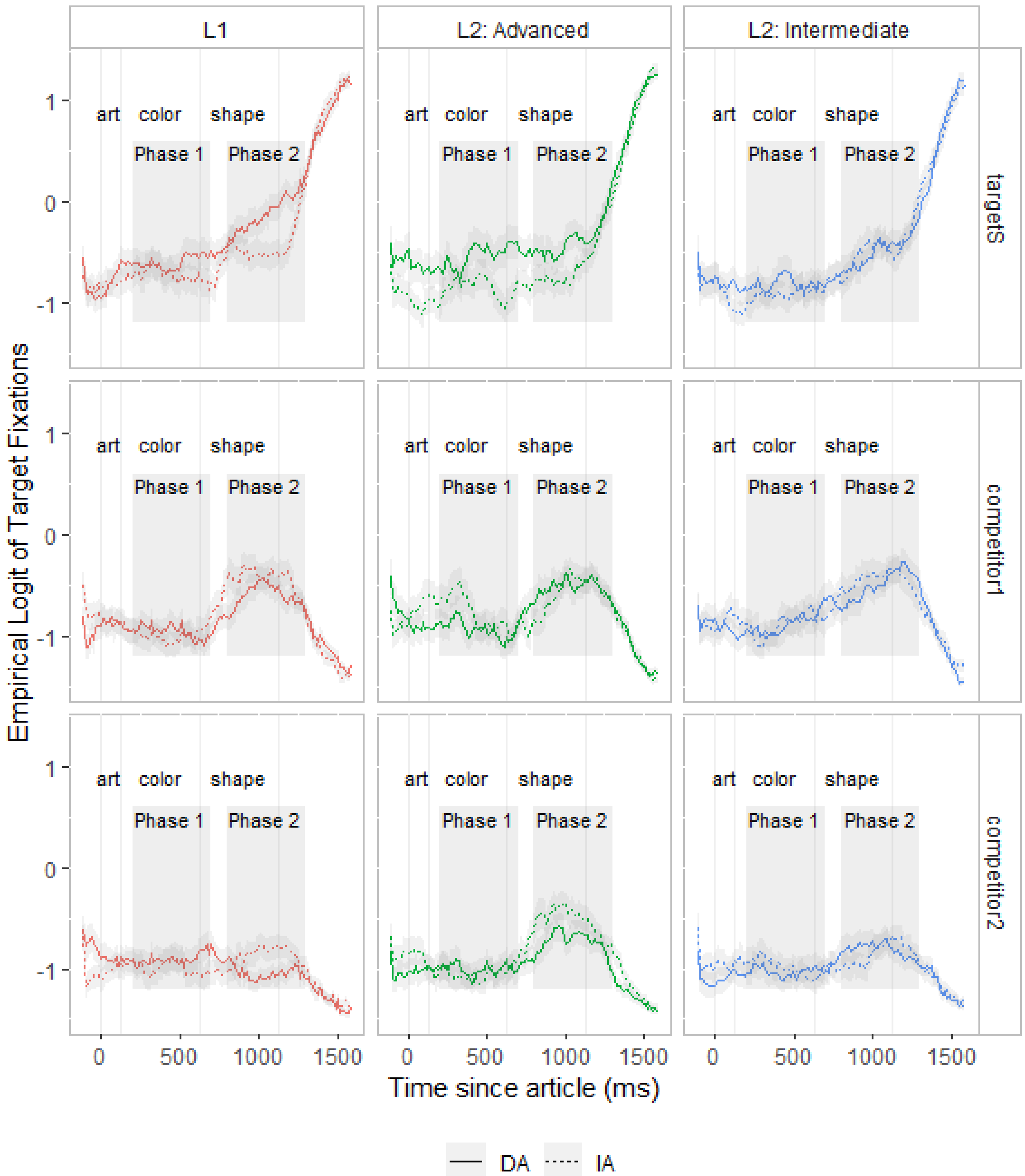


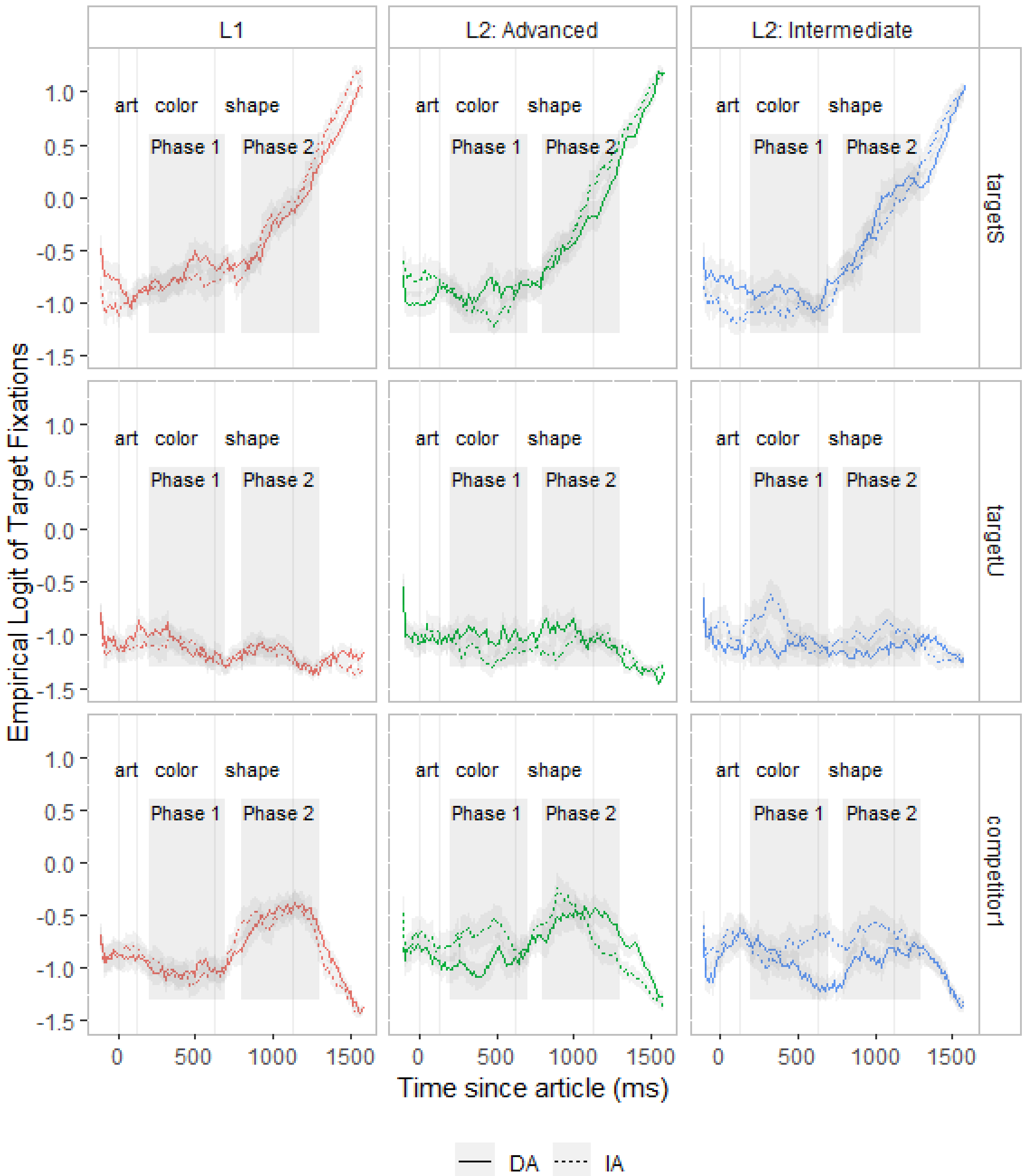


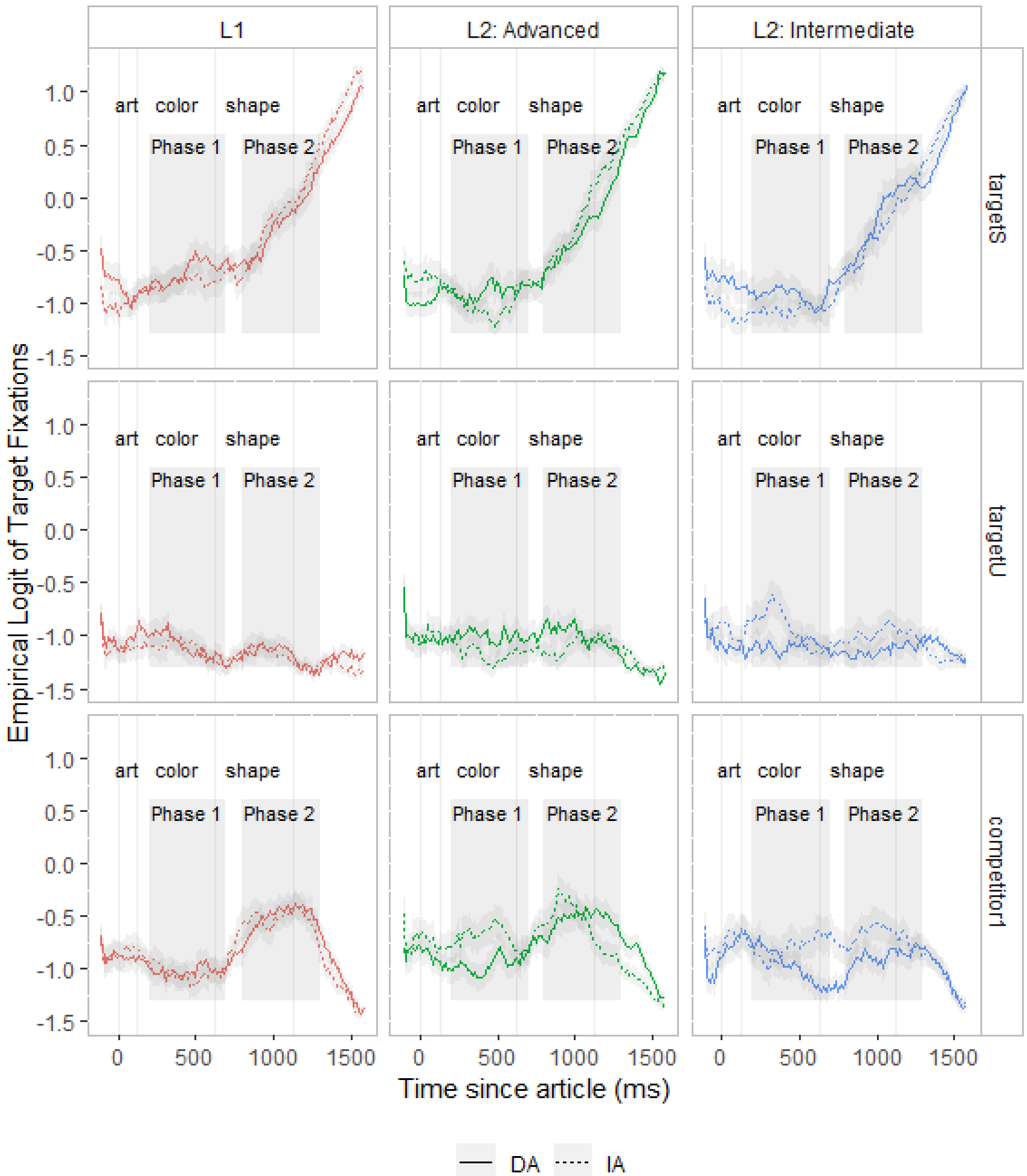





L1
L2: Advanced
L2: Intermediate
art
color
shape
Phase 1
Phase 2
targetS
competitor1
competitor2
Empirical Logit of Target Fixations
1
0
-1
0
500
1000
1500
Time since article (ms)
DA
IA


L1
L2: Advanced
L2: Intermediate
targetS
competitor1
competitor2
art
color
shape
Phase 1
Phase 2
Empirical Logit of Target Fixations
Time since article (ms)
DA
IA


L1
L2: Advanced
L2: Intermediate
targetS
targetU
competitorI
art
color
shape
Phase 1
Phase 2
Empirical Logit of Target Fixations
1.0
0.5
0.0
-0.5
-1.0
-1.5
0
500
1000
1500
Time since article (ms)
DA
IA


L1
L2: Advanced
L2: Intermediate
targetS
targetU
competitor1
art
color
shape
Phase 1
Phase 2
Empirical Logit of Target Fixations
1.0
0.5
0.0
-0.5
-1.0
-1.5
0
500
1000
1500
Time since article (ms)
DA
IA