Title
Anticipating object shapes using world knowledge and classifier information: Evidence from eve-movements in L1 and L2 processing

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 46(0)

Authors
Xu, Baorui
Feng, Shuo
Wei, Yipu

Publication Date
2024

Peer reviewed

Anticipating object shapes using world knowledge and classifier information: Evidence from eye-movements in L1 and L2 processing of Mandarin Chinese

Baorui Xu (baorui@hawaii.edu)
School of Chinese as a Second Language, Peking University, Beijing, China
Department of Second Language Studies, University of Hawaiʻi at Mānoa, Honolulu, HI, USA

Shuo Feng (shuo.feng@pku.edu.cn)
School of Foreign Languages, Peking University, Beijing, China

Yipu Wei (weiyipu@pku.edu.cn)
School of Chinese as a Second Language, Peking University, Beijing, China

Abstract

This study explores how L1 and L2 Chinese speakers use world knowledge and classifier information to predict fine-grained referent features. In a visual-world-paradigm eye-tracking experiment, participants were presented with two visual objects that were denoted by the same noun in Chinese but matched different shape classifiers. Meanwhile, they heard sentences containing world knowledge triggering context and classifiers. The effect of world knowledge has been differentiated from word-level associations. Native speakers generated anticipations about the shape/state features of the referents at an early processing stage and quickly integrated linguistic information with world knowledge upon hearing the classifiers. In contrast, L2 speakers show delayed, reduced anticipation based on world knowledge and minimal use of classifier cues. The findings reveal different cue-weighting strategies in L1 and L2 processing. Specifically, L2 speakers whose first languages lack obligatory classifiers do not employ classifier cues in a timely manner, even though the semantic meanings of shape classifiers are accessible to them. No evidence supports over-reliance on world knowledge in L2 processing. This study contributes to the understanding of L2 real-time processing, particularly in L2 speakers' utility of linguistic and non-linguistic information in anticipating fine-grained referent features.

Keywords: L2 processing; world knowledge; Mandarin Chinese; classifier; prediction; eye-tracking

Introduction

Predictive processing prevails in native language comprehension. Native language users generate expectations about upcoming content in online processing (Kuperberg & Jaeger, 2016; Pickering & Gambi, 2018; Smith & Levy, 2013). They anticipate the gender of upcoming nouns (Van Berkum et al., 2005), arguments of verbs (Boland, 2005), syntactic structures (Ferreira & Qiu, 2021) and even words' phonological forms (DeLong et al., 2005). Predictions are constrained by various sources of information (Pickering & Gambi, 2018; Pickering & Garrod, 2013), including linguistic cues such as gender-marked articles, verb properties and prosodic cues (Boland, 2005; Kamide et al., 2003; Lew-Williams & Fernald, 2010; Nakamura et al., 2012; Weber et al., 2006), visual contexts (Knoeferle et al., 2005; Knoeferle & Crocker, 2007), as well as real-world knowledge (Altmann & Kamide, 1999; Kamide et al., 2003).

The current visual world eye-tracking study aims to explore whether and how native speakers and L2 speakers of Mandarin Chinese anticipate fine-grained perceptual features of upcoming referents. We investigate the use of world knowledge and classifier information in making such anticipation and how these two types of information interplay in real-time language processing.

Predicting fine-grained features of referents

Comprehenders anticipate the upcoming information in processing. World knowledge, together with other cues, influences the probability of a referent appearing in certain contexts. For example, when hearing *The day was breezy, so the boy went outside to fly...*, participants proactively anticipate the critical noun *kite* (instead of *airplane*) based on verb information and world knowledge (DeLong et al., 2005). Informed by world knowledge information, language users not only anticipate forthcoming referents but also represent and anticipate fine-grained perceptual features of the referents, such as object shape (Zwaan et al., 2002), orientation (Stanfield & Zwaan, 2001), color (Hoeben-Mannaert et al., 2017), and state change (Kang et al., 2017). Online evidence from eye-tracking and ERP studies supports the predictive processing of such fine-grained object features (e.g., Bobb et al., 2016; Rommers et al., 2013). In Rommers et al. (2013), participants listened to sentences such as *In 1969 Neil Armstrong was the first man to set foot on the moon/tomato/rice*, and were presented with one of the three types of visual objects: the target (moon), a shape competitor (tomato), and a control object (rice), together with other unrelated objects. Both targets and shape competitors (but not the control object) attracted more fixations than unrelated items, demonstrating the pre-activation of shape information along with the predicted noun.

Information about fine-grained perceptual features is not just cued by world knowledge. Some numeral classifiers in Mandarin Chinese encode information on fine-grained features such as animacy, size, and shape ("shape classifier"; see Gao & Malt, 2009; Tai, 1990, 1994; Tai & Wang, 1990). An object that has a flat surface (e.g., *liǎn* 'face' and *zhuōzi*

In L. K. Samuelson, S. L. Frank, M. Toneva, A. Mackey, & E. Hazeltine (Eds.), *Proceedings of the 46th Annual Conference of the Cognitive Science Society.*

'desk') can often be preceded by the classifier *zhāng*, while *xuěqiú* 'snowball,' which has a round shape, can be modified by the classifier *tuán*. Importantly, the same noun may endorse referents of different shapes/states indicated by different preceding classifiers. For instance, *yī zhāng zhǐ* 'one CL_{zhang} paper' indicates that this piece of paper is spread out, whereas *yī tuán zhǐ* 'one CL_{tuan} paper' refers to a crumpled piece of paper.

Classifiers, like gender-marked articles *la/le* in French (Lew-Williams & Fernald, 2010), can be used by native speakers to generate predictions for upcoming nouns (Huettig et al., 2010; Klein et al., 2012; Mitsugi, 2018; Tsang & Chambers, 2011), or used for probing the pre-activation of nouns in highly-constraining contexts (Kwon et al., 2017). In Huettig et al.'s (2010) visual world eye-tracking study, for example, native Mandarin-speaking participants directed more attention to visual objects that match the classifier than to distractors upon hearing the classifier being uttered.

When provided with various sources of cues such as linguistic classifier information and world knowledge, language users incorporate these cues in processing, forming online expectations throughout the process. Theoretical approaches such as *the Coordinated interplay account* (Knoeferle & Crocker, 2006; 2007) hypothesize that linguistic expectations are formed on the basis of linguistic constraints stored in long-term memory, including world knowledge and previous linguistic input; as the utterance unfolds, comprehenders update sentence interpretation by integrating current linguistic input and scene information. Evidence for the interplay of linguistic and non-linguistic information has been reported at various levels of language processing (for a recent review, see Warren & Dickey, 2021). Native speakers are highly proficient in integrating world knowledge information and other linguistic information in language processing (e.g., Altmann & Mirković, 2009; Knoeferle & Crocker, 2006; 2007). In an eye-tracking study by Chow and Chen (2020), L1 Mandarin speakers immediately revise and update their expectations on upcoming referents when linguistic input (classifier) contradicts the world knowledge information given by the prior context.

Prediction in L2 processing

Compared to L1 processing, L2 processing is often found to be less automatic and more resource-demanding. For instance, numerous research studies have claimed qualitative L1 vs. L2 differences in predictive processing (i.e., prediction in L1 while absence of prediction in L2; Hopp, 2015; Mitsugi & MacWhinney, 2016), proposed that L2 speakers have no or reduced ability to generate expectations of upcoming referents (Grüter et al., 2014, 2016), or reported quantitative L1 vs. L2 differences in the timing of prediction (Kim & Grüter, 2021; Martin et al., 2016) and strength of pre-activation effects (Schlenter & Felser, 2021; see a recent review in Schlenter, 2023). Nevertheless, L2 speakers ***do*** exhibit the ability to make native-like predictions in some circumstances, even when the linguistic properties at use are absent in learners' L1 (e.g., Foucart et al., 2014; 2016; Hopp, 2013; Trenkic et al., 2014).

In this context of discussion on L2 prediction effects on upcoming referents, less is known, however, on whether and to what extent L2 speakers anticipate fine-grained features of referents in real-time comprehension and the types of cues they might employ for such predictions. According to Ahn and Jiang (2018), L2 speakers can form conceptualization of object features as L1 speakers (*cf.* Zwaan et al., 2002). After reading a sentence (e.g., *The professor placed the book on the copy machine*. / *The professor put the book into the backpack.*), L2 speakers were probed to answer whether or not the object (e.g., a spread-out book) presented in a follow-up picture was mentioned in the sentence. A congruency effect similar to L1 speakers was found for L2 speakers: L2 speakers' picture recognition time was significantly faster when the direction or shape of the object was congruent with the implied meaning of the sentence (a spread-out book following the sentence with *book on the copy machine*) and slower in the incongruent condition, indicating a similar semantic integration process between L1 and L2 speakers (Ahn & Jiang, 2018).

Despite the valuable insights into how L2 speakers integrate the semantics of object features in previous research, no evidence, to our best knowledge, shows whether L2 speakers can generate expectations of such fine-grained perceptual features *during* online comprehension and what types of cues are employed to make such anticipations. This includes anticipating the shape or state of referents that can be inferred from world knowledge or classifiers.

Theories on L2 processing, such as the *Shallow Structure Hypothesis* (Clahsen & Felser, 2006; 2018), claim that L2 speakers use fewer syntactic details than natives in processing, and instead, they depend on lexico-semantic, pragmatic and world-knowledge cues which guide "shallow processing" (Clahsen & Felser, 2006, p. 117). In line with this hypothesis, Grüter et al. (2020) found that L2 speakers relied more on the semantics of prenominal classifiers (shape information) than information on the grammatical association of classifier-noun pairs to predict upcoming nouns.

Competition Model (Bates & MacWhinney, 1989) also suggests that acoustic-articulatory channel resources are limited in language processing and comprehenders prioritize certain types of cues in processing to optimize efficiency. Nevertheless, the *Competition Model* posits that comprehenders assign weight to various cues by assessing the availability and reliability of information sources, not necessarily by linguistic categories as the *Shallow Structure Hypothesis* assumes. According to the *Competition Model*, linguistic cues that are absent or less reliable in learners' L1 while present in their L2 might be less used in L2 processing. Now consider the distribution of classifiers across languages: English does not have a close set of classifiers — classifiers are optionally used to individualize mass nouns, as in *I want* ***a box of*** *candy*. For languages such as Vietnamese, Mandarin Chinese, and Korean, numeral classifiers are obligatory at the prenominal position when count nouns are used in combination with numerals or demonstratives. The *Competition Model* would predict that adult learners of a classifier-obligatory language might be less proficient in employing classifier information in L2 processing if classifiers are not reliably used in their L1. On the other hand, world knowledge, particularly in basic and universal categories, is supposed to be a reliable cue across languages and acquired prior to a second language. Thus, world knowledge may serve as a more reliable source of information in L2 processing compared to L1-absent linguistic cues such as classifiers.

A recent study by Ahn (2021) supports the prioritized use of world knowledge by L2 speakers. L1-Korean L2-English participants were prompted to identify the target given the linguistic cue of definiteness (English *the/a*; Korean is article-less) and non-linguistic information (world knowledge shared by both L1 and L2 speakers). The target was either predictable or unpredictable based on world knowledge. For example, participants heard *The man will want to use a/the stethoscope/laptop* when viewing a display depicting a doctor (who is more likely to use stethoscopes according to basic world knowledge) being surrounded by three stethoscopes and one laptop, or one stethoscope and three laptops. The results showed that for L2 speakers when the target aligns with one's world knowledge (e.g., doctor – stethoscope), article types did not interact with the uniqueness of visual objects as in the L1 results. This finding was interpreted as evidence for L2 speakers' prioritized use of non-linguistic information compared to L1 speakers (but see counter-argument in Wiener & Rohde, 2021).

It is noteworthy that the world knowledge effect found in Ahn (2021) and Wiener and Rohde (2021) cannot be differentiated from word association effects – participants might anticipate *stethoscope* rather than other objects to be the target because of a stronger word association between the word *doctor* in the context and *stethoscope*, instead of world knowledge per se. Thus, this effect can also be attributed to lexical activation instead of genuine prediction (see Ferreira & Chantavarin, 2018; *cf.* Kuperberg & Jaeger, 2016). In the current study, instead of testing predictive effects on different target nouns, we evaluate whether participants can predict different shapes/states (spread-out or crumpled) of the same upcoming referent (e.g., *báizhǐ* 'white paper'). This fine-grained object feature can be predicted by world knowledge given the prior context or cued by different classifiers at a later point (*zhāng* for spread-out paper and *tuán* for crumpled paper). This design mitigates the effects of word association between context words and target nouns, making it more appropriate for assessing predictions based on world knowledge.

The present study aims to examine how world knowledge and linguistic information of shape classifiers affect L1/L2 predictive processing of fine-grained object features with a more time-sensitive measure (visual world paradigm eye-tracking). If L2 speakers over-rely on world knowledge that is (presumably) readily available, as indicated by the *Competition model* and revealed in Ahn's (2021) finding, it is expected that L2 learners would exhibit a greater tendency to gaze continuously at objects whose shapes are anticipated based on world knowledge. Linguistic cues absent in learners' L1, such as classifiers, might not overwrite the effect of world knowledge. If, by contrast, L2 speakers are sensitive to the semantic details conveyed by shape classifiers (Grüter et al. 2020) as well as world knowledge, both contributing to the "shallow processing" of L2 as suggested by the *Shallow Structure Hypothesis*, they should be capable of adjusting their initial world-knowledge-based predictions as soon as they access classifier information. This process would mirror the approach of L1 speakers.

Method

Design The eye-tracking experiment is conducted with a visual world paradigm. We used a 2 × 2 design, including two within-participant variables: world knowledge expectancy (expected vs. unexpected) and classifier match (match vs. mismatch).

Stimuli The experiment contained 12 experimental trials and 12 filler trials (with adjectival modifiers instead of classifiers). For each trial, participants viewed a visual display (see Figure 1) as they heard a spoken Chinese sentence (example (1)).

Figure 1: Example visual display of the experiment. The left object was the critical interest area for analysis.

(1) a. WK expected; CL match
jiàoshì de chōuti lǐ yǒu yī zhāng báizhǐ
classroom MOD drawer in have a CL$_{zhang}$ paper
b. WK expected, CL mismatch
jiàoshì de chōuti lǐ yǒu yī tuán báizhǐ
classroom MOD drawer in have a CL$_{tuan}$ paper
c. WK unexpected, CL match
jiàoshì de lājītǒng lǐ yǒu yī zhāng báizhǐ
classroom MOD dustbin in have a CL$_{zhang}$ paper
d. WK unexpected, CL mismatch
jiàoshì de lājītǒng lǐ yǒu yī tuán báizhǐ
classroom MOD dustbin in have a CL$_{tuan}$ paper
"In the classroom drawer/dustbin, there is a CL$_{zhang/tuan}$ white paper."

The auditory input contained a context that triggers world knowledge, either biasing expectations towards the interest-area object (IA object) or not (WK expected vs. WK unexpected). For instance, when the context is *in the classroom drawer* (WK expected condition), the paper is more likely to appear in a spread-out state, in comparison to the *dustbin* condition (WK unexpected), where one is more likely to find crumpled paper. The classifiers either matched or mismatched the shape/state of the IA object (CL match vs. CL mismatch). For example, the classifier *zhāng* is consistent with the spread-out paper (IA object). The other classifier *tuán*, though also compatible with the noun *báizhǐ* 'white paper,' is inconsistent with the spread-out shape of the IA object. The four versions of sentences were distributed to four lists arranged by a Latin square design.

Before the experiment, we conducted two norming tests to validate the materials. In the first norming test, we chose 36 referents that are frequently encountered in daily life (e.g., white paper), and for each referent, we created a pair of images featuring two possible states/shapes of it. We then made two versions of sentences without classifiers, each biasing the interpretation towards one of the states/shapes based on world knowledge inference (e.g., *In the classroom drawer/dustbin, there is white paper*). In an online questionnaire, 124 native Chinese participants read those sentences and were asked to select the visual object with the state/shape that matched the sentence description. A bias ratio (the number of participants who chose the expected object divided by the total number of participants) was calculated as a measure of world knowledge predictability. The stimuli (12 sets of sentences with 24 images) used in the experiment all had a clear-cut bias (bias ratio: Mean = 0.73, SD = 0.13, range = [0.57, 0.97]). The context that was biased towards one state/shape did not make the other implausible.

In the second norming test, we assessed the degree of match between the semantics of the classifier and the state/shape of the referent. In an online questionnaire, the 24 images from the first norming test were accompanied by classifier-noun phrases that either matched or mismatched the state/shape of the visual object. All classifiers were selected from beginner to intermediate Chinese textbooks. 18 native Chinese-speaking participants rated the degree of match between the phrase and the visual object on a 6-point scale. Mann-Whitney U test showed that the ratings for the classifier-match condition (Mean = 5.42, SD = 1.31) were significantly higher than those for the classifier-mismatch condition (Mean = 1.33, SD = 0.95).

The audio source was from an adult female speaker of Mandarin Chinese. Each sentence has three parts: (a) context that triggers world knowledge inference, (b) numeral and classifier, and (c) noun. The numeral is either *yī* 'one' or *jǐ* 'several'. To avoid the potential influence of tone on the processing of classifiers (*cf.* Grüter et al., 2020), *yī* 'one' in the experimental materials was always the original tone. To prevent the coarticulation effect on auditory language processing (Magnuson et al., 2003), each segment was recorded independently and then reassembled to make a complete sentence. The naturalness of auditory sentences was evaluated and double-checked by native speakers of Mandarin Chinese.

The onset of the classifier and the onset of the noun were aligned across items (see the timeline in Figure 2). As the context part varied in length, we added extra white noise to align the offset of context information. Each sentence had a one-second minimum preview time.

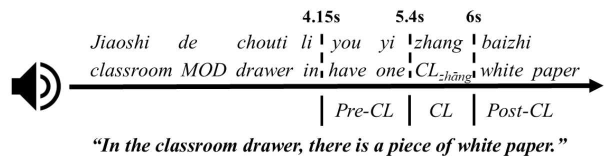

Figure 2: Sample of the auditory sentence. Acoustic onsets of three regions marked by solid lines; "CL" represents "classifier."

Participants Thirty-four L1 Mandarin speakers (mean age = 22, 24 females) and 30 advanced-level L2 speakers (mean age = 23, 14 females) participated in the experiment. All L2 participants self-reported proficiency in Mandarin Chinese at HSK levels 5–6, aligning with C1–C2 on the scale of the Common European Framework of Reference for Languages (CEFR). L2 learners' L1s included Russian (n = 3), Albanian (n = 1), English (n = 5), German (n = 4), Spanish (n = 2), Mongolian (n = 1), Danish (n = 1), Finnish (n = 1), French and German (n = 1), Portuguese (n = 1), Armenian (n = 3), Slovene (n = 1), Italian (n = 2), and Hebrew (n = 3). None of these languages has obligatory numeral classifiers. One early bilingual speaker in the L2 group was excluded from the analysis. Informed consent was collected from all subjects before the experiment.

Apparatus The task was implemented with Experiment Builder (version 2.4.913, SR Research). We used EyeLink 1000 Plus (SR Research) to record eye movements, sampling at 1000 Hz.

Procedure During the experiment, participants were prompted to answer a comprehension question in one-third of the trials (all were filler items). L2 participants completed a language exposure survey and a world knowledge questionnaire in English after the eye-tracking experiment. The questionnaire was designed to validate the L2 participants' world knowledge inference given the contexts.

Results

Coding and analysis All participants completed the task with reasonably high accuracy rates in the comprehension questions (L1 mean accuracy = 97%; L2 mean accuracy = 81%). To ensure that both groups were similarly influenced by world knowledge, we included only those test items for which the preferences expressed by L2 participants in the world knowledge questionnaire matched the responses of L1 speakers in the world knowledge norming test. Cases in which L1 and L2 groups differed in the direction of world knowledge preference, proportional looking time to both object areas fell below 50%, or a major part of the data was missing were excluded from the analysis (15.8% of L1 data and 16.5% of L2 data).

The critical time windows for analysis include (a) the pre-classifier window (from the point where the context information that triggers world knowledge inference is complete till the average onset of classifiers), (b) the classifier window (from the average onset of classifiers to the onset of the noun), and (c) the post-classifier window, from the onset of the noun to 1000ms afterward. Figure 3 exhibits the change of looks to the interest-area object under the four conditions in each language group.

The proportion of fixations on the interest area in each time window (dependent variable) was analyzed respectively using a linear mixed-effect model with *lme4* (Bates et al., 2015) in R. For the pre-classifier window, the analytical models included world knowledge expectancy as a fixed effect, and random intercepts of subject and item as random effects. For the classifier window and the post-classifier window, the basic models included world knowledge expectancy and classifier as fixed factors and random intercepts of subject and item as random effects. We then added interaction terms of the two fixed factors to the basic model and compared the goodness-of-fit of the new model against the basic one.

For the analysis of L1 data, we included mutual information scores between classifiers and nouns (Church et al., 1991) as a covariate, which reflects the association strength of classifiers and nouns. For L2 data, three standardized measures of language exposure were added to the models for analysis, including the total learning time of Chinese, formal instruction time, and informal exposure time to Chinese.

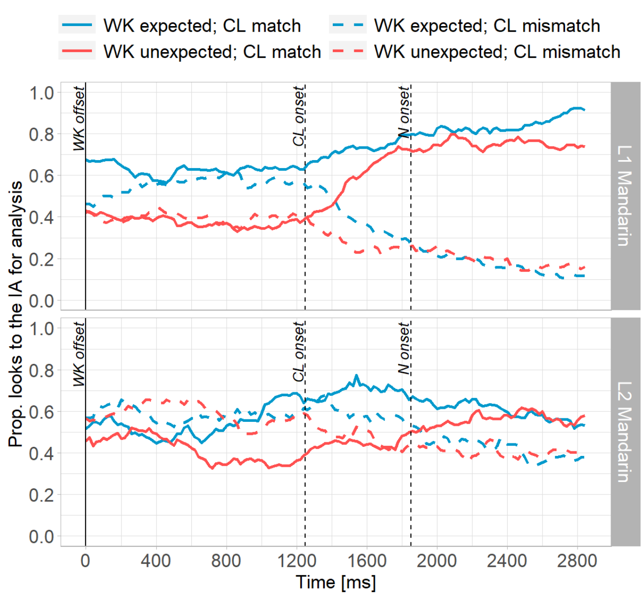

Figure 3: Proportion of looks to the IA object after the offset of world-knowledge triggering context (t = 0ms) under four conditions. Dotted lines represent the onset (200ms added) of the classifier and noun.

L1 Group In the pre-classifier window, when the world knowledge, as the only possible source of information that can be used to infer the object shape/state, was provided in completion, a significant world knowledge effect was observed (β = -.21, *SE* = .04, $t(40)$ = -4.7, p < .001). Sentential contexts with world knowledge that make the shape/state of IA objects expected to lead to more looks at the objects.

Significant world knowledge effects were sustained throughout the classifier window (β = -.14, *SE* = .05, $t(38)$ = -2.97, p < .05). However, we also observed a significant classifier effect after the classifier was uttered (β = -.29, *SE* = .05, $t(38)$ = -6.12, p < .001), demonstrating clear influence of classifiers on predicting the shape/state of the forthcoming objects. Classifiers that matched the shape/state of the IA object immediately directed more attention to the IA objects than the mismatch classifiers did. Adding the interaction term between world knowledge and classifier did not significantly enhance the model fit in the analysis of the classifier window (χ^2 = 0.34, p = .56). The analysis did not support any interaction effect between world knowledge and classifier type.

In the post-classifier window, the impact of world knowledge on gaze patterns was no longer significant (β = -.04, SE = .03, $t(36)$ = -1.22, p = .23), while the influence of classifiers remained prominent (β = -.60, *SE* = .03, $t(36)$ = -18.23, p < .001), showing L1 speakers' reliance on classifier information to identify the intended objects.

The results provided a clearer picture of how L1 speakers used world knowledge and classifier information to predict fine-grained features of the forthcoming referents. These findings showed that they are able to anticipate the shape/state of referents based on world knowledge information well before the referent was mentioned. Moreover, upon hearing the classifier, L1 speakers immediately integrated world knowledge with the shape/state information provided by the classifier and revised the prior world-knowledge-based expectation (see also in Altmann & Mirković, 2009; Knoeferle & Crocker, 2006; 2007; Chow & Chen, 2020).

L2 Group In the pre-classifier window, we found no significant world knowledge effect (β = -.04, *SE* = .04, $t(38)$ = -.98, p = .33), unlike the L1 results. L2 speakers showed no consistent patterns of using world knowledge to predict the object shape/state, even though their preference for the object shape/state was comparable to the L1 group, according to the data from the post-experiment questionnaire.

In the classifier window, the world knowledge effect was significant (β = -.14, *SE* = .05, $t(40)$ = -2.95, p < .05). More looks were directed to the IA object under the world-knowledge-expected conditions compared to the world-knowledge-unexpected conditions as the classifiers were heard. The classifier effect, however, was not observed during the classifier window (β = -.03, *SE* = .05, $t(41)$ = -.67, p = .5).

We only observed a significant classifier effect for the L2 group in the post-classifier window ($\beta = -.16$, $SE = .04$, $t(39) = -3.81$, $p < .001$) when classifier information is necessary for referent resolution at the noun.

In sum, L2 speakers were able to predict the fine-grained features of referents to some extent based on context cues, as manifested by the world knowledge effect in the classifier window before the referent noun was heard. However, they did not do so immediately when the world knowledge information was complete, and neither did they integrate world knowledge with the current linguistic information (i.e., classifier) incrementally as L1 speakers did. This finding is in line with the general delayed pattern of L2 processing in previous findings (Corps et al., 2023; Foltz, 2021). For the processing of classifiers, though L2 speakers were sensitive to the difference between the two classifiers encoding distinct shapes/states (significant classifier effect at the post-classifier region), they fell short of reliably employing this classifier information to predict the shape/state of referents before the noun.

General Discussion & Conclusion

The current study provides important evidence for expectation-driven processing of fine-grained referent features in online comprehension. Native speakers of Chinese anticipated the shape/state of upcoming referents based on world knowledge triggered by the context at an early processing stage. Different from previous studies (Ahn, 2021; Wiener & Rohde, 2021), this effect of world knowledge inference is not due to word-level associations between the context word and target noun because, in all conditions, the referents were denoted by the same noun form in the linguistic input with difference only in their shapes or states of appearance. Therefore, any difference between conditions in anticipating the shape/state of the referent before the classifiers can be ascribed to the world-knowledge-based inference. The world-knowledge-driven expectation can be updated immediately by classifiers' linguistic information: L1 speakers shifted their attention to the object with the shape/state that matched the classifier information as soon as the classifier was uttered. This finding is in line with the *Coordinated interplay account* (Knoeferle & Crocker, 2006; 2007), which highlights the interplay of various sources of information in processing. World knowledge that is stored in language users' long-term memory (e.g., paper in the classroom drawer is more likely to be in a spread-out shape) forms the basis of interpretation (expectation) during processing, and the current linguistic input — the classifiers, provide information that either matches (*zhāng*) or mismatches (*tuán*) the expectation. L1 comprehenders incrementally reconcile the classifier information with the world-knowledge-based expectation and for new interpretations of the sentence.

L2 speakers could, to some extent, predict the fine-grained features of referents based on world knowledge, though such anticipation effects were delayed and reduced. They barely used classifier cues to anticipate such fine-grained features of referents on hearing them, though other studies reported that semantic information encoded in classifiers is accessible to them in processing (Grüter et al., 2020). The information on the shapes/states of the referents encoded in classifiers only played a role at the post-classifier window when the noun was uttered (see also in Ahn & Jiang, 2018). That is to say, we did not find the same level of appliance of semantic information provided by classifiers as reported in Grüter et al. (2020) in predicting the fine-grained features of the referents. Even though the semantics of classifiers are supposed to be part of the shallow processing, as the *Shallow Structure Hypothesis* would propose, advanced second language users did not employ such information to predict object shape/state with ease. One explanation is that for L2 speakers, the utility of classifier cues may be reduced (Grüter & Rohde, 2021; *cf.* Grüter et al., 2014, 2016). In other words, the costs of predicting fine-grained features using classifiers in online processing outweigh the benefits, at least with the presence of other contextual information.

From the perspective of the *Competition model* (Bates & MacWhinney, 1989), language users may weigh various cues differently to optimize the use of limited cognitive resources. Classifiers specifying fine-grained referent features are absent in L2 speakers' first languages; therefore, although classifiers carry significant weight in L1 Chinese, they are not assigned the same level of importance in L2 processing. Accordingly, L2 speakers tend not to make fine-grained predictions based on classifiers when processing their second language (*cf.* van Bergen & Flecken, 2017). As a substitute, L2 speakers may partially rely on world knowledge in processing the fine-grained features of upcoming referents. However, we did not find evidence showing over-reliance on world knowledge information in L2 processing, as indicated by the *Competition model* and previously reported in Ahn (2021).

Another explanation for the reduced reliance on world knowledge is that participants might have learned very quickly over the course of the experiment that world knowledge was not a reliable cue to pre-identify the target (in half of the cases, it gave the wrong prediction). Given that wrong predictions are particularly costly for L2 processing, L2 participants may cease to predict based on world knowledge as the experiment proceeded. In contrast, L1 speakers who have spare cognitive resources in online processing kept trying to predict the target using their world knowledge, despite recognizing the unreliability of these cues.

To conclude, this study sheds light on a more in-depth exploration of the predictive processing of fine-grained features. Future research could broaden the scope of this line of research by encompassing other cues (e.g., visual elements and speaker details) and examining how L1 backgrounds and L2 proficiency might affect utilizing different cues in L2 prediction.

Acknowledgments

The study is supported by the Ministry of Education in China - Project of Humanities and Social Science (grant number 21YJC740062). We thank Xiaoyue Duan for her assistance with participant recruitment and data collection.

References

Ahn, H. (2021). L2 processing of linguistic and nonlinguistic information: L2 speakers use definiteness if real-world knowledge is unusable. *Studies in Second Language Acquisition*, 1-29.

Altmann, G. T. M., & Kamide, Y. (1999). Incremental interpretation at verbs: restricting the domain of subsequent reference. *Cognition*, *73*(3), 247-264.

Altmann, G. T. M., & Kamide, Y. (2007). The real-time mediation of visual attention by language and world knowledge: Linking anticipatory (and other) eye movements to linguistic processing. *Journal of Memory and Language*, *57*(4), 502-518.

Altmann, G. T. M., & Mirković, J. (2009). Incrementality and prediction in human sentence processing. *Cognitive Science*, *33*(4), 583-609.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1 - 48.

Bates, E., & MacWhinney, B. (1989). Functionalism and the competition model. In B. MacWhinney & E. Bates (Eds.), *The crosslinguistic study of sentence processing* (Vol. 3, pp. 73–112). Cambridge University Press.

Bobb, S. C., Huettig, F., & Mani, N. (2016). Predicting visual information during sentence processing: Toddlers activate an object's shape before it is mentioned. *Journal of Experimental Child Psychology*, *151*, 51-64.

Boland, J. E. (2005). Visual arguments. *Cognition*, 95(3), 237-274.

Chow, W.-Y., & Chen, D. (2020). Predicting (in)correctly: listeners rapidly use unexpected information to revise their predictions. *Language, Cognition and Neuroscience*, *35*(9), 1149–1161.

Church, K., Gale, W., Hanks, P., & Hindle, D. (1991). Using statistics in lexical analysis. In U. Zernik (Ed.), *Lexical acquisition: exploiting online resources to build a lexicon*. Psychology Press.

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.

Corps, R. E., Liao, M., & Pickering, M. J. (2023). Evidence for two stages of prediction in non-native speakers: A visual-world eye-tracking study. *Bilingualism: Language and Cognition*, 26(1), 231–243.

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.

Ferreira, F., & Chantavarin, S. (2018). Integration and prediction in language processing: A synthesis of old and new. *Current Directions in Psychological Science*, *27*(6), 443-448.

Ferreira, F., & Qiu, Z. (2021). Predicting syntactic structure. *Brain Research*, 1770, 147632.

Foltz, A. (2021). Using prosody to predict upcoming referents in the L1 and the L2: The role of recent exposure. *Studies in Second Language Acquisition*, 43(4), 753–780.

Foucart, A., Martin, C., Moreno, E., & Costa, A. (2014). Can bilinguals see it coming? Word anticipation in L2 sentence reading. *Journal of experimental psychology. Learning, memory, and cognition*, *40*.

Foucart, A., Romero-Rivas, C., Gort, L., & Costa, A. (2016). Discourse comprehension in L2: Making sense of what is not explicitly said. *Brain and Language*, *163*.

Gao, M. Y., & Malt, B. C. (2009). Mental representation and cognitive consequences of Chinese individual classifiers. *Language and Cognitive Processes*, *24*(7-8), 1124–1179.

Garrod, S., & Pickering, M. J. (2013). An integrated theory of language production and comprehension. *Behavioral and Brain Sciences*, 36(4), 329–347.

Grüter, T., Lau, E., & Ling, W. (2020). How classifiers facilitate predictive processing in L1 and L2 Chinese: the role of semantic and grammatical cues. *Language, Cognition and Neuroscience*, *35*(2), 221-234.

Grüter, T., & Rohde, H. (2021). Limits on expectation-based processing: Use of grammatical aspect for co-reference in L2. *Applied Psycholinguistics*, *42*(1), 51-75.

Grüter, T., Rohde, H., & Schafer, A. (2014). The role of discourse-level expectations in non-native speakers' referential choices. In W. Orman & M. J. Valleau (Eds.) *Proceedings of the 38th annual Boston University Conference on Language Development* (pp. 179–191). Cascadilla Press.

Grüter, T., Rohde, H., & Schafer, A. J. (2017). Coreference and discourse coherence in L2: The roles of grammatical aspect and referential form. *Linguistic Approaches to Bilingualism*, *7*(2), 199-229.

Hoeben Mannaert, L. N., Dijkstra, K., & Zwaan, R. A. (2017). Is color an integral part of a rich mental simulation? *Memory & Cognition*, *45*(6), 974-982.

Hopp, H. (2013). Grammatical gender in adult L2 acquisition: Relations between lexical and syntactic variability. *Second Language Research*, *29*(1), 33-56.

Hopp, H. (2015). Semantics and morphosyntax in predictive L2 sentence processing. *International Review of Applied Linguistics in Language Teaching*, 53(3), 277–306.

Huettig, F., Majid, A., Chen, J., & Bowerman, M. (2010). Do language-specific categories shape conceptual processing? Mandarin classifier distinctions influence eye gaze behavior, but only during linguistic processing. *Journal of Cognition and Culture*, *10*(1-2), 39–58.

Huettig, F., Rommers, J., & Meyer, A. S. (2011). Using the visual world paradigm to study language processing: A

review and critical evaluation. *Acta Psychologica*, *137*(2), 151–171.
Kaan, E. (2014). Predictive sentence processing in L2 and L1: What is different? *Linguistic Approaches to Bilingualism*, *4*(2), 257–282.
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* (pp. 2–24). John Benjamins Publishing Company.
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.
Kang, X., Eerland, A., Joergensen, G. H., Zwaan, R. A., & Altmann, G. T. M. (2020). The influence of state change on object representations in language comprehension. *Memory & Cognition*, *48*(3), 390-399.
Kim, H., & Grüter, T. (2021). Predictive processing of implicit causality in a second language: A visual-world eye-tracking study. *Studies in Second Language Acquisition*, 43(1), 133-154.
Klein, N. M., Carlson, G. N., Li, R., Jaeger, T. F., & Tanenhaus, M. K. (2012). Classifying and massifying incrementally in Chinese language comprehension. In D. Massam (Ed.), *Count and Mass Across Languages* (pp. 261-282). Oxford University Press.
Knoeferle, P., & Crocker, M. W. (2006). The coordinated interplay of scene, utterance, and world knowledge: Evidence from eye tracking. *Cognitive Science*, *30*(3), 481-529.
Knoeferle, P., & Crocker, M. W. (2007). The influence of recent scene events on spoken comprehension: Evidence from eye movements. *Journal of Memory and Language*, *57*(4), 519-543.
Knoeferle, P., Crocker, M. W., Scheepers, C., & Pickering, M. J. (2005). The influence of the immediate visual context on incremental thematic role-assignment: evidence from eye movements in depicted events. *Cognition*, 95(1), 95-127.
Kuperberg, G. R., & Jaeger, T. F. (2016). What do we mean by prediction in language comprehension? *Language, Cognition and Neuroscience*, *31*(1), 32-59.
Kwon, N., Sturt, P., & Liu, P. (2017). Predicting semantic features in Chinese: Evidence from ERPs. *Cognition*, *166*, 433–446.
Magnuson, J. S., McMurray, B., Tanenhaus, M. K., & Aslin, R. N. (2003). Lexical effects on compensation for coarticulation: the ghost of Christmash past. *Cognitive Science*, 27(2), 285–298.
Martin, C. D., Garcia, X., Breton, A., Thierry, G., & Costa, A. (2016). World knowledge integration during second language comprehension. *Language, Cognition and Neuroscience*, *31*(2), 206-216.
Martin, C. D., Thierry, G., Kuipers, J.-R., Boutonnet, B., Foucart, A., & Costa, A. (2013). Bilinguals reading in their second language do not predict upcoming words as native readers do. *Journal of Memory and Language*, *69*(4), 574-588.
Mitsugi, S. (2020). Generating predictions based on semantic categories in a second language: A case of numeral classifiers in Japanese. *International Review of Applied Linguistics in Language Teaching*, 58(3), 323–349.
Mitsugi, S., & Macwhinney, B. (2016). The use of case marking for predictive processing in second language Japanese. *Bilingualism: Language and Cognition*, 19(1), 19–35.
Nakamura, C., Arai, M., & Mazuka, R. (2012). Immediate use of prosody and context in predicting a syntactic structure. *Cognition*, 125(2), 317-323.
Pickering, M. J., & Gambi, C. (2018). Predicting while comprehending language: A theory and review. *Psychological Bulletin*, *144*(10), 1002–1044.
Rommers, J., Meyer, A. S., Praamstra, P., & Huettig, F. (2013). The contents of predictions in sentence comprehension: Activation of the shape of objects before they are referred to. *Neuropsychologia*, *51*(3), 437-447.
Schlenter, J. (2023). Prediction in bilingual sentence processing: How prediction differs in a later learned language from a first language. *Bilingualism: Language and Cognition*, 26(2), 253-267.
Schlenter, J., & Felser, C. (2021). Second language prediction ability across different linguistic domains. In E. Kaan & T. Grüter (Eds.), *Prediction in Second-Language Processing and Learning*. (pp. 48-68). John Benjamins Publishing Company.
Smith, N. J., & Levy, R. (2013). The effect of word predictability on reading time is logarithmic. *Cognition*, *128*(3), 302–319.
Stanfield, R. A., & Zwaan, R. A. (2001). The effect of implied orientation derived from verbal context on picture recognition. *Psychological Science*, *12*(2), 153-156.
Tai, J. H.-Y. (1990). Variation in classifier systems across Chinese dialects: Towards a cognition-based semantic approach. The first international symposium on Chinese languages and linguistics, Taipei: Academia Sinica.
Tai, J. H.-Y. (1994). Chinese classifier systems and human categorization. In *In honor of William S.-Y. Wang:Interdisciplinary studies on language and language change* (pp. 479-494).
Tai, J. H.-Y., & Wang, L. (1990). A semantic study of the classifier tiao. *Journal of the Chinese Language Teachers Association*, *25*, 35-56.
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.
Tsang, C., & Chambers, C. G. (2011). Appearances aren't everything: Shape classifiers and referential processing in Cantonese. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *37*(5), 1065-1080.
van Bergen, G., & Flecken, M. (2017). Putting things in new places: Linguistic experience modulates the

predictive power of placement verb semantics. *Journal of Memory and Language*, 92, 26-42.
Van Berkum, J. J. A., Brown, C. M., Zwitserlood, P., Kooijman, V., & Hagoort, P. (2005). Anticipating Upcoming Words in Discourse: Evidence From ERPs and Reading Times. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 31(3), 443-467.
Warren, T., & Dickey, M. W. (2021). The use of linguistic and world knowledge in language processing. *Language and Linguistics Compass*, *15*(4), e12411.
Weber, A., Grice, M., & Crocker, M. W. (2006). The role of prosody in the interpretation of structural ambiguities: A study of anticipatory eye movements. *Cognition*, 99(2), B63-B72.
Wiener, S., & Rohde, H. (2021). *What did you expect? Linguistic prediction occurs in both a first and second language but there are differences* Poster at the 3rd International Symposium on Bilingual and L2 Processing in Adults and Children (ISBPAC).
Zwaan, R. A., Stanfield, R. A., & Yaxley, R. H. (2002). Language comprehenders mentally represent the shapes of objects. *Psychological Science*, *13*(2), 168–171.