*Bilingualism: Language and Cognition*

**cambridge.org/bil**

# Native speakers and learners of Mandarin predict upcoming arguments in dative constructions based on categorical and gradient verb constraints

Yanxin (Alice) Zhu and Theres Grüter

Second Language Studies, University of Hawai'i at Mānoa, Honolulu, HI, USA

**Research Article**

**Cite this article:** Zhu, Y. and Grüter, T. (2025). Native speakers and learners of Mandarin predict upcoming arguments in dative constructions based on categorical and gradient verb constraints. *Bilingualism: Language and Cognition* **28**, 601–612. https://doi.org/10.1017/S1366728924000749

Received: 12 February 2024
Revised: 1 September 2024
Accepted: 3 September 2024
First published online: 12 December 2024

**Keywords:**
prediction; dative alternation; Mandarin; adult L2 learners; heritage speakers

**Corresponding author:**
Yanxin (Alice) Zhu;
Email: yanxinz@hawaii.edu

This research article was awarded Open Data and Open Materials badges for transparent practices. See the Data Availability Statement for details.

## Abstract

This study investigated the predictive use of dative verb constraints in Mandarin among home-country-raised native speakers and classroom learners (including both sequential L2 learners and heritage speakers). In a visual world eye-tracking experiment, participants made anticipatory looks to the upcoming argument (recipient versus theme) following categorical restrictions of non-alternating verbs and gradient bias of alternating verbs before the acoustic onset of the disambiguating noun. Crucially, no delay or reduction in the prediction effects was observed among L2 learners and heritage speakers in comparison with home-country-raised native speakers. Mandarin proficiency and dominant language (English versus other) did not modulate prediction effects among classroom learners. These findings provide direct support for the assumption of error-driven learning accounts of the dative alternation, that is, language users actively predict upcoming arguments based on verb information during real-time sentence processing.

## Highlights

- L1, L2 and heritage speakers of Mandarin use dative verb constraints for prediction.
- L2 and heritage speakers predict as quickly and to the same extent as L1 speakers.
- Mandarin proficiency and dominant language did not modulate learners' prediction.

## 1. Introduction

Error-driven accounts of language learning (e.g., Chang et al., 2006; Goldberg, 2019) propose that language users acquire the dative alternation in English by making predictions about the type of upcoming argument (recipient versus theme) immediately following the verb, and updating probabilistic information in their mental representations in light of prediction errors to improve future prediction accuracy. As a result of repeated prediction errors and subsequent adjustments, learners are argued to acquire constraints of dative verbs, including (i) which verbs can or cannot occur in certain dative constructions (i.e., categorical constraints), for instance, the verb "donate" cannot alternate between the double-object (DO) and prepositional (PO) dative constructions in English while the verb "give" can alternate, as indicated in examples (1)–(4); (ii) finer-grained frequency distributions of alternating verbs in each dative construction (i.e., gradient constraints, also known as verb bias), for example, the alternating verb "give" is DO-biased as it occurs more often in DO than in PO constructions in English (Jaeger & Snider, 2013).

(1) The businessman will give the money to the nun *(PO with an alternating verb)*
(2) The businessman will give the nun the money *(DO with an alternating verb)*
(3) The businessman will donate the money to the nun *(PO with a non-alternating verb)*
(4) *The businessman will donate the nun the money

Supporting evidence for error-driven learning has come primarily from production data in structural priming studies. Of particular relevance is the well-established *inverse frequency effect* (e.g., Kaschak et al., 2011), whereby priming effects tend to be larger when prime items are less frequent. Less frequent primes are expected to give rise to greater prediction error, and thus lead to greater adjustments, as predicted by error-driven learning accounts. Specifically, some studies have found inverse frequency effects based on the gradient bias of dative verbs, that is, priming effects are stronger when the prime dative construction is paired with a verb that is more likely to occur in an alternative construction (i.e., Bernolet & Hartsuiker, 2010; Fazekas et al., 2020; Jaeger & Snider, 2008, 2013; Peter et al., 2015), indirectly suggesting that language users make use of dative verb bias for prediction. Only a handful of studies have provided more direct evidence to

support the crucial prerequisite for error-driven learning of the dative alternation, namely, that language users predict based on the distribution information of dative verbs during real-time sentence processing. In these visual world eye-tracking studies (discussed in more detail below), native English speakers made anticipatory looks to the theme or recipient following categorical and gradient constraints of dative verbs (Scheepers et al., 2007; Tily et al., 2008). More recently, second language (L2) learners of English have been shown to make similar predictions, although modulated by language background or proficiency (Şafak & Hopp, 2023; Wolk et al., 2011). In the present study, we seek further evidence to assess the validity of this critical assumption underlying error-driven learning by examining directly to what extent learners and speakers of Mandarin engage in prediction while processing dative sentences in real time.

In the first study on the predictive processing of dative verbs in Mandarin, Chen et al. (2022) employed a visual world structural priming paradigm and found that native Mandarin speakers' predictions of upcoming arguments were influenced by both prime type and gradient verb bias: participants tended to predict another DO dative following a DO prime (relative to a PO prime), and vice versa. Furthermore, priming effects were stronger after a DO prime paired with a PO-biased verb (the unexpected combination) than with a DO-biased one. Yet it remains unclear whether verb-bias effects can be observed among native Mandarin speakers without structural priming, and whether these effects extend to Mandarin learners. The present study further investigates the role of prediction in the processing of Mandarin dative constructions by (i) examining processing in the absence of structural priming, (ii) including non-alternating in addition to alternating dative predicates, and (iii) testing both native speakers (L1) and classroom learners (CL) of Mandarin.

## 2. L1 and L2 prediction based on dative verb restrictions

### 2.1. Evidence of prediction from production

The claim that language users make use of dative verbs' distributional statistics to predict upcoming alternative constructions has received support from L1 production data in structural priming studies. Jaeger and Snider (2008) reanalyzed dative constructions in an English speech corpus from Bresnan et al. (2007) and found that priming effects were larger for PO primes if the verbs in those primes were DO-biased. Bernolet and Hartsuiker (2010) replicated this result with adult native speakers of Dutch in a structural priming experiment. Jaeger and Snider (2013) extended the finding to adult native speakers of English in experimental settings. Peter et al. (2015) found that English-speaking children aged 3–4 and 5–6 showed stronger verb bias effects than adults in productions immediately following prime sentences. Fazekas et al. (2020) indicated longer-term verb bias effects in a posttest immediately after a priming phase among 5 to 6-year-olds and adult English speakers. Collectively, these studies support the error-driven learning hypothesis of the dative alternation: L1 speakers' higher rates of production change in response to less predictable verb-dative combinations implies that they are aware of the bias of various dative verbs, use this bias for prediction, and adjust their representations as a result of prediction error.

Despite the robust verb-bias effects observed among child and adult L1 speakers, a still limited number of studies on L2 learners have shown mixed effects. Kaan and Chun (2018) reported a construction-wise inverse frequency effect among L1 Korean L2 learners of English, that is, the size of cumulative priming was larger for DO, the dative the L2 learners initially produced less. However, there were no verb-specific inverse frequency effects with these learners. On the other hand, in Zhu and Grüter (2024), CL of Mandarin produced significantly more PO datives and numerically fewer DO datives with alternating verbs in an immediate posttest following a priming phase, where they had encountered an equal number of DO and PO primes with these verbs. Zhu and Grüter interpreted this adjustment as an inverse frequency effect, on the assumption that PO datives caused greater surprisal given L2 learners' initial preference to use these verbs in DO datives, as evidenced in their productions during a baseline phase. More research on the predictive use of verb bias among L2 learners is needed for a better understanding of prediction and error-driven learning in this population.

### 2.2. Evidence of prediction during real-time processing

Instead of making inferences on prediction from production data, several studies have directly examined whether L1 and L2 speakers make use of dative verb constraints for prediction during real-time sentence comprehension.

Scheepers et al. (2007) found that *native English speakers* used categorical restrictions of non-alternating dative verbs for prediction. In a visual world eye-tracking experiment, Scheepers et al. recorded participants' fixations on three arguments (agent, recipient and theme) in a visual scene while they were listening to sentences as in (1)–(3) above. Participants showed a preference to look at the theme (the money) upon hearing a non-alternating PO-only verb such as "donate" in (3) versus an alternating verb like "give" in (1) and (2) before they heard the first post-verbal noun. In a similar visual world study, Tily et al. (2008) observed that native English speakers also utilized gradient bias of alternating verbs: participants predicted PO datives after a PO-biased verb such as "send" by launching earlier looks to the theme argument immediately after the verb, and showed earlier looks to the recipient argument immediately following a DO-biased verb like "serve".

Chen et al. (2022) is the first study that tapped into the predictive processing of dative verbs in Mandarin. Different from the studies described above, Chen et al. employed a visual world structural priming paradigm: *native Mandarin speakers* first read aloud a DO or PO prime sentence; then they listened to a DO or PO target sentence while looking at a visual scene with three entities (agent, recipient and theme). Chen et al. manipulated biases of verbs between prime and target, with DO- and PO-biased verbs in prime trials, and neutral-biased verbs in target trials, aiming to test the inverse frequency effects of structural priming in predictive processing. Their results showed priming effects in anticipation of dative constructions as well as an inverse frequency effect: Priming effects were greater after a DO prime with a PO-biased verb than with a DO-biased one. In their analysis, Chen et al. compared anticipatory gazes following neutral-biased verbs between DO and PO prime conditions. The study did not include baseline trials to test whether native speakers actively predict upcoming dative constructions according to verbs of different categorical and gradient constraints without the influence of structural priming.

Two studies have extended this line of research from native speakers to *adult L2 learners*. Wolk et al. (2011) found that German-speaking L2 learners of English with higher proficiency relied on gradient restrictions of DO- and PO-biased verbs to make predictions, like the native speakers in Tily et al. (2008). However, L2 learners with lower proficiency tended to anticipate a recipient

https://doi.org/10.1017/S1366728924000749 Published online by Cambridge University Press

regardless of verb bias. Wolk et al. (2011) argued that lower-proficiency learners' general tendency to expect a recipient might result from the influence of a prevalent recipient-before-theme order in their L1.

Şafak and Hopp (2023) further explored the role of cross-linguistic influence in the predictive processing of the dative alternation among L2 English learners. They analyzed data from German- and Turkish-speaking learners at similar proficiency levels yet found contrasting performances. German speakers were sensitive to both categorical and gradient verb bias and were able to make real-time predictions accordingly, similar to native English speakers. Turkish speakers were also able to use categorical verb restrictions for anticipation, yet not as rapidly as native speakers and L1 German learners did. However, they tended to be selective regarding the recruitment of gradient verb bias. On one hand, they performed similarly to native speakers and L1 German learners as they launched more anticipatory gazes towards the theme upon hearing PO-biased verbs than DO-biased ones while processing PO constructions. On the other hand, they did not distinguish between DO- versus PO-biased verbs in the processing of DO constructions. Şafak and Hopp attributed the differential processing patterns observed in the two L2 groups to the Turkish speakers' reduced experience with making predictions based on verb information due to the verb-final structure of Turkish.

There has been only a limited amount of research so far on predictive processing among *heritage speakers* (HSs), that is early bilinguals who received reduced input in the heritage language compared to home-country-raised native speakers yet earlier and typically more naturalistic input compared to late L2 learners. Research on HSs' predictive processing will thus enable us to better understand how language experience influences real-time prediction and its connection to language learning. Predictive processing among HSs is still an area in infancy, with a limited number of existing studies mostly focusing on the predictive use of morpho-syntactic cues (e.g., Fuchs, 2022a, 2022b on gender markers; Karaca et al., 2023; Özsoy et al., 2023, on case markers). These studies have revealed mixed results: Fuchs (2022a) found that Polish HSs could use gender markers for prediction as quickly as home-country-raised speakers, yet Fuchs (2022b) showed that Spanish HSs tended to be slower than a home-country-raised comparison group; Karaca et al. (2023) and Özsoy et al. (2023) indicated that HSs could use case markers predictively in certain circumstances but not as extensively and quickly as home-country-raised speakers. More recently, Ito et al. (2023) tested Vietnamese HSs' predictive use of semantic constraints of transitive verbs and revealed that they could anticipate verb-compatible objects as rapidly as L1 speakers raised in Vietnam, but only when there was no verb constraint conflict from their dominant language German. No previous research has examined the predictive use of dative verb constraints among HSs.

## 3. The present study

The primary goal of the present study was to extend the investigation of dative verb constraints in predictive processing to L2 learners of Mandarin, building on the existing work on dative verb bias as a predictive cue in the processing of L2 English (Wolk et al., 2011; Şafak & Hopp, 2023) and L1 Mandarin (Chen et al., 2022), and contributing to the broader understanding of the role of prediction in L2 processing (for recent reviews, see Bovolenta & Marsden, 2021; Kaan & Grüter, 2021; Schlenter, 2023). In light of the fact that a substantial portion of students enrolled in college-level Mandarin classes in the U.S. are HSs of a Chinese language (Luo et al., 2019), we took this opportunity to further extend the investigation to HSs. Given that L2 learners (L2ers) and HSs are typically enrolled in the same classes, we collectively refer to them here as *CLs* of Mandarin. We will look at the performance of the CLs as a group, as well as that of the L2ers and HSs separately.

### 3.1. Research questions and hypotheses

This study addresses three research questions (RQs):

(RQ1): Do home-country-raised native speakers (hereafter L1 speakers or L1ers) and CLs of Mandarin actively predict upcoming arguments based on *categorical restrictions of non-alternating dative verbs*?

(RQ2): Do L1 speakers and CLs of Mandarin actively predict upcoming arguments based on *gradient bias of alternating dative verbs*?

(RQ3): Are there any differences between L2 learners and heritage learners of Mandarin in their use of categorical and gradient constraints of dative verbs for prediction?

In view of error-driven learning accounts and previous research on L1 and L2 predictive use of dative verb information reviewed above, we hypothesize for RQ1 and RQ2 that *L1 speakers* will be able to anticipate upcoming arguments following categorical restrictions of non-alternating dative verbs as well as gradient bias of alternating dative verbs. We expect that *CLs* will also be able to use verb bias information predictively, yet they may not do so to the same extent as L1 speakers. As for RQ3, given the limited amount of previous research on the comparison between HSs and L2 learners in predictive processing, we treat this question as exploratory.

### 3.2. Participants

Sixty-five L1 speakers and 65 CLs of Mandarin participated in this study. Participants all reported having normal hearing as well as normal or corrected-to-normal vision. One classroom learner was excluded from all analyses as they rated their proficiency in Mandarin higher than any other languages they speak. Participants in the L1 group included students, scholars and their family members from the University of Hawai'i (UH) community, as well as students from Peking University (PKU) in China. The CLs were mostly current students enrolled in Chinese classes at Pennsylvania State University (PSU), UH, and PKU, and the rest had previously learned Mandarin in classroom settings. All learners must have been learning Mandarin for at least 1.5 years so that they would likely be familiar with at least the pronunciation and basic semantics of the dative verbs used in this experiment. Participants received financial compensation commensurate with local norms and the time needed to complete the experiment.

The 64 CLs of Mandarin included 42 sequential L2 learners (L2ers) and 22 HSs who were exposed to a variety of Chinese in their childhood homes. HSs were part of the CL samples at all 3 testing sites (11/28 in Honolulu, 6/15 in State College, and 5/21 in Beijing). Among the 22 HSs, 7 were exposed to Cantonese during childhood, 6 to Mandarin, 2 to Fuzhounese, 1 to Hokkien, 1 to Qingdao dialect, 1 to Teochew, 1 to Mandarin and Cantonese, 1 to Mandarin and Fuzhounese, 1 to Mandarin and Northeast dialect, and 1 to Mandarin and Shanghainese. Their reported dominant

https://doi.org/10.1017/S1366728924000749 Published online by Cambridge University Press

**Table 1.** Participant demographics (means, standard deviations and ranges)

| | L1ers | Classroom learners (CLs) |
|---|---|---|
| | (*N* = 65, *Female* = *48*) | (*N* = 64, *Female* = 43) |
| Age | 27.2 (*SD* = 5.4) | 22.8 (*SD* = 5.7) |
| | (19–43) | (18–55) |
| Length of stay in a Mandarin-speaking environment | N.A. | 26.0 (*SD* = 44.8) |
| (in months) | | (0–220) |
| LexTALE_CH score | 43.7 (*SD* = 6.4) | 19.7 (*SD* = 13.8) |
| (Corrected Accuracy, range: −60–60) | (26–57) | (−19–42) |
| Self-rating of overall Mandarin language ability (1–10) | 9.1 (*SD* = 0.9) | 6.0 (*SD* = 1.5) |
| | (7–10) | (3–10) |

*Note:* Corrected Accuracy is a scoring method recommended by Chan & Chang (2018): number of correct hits - 2 * number of false alarms.

languages included English (*N* = 16), Burmese (*N* = 1), German (*N* = 1), Indonesian (*N* = 1), Japanese (*N* = 1), Russian (*N* = 1) and Spanish (*N* = 1). As for the 42 L2ers, their dominant languages were English (*N* = 21), Korean (*N* = 5), Japanese (*N* = 3), Mongolian (*N* = 3), Russian (*N* = 3), Thai (*N* = 3), German (*N* = 1), Hungarian (*N* = 1), Spanish (*N* = 1), and Tajik (*N* = 1).[1] Two Sample *t* tests showed that L2ers and HSs were comparable in their scores on the Chinese LexTALE Test (LexTALE_CH, a character-based lexical test for proficiency, Chan & Chang, 2018), $t(62) = -1.31$, $p = .19$, self-rated overall proficiency $t(62) = -.71$, $p = .48$, and age, $t(62) = .32$, $p = .75$, yet different with regard to length of stay in Mandarin-speaking environment, $t(22.9) = -2.79$, $p = .01$ (see Table S1 in supplementary materials for separate demographics for the L2 and HS groups). Table 1 below presents the demographics of the L1 and CL groups. The two groups differed significantly in their LexTALE_CH scores, $t(88.5) = 12.64$, $p < .001$, self-rated overall proficiency, $t(99.6) = 14.7$, $p < .001$, as well as age, $t(127) = 4.49$, $p < .001$.

### 3.3. Linguistic materials (verbs)

This experiment focused on the processing of the Mandarin DO and PO constructions as illustrated in example sentences (5) and (6).[2] A total of 10 verbs (see Table 2), selected from *Integrated Chinese* (Liu et al., 2021), a widely used set of textbooks in American universities (Ye, 2019), were included to ensure they were familiar to the CLs in this study. To address RQ1 on categorical constraints, we included 6 non-alternating dative verbs: 3 *MAKE* verbs, i.e., verbs of creation, which can only appear in PO; and 3 *TELL* verbs that represent verbs of communication and can only occur in DO. For RQ2 on gradient constraints, 4 alternating *GIVE* verbs that signify acts of giving were used, among which 1 was DO-biased and 3 were PO-biased.

**Table 2.** Experimental dative verbs

| Non-alternating | | Alternating | |
|---|---|---|---|
| MAKE (PO-only) | TELL (DO-only) | GIVE (PO-biased) | GIVE (DO-biased) |
| 做 (*zuo*, make) | 教 (*jiao*, teach) | 借 (*jie*, lend) | 送 (*song*, give) |
| 画 (*hua*, draw) | 问 (*wen*, ask) | 租 (*zu*, rent) | |
| 炒 (*chao*, fry) | 告诉 (*gaosu*, tell) | 分 (*fen*, share) | |

*Note.* Bias information of the GIVE verbs is from a written production task reported in Chen et al. (2022). According to Chen et al., Mandarin is strongly PO-biased. There are only a limited number of DO-biased verbs, and their bias tends to be weak. The study thus included the only DO-biased verb found in the textbooks (i.e., *Integrated Chinese*) that CLs used for the first 3 semesters of Mandarin learning.

(5) The double-object (DO) dative
Mali **song** le Dawei yi ge dangao.
Mary **give** PFV David one CLF cake
'Mary **gave** David a cake.' (PFV = perfective; CLF = general classifier)

(6) The prepositional object (PO) dative
Mali **song** le yi ge dangao gei Dawei.
Mary **give** PFV one CLF cake GEI David
'Mary gave a cake to David.'

### 3.4. Methods and procedure

The experiment was completed online and in our laboratory. At least 1 day before the lab session, participants completed two tasks online: (1) a language background questionnaire (5–10 minutes), and (2) a sentence completion task (20–30 minutes) that probed participants' biases when using the experimental dative verbs in written production. For the purpose of the present study, the primary goal of the sentence completion task was to ascertain that participants were aware of the different constraints of the verbs used in this study. A more detailed analysis of this task is beyond the scope of this paper and will be reported elsewhere. Relevant to the goals of the present study, results from this task showed that both L1ers and CLs respected the constraints of MAKE and TELL verbs, with no or very few DO continuations following MAKE verbs (L1: 0%, CL: 7%) and similarly very few PO continuations following TELL verbs (L1: 9%, CL: 5%). For GIVE verbs, both L1ers and CLs

[1]Since the investigation of L1 influence was not a goal of this study, we did not exclude CL participants based on L1. To explore the potential role of the L1 within the CL group, we report additional analyses with L1 as a predictor (see Results; no significant effects for this predictor emerged).

[2]Mandarin is argued to have more than two dative constructions (Liu, 2006). To simplify the research design and following Chen et al. (2022), this study focuses on the alternation between two of them, the DO and PO constructions most similar to those in English. Example sentence (i) illustrates a third dative construction V-GEI DO, in which the verb ***song*** (give) combines with **GEI** to form a compound verb that takes double objects.

(i) The V-GEI double-object (DO) dative
Mali **song-GEI** le Dawei yi ge dangao.
Mary **give-GEI** PFV David a CLF cake
'Mary **gave** David a cake.'

The possibility that participants in our experiment might predict a V-GEI DO is ruled out by the design that all experimental sentences have the perfective marker *le* following the (simple) dative verb so that a compound verb is no longer possible when *le* is heard. As shown in (i), *le* must come after the entire compound verb and cannot precede GEI.

https://doi.org/10.1017/S1366728924000749 Published online by Cambridge University Press

| | CR1 | | CR2 | | | | | |
|---|---|---|---|---|---|---|---|---|
| Nanren | song | le | yi | ge | daxuesheng | yi | ge | shafa. |
| Man | give | PFV | one | CLF | college student | one | CLF | sofa |
| 'The man gave a college student a sofa. ' | | | | | | | | |

**Figure 1.** Illustration of the experimental item in VWP (PFV = perfective marker; CLF = general classifier).
*Note.* The critical regions (CRs) for analysis are framed in the target sentence.

provided DO as well as PO continuations (L1: 70% PO, 19% DO; CL: 39% PO, 22% DO).[3]

During the lab session, participants first completed a visual world eye-tracking (VWP) task (~15 minutes, described below) and then the LexTALE_CH test (5–10 minutes, Chan & Chang, 2018), in which they judged whether the given characters were real Chinese characters or not.

### 3.5. Visual world eye-tracking (VWP) task

In the VWP task, participants listened to recordings of a native speaker of Mandarin describing visual scenes with three entities (agent, theme, recipient, for experimental trials) while looking at the scenes (Figure 1).

There were a total of 30 experimental trials: 3 MAKE verbs each occurring in 3 PO sentences, 3 TELL verbs each occurring in 3 DO sentences, 3 PO-biased GIVE verbs each occurring twice in PO and once in DO sentences, and 1 DO-biased GIVE verb occurring twice in DO sentences and once in a PO sentence (see Table 2 for the verbs). Each auditory stimulus was paired with a visual scene as in Figure 1, with 3 entities (i.e., agent, theme, recipient) arranged in a triangular fashion and counterbalanced. The 30 experimental items were interspersed with 30 filler trials containing intransitive, transitive, existential, passive and BA sentences.[4] Participants were randomly assigned to one of two versions of the task differing only in the pseudo-randomised order of items. All materials are available at https://osf.io/pvymf/.

The auditory stimuli were produced by a female native speaker of Mandarin in a sound-proof booth and recorded at 44.1 k Hz in Praat (Boersma & Weenink, 2022). In order to minimize variability in the duration of the critical regions for analysis (see Figure 1), small segments of silence were inserted or removed between the perfective marker *le* and the numeral *yi* ('one') in some items. The mean duration of the verb was 520 ms (SD = 56.4 ms), while the mean duration of the ambiguous prediction window from the onset of the perfective marker *le* to the onset of the first post-verbal noun was 1620 ms (SD = 59.6 ms). Two native speakers of Mandarin confirmed the naturalness of the final auditory stimuli.

The experiment began with instructions and a practice session of two items. The presentation of the auditory stimuli started after a 1500 ms preview of the visual scene, and the visual displays remained visible for 2000 ms after the sentence offset. To examine whether participants were attentive to the task, they were asked, in half of the filler trials, to judge whether the speaker had mentioned every entity in the visual scene. Participants' eye gaze data were collected through the SMI Experiment Suite on an SMI remote eye tracker sampling at 60 Hz. Following the practice items, the eye tracker was calibrated with a 9-point calibration grid for both eyes.

### 3.6. Data analysis

Of critical interest is the likelihood of participants looking at the theme versus the recipient during two critical ambiguous time windows (CR1 = verb, CR2 = perfective aspect marker + numeral + general classifier, see Figure 1). These time windows were offset by 200 ms from their respective acoustic onsets to allow for the execution of ballistic eye movements. Thus, CR1 extended from 200 ms after verb onset to 200 ms after the mean onset of the perfective marker *le*, for a duration of 520 ms; CR2 started from the end point of CR1 and extended to the mean onset of the first post-verbal noun, for a duration of 1420 ms. Note that the general classifier *ge* in CR2 is compatible with both the recipient and the theme, and thus provides no cues to the upcoming argument. The signal thus remains fully ambiguous during both analysis regions, allowing us to probe for effects of prediction in the strict sense, that is, before the acoustic onset of any disambiguating information. We interpret changes in the likelihood of looks to the theme versus recipient from CR1 to CR2 as evidence of prediction based on verb information. From CR1 to CR2, we expect participants to increase

[3]Percentages do not add up to 100% because participants also produced responses that could not be classified as PO or DO. The vast majority of these "other" responses were transitive sentence (e.g., 老师租了一个律师的一个办公室。'The teacher rented a lawyer's office.'). There were also sentences with purposive GEI (e.g., 小朋友送了一个苹果给一个病人吃。'The kid gave an apple GEI a patient eat.'), serial VP sentences (e.g., 姐姐分了一个西瓜送给一个小孩。'The sister divided a watermelon give-GEI a friend.'), V-GEI double object sentences (e.g., 女人租给一个律师一个办公室。'The woman rent-GEI a lawyer an office.'), sentences with a preverbal argument preceded by GEI (e.g., 女人给一个律师租了一个办公室。'The woman GEI a lawyer rent an office.'), as well as some ungrammatical sentences (e.g., *女人租了一个办公室一个律师。'*The woman rented an office a lawyer.') and sentences with other verbs instead of the experimental dative verbs.

[4]The BA-construction conveys an event where an object is affected, dealt with, or disposed of (Huang et al., 2009). For instance, an example filler BA-sentence is as follows:

(ii) Nanhai ba xigua chi-le.
boy BA watermelon eat-LE
'The boy ate the watermelon.'

https://doi.org/10.1017/S1366728924000749 Published online by Cambridge University Press

fixations on the recipient (i.e., predicting DO) upon hearing a TELL verb or DO-biased GIVE verb, and increase looks to the theme (i.e., predicting PO) following a MAKE verb or PO-biased GIVE verb.

Prior to analysis, we excluded eye movement data from 7 participants with overall tracking ratios lower than 70%, and 3 participants with accuracy rates lower than 80% in the secondary judgment task, which left us data from 59 L1 and 60 CL participants. After these exclusions, accuracy on the judgment task was high for the remaining participants (98.3% for L1ers; 99.5% for CLs), suggesting they were attentive during the VWP task.

We recorded fixations within a rectangular area of interest (AOI) around the corresponding entities (agent, recipient, theme) for the target trials. Since fixations at different AOIs are not independent of one another (i.e., when an AOI is fixated, another cannot be), we used a log ratio between fixations to the theme versus recipient as the dependent variable (Ito & Knoeferle, 2022). Within each trial and critical region (CR), the log ratio was calculated according to the following formula: log((fixation proportion on theme + .5)/(fixation proportion on recipient + .5)). The log-ratios thus represent fixation bias, with a positive number reflecting more looks to the theme than to the recipient, and a negative number indicating more looks to the recipient than to the theme. We interpret the change in fixation biases from CR1 to CR2 as evidence of prediction based on verb constraints. All analyses were conducted in R 4.2.2 (R Core Team, 2022) using the lmerTest package (Kuznetsova et al., 2017).

### 3.7. Results

For each research question, we first present results from the L1 and CL groups, aggregating over L2ers and HSs. We then present comparisons between the L2 and HS groups. All data and scripts are available at https://osf.io/pvymf/.

#### RQ1: Categorical constraints of non-alternating verbs

Figure 2 displays participants' mean proportions of looks to the 3 AOIs (agent, recipient, and theme) in trials with PO-only MAKE verbs and DO-only TELL verbs. For both trial types, the L1 and CL groups showed an initial preference for looking at the recipient in CR1, which is unsurprising as the recipient was consistently animate whereas the theme was inanimate. However, fixation biases diverged in CR2: both groups increased looks to the theme following MAKE verbs and to the recipient following TELL verbs, consistent with the expected pattern for prediction based on categorical constraints of these non-alternating verbs. Notably, the CLs' anticipatory fixations do not appear delayed or reduced compared to the L1ers.

LMER models confirmed the prediction patterns evident in Figure 2. The initial model (see Table 3) included critical region (CR, −.5 = CR1, .5 = CR2), verb type (−.5 = MAKE, .5 = TELL) and group (Helmert coded; the first comparison: HS = −.33, L2 = −.33, L1 = .67; the second comparison: HS = −.5, L2 = .5, L1 = 0) and their interactions as fixed effects. The maximal random effects for the model to converge included random participant and item intercepts, as well as random participant slopes for verb type and random item slopes for CR. Crucially, the model returned a significant interaction between CR and verb type ($b = -.36$, $p < .001$). We then ran models for the two CRs separately (see Table S2 in supplementary materials), which indicated that participants tended to look more to the recipient in CR1 (intercept, $b = -.10$, $p = .06$) irrespective of verb type ($b = .01$, $p = .92$), but verb type significantly modulated fixation bias in CR2 ($b = -.35$, $p = .003$). We also divided the data by verb type and ran separate models (see Table S3), and the results indicated that from CR1 to CR2 participants looked more to the theme (i.e., predicting a PO) following MAKE verbs ($b = .25$, $p < .001$), and more to the recipient (i.e., predicting a DO) following TELL verbs ($b = -.11$, $p = .003$). There were no effects of or relevant interactions with group (L1 versus CL) in any of these models, and thus no evidence that L1ers and CLs made use of categorical constraints to predict to different extents.

Figure 3 presents fixation proportions for the L2 and HS groups separately, showing very similar-looking patterns in these two subgroups. The LMER models (see Table 3, and Tables S2–S3 in supplementary materials) confirmed our observation from Figure 3. While L2ers were overall less likely to look at the theme

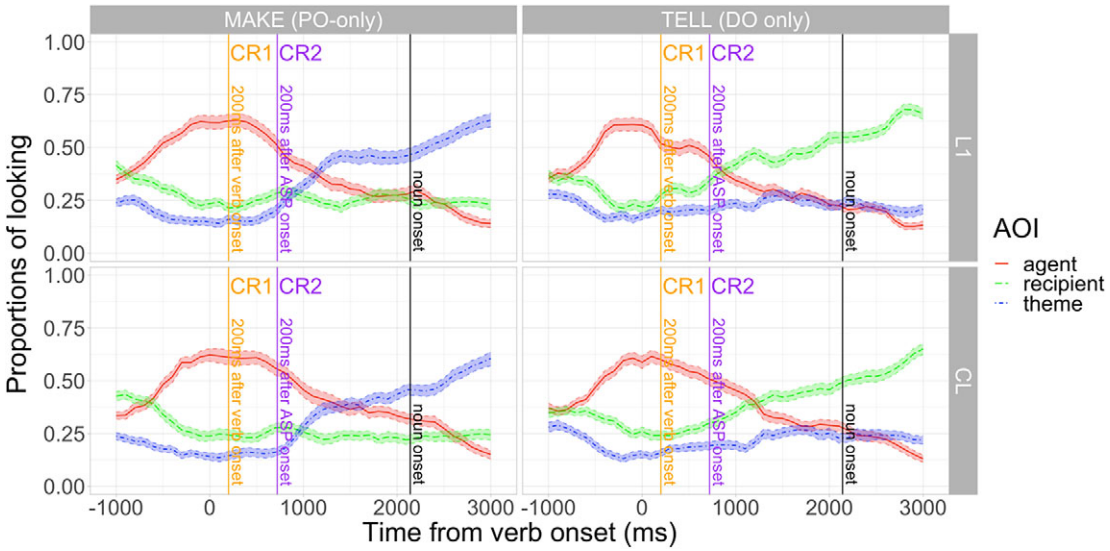


**Figure 2.** RQ1: Proportion of looks to 3 AOIs by verb type and group. 0 ms is verb onset. Error bands represent a 95% confidence interval.

https://doi.org/10.1017/S1366728924000749 Published online by Cambridge University Press

**Table 3.** RQ1: Overall model output for fixation bias (L1: n = 59; CL: n = 60, L2: n = 38, HS: n = 22)

| Predictors | Estimate | *SE* | *df* | *t* | *p* |
|---|---|---|---|---|---|
| (Intercept) | −0.07 | 0.05 | 31.54 | −1.36 | 0.18 |
| CR | 0.07 | 0.03 | 35.75 | 2.58 | **0.01** |
| Verb type | −0.17 | 0.10 | 32.54 | −1.62 | 0.12 |
| Group (CL versus L1) | −0.01 | 0.02 | 113.83 | −0.61 | 0.54 |
| Group (HS vs. L2) | −0.06 | 0.03 | 112.25 | −2.22 | **0.03** |
| CR * Verb type | −0.36 | 0.05 | 35.75 | −6.93 | **<0.001** |
| CR * Group (CL vs. L1) | −0.03 | 0.04 | 3639.58 | −0.88 | 0.38 |
| CR * Group (HS vs. L2) | −0.02 | 0.05 | 3638.15 | −0.32 | 0.75 |
| Verb type * Group (CL vs. L1) | −0.04 | 0.05 | 112.91 | −0.85 | 0.40 |
| Verb type * Group (HS vs. L2) | 0.07 | 0.07 | 111.74 | 1.06 | 0.29 |
| CR * Verb type * Group (CL vs. L1) | −0.05 | 0.08 | 3639.06 | −0.71 | 0.48 |
| CR * Verb type * Group (HS vs. L2) | −0.00 | 0.11 | 3638.02 | −0.00 | 1.00 |

Formula: lmer (log ratio fixation bias ~ CR * Verb type * Group + (1 + Verb type | participant) + (1 + CR | item)).

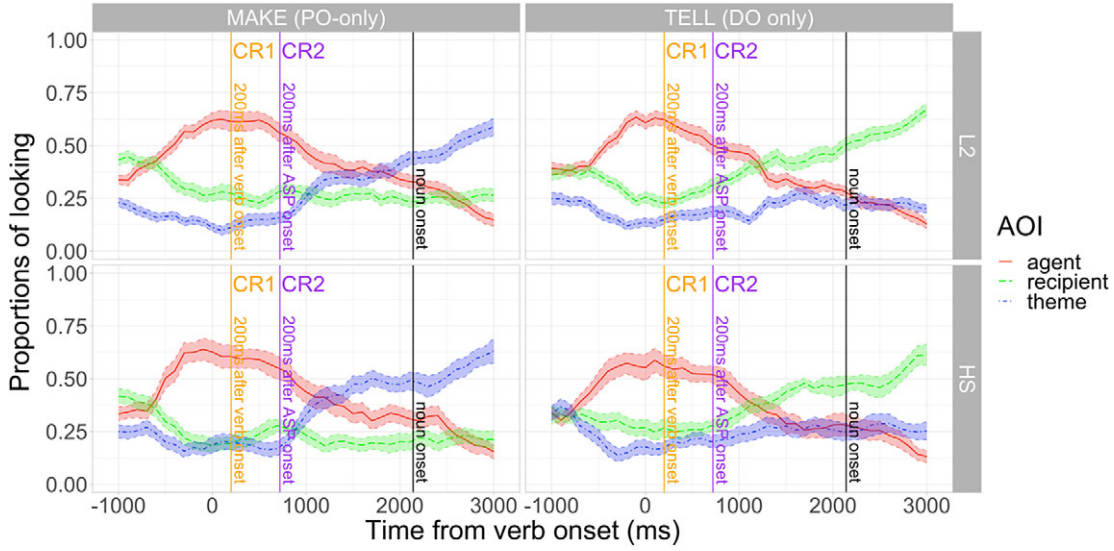


**Figure 3.** RQ1: Proportion of looks to 3 AOIs by verb type and group (L2 versus HS). 0 ms is verb onset. Error bands represent a 95% confidence interval.

than HSs ($b = -.06$, $p = .03$), there were no interactions with the group, and thus no evidence for differences in the predictive looking patterns between the two learner groups. Both the L2ers and HSs, like the L1ers, predicted upcoming arguments based on categorical constraints of dative verbs.

Additional exploratory models of the CL data with LexTALE scores as a continuous predictor showed that the CLs of higher proficiency looked marginally more to the theme overall regardless of CR and verb type (LexTALE scores: $b = .03$, $p = .05$), yet there were no significant interactions between LexTALE scores and verb type and critical window (see Table S4). We thus see no evidence for a modulation of L2 proficiency on prediction following categorical verb constraints. In light of the heterogeneity in CL participants' dominant languages, we also conducted additional analyses of the CL data with dominant language as a binary predictor (−.5 = other, .5 = English) to explore the potential role of cross-linguistic influence. No significant effects involving this factor were observed (see Table S5).

### RQ2: Gradient constraints of alternating verbs

Figure 4 visualizes participants' mean proportion of looks in trials with the alternating GIVE verbs. Both the L1 and CL groups initially looked more to the animate recipient in CR1 regardless of verb bias. In CR2, both groups gradually increased fixations to the theme following PO-biased verbs, as expected if they predicted based on the bias of these verbs. However, their fixations bias did not change following the DO-biased GIVE verb.

LMER models corroborated the pattern we saw in Figure 4. The initial model (see Table 4) had CR (−.5 = CR1, .5 = CR2), verb bias (−.5 = DO, .5 = PO) and group (Helmert coded; the first comparison: HS = −.33, L2 = −.33, L1 = .67; the second comparison:

https://doi.org/10.1017/S1366728924000749 Published online by Cambridge University Press

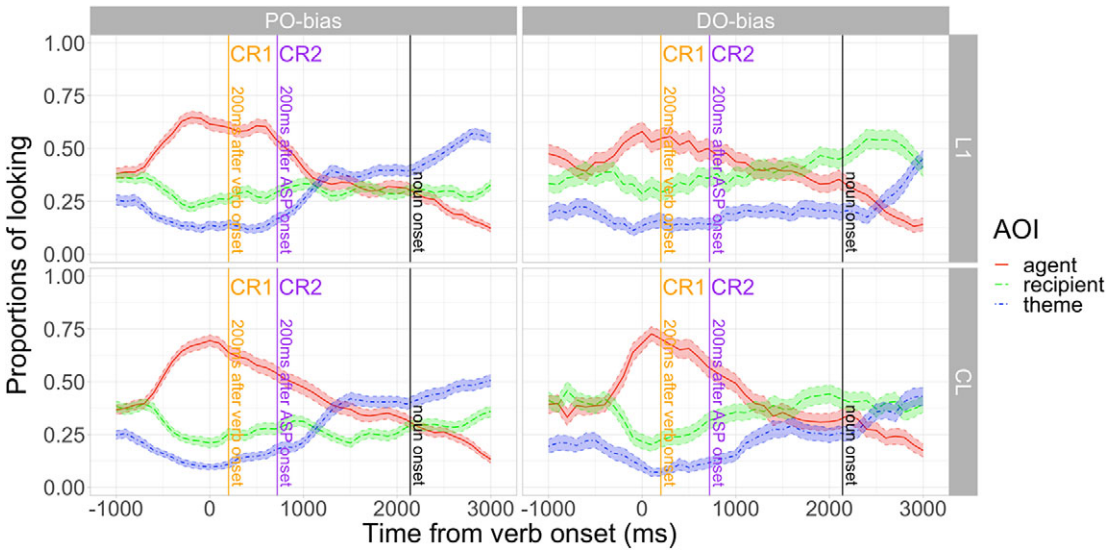


**Figure 4.** RQ2: Proportion of looks to 3 AOIs by verb bias and group. 0 ms is verb onset. Error bands represent a 95% confidence interval.

**Table 4.** RQ2: Overall model output for fixation bias (L1: n = 59; CL: n = 60, L2: n = 38, HS: n = 22)

| Predictors | Estimate | *SE* | *df* | *t* | *p* |
|---|---|---|---|---|---|
| (Intercept) | −0.14 | 0.05 | 22.17 | −2.93 | **0.008** |
| CR | 0.09 | 0.03 | 22.73 | 2.78 | **0.01** |
| Verb bias | 0.27 | 0.06 | 180.11 | 4.33 | **<0.001** |
| Group (CL vs. L1) | −0.03 | 0.03 | 167.06 | −0.94 | 0.35 |
| Group (HS vs. L2) | 0.01 | 0.04 | 165.38 | 0.14 | 0.89 |
| CR * Verb bias | 0.26 | 0.06 | 33.47 | 4.16 | **<0.001** |
| CR * Group (CL vs. L1) | −0.00 | 0.05 | 2390.27 | −0.08 | 0.93 |
| CR * Group (HS vs. L2) | 0.19 | 0.07 | 2390.91 | 2.50 | **0.01** |
| Verb bias * Group (CL vs. L1) | −0.01 | 0.05 | 2391.21 | −0.13 | 0.90 |
| Verb bias * Group (HS vs. L2) | −0.05 | 0.07 | 2390.20 | −0.66 | 0.51 |
| CR * Verb bias * Group (CL vs. L1) | 0.02 | 0.10 | 2390.33 | 0.22 | 0.82 |
| CR * Verb bias * Group (HS vs. L2) | −0.26 | 0.15 | 2390.73 | −1.76 | 0.08 |

Formula: lmer (log ratio fixation bias ~ CR * Verb Bias * Group + (1 | participant) + (1 + CR | item)).

HS = −.5, L2 = .5, L1 = 0) and their interactions as fixed effects. The maximal random effects for the model to converge included random participant and item intercepts, as well as random item slopes for CR. Due to a significant interaction between CR and verb bias ($b = .26$, $p < .001$), we ran separate models for the two CRs (see Table S6). These models indicated a general looking preference for the recipient in CR1 (intercept, $b = -.17$, $p = .002$) regardless of verb bias ($b = .06$, $p = .43$), yet verb bias showed a significant modulating effect in CR2 ($b = .41$, $p < .001$). Separate models by verb bias (Table S7) showed that from CR1 to CR2 participants increased looks to the theme (i.e., predicting a PO) following PO-biased verbs ($b = .21$, $p < .001$), yet their looking pattern did not change following the DO-biased verb ($b = -.02$, $p = .62$). There were no effects or relevant interactions with group (L1 versus CL), and thus no evidence that L1ers and CLs made use of gradient constraints to predict to different extents.

Again, we also inspected the data from L2ers and HSs separately. As visualized in Figure 5, the L2 and HS groups appeared to show similar predictions based on the PO-biased alternating verbs (left in figure). However, the two groups seemed to behave differently regarding the DO-biased verb (right in the figure): the L2ers looked more to the theme in CR2 (i.e., predicting a PO), while the HSs looked more to the recipient (i.e., predicting a DO).

The omnibus LMER model (see Table 4) indicated a CR by group (HS versus L2) interaction ($b = .19$, $p = .01$), and the follow-up models by verb bias (see Table S7 in supplementary materials) revealed a CR by group (HS versus L2) interaction only for the DO-biased verb ($b = .32$, $p = .01$), not for the PO-biased ones ($b = .06$, $p = .48$), suggesting that L2ers and HSs anticipated the theme to the same degree based on the PO-biased verbs, yet were different regarding the DO-biased verb. We then further split the data of the DO-biased verb by CR and group (see Tables S8–S9 in

https://doi.org/10.1017/S1366728924000749 Published online by Cambridge University Press

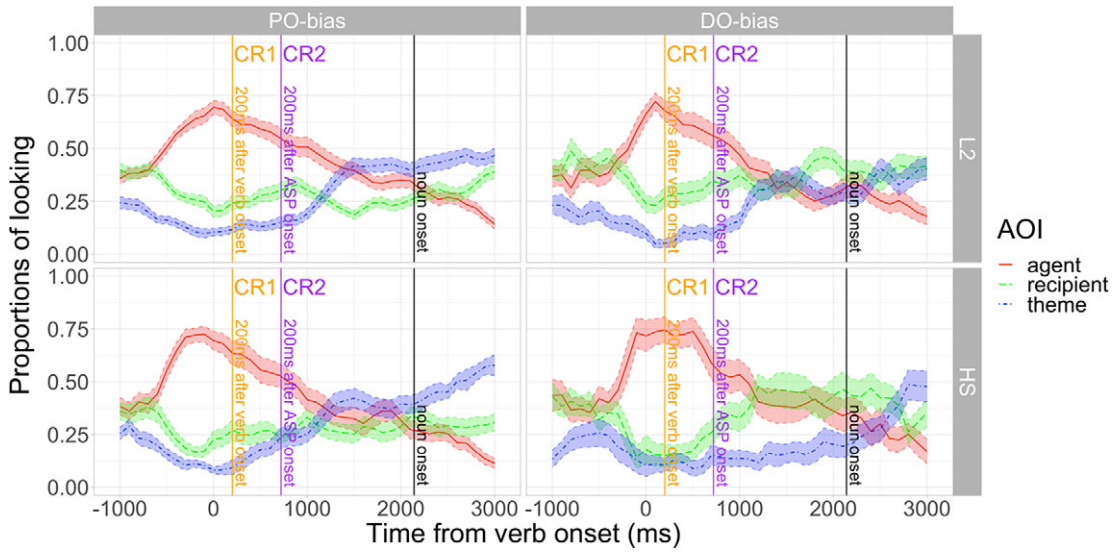


**Figure 5.** RQ2: Proportion of looks to 3 AOIs by verb type and group (L2 versus HS). 0 ms is verb onset. Error bands represent a 95% confidence interval.

supplementary materials). As observed from Figure 5, L2ers had marginally more looks to the theme in CR2 following the DO-biased verb ($b = .14$, $p = .05$), whereas HSs looked marginally more to the recipient ($b = -.17$, $p = .06$).

Again, we ran additional models of the CL data to explore effects of proficiency. The overall model with LexTALE scores as a predictor (see Table S10) revealed a marginal 3-way interaction between CR, verb bias and LexTALE scores ($b = .14$, $p = .05$). We thus split the data by CR and verb bias for further analyses (see Tables S11–S13). The sub-models by verb bias indicated that LexTALE scores had no effects for PO-biased verbs, but there was a CR by LexTALE scores interaction for the DO-biased verb ($b = -.13$, $p = .02$). We then split the data for the DO-biased verb and follow-up models showed that this was because with increasing LexTALE scores the CLs' fixations bias at the theme significantly increased in CR1 ($b = .08$, $p = .04$) but numerically decreased in CR2 ($b = -.05$, $p = .22$). In light of the sparsity of data on which these additional models are based, we do not attempt to interpret this pattern further and conclude that the data from this experiment do not provide clear evidence for modulating effects of LexTALE scores on CLs' predictive use of gradient verb bias. Again, additional models of the CL data with dominant language as a binary predictor (−.5 = other, .5 = English) indicated no significant effects involving this factor (see Table S14).

## 4. Discussion

Studies on the real-time predictive use of verb bias in dative verbs are critical to examine a key assumption of error-driven learning accounts for the dative alternation, namely that language users actively use dative verb constraints to predict upcoming arguments/constructions during real-time sentence processing, a claim that has received substantial indirect support from priming effects in production data. As part of the endeavour to better understand the real-time predictive use of dative verbs, the current study investigated to what extent home-country-raised L1 speakers and CLs (including L2 learners and HSs) of Mandarin make use of (RQ1) categorical and (RQ2) gradient constraints of dative verbs for prediction. Participants completed a visual world eye-tracking experiment, during which they listened to dative sentences while looking at visual scenes with the three arguments (agent, recipient and theme) included in the sentences.

With categorical constraints of non-alternating dative verbs (RQ1), CLs predicted upcoming arguments in a manner not distinguishable from L1 speakers. They launched anticipatory looks to the theme upon hearing PO-only MAKE verbs, and to the recipient immediately following DO-only TELL verbs before they heard the first post-verbal noun. No differences were observed between L2 learners and HSs.

Similarly, with gradient constraints of alternating dative verbs (RQ2), CLs predicted based on *PO-biased verbs* by launching anticipatory fixations to the theme upon hearing the verb, with no delay or reduction compared with L1 speakers. With the only DO-biased verb included in this study, the CLs as a group also appeared to perform similarly to the L1ers, with no fixation change upon hearing the DO-biased verb in either group. When the data from the L2ers and HSs was examined separately, however, we found that the two learner groups seemed to make opposite predictions: the L2ers looked marginally more to the theme (i.e., predicting PO) after hearing the verb, whereas the HSs looked marginally more to the recipient (i.e., predicting a DO).

Mandarin was found to be strongly PO-biased in terms of the dative alternation in a norming study in Chen et al. (2022), who recorded native Mandarin speakers' written productions of DO and PO sentences (as well as sentences other than DO and PO) using 48 alternating dative verbs, and calculated biases of these verbs via log ratio between the numbers of DO and PO sentences produced (formula: log[(#DO + 1)/(#PO + 1)]), such that the result values larger than 0 indicated a DO-bias and values below 0 indicated a PO-bias. Out of the 48 alternating dative verbs, only 7 were DO-biased, and all of them tended to have only weak biases compared to PO-biased verbs. In the current study, we included the only DO-biased verb found in CLs' textbooks (*Integrated Chinese*), "*song* (to give)", with a weak bias of 0.34. Regarding the trials with this DO-biased verb, L1 speakers showed fixation bias to the recipient already in the verb window (CR1), as evidenced by a significant intercept ($b = -.21$, $p = .02$, Table S9). Therefore, it is

https://doi.org/10.1017/S1366728924000749 Published online by Cambridge University Press

possible that L1 speakers' fixation bias did not increase further in the second ambiguous window (CR2) due to the weak bias of this verb. By contrast, HSs had a weaker initial bias to fixate the recipient in CR1 ($b = -.18$, $p = .18$, Table S9), thus leaving more room for change, and showed a (marginal) increase in looks to the recipient in CR2 ($b = -.17$, $p = .06$, Table S9) as would be expected based on prediction using verb bias. On the other hand, the L2ers looked (marginally) more to the theme in CR2, a pattern for which we have no ready explanation. However, given that these effects were marginal and based on a single lexical item, no strong conclusions can be drawn. Future studies may consider including more DO-biased verbs to test prediction among L1 speakers, and L2 and HSs of more advanced proficiency.

Similar to the findings from Wolk et al. (2011), Şafak and Hopp (2023), learners in this study were shown to anticipate upcoming arguments on the basis of selectional constraints of dative verbs. Moreover, learners in the current study demonstrated active prediction in a fully ambiguous region before the first postverbal noun was heard, contrasting with Wolk et al. (2011), Şafak and Hopp (2023), which observed effects in the first complement region after the argument was encountered. Therefore, our findings present clearer evidence for active prediction in the strict sense, that is, activation of linguistic information before input that carries the information is processed (Pickering & Gambi, 2018).

The present study extends the findings from previous L2 studies of the predictive use of dative verb constraints in English to Mandarin, a language in which the DO and PO dative constructions in question share similar word orders yet different semantic constraints from those in English (for instance, MAKE and TELL verbs can alternate between DO and PO in English, but not in Mandarin). Crucially, the learners in our study showed effects of prediction comparable in timing and degree to those observed in home-country-raised native speakers. These results align with the native-like prediction in the German-speaking L2 learners of English on dative verb restrictions reported in Şafak and Hopp (2023), but differ from recent studies that have shown L2 learners make predictions based on verb constraints or situational context yet their predictions tend to be slower and weaker than those of L1 speakers (e.g., Dijkgraaf et al., 2019; Ito et al., 2018).

There was no clear evidence showing that proficiency modulated learners' engagement in prediction in this study. Lower-level proficiency learners (who had been taking Mandarin courses in U.S. colleges for a year and a half only) performed similarly to advanced learners (who were taking Mandarin courses in a university in China for preparation of an undergraduate or graduate program fully instructed in Mandarin), and they both were not distinguishable from L1 speakers. These results contrast with those from Wolk et al. (2011), where only L1 German L2 learners of English of the higher proficiency group predicted based on gradient verb bias, whereas the lower proficiency group tended to predict a recipient immediately after the verb regardless of verb bias with the DO constructions. However, the absence of proficiency effects in this experiment conforms with the results from a series of recent studies that have directly probed proficiency effects on L2 prediction yet found none (e.g., Kim & Grüter, 2021; Mitsugi, 2020; Perdomo & Kaan, 2021, see Kaan & Grüter, 2021 for a discussion).

One possible reason for native-like prediction among CLs and the lack of modulation by proficiency that we observed in this study may be positive transfer from analogous verb constraints in the first/dominant language. Recall that while the CL group included participants with a variety of first/dominant languages, the majority (37 out of 60) were English-dominant, and additional analyses showed the same pattern of results between the English-dominant CLs and the non-English-dominant ones. Like Mandarin, English has DO and PO datives as well as rules of alternation between them. Let us tentatively consider if it is possible that the native English speakers had benefited from dative verb constraints and their corresponding processing routines in English. The English translation equivalents of TELL verbs used in this study such as *tell* and *teach* can alternate yet are biased to DO constructions (bias information from a norming study using magnitude estimation in Jaeger & Snider, 2013), while English MAKE verbs such as *make* and *fry* are alternating but also strongly biased to be followed by a theme rather than a recipient. In other words, the gradient constraints of English TELL and MAKE verbs align with the categorical constraints of these verbs in Mandarin. Thus English speakers may predict in the correct directions by relying on the verb constraints and predictive processes they are familiar with from English. Let us then look at the alternating GIVE verbs: The translation equivalents of the DO-biased verb *song* could be *give* or *send* in English, and both are suggested to have DO-bias (Jaeger & Snider, 2013). As for the translation equivalents of the PO-biased verbs, *zu* is *rent* in English, which is suggested to be PO-biased as well (Jaeger & Snider); *fen* is *share* (with the meaning of dividing food, money, goods, and so forth and giving part of it to someone else), which tends to be followed by an inanimate theme. Therefore, English speakers can also employ the gradient constraints of these translation equivalents during predictive processing. It is intriguing to explore the possibility of cross-linguistic influence on predictive use of dative verb constraints. For instance, future studies can manipulate the variable of congruent versus incongruent verb constraints between target language and learners' first/dominant languages.

Although we cannot rule out the possibility of facilitation through cross-linguistic influence among English-dominant learners, note that the 23 CLs with other dominant languages also demonstrated similar patterns of performance. These non-English speakers had diverse first/dominant languages that differed in word order (e.g., SVO, SOV, etc.) and whether or not they had alternating dative verbs (e.g., German, Thai, versus Japanese). This suggests that the effect in the overall CL group cannot be solely attributed to facilitation from congruent phenomena in English.

Interestingly, Şafak and Hopp (2023) attributed the less native-like prediction patterns in their L1 Turkish (versus the L1 German) learner group to differences in L1 word order. More specifically, they argued that a verb-final L1 affords less experience using the verb as a predictive cue, leading to less use of this cue in L2 processing. In the current study, we had only a subset of 11 CLs with a verb-final L1, therefore could not perform a statistical analysis to test this hypothesis more directly. A follow-up study with L1 Japanese and Korean learners of Mandarin is currently underway in our lab to further probe the relation between L1 word order and verb-based prediction in the L2.

Although not originally conceived to do so, the present study also contributes to the emerging line of research on predictive processing among HSs. HSs' predictive use of dative verbs, comparable to the pattern observed among home-country-raised speakers, aligns with the results from Ito et al. (2023) on prediction based on semantic constraints of transitive verbs among German-dominant Vietnamese HSs. On the other hand, recent studies on predictive use of morphosyntactic cues among HSs (e.g., Fuchs, 2022b, on gender markers; Karaca et al., 2023, on case markers) have shown reduced, slower or different predictions between HSs and home-country-raised speakers. This emerging pattern of asymmetrical performance on verb constraints versus morphosyntactic

https://doi.org/10.1017/S1366728924000749 Published online by Cambridge University Press

cues suggests HSs' engagement in prediction may vary differentially depending on the nature of the linguistic cues involved. More research on predictive processing among HSs is needed to further examine this claim.

As an initial endeavour to investigate Mandarin users' real-time predictive use of dative verbs, a limitation of this study is that only a few verbs of each category of constraints were tested. Future studies can increase the diversity of dative verbs to see whether the prediction patterns found here still hold with an increased number of verbs in each type of constraint.

Another consideration that must be kept in mind in the interpretation of eye gaze data from this and other VWP experiments on the dative alternation is that these data reflect participants' expectation of an (inanimate) theme or an (animate) recipient following the verb, which does not necessarily map to a PO or DO construction, respectively, in a one-to-one manner. As a reviewer has pointed out, for MAKE verbs that can appear in a ditransitive (PO) construction but also take a theme argument only and form a simple transitive structure (e.g., 玛丽做了一个蛋糕。'Mary made a cake.'), there is a possibility that participants might anticipate a simple transitive sentence, which would also be reflected in predictive looks to the theme. Within the design of the present study, we cannot be sure to what extent these predictive looks are driven by expectations for a ditransitive PO construction or a simple transitive sentence. Future studies will be needed to tease apart these different expectations, as well as the consequences of such expectations for error-driven learning.

In addition, as pointed out by a reviewer, it is worth noting that all VWP studies to date that have probed dative-verb-based predictions (Şafak & Hopp, 2023; Scheepers et al., 2007; Tily et al., 2008; Wolk et al., 2011), including the current one, relied on sentences in which the recipient was always prototypically animate and the theme was always prototypically inanimate. This may have facilitated predictive looks in that participants could quickly assign the recipient argument to the animate entity, and the theme argument to the inanimate entity. In other words, these VWP studies have tested the most prototypical scenario regarding the animacy of the two arguments, which language users are mostly likely to encounter in the routine processing of dative constructions. The most important motive for this design was to minimize confounding effects from animacy, which can independently affect the allocation of eye gaze (e.g., Fernald et al., 2008) as well as influence the choices of dative constructions (e.g., Bresnan et al., 2007). Further investigation of the role of animacy in the predictive processing of the dative alternation thus remains an intriguing topic for future studies to explore.

## 5. Conclusion

In conclusion, this study presents new evidence of active prediction of the dative alternation in Mandarin for both categorical and gradient constraints and demonstrates these effects generalize from home-country-raised native speakers to L2 learners and HSs. Notably, no delay or reduction in the prediction effects was observed among the latter two groups. These findings provide direct support for the assumption of error-driven learning accounts of the dative alternation, that is, language users actively predict upcoming arguments on the basis of verb information during real-time sentence processing. To examine whether these predictions lead to learning, the critical claim of error-driven learning accounts, future studies will have to test whether language users adjust their real-time predictions on the dative alternation according to recent exposure to specific dative constructions.

**Supplementary material.** To view supplementary material for this article, please visit http://doi.org/10.1017/S1366728924000749.

**Data availability statement.** The data that support the findings of this study are openly available in the Open Science Framework at https://osf.io/pvymf/.

**Acknowledgements.** This research was funded by a Doctoral Dissertation Research Improvement Grant from the U.S. National Science Foundation (BCS-2213581). We would like to thank Shin Fukuda, Carrie N. Jackson, Li (Julie) Jiang, William O'Grady, Bonnie D. Schwartz & Yang Zhao for assistance with various aspects of the research. We thank Jia Liu, Jue Wang, and Baorui (Louis) Xu for their comments on experimental stimuli. We also extend our thanks to the audiences at HSP 2023 at the University of Pittsburgh, BUCLD 48 at Boston University, and AMLaP Asia 2 at the Chinese University of Hong Kong for their constructive comments. We are grateful to all our participants. We are also thankful to three anonymous reviewers for their very helpful feedback.

**Competing interests.** The authors declare none.

## References

**Bernolet, S.**, & **Hartsuiker, R. J.** (2010). Does verb bias modulate syntactic priming? *Cognition*, **114**(3), 455–461. https://doi.org/10.1016/j.cognition.2009.11.005.

**Boersma, P.** & **Weenink, D.** (2022). *Praat: doing phonetics by computer [Computer program]*. Version 6.2.12, retrieved 17 April 2022 from http://www.praat.org/.

**Bovolenta, G.**, & **Marsden, E.** (2021). Prediction and errorbased learning in L2 processing andacquisition: A conceptual review *Studies in Second Language Acquisition*, **44**(5), 1384–1409. https://doi.org/10.1017/S0272263121000723.

**Bresnan, J.**, **Cueni, A.**, **Nikitina, T.**, & **Baayen, H.** (2007). Predicting the dative alternation. In *Proceedings of the royal Netherlands academy of science workshop on foundations of interpretation* (pp. 69–94).

**Chan, I. L.**, & **Chang, C. B.** (2018). LEXTALE_CH: A quick, character-based proficiency test for Mandarin Chinese. In A. B. Bertolini & M. J. Kaplan (Eds.), *Proceedings of the 42nd annual Boston University conference on language development* (pp. 114–130). Somerville, MA: Cascadilla Press.

**Chang, F.**, **Dell, G. S.**, & **Bock, K.** (2006). Becoming syntactic. *Psychological Review*, **113**(2), 234–272. https://doi.org/10.1037/0033-295x.113.2.234.

**Chen, X.**, **Wang, S.**, & **Hartsuiker, R. J.** (2022). Error-based structure prediction in language comprehension: Evidence from verb bias effects in a visual-world structural priming paradigm for Mandarin Chinese. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, **48**(1), 60–71. https://doi.org/10.1037/xlm0001048.

**Dijkgraaf, A.**, **Hartsuiker, R. J.**, & **Duyck, W.** (2019). Prediction and integration of semantics during L2 and L1 listening. *Language, Cognition and Neuroscience*, **34**(7), 881–900. https://doi.org/10.1080/23273798.2019.1591469.

**Fazekas, J.**, **Jessop, A.**, **Pine, J.**, & **Rowland, C.** (2020). Do children learn from their prediction mistakes? A registered report evaluating error-based theories of language acquisition. *Royal Society Open Science*, **7**(11), 180877–180877. https://doi.org/10.1098/rsos.180877.

**Fernald, A.**, **Zangl, R.**, **Portillo, A. L.**, & **Marchman, V. A.** (2008). Looking while listening: Using eye movements to monitor spoken language comprehension by infants and young children. In I. A. Sekerina, E. M. Fernández, & H. Clahsen (Eds.), *Developmental psycholinguistics: On-line methods in children's language processing* (pp. 97–135). John Benjamins Publishing Company. https://doi.org/10.1075/lald.44.06fer.

**Fuchs, Z.** (2022a). Eyetracking evidence for heritage speakers' access to abstract syntactic agreement features in real-time processing. *Frontiers in Psychology*, **13**, 960376. https://doi.org/10.3389/fpsyg.2022.960376.

**Fuchs, Z.** (2022b). Facilitative use of grammatical gender in Heritage Spanish. *Linguistic Approaches to Bilingualism*, **12**(6), 845–871. https://doi.org/10.1075/lab.20024.fuc.

**Goldberg, A. E.** (2019). *Explain me this: Creativity, competition, and the partial productivity of constructions*. Princeton: Princeton University Press. https://doi.org/10.1515/9780691183954.

**Huang, C.-T. J., Li, Y.-H. A.**, & **Li, Y.** (2009). *The syntax of Chinese* (1st ed.). Cambridge University Press. https://doi.org/10.1017/CBO9781139166935

**Ito, A.**, & **Knoeferle, P.** (2022). Analysing data from the psycholinguistic visual-world paradigm: Comparison of different analysis methods. *Behavior Research Methods*, **55**(7), 3461–3493. https://doi.org/10.3758/s13428-022-01969-3.

**Ito, A., Nguyen, H. T. T.**, & **Knoeferle, P.** (2023). *German-dominant Vietnamese heritage speakers use semantic constraints of German for anticipation during comprehension in Vietnamese* (pp. 1–18). Bilingualism: Language and Cognition. https://doi.org/10.1017/S136672892300041X.

**Ito, A., Pickering, M. J.**, & **Corley, M.** (2018). Investigating the time-course of phonological prediction in native and non-native speakers of English: A visual world eye tracking study. *Journal of Memory and Language*, **98**, 1–11. https://doi.org/10.1016/j.jml.2017.09.002.

**Jaeger, T. F.**, & **Snider, N. E.** (2008). Implicit learning and syntactic persistence: Surprisal and cumulativity. In Proceedings of the 29th annual Cognitive Science Society (CogSci09*)* (pp. 1061–1066). Austin, TX: Cognitive Science Society.

**Jaeger, T. F.**, & **Snider, N. E.** (2013). Alignment as a consequence of expectation adaptation: Syntactic priming is affected by the prime's prediction error given both prior and recent experience. *Cognition*, **127**(1), 57–83. https://doi.org/10.1016/j.cognition.2012.10.013.

**Kaan, E.**, & **Chun, E.** (2018). Priming and adaptation in native speakers and second-language learners. *Bilingualism: Language and Cognition*, **21**(2), 228–242. https://doi.org/10.1017/S1366728916001231.

**Kaan, E.**, & **Grüter, T.** (2021). Prediction in second language processing and learning. In E. Kaan & T. Grüter (Eds.), *Prediction in second language learning and processing* (pp. 1–24). John Benjamins. https://doi.org/10.1075/bpa.12.01kaa.

**Karaca, F., Brouwer, S., Unsworth, S.**, & **Huettig, F.** (2023). Morphosyntactic predictive processing in adult heritage speakers: Effects of cue availability and spoken and written language experience. *Language, Cognition and Neuroscience*, 1–18. https://doi.org/10.1080/23273798.2023.2254424.

**Kaschak, M. P., Kutta, T. J.**, & **Jones, J. L.** (2011). Structural priming as implicit learning: Cumulative priming effects and individual differences. *Psychonomic Bulletin and Review*, **18**(6), 1133–1139. https://doi.org/10.3758/s13423-011-0157-y.

**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. https://doi.org/10.1017/S0272263120000443.

**Kuznetsova, A., Brockhoff, P. B.**, & **Christensen, R. H. B.** (2017). lmerTest Package: Tests in linear mixed effects models. *Journal of Statistical Software*, **82**(13), 1–26. https://doi.org/10.18637/jss.v082.i13.

**Liu, F.** (2006). Dative constructions in Chinese. *Language and Linguistics*, **7**(4), 863–904.

**Liu, Y., Yao, T., Bi, N., Ge, L.** & **Shi, Y.** (2021). *Integrated Chinese* (4th ed., Vol. **1**). Boston, MA: Cheng & Tsui.

**Luo, H., Li, Y.**, & **Li, M.** (2019). Heritage language education in the United States: A national survey of college-level Chinese language programs. *Foreign Language Annals*, **52**(1), 101–120. https://doi.org/10.1111/flan.12378.

**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. https://doi.org/10.1515/iral-2017-0118.

**Özsoy, O., Çiçek, B., Özal, Z., Gagarina, N.**, & **Sekerina, I. A.** (2023). Turkish-German heritage speakers' predictive use of case: webcam-based vs. in-lab eye tracking. *Frontiers in Psychology*, **14**, 1155585. https://doi.org/10.3389/fpsyg.2023.1155585.

**Perdomo, M.**, & **Kaan, E.** (2021). Prosodic cues in second-language speech processing: A visual world eye-tracking study. *Second Language Research*, **37**(2), 349–375. https://doi.org/10.1177/0267658319879196.

**Peter, M., Chang, F., Pine, J. M., Blything, R.**, & **Rowland, C. F.** (2015). When and how do children develop knowledge of verb argument structure? Evidence from verb bias effects in a structural priming task. *Journal of Memory and Language*, **81**, 1–15. https://doi.org/10.1016/j.jml.2014.12.002.

**Pickering, M. J.**, & **Gambi, C.** (2018). Predicting while comprehending language: A theory and review. *Psychological Bulletin*, **144**(10), 1002–1044. https://doi.org/10.1037/bul0000158.

**Şafak, D. F.**, & **Hopp, H.** (2023). Cross-linguistic differences in predicting L2 sentence structure: The use of categorical and gradient verb constraints. *Studies in Second Language Acquisition*, **45**, 1234–1260. https://doi.org/10.1017/S0272263123000256.

**Scheepers, C., Williams, R.S., Mohr, S., Arai, M.**, & **van Gompel, R.P.G.** (2007). Sometimes it's better to donate than to give: Syntactic projections in on-line sentence comprehension. In *Poster presented at the 20th annual CUNY conference on human sentence processing*, San Diego, USA, March 29–31.

**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. https://doi.org/10.1017/S1366728922000736.

**R Core Team** (2022). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/.

**Tily, H., Hemforth, B., Arnon, I., Shuval, N., Snider, N.**, & **Wasow, T.** (2008). Eye movements reflect comprehenders' knowledge of syntactic structure probability. In Talk presented at the 14th architectures and mechanisms for language processing (AMLaP*)*, Cambridge, UK, September 04–06.

**Wolk, C., Wolfer, S., Baumann, P., Hemforth, B.**, & **Konieczny, L.** (2011). Acquiring English dative verbs: proficiency effects in German L2 learners. In Carlson, C. Hölscher, & T. Shipley (Eds.), *Proceedings of the 33rd annual conference of the cognitive science society*. Cognitive Science Society.

**Ye, S.** (2019). Jiyu pinggu he xuanyong yuanze de "Zhongwen Ting Shuo Du Xie" jiaocai pingxi: Jian lun "muyu zhiru" xianxiang (An analysis on Integrated Chinese under Evaluation and Selection Principles: As well as the phenomenon of First Language Implantation). *Haiwai Huawen Jiaoyu (Overseas Chinese Education)*, **4**, 46–54.

**Zhu, Y. A.**, & **Grüter, T.** (2024). Structural priming facilitates L2 learning of the dative alternation in Mandarin. *Studies in Second Language Acquisition*. Advance online publication. https://doi.org/10.1017/S027226312400041X