Studies in Second Language Acquisition (2025), **47**, 136–156
doi:10.1017/S0272263124000639

RESEARCH ARTICLE

The role of processing goals in second language predictive processing

A visual–world eye–tracking study of Korean honorific agreement

Hyunwoo Kim[1], Kitaek Kim[2] and Joonhee Kim[3]

[1]Yonsei University, Seoul, Republic of Korea; [2]Seoul National University, Seoul, Republic of Korea and [3]University of Hawaii at Manoa, Manoa, HI, USA
Corresponding author: Kitaek Kim; Email: kitaek@snu.ac.kr

(Received 26 August 2023; Revised 09 September 2024; Accepted 25 September 2024)

Abstract

This study investigates how second language (L2) learners engage in prediction based on their processing goals. While prediction is a prominent feature of human sentence comprehension in first–language speakers, it remains less understood when and how L2 learners engage in predictive processing. By conducting a visual–world eye–tracking experiment involving Chinese–speaking L2 learners of Korean, we tested the hypothesis that L2 learners determine whether to engage in prediction by evaluating the costs and benefits of anticipatory processing. The experiment specifically focused on the impact of a top–down comprehension goal for L2 learners' predictive use of an honorific form in Korean by providing them with different types of task instruction. Our results indicated that all groups engaged in predictive processing in early and entire predictive regions. However, in the late predictive region, L2 learners presented with a prediction–oriented task, but not those with a simple comprehension task, actively generated expectations about the honorific status of an upcoming referent. These findings lend support to the utility account of L2 prediction, suggesting that L2 learners' engagement in prediction depends on their current goals and strategies for processing efficiency.

Keywords: Korean honorific agreement; L2 prediction; processing goals; utility of prediction; visual–world eye-tracking

Introduction

Anticipating forthcoming information during sentence processing is a prominent characteristic of human sentence comprehension (Huettig, 2015; Kuperberg & Jaeger, 2016; Pickering & Gambi, 2018). In first language (L1) processing, comprehenders can preactivate various aspects of the linguistic representation of a word or phrase before

encountering the actual expression. However, the status of prediction as a basic aspect of second language (L2) processing mechanisms remains unclear.

Recent research has attempted to characterize specific features associated with prediction and elucidate their impact on L2 predictive processing, as well as on L2 acquisition and learning more broadly (for comprehensive reviews, see Bovolenta & Marsden, 2022; Schlenter, 2023). In line with these efforts, several studies have explored various factors to characterize L2 processing mechanisms (e.g., Hopp & Lemmerth, 2018; Ito, Corley, & Pickering, 2018; Kim & Grüter, 2021; Mitsugi, 2020). Notably, Kaan and Grüter (2021) proposed the concept of *utility* as a function that shapes L2 predictive processing. The utility account of L2 prediction posits that L2 learners determine whether to engage in prediction by assessing the costs and benefits of anticipatory processing. This account predicts that L2 learners are more likely to engage in prediction when the benefits of generating expectations outweigh the costs while reducing prediction when the costs outweigh the benefits.

The utility account presents a compelling framework for understanding the dynamic and adaptive nature of L2 predictive processing, highlighting the importance of considering multiple influential factors (Bovolenta & Marsden, 2022; Kaan, 2014; Kaan & Grüter, 2021; Schlenter, 2023). However, this proposal is yet to be explicitly tested, leaving many unanswered questions about when and how L2 learners engage in predictive processing (Grüter & Rohde, 2021; Jackson & Hopp, 2020; Kaan & Grüter, 2021). To address this gap, the current study aims to contribute to the understanding of L2 predictive processing by investigating the role of processing goals as a potential factor that may enhance the utility of prediction. Specifically, this study employs the visual–world eye–tracking paradigm to explore the impact of top–down comprehension goals on L2 learners' predictive use of an honorific form in Korean.

The influence of comprehension goals on modulating prediction has been well-documented in L1 processing (for a review, see Kuperberg & Jaeger, 2016). For example, Brothers, Swaab, and Traxler (2017) found that English speakers exhibited stronger prediction when they were instructed to actively anticipate the final word of passages, compared with when they read passages simply for comprehension. This significant effect emerged early, approximately 300–500 milliseconds (ms) after the onset of the final critical word. However, the extent to which processing goals similarly affect L2 predictive processing remains less clear. According to the utility account of L2 prediction, tasks that emphasize top–down goals promoting prediction should render prediction a beneficial strategy for efficient processing, thereby encouraging L2 learners to engage in stronger predictive behaviors. To test this hypothesis, we assigned L2 participants to two groups, each receiving distinct instructions during the task. Based on Brothers et al. (2017), one group was instructed to simply listen to a sentence and answer a question, while the other group was directed to actively anticipate a specific target referent that would be mentioned in the audio stimuli.

The role of utility in prediction

In recent decades, prediction has become a topic of growing interest in the field of psycholinguistics. While definitions of prediction may vary across studies, a prevailing perspective, particularly in the domain of language processing, characterizes it as the "activation of a linguistic representation of a word before the comprehender encounters that word" (Pickering & Gambi, 2018, p. 1005). Anchored within this conceptual framework of prediction, numerous studies have proffered substantial evidence of

https://doi.org/10.1017/S0272263124000639 Published online by Cambridge University Press

predictive processing among language users, employing temporally sensitive methodologies such as event–related potential (ERP, e.g., Brothers et al., 2017; Kutas, DeLong, & Smith, 2011) and visual–world eye–tracking techniques (e.g., Altmann & Kamide, 1999; Altmann & Mirkovic, 2009; Tanenhaus, Spivey-Knowlton, Eberhard, & Sedivy, 1995).

For example, Brothers et al. (2017) conducted an EPR experiment using the rapid serial visual presentation paradigm. In their study, English speakers were presented with passages containing predictable final critical words (e.g., *Thomas didn't like the temperature of his drink. He thought it was much too hot*) and passages with unpredictable final words (e.g., *Thomas didn't like the look of the water. He thought it was much too hot*). The analysis of electroencephalogram recordings during the comprehension task revealed a larger N400 amplitude—typically associated with processing difficulty arising from semantic anomalies—for unexpected continuations compared with expected ones. In another study, Altmann and Kamide (1999) utilized the visual–world eye–tracking paradigm to explore predictive eye movements based on lexical information among English speakers. During the task, participants listened to auditory stimuli such as "The boy will eat…" while simultaneously viewing images of a ball, toy car, toy train, and cake. The results indicated that participants initiated fixations on the image of the cake, the only edible item within the visual display, even before the onset of the target item in the audio stimuli (See Kamide, Altmann, & Haywood, 2003, for similar findings). Collectively, these studies suggest that comprehenders preactivate certain aspects of the target word before encountering it, supporting the notion of prediction as a fundamental component of language processing.

Despite the ample evidence supporting predictive processing, recent research indicates that prediction may not always occur but is instead influenced by various factors (Huettig, 2015; Kuperberg, 2007, 2013; Kuperberg & Jaeger, 2016; Pickering & Gambi, 2018). Several studies have explored how comprehenders adjust the extent of their predictive processing based on available cues and the strategies they employ (see Pickering & Gambi, 2018, for a review). These investigations suggest that comprehenders are more likely to engage in prediction when it proves beneficial for the task at hand (e.g., Brothers et al., 2017; Fine, Jaeger, Farmer, & Qian, 2013; Kleinschmidt & Jaeger, 2015). For instance, in Brothers et al.'s (2017) study, the N400 component, associated with processing unexpected words, was more pronounced when participants were instructed to actively predict upcoming information than when they were simply asked to read each passage for comprehension.

The variability of prediction contingent on processing goals underscores its close connection with the efficiency of language processing, where comprehenders strategically modulate the extent of their predictive efforts to optimize their cognitive resources (Kuperberg, 2013; Kuperberg & Jaeger, 2016; Pickering & Gambi, 2018). In this cognitive process, comprehenders are hypothesized to evaluate the potential benefits and costs of prediction, allocating varying levels of cognitive resources depending on the perceived utility of prediction in a given context. For instance, when comprehenders recognize that prediction will significantly enhance processing efficiency, such as by facilitating comprehension or reducing cognitive load, they are more inclined to engage in anticipatory processing. Conversely, if prediction is cognitively taxing or introduces processing conflicts, they may reduce their reliance on predictive strategies. This perspective aligns with Kuperberg and Jaeger's (2016) conceptualization of prediction as having a utility function. According to this approach, prediction is a flexible mechanism that responds to the demands of the current context, rather than being automatically triggered by bottom–up input alone.

https://doi.org/10.1017/S0272263124000639 Published online by Cambridge University Press

Extending the utility–based perspective on prediction, several studies have investigated the role of the utility function in L2 prediction and the factors that may influence this process (e.g., Bovolenta & Marsden, 2022; Kaan & Grüter, 2021; Schlenter, 2023). For example, Kaan and Grüter (2021) proposed that cue reliability and utility play crucial roles in L2 predictive processing. According to their proposal, the individual variability observed in L1 predictive processing may also account for differences in L2 prediction, especially considering the limited cognitive and linguistic resources available to L2 learners. Moreover, they suggest that unique factors influence L2 learners, such as crosslinguistic influence, L2 proficiency, and L2 learning experience. These factors can impact linguistic cues' reliability, thus influencing the prediction's utility. For example, when predictive cues exhibit crosslinguistic similarities, L2 learners might perceive these cues as more reliable, prompting them to engage more fully in predictive processing. In contrast, if the target cues are absent or significantly different in the learners' L1, relying on these cues for prediction may impose substantial cognitive demands, hindering learners' ability to generate expectations based on such unfamiliar cues (e.g., Hopp & Lemmerth, 2018).

The utility–based account of L2 predictive processing also suggests that comprehension goals and awareness can significantly influence the degree of L2 prediction. Previous studies have shown that language users are more likely to engage in predictive processing when they recognize its importance within a given task (e.g., Andringa, 2020; Brothers et al., 2017; Curcic, Andringa, & Kuiken, 2019; Ito, 2023; Koch, Bulté, Housen, & Godfroid, 2021). For example, Ito (2023) conducted two types of visual world eye–tracking experiments involving Chinese–speaking learners of English. In one experiment, participants were asked to click on an object mentioned in the audio stimuli (comprehension task), while in another, they were instructed to anticipate how the sentence would likely unfold before selecting an object (prediction task). The results revealed that participants engaged in stronger predictive processing during the prediction task compared with the comprehension task. Based on these findings, Ito (2023) concluded that L2 prediction can be actively controlled by learners based on their comprehension goals.

Some studies found a modulating effect of awareness in L2 prediction. Koch et al. (2021) observed that Dutch–speaking L2 learners of German effectively used verb morphology to predict subject numbers in sentence processing. Notably, debriefing interviews revealed that these learners were aware of the predictive function of the verb form in completing the task. Similarly, Curcic et al. (2019) conducted a visual–world experiment with Dutch speakers learning a novel determiner–noun agreement pattern in an artificial language. Their findings indicated that the extent of prediction varied depending on learners' awareness of the determiner's role in the task. Specifically, only those learners who recognized the importance of determiners exhibited anticipatory eye movements toward the target noun based on the determiner prior to the noun's onset (see Andringa, 2020, for similar findings).

The increase in predictive processing as a function of comprehenders' awareness of the importance of prediction is in line with the utility account of prediction. However, despite the significance of processing goals and awareness, their potential influence has been primarily explored in the context of artificial language learning (e.g., Andringa, 2020; Curcic et al., 2019) or L1 processing (e.g., Brothers et al., 2017), with the exception of Ito (2023). To achieve a more comprehensive understanding of the role of processing goals and awareness in L2 prediction, further research is required to examine their roles in authentic L2 acquisition settings. To address this gap, the current study aims to explore the role of processing goals in L2 prediction by examining honorific agreement

https://doi.org/10.1017/S0272263124000639 Published online by Cambridge University Press

in Korean. Employing the visual–world eye–tracking paradigm, we investigate how L2 learners adjust their predictive use of an honorific form based on different types of task instructions. Before presenting our study, we will provide a brief overview of the linguistic features of Korean honorific agreement and discuss how this phenomenon is relevant to our research objectives.

Honorific agreement in Korean

In Korean, diverse linguistic devices are employed to signify the social status of a human referent (Sohn, 1999). One notable example is the use of honorific markers. As illustrated in (1), honorification is conveyed morphologically through markers such as the honorific nominative marker *–kkeyse* and the verbal suffix *–si* (including its phonetic variant *sy*). These honorific markers function as grammatical agreement morphemes (Koopman, 2005; Pak, 2006), specifically modifying referents of high social status. For instance, in contexts like (1), using the honorific suffix *–si* within a verb associated with a nonhonorable referent, such as *haksayng* 'student,' results in an agreement violation.

(1) Sensayngnim-kkeyse / Sensayngnim-i kyosil-ey tuleo-sy-ess-ta.
선생님-께서 / 선생님-이 교실-에 들어오-셨-다
teacher-NOM$_{HON}$ / teacher-NOM[1] classroom-to enter-HON-PAST-DECL
"The teacher entered the classroom."

Moreover, honorific agreement between the subject and the predicate is not only a morphological process but also involves semantic and pragmatic considerations (Choi, 2003). Honorific markers convey contextual information, reflecting the speaker's intent to show respect. Therefore, in neutral contexts where no such intention is indicated, the honorific suffix can be omitted, allowing agreement between an honorable subject and a predicate lacking an honorific marker.

Among the various types of honorific forms, this study focuses on the auxiliary verb *–kyeysi*, a word combined with the verb root *–kyey* 'exist' and the honorific suffix *–si*. It is a specific honorific form of the suppletive *–iss* 'be, exist, have' (Kim & Sells, 2007). While the nonhonorific form *–iss* can be used with any type of human referent, whether honorable or not, the honorific *–kyeysi* is exclusively used for referring to honorable individuals, as exemplified in (2). In (2a), the noun *sensayngnim* 'teacher,' typically interpreted as referring to an honorable person, can be modified by both the neutral predicate *wus-ko iss-nun* and the honorific-marked predicate *wus-ko kyeysi-nun*. In contrast, in (2b), the noun *haksayng* 'student,' which does not inherently denote an honorable status, is generally not appropriate to be modified by the honorific predicate *wusko-kyeysi-nun* under normal circumstances.

(2) a. Wus-ko iss-nun / Wus-ko kyeysi-nun sensayngnim
웃-고 있-는 / 웃-고 계시-는 선생님
laugh-PROG-ADN / laugh- PROG$_{HON}$-ADN teacher
'A teacher who is laughing'

[1]Abbreviations used in the glosses: ACC = Accusative marker, ADN = Adnominal marker, DECL = Declarative marker, HON = Honorific form, LOC = Locative marker, NOM = Nominative marker, PAST = Past tense marker, PRES = Present tense marker, PROG = Progressive marker, TOP = Topic marker.

https://doi.org/10.1017/S0272263124000639 Published online by Cambridge University Press

b. Wus-ko iss-nun / *Wus-ko kyeysi-nun haksayng
웃-고 있-는 / 웃-고 계시-는 학생
laugh-PROG-ADN / laugh- $PROG_{HON}$-ADN student
'A student who is laughing'

Crucial to the purpose of the present study, the suppletive type within the relative clause predicate in examples like (2) serves as a predictive cue for the honorific status of the subsequent noun. A predicate containing the honorific *–kyeysi* signals the appearance of a referent with an honorific status. In contrast, a predicate lacking the honorific form indicates the occurrence of any type of human referent without necessarily signaling an honorable person.

The predictive function of the honorific *–kyeysi* allows for assessing comprehenders' ability to use this cue to anticipate the honorific status of the subsequent noun. In a visual–world eye–tracking task, where participants listen to phrases like (2) while viewing images of a teacher and a student, we expect that participants who are engaging in predictive processing will orient their eye movements more frequently toward the image of a teacher when encountering the predicate with the honorific *–kyeysi*, compared with when they encounter the predicate without it.

Current study

Using a visual–world eye–tracking task, this study investigated whether L2 learners of Korean can use the honorific *–kyeysi* in a relative–clause predicate to anticipate the honorific status of the following noun. Importantly, we aimed to assess whether the extent of L2 predictive processing is influenced by specific comprehension goals. Building on previous research demonstrating that task instructions promoting prediction facilitate predictive processing in L1 speakers (Brothers et al., 2017), we sought to explore the modulating role of top–down processing goals in L2 learners' predictive use of the honorific form. To achieve these objectives, we provided participants with different types of task instructions. We formulated the following research questions to guide our investigation:

1. Do L2 learners demonstrate the ability to use the honorific form to predict the following noun?
2. Do processing goals modulate the extent of predictive processing among L2 learners?

Methods

Participants

The study involved 54 L2 learners of Korean (7 male and 47 female), recruited from the undergraduate and graduate student population at a university in South Korea. The L2 participants' mean age was 26.6 (*SD* = 2.9), ranging from 23 to 39. They studied Korean for an average of 4.9 years (*SD* = 2.6), with a mean onset age of 20.3 (*SD* = 2.9).

We established two criteria for selecting L2 participants for the visual–world eye–tracking task. First, to control for crosslinguistic influence in L2 predictive processing (e.g., Hopp & Lemmerth, 2018), we specifically targeted speakers of Mandarin Chinese, a language lacking honorific agreement between subjects and predicates. This selection minimized potential confounding effects from crosslinguistic similarities. Second, to

https://doi.org/10.1017/S0272263124000639 Published online by Cambridge University Press

ensure an accurate assessment of participants' predictive use of honorific cues, it was crucial that they correctly associate predicates containing the honorific form with an honorable referent. For this purpose, we conducted a fill–in–the–blank task, where participants selected appropriate predicates for given referents (i.e., honorable or nonhonorable persons). Based on the results, we excluded data from 8 participants, leaving 46 participants for analysis. Detailed information about the fill–in–the–blank task is provided in the *Fill–in–the–blank task* sections under the *Materials and procedure* and *Results* sections.

To examine the influence of processing goals on L2 predictive processing, participants were randomly assigned to one of two groups, each receiving different instructions for the eye–tracking task. Following the methodology used by Brothers et al. (2017), one group, designated as the prediction group (n = 23), was instructed to actively predict the referent mentioned in the auditory input. The other group, referred to as the comprehension group (n = 23), was directed to simply listen to each passage for comprehension without any emphasis on predicting the target referent. Detailed procedural information is provided in the *Visual–world eye–tracking task* section.

To ensure comparability between the two groups in terms of Korean learning experience and proficiency, we compared their onset age of L2 learning, duration of studying Korean, duration of stay in Korea, self–rated Korean proficiency, and scores from a proficiency test. The self–rated proficiency was obtained by asking participants to rate their Korean fluency in reading, writing, listening, and speaking on a ten–point scale. For the proficiency test, we used a modified version of the Test of Proficiency in Korean (Jeong, 2017), where participants selected the appropriate morpheme or word from four options to complete 20 sentences. Independent samples t-tests indicated no significant differences between the two groups in onset age of L2 learning ($t(44) = 0.993$, p = .326), duration of studying Korean ($t(44) = –0.854$, p = .398), duration of stay in Korea ($t(44) = 0.425$, p = .673), self–rated proficiency averaged across the four domains ($t(44) = 0.011$, p = .991), and proficiency test scores ($t(44) = 0.524$, p = .603). These results confirm that the two groups were closely matched in their Korean learning experience and proficiency. Table 1 presents detailed information regarding the background of the two learner groups.

In addition to the L2 learners, the study included a baseline group of 24 native speakers of Korean (NS group; 8 male and 16 female). These native speakers were recruited from the graduate and undergraduate student populations at a university in South Korea, with ages ranging from 20 to 33 years.

All participants reported normal or corrected–to–normal vision and had no history of visual or auditory deficits. Prior to the experiment, they provided written consent. As compensation for their participation, they received an equivalent of $10 in Korean

Table 1. Background information of L2 participants

	Prediction group		Comprehension group	
	Mean	SD	Mean	SD
Age	26.2	2.5	27.6	3.3
Onset age	20.0	2.3	20.9	3.6
Years of studying Korean	5.3	3.2	4.7	2.1
Months of residence in Korea	33.0	37.6	37.5	33.8
Self-ratings (1–10)	6.8	1.5	6.8	1.0
Proficiency test scores (0–20)	13.8	3.1	14.3	3.1

https://doi.org/10.1017/S0272263124000639 Published online by Cambridge University Press

currency. The study was reviewed and approved by the Ethics Board at Seoul National University.

Materials and procedure

A key prerequisite for assessing L2 learners' predictive processing in this study was ensuring their ability to associate an honorific form with a referent of honorific status. To illustrate, to preactivate the honorific feature of an upcoming referent (e.g., *teacher*) based on an honorific–marked predicate (e.g., *laugh*-HON), as exemplified in (2a), comprehenders must recognize that the predicate agrees with an honorable person and not with a nonhonorable person such as *student*. Therefore, after the visual–world eye–tracking task, we administered a fill–in–the–blank task to the L2 learners to assess their understanding of honorific agreement in Korean.

Fill–in–the–blank task

The task comprised 20 experimental items, divided into two conditions, each containing 10 items. Each item presented a blank space to be filled with a relative–clause predicate that modifies the subsequent referent, as illustrated in (3). The honorific condition (3a) included a noun with honorific status (e.g., *elusin* 'elder'), while the nonhonorific condition (3b) included a noun without honorific status (e.g., *sonye* 'girl') following the blank.

(3) a. Honorific condition

Cikum	ceki	chwum-ul	( )	elusin-un
지금	저기	춤-을		어르신-은
now	there	dance-ACC		elder-TOP

cey halmeni-eyyo.
제 할머니-에요
my grandmother-DECL
'The elder who is dancing over there now is my grandmother.'

b. Nonhonorific condition

Cikum	ceki	chwum-ul	( )	sonye-nun
지금	저기	춤-을		소녀-는
now	there	dance-ACC		girl-TOP

cey tongsayngi-eyyo.
제 동생이-에요
my little sister-DECL
'The girl who is dancing over there now is my little sister.'

In each item, participants were presented with four options, each containing a target predicate marked by different types of suffixes. Among these options, two included a contextually appropriate tense marker, while the other two choices included an inappropriate tense marker. Within each tense marker option, one choice included an honorific marker, and the other included a neutral marker. For example, in item (3), the four options were *chwu-si-n* (dance-HON-ADN_{PAST}), *chwu-n* (dance-ADN_{PAST}),

https://doi.org/10.1017/S0272263124000639 Published online by Cambridge University Press

chwu-si-nun (dance-HON-ADN_{PRES}), *chwu-nun* (dance-ADN_{PRES}). Given that (3a) included the temporal adverb *cikum* 'now,' indicating present tense, and the honorable referent *elusin* 'elder,' the correct answer was *chwu-si-nun* (dance-HON-ADN_{PRES}). While *chwu-nun* (dance-ADN_{PRES}) was also a plausible option, it was deemed incorrect for this condition because it does not demonstrate the participants' ability to associate an honorific-marked predicate with an honorable person. On the other hand, the correct answer for (3b) was *chwu-nun* (dance-ADN_{PRES}) since the following referent *sonye* 'girl' does not carry an honorific status.

The experimental items were interspersed with 20 filler items that contained a blank space in various grammatical positions within a sentence. The experimental items were counterbalanced across two lists, with each participant encountering a single condition type for each item.

Participants completed the fill–in–the–blank task individually through a web–based interface in a quiet lab setting. Prior to the task, they received both oral and written instructions in Korean. Each trial presented a target sentence with a blank on the computer screen, and participants were instructed to choose the most appropriate option from the four choices provided below the sentence. Each item was displayed on a separate page, and participants were not allowed to navigate back to previous items to alter their responses.

Visual–world eye–tracking task

The visual–world eye–tracking task included auditory stimuli paired with visual images. The auditory stimuli consisted of 20 critical items, with 10 in honorific and 10 in nonhonorific conditions. Each item comprised a critical sentence followed by a question, as illustrated in (4). These passages were counterbalanced across two conditions in which the relative–clause predicate was manipulated. Specifically, in the honorific condition, the relative–clause predicates contained the honorific *–kyeysi*, while in the nonhonorific condition, the predicates included the neutral *–iss*.

(4) Critical sentence

Ecey kongwen-eyse peynchi-ey
어제 공원-에서 벤치-에
yesterday park-at bench-LOC

anca-kyeysi-ess-ten / anca-iss-ess-ten palo ce
앉아-계셨-던 / 앉아-있었-던 바로 저
sit-exist$_{HON}$-PAST-ADN / sit-exist-PAST-ADN very that

halapeci-nun aisukhulim-ul cohahanta.
할아버지-는 아이스크림-을 좋아한다
old man-TOP ice cream-ACC like
'The old man right there who was sitting on the bench at the park yesterday likes ice cream.'

Question
Nwu-ka aisukhulim-ul cohahana-yo?
누-가 아이스크림-을 좋아하나-요
who-NOM ice cream-ACC like-Q
'Who likes ice cream?'

https://doi.org/10.1017/S0272263124000639 Published online by Cambridge University Press

After the relative clause predicate and before the target referent, we inserted a demonstrative adjectival phrase *palo ce* 'very that.' This phrase served as a padding element, creating a time window for participants to generate expectations about the subsequent referent. This interval, designated as a critical region for analysis, allowed us to investigate any potential preactivation of honorific information. During this period, we closely examined whether participants shifted their eye gaze toward the expected target image.

Following the padding phrase, the target referent appeared as the subject of the main clause. The target referents were carefully selected to represent entities that could logically receive honorific treatment and thus be modified by an honorific–marked predicate. These referents included various nouns, such as professional titles (e.g., *sensayngnym* 'teacher,' *kyoswunim* 'professor,' *uysasensayngnim* 'doctor,' *pyenhosanim* 'lawyer,' *kamtoknim* 'director or supervisor'), kinship terms (e.g., *hyengnim* 'older brother,' *apenim* 'father,' *emenim* 'mother,' *halapeci* 'grandfather or male elder person,' *halmeni* 'grandmother or female elder person,' *samchon* 'uncle'), and nouns denoting specific social statuses (*samonim* 'madam,' *sacangnim* 'boss or sir,' *acwumeni* 'madam or ma'am,' *senpaynim* 'senior,' *elusin* 'elder').

Since the critical sentences consistently included referents representing honorable individuals, it was necessary to prevent participants from developing strategies for expecting honorable referents before encountering the honorific–marked relative clause predicate. Therefore, we included 10 distractor items that resembled the critical sentences but used nonhonorific predicates and nonhonorable referents. The referents in the distractor items included *chengnyen* 'young man,' *congepwen* 'employee or worker,' *elini* 'child,' *kkomaai* 'kid,' *namcaai* 'boy,' *namhaksayng* 'male student,' *yehaksayng* 'female student,' *cwunghaksayng* 'middle–school student,' *yetaysayng* 'female college student,' and *sinipsayng* 'freshman.'

To ensure that the selected honorable and nonhonorable referents were clearly distinguished in terms of their compatibility with the honorific *–kyeysi*, we conducted a norming task to assess the acceptability of combining each noun with honorific and nonhonorific predicates. Thirty native speakers of Korean, who did not participate in the main experiment, were asked to read sentences and rate their naturalness on a scale from 1 (very unnatural) to 4 (very natural). A total of 30 sets of items were created by matching each referent (honorable and nonhonorable) with either an honorific–marked predicate (e.g., *anca-kyeysy-ess-ten* 'who was sitting') or a nonhonorific–marked predicate (e.g., *anca-iss-ess-ten* 'who was sitting') in sentences like (4) (excluding the question sentence).

Analysis of the norming task revealed that sentences were highly acceptable when a nonhonorific predicate modified a nonhonorable referent ($M = 3.90$, $SD = 0.31$), when it modified an honorable referent ($M = 3.67$, $SD = 0.48$), and when an honorific predicate modified an honorable referent ($M = 3.83$, $SD = 0.38$). In contrast, the acceptance rates were low when an honorific predicate modified a nonhonorable referent ($M = 1.10$, $SD = 0.31$). The acceptance ratings for the sentences in the four conditions were significantly different ($F(3, 87) = 403.249$, $p < .001$, $\eta^2 = 0.933$), with significantly lower ratings for sentences matching an honorific predicate with a nonhonorable referent than those in the other three conditions (all p-values $< .001$). These results confirmed that the honorable referents selected for this study were fully compatible with the honorific *–kyeysi*, while the nonhonorable referents were deemed unacceptable when modified by the honorific form. These findings establish that the honorific *–kyeysi* can function as a predictive cue for the honorable referents used in the eye–tracking experiment.

https://doi.org/10.1017/S0272263124000639 Published online by Cambridge University Press

In addition to the 10 distractor items that paired a nonhonorific predicate with a nonhonorable referent, we included 36 fillers describing diverse events that involved interactions between two human referents. Each item was followed by a question, prompting participants to select one of the referents presented on the display.

The auditory stimuli were recorded by a female native speaker of Korean, who delivered the utterances at a natural speech rate while maintaining standard intonation. To ensure consistency between the honorific and nonhonorific conditions in terms of the duration of the recording frames, we edited the recordings of the critical sentences using Praat (Boersma & Weenink, 2017). This editing process resulted in comparable lengths of critical sentences in both conditions, with no significant differences observed in the overall sentence length ($t(38) = 0.755$, $p = .455$), the offset of the predicate in the relative clause ($t(38) = –0.058$, $p = .954$), the onset of the padding phrase ($t(38) = 0.073$, $p = .942$), and the onset of the target referent ($t(38) = –0.123$, $p = .902$).[2]

The experimental stimuli were presented with visual scenes that included images of human faces paired with printed Korean words, as illustrated in Figure 1. No same image was reused for targets and competitors across trials. The major distinction between the two pictures was in their honorific status, with one portraying a person of honorable standing, such as an elderly man, and the other depicting a person of nonhonorable status, such as a child. In each visual scene, the audio recordings of the critical sentences consistently referred to honorable individuals, making the image representing the honorable referent the target for that trial. Conversely, the image representing the nonhonorable referent was designated as the competitor. The position of the target image within each visual scene was counterbalanced across items.

Figure 1. Example of a visual stimulus illustrating two images representing an old man (left) and a kid (right), along with the Korean words associated with each image printed below.

[2]Although we did not quantify the speech rate of our recordings, we had feedback from four Korean speakers who did not participate in the experiment. When asked to check the naturalness of the speech, they responded that the speech rate was natural and easy to understand for both experimental and distractor items. Additionally, during debriefing interviews, participants mentioned that the speed of recordings was natural.

https://doi.org/10.1017/S0272263124000639 Published online by Cambridge University Press

The visual images for the fillers featured two human characters with varying combinations of honorific status, i.e., both honorable characters, both nonhonorable characters or one honorable and one nonhonorable character.

Each L2 participant completed both the eye–tracking experiment and the fill–in–the–blank task, whereas the NS group only participated in the eye–tracking experiment. The participants' eye movements were tracked using an EyeLink Portable Duo system (SR Research Ltd., Mississauga, Ontario, Canada), which was connected to a 16–inch monitor. The system remotely captured eye–movement data from the right eye at a sampling rate of 1000 Hz, resulting in one sample taken every millisecond.

Prior to the experiment, participants received both written and oral instructions in Korean. To investigate the impact of top–down processing goals on L2 predictive processing, we provided distinct instructions to the two groups of participants. The Prediction group was directed to actively anticipate the referent that the sentence in the audio stimuli would describe. The specific instruction given to this group was "(English translation) Look at the picture on the screen and click on the character as quickly as possible that is mentioned in the question sentence. Try to predict who the question is likely to ask about." In contrast, the Comprehension group was instructed to listen to the sentences for comprehension without any specific emphasis on predicting a target referent. Specifically, they were instructed: "(English translation) Look at the picture on the screen and click on the character as quickly as possible that is mentioned in the question sentence." Similarly, the NS group, serving as a reference point, was instructed to perform the task solely for comprehension purposes. The rest of the experimental procedure remained consistent across the two learner groups and the native speaker group. Following the instruction phase, participants completed a calibration and validation process to ensure accurate eye tracking during the experiment. Subsequently, three practice trials were presented to help participants familiarize themselves with the task procedures.

Each trial began with a fixation cross appearing at the screen center. Participants initiated the trial by maintaining their gaze on the cross for a duration of 500 ms. If the tracker did not detect a fixation within the initial 5 seconds after the cross appeared, the screen automatically transitioned to the calibration page for recalibration. Once participants successfully fixated on the cross, a visual scene featuring two human faces, as depicted in Figure 1, was displayed on the screen for a duration of 2000 ms without any auditory signal. This silent period was included to provide participants with an opportunity to preview the visual context, including the face images and the printed words (e.g., Ferreira, Foucart, & Engelhardt, 2013). After the preview, the audio recording of a critical sentence was played. Subsequently, there was a 1000–ms pause, after which the recording of a question was played. Throughout this phrase, the visual scene remained on the screen. Upon listening to the question, participants were prompted to respond by clicking on one of the images displayed on the screen. There was no feedback provided after their response, and the next trial automatically began 1000 ms after the participants' mouse–click response.

Supplementary materials including experimental stimuli, data files, analysis scripts, and information on the normality assumptions of the models can be found at https://osf.io/erx8p/?view_only=764643127248490cb7bfc320eca547d4.

Analysis of eye–tracking data

We identified participants' fixations on the target and the competitor and aggregated them into 20–ms bins using the Eyelink DataViewer software. To assess participants' anticipatory–looking behavior, we calculated the proportions of looks directed toward

https://doi.org/10.1017/S0272263124000639 Published online by Cambridge University Press

the target, the competitor, and the remaining space for each trial. We hypothesized that any evidence of predictive processing would be reflected in an increased proportion of looks to the target in the honorific condition (sentences with the honorific *–kyeysi*) compared with the nonhonorific condition (sentences with the neutral *–iss*) during specific critical time windows.

As the critical region of analysis, we focused on the segment referred to as the 'predictive window.' This region spanned from the offset of the relative clause predicate to the onset of a target referent, which included the demonstrative adjectival phrase (*palo ce* 'very that') inserted as padding material. We hypothesized that any modulating effect of top–down goals would enhance participants' predictive use of the honorific *–kyeysi*, resulting in the Prediction group exhibiting a larger and/or more consistent effect of predictive processing within this time window.

For exploratory analyses, we further divided the predictive region into two distinct time windows. The first window, called the relative–clause boundary region, extended from the offset of the relative clause predicate to the onset of the padding phrase. The second window, referred to as the main clause region, covered the duration from the onset of the padding phrase to the onset of the target referent. Through the analysis of these narrowly defined time windows, we aimed to pinpoint the exact timing of the effect of top–down goals on participants' anticipatory processing. To accommodate for the saccadic lag in eye movements (Matin, Shao, & Boff, 1993), we applied an offset of 200 ms in the analysis of fixations within each time window. As a result, the duration for analysis was 1357.5 ms for the relative–clause boundary region (3389.9 to 4747.4 ms after the critical sentence onset), 913.0 ms for the main–clause region (4747.4 to 5660.4 ms after the critical sentence onset), and 2270.5 ms for the combined windows (3389.9 to 5660.4 ms after the critical sentence onset).

Before conducting statistical analyses, we processed data to control for autocorrelation, a common issue in eye–tracking data analysis where fixations at one time point strongly correlate with those in adjacent time bins (Ito & Knoeferle, 2023; Rayner, 2009). To manage this issue, we aggregated the proportion of looks to the target into larger time bins of 250 ms and transformed these proportions into empirical logits (Ito & Knoeferle, 2023). These logit–transformed proportions were analyzed using linear mixed–effects regression (Bates, Maechler, Bolker, & Walker, 2015) in R version 4.3.3 (R Core Team, 2024).

We constructed three mixed–effects models corresponding to the three critical regions. Each model included *Group* (NS, Comprehension, Prediction), *Condition* (Honorific, Nonhonorific), and their interaction as fixed effects. For *Group*, Helmert coding was applied, with one contrast comparing the NS group with the two L2 groups and another contrast comparing the Prediction group with the Comprehension group. The fixed effect of *Condition* was contrast-coded, with –0.5 assigned to the Honorific condition and 0.5 to the Nonhonorific condition. In addition to the fixed effects, participant and item were included as random effects. Initially, we constructed the maximal random–effects structure permitted by the design, including random intercepts for both participants and items, as well as random slopes for *Condition* and *Group* (Barr et al., 2013). However, due to issues related to model convergence, we modified the random effects structure by removing the by–item random slope for *Condition* (e.g., Hopp & Grüter, 2023; Kim & Grüter, 2021), resulting in a final model that included a by–participant random slope for *Condition*, a by–item random slope for *Group*, and random intercepts for participants and items. In the case of a significant interaction emerging between *Group* and *Condition*, we conducted separate analyses for each group, generating linear mixed–effects models that contained the fixed effect of

https://doi.org/10.1017/S0272263124000639 Published online by Cambridge University Press

Condition (contrast-coded) and the random effects of participant and item. The models also included by–participant and by–item random intercepts, as well as a by–participant random slope for *Condition*. To account for multiple comparisons, we adjusted the alpha level to .017 (.05 divided by 3). Details on the normality of these models are provided in the supplementary materials.

Results

Fill–in–the–blank task

Participants' responses were coded as either correct or incorrect. Correct responses included selecting an honorific predicate for an honorable referent and a nonhonorific predicate for a nonhonorable referent. In contrast, responses were considered incorrect when participants chose an honorific predicate for a nonhonorable referent or a nonhonorific predicate for an honorable referent.

Among the initial 54 L2 participants, eight scored 10 or below out of the 20 items. These participants' low scores suggest that they did not have stable knowledge regarding honorific agreement in Korean, which might indicate their potential difficulty in recognizing an honorific–marked relative clause predicate as a predictive cue for an upcoming referent in the eye–tracking experiment. Therefore, we excluded the eye gaze data of these eight participants from further analysis. The remaining L2 participants achieved an accuracy rate of at least 16 out of the 20 items. The accuracy rates between the Prediction group (M = 17.9, SD = 1.5) and the Comprehension group (M = 18.4, SD = 1.1) were not significantly different ($t(44) = 1.233$, $p = .226$), suggesting that both groups had comparable levels of accuracy in associating honorific–marked predicates with honorable referents.

Visual–world eye–tracking task

We first examined mouse–click accuracies for responses to the questions for the experimental items. All three groups demonstrated high accuracy rates. The NS group achieved a perfect accuracy score (M = 100%), followed by the Comprehension group (M = 98.8%, SD = 0.1) and the Prediction group (M = 97.9%, SD = 0.1). These results indicate that participants across all groups paid close attention to the sentence meanings throughout the task. For the subsequent analyses of eye gaze data, we included only those trials in which participants provided correct responses to the questions.

Figure 2, 3, and 4 illustrate the participants' fixations to the target and the competitor throughout the predictive window, including the relative-clause boundary and main–clause regions, for the NS group (Figure 2), the Prediction group (Figure 3), and the Comprehension group (Figure 4), respectively. Visual inspection reveals that the NS group and the Prediction group exhibited similar patterns of eye movements by demonstrating increased fixations toward the target relative to the competitor in the honorific condition compared with the nonhonorific condition within the critical time windows.[3] These results suggest that both groups shifted their gaze toward the anticipated target referent upon encountering the honorific *–kyeysi*. In contrast, the Comprehension group showed increased looks to the target only during the first time

[3]When we further examined the data for the 10 distractor items, which included nonhonorific relative predicates and nonhonorable characters, we observed that eye-gazes were almost evenly distributed between the target and nontarget characters for both the Prediction group and the Comprehension group. We thank an anonymous reviewer for suggesting this additional analysis.

https://doi.org/10.1017/S0272263124000639 Published online by Cambridge University Press

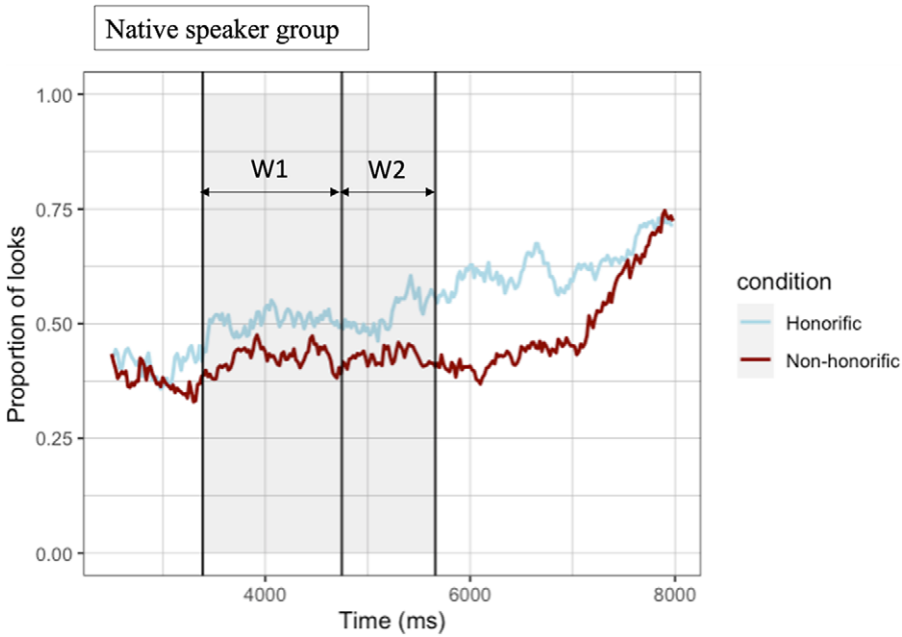

Figure 2. Mean proportion of looks to the target in honorific (blue line) and nonhonorific (red line) conditions for the native speaker group. W1: relative-clause boundary region (from the offset of the predicate in the relative clause to the onset of the padding phrase); W2: main–clause region (from the onset of the padding phrase to target onset).

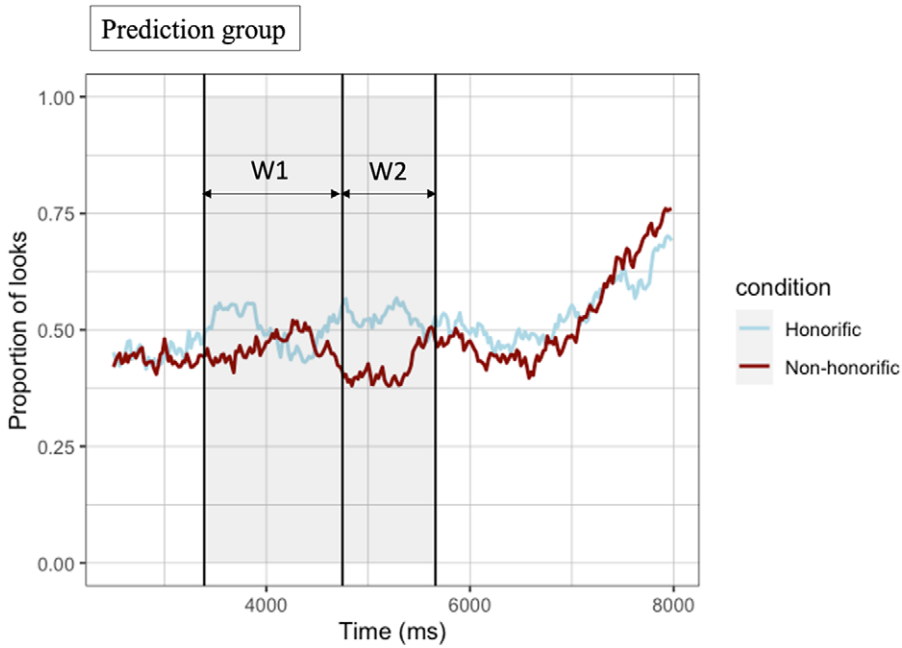

Figure 3. Mean proportion of looks to the target in honorific (blue line) and nonhonorific (red line) conditions for the Prediction group. W1: relative-clause boundary region (from the offset of the predicate in the relative clause to the onset of the padding phrase); W2: main–clause region (from the onset of the padding phrase to target onset).

https://doi.org/10.1017/S0272263124000639 Published online by Cambridge University Press

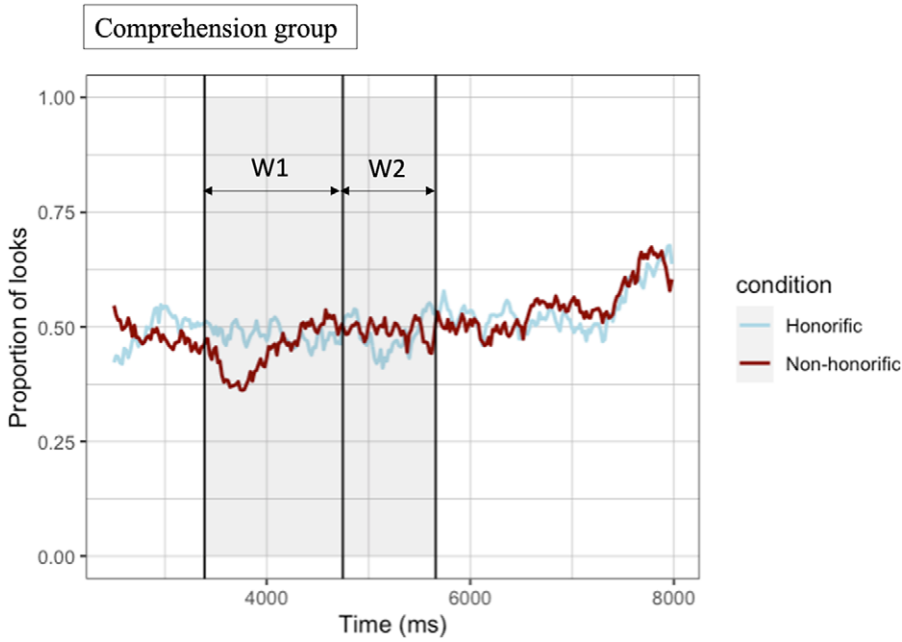

Figure 4. Mean proportion of looks to the target in honorific (blue line) and nonhonorific (red line) conditions for the Comprehension group. W1: relative-clause boundary region (from the offset of the predicate in the relative clause to the onset of the padding phrase); W2: main–clause region (from the onset of the padding phrase to target onset).

window, with their gaze behavior remaining relatively constant across the honorific and nonhonorific conditions during the subsequent windows.

To compare the processing patterns of the three groups in detail, we analyzed the results from the mixed–effects models. Table 2 presents the outcomes of these models across the three time windows.

In the predictive region, we observed a main effect of *Condition* driven by greater fixations on the target in the honorific than the nonhonorific condition. Similarly, in the relative–clause boundary region, our analysis only revealed a significant effect of *Condition*. However, in the main–clause region, we found a significant effect of *Condition* and its interaction with *Group* in the comparison between the Prediction and Comprehension groups. Separate analyses for each group demonstrated a significant effect of *Condition* in the NS group (b = –0.120, SE = 0.034, p = .002) and the Prediction group (b = –0.143, SE = 0.051, p = .009) but not in the Comprehension group (b = 0.024, SE = 0.054, p = .665). These results indicate that both the NS group and the Prediction group engaged in predictive processing across the critical regions, whereas the Comprehension group exhibited predictive processing only in the relative–clause boundary region.

Discussion and conclusion

The primary objective of this study was to assess the influence of top–down goals intended to enhance predictive abilities in the context of anticipatory processing among

https://doi.org/10.1017/S0272263124000639 Published online by Cambridge University Press

Table 2. Model outcomes from the analysis of eye–movement data

Region	Factor	ß	*SE*	*p*
Predictive region (i.e. relative-clause boundary region + main-clause region)	Intercept	–3.198	0.028	< .001***
	Group (NS vs. L2 groups)	–0.0001	0.033	.997
	Group (Comprehension vs. Prediction)	–0.013	0.037	.724
	Condition	–0.069	0.020	.001**
	Group (NS vs. L2 groups) × Condition	–0.026	0.043	.555
	Group (Comprehension vs. Prediction) × Condition	–0.054	0.051	.289
Relative-clause boundary region	Intercept	–3.190	0.031	< .001***
	Group (NS vs. L2 groups)	–0.003	0.039	.935
	Group (Comprehension vs. Prediction)	0.043	0.038	.252
	Condition	–0.061	0.025	.017*
	Group (NS vs. L2 groups) × Condition	0.023	0.053	.667
	Group (Comprehension vs. Prediction) × Condition	0.016	0.062	.801
Main–clause region	Intercept	–3.198	0.030	< .001***
	Group (NS vs. L2 groups)	–0.044	0.043	.301
	Group (Comprehension vs. Prediction)	–0.045	0.054	.406
	Condition	–0.078	0.027	.005**
	Group (NS vs. L2 groups) × Condition	–0.093	0.057	.106
	Group (Comprehension vs. Prediction) × Condition	–0.144	0.067	.035*

***$p < .001$
**$p < .01$
*$p < .05$

L2 Korean speakers. Employing the visual–world eye–tracking paradigm, we investigated how two groups of L2 learners with matched proficiency, who received different types of task instructions, utilized an honorific form to preactivate the honorific status of a subsequent referent.

During the time windows spanning the entire predictive region and specifically focusing on the relative–clause boundary region, we found a global effect of the honorific condition. This effect suggests that all groups, including native speakers and L2 learners, engaged in predictive processing. The predictive processing observed among the L2 learners is consistent with previous findings indicating that L2 learners can generate expectations during online comprehension (Dijkgraaf, Hartsuiker, & Duyck, 2017; Foucart, Romero-Rivas, Gort, & Costa, 2016; Ito et al., 2018; Kim & Grüter, 2021). However, unlike earlier research, which found a delayed prediction effect among L2 learners, our results revealed similar early processing patterns between native and L2 speakers. This early effect may be attributed to the L2 learners' high proficiency and the frequent occurrence of honorific markers in Korean. Moreover, the simplicity of the task, which required participants to make a binary choice (i.e., honorable versus nonhonorable character) based on a single honorific form, may have facilitated this early prediction. Further research should investigate whether predictive processing is observed in learners with lower proficiency and in contexts involving more complex linguistic phenomena.

In addition, the predictive processing observed in the L2 groups may be attributed to the presence of honorific forms in Mandarin Chinese. Although the language lacks honorific agreement between a subject and a predicate, it does employ an honorific system that features various forms expressing social respect and politeness, such as *nin* (a second–person honorific pronoun). As noted by a reviewer, this honorific system in

https://doi.org/10.1017/S0272263124000639 Published online by Cambridge University Press

Chinese may have facilitated the predictive processing of L2 participants. To further explore the potential influence of L1 knowledge, additional research is required to compare the current findings with those of L2 learners whose L1 lacks an honorific system, such as English.

Notably, we found an interactive effect of instruction in the main–clause region. The L2 group explicitly instructed to anticipate a target referent during the task (Prediction group) exhibited predictive processing patterns. In contrast, the L2 group who performed the task primarily for comprehension purposes (Comprehension group) did not display clear signs of predictive processing. These findings underscore the instrumental role of processing goals that foster prediction in facilitating L2 predictive processing. These results lend support to the utility account of L2 predictive processing (Bovolenta & Marsden, 2022; Kaan & Grüter, 2021; Schlenter, 2023), which posits that the degree of L2 prediction is modulated by how L2 learners evaluate the benefits and costs of engaging in anticipatory processing.

In the case of the Comprehension group, it is conceivable that they did not maximize the utility of their predictions for the task, due to their prior extensive experience with L1 Chinese, which is linguistically distinct from L2 Korean in the relevant morphology domain. Consequently, using an honorific form to anticipate an upcoming referent may have placed a substantial cognitive burden on these learners. In contrast, the findings from the Prediction group support the facilitative role of top–down goals in predictive processing. Despite the cognitive demands imposed by predictive processing, the explicit instruction provided to this group likely focused their attention on top–down comprehension goals. These findings are consistent with the notion that L2 prediction can be modulated by various factors, including top–down comprehension strategies (Ito, 2023; Kaan & Grüter, 2021).

It is also plausible that the specific instructions given to the Prediction group may have facilitated their predictive processing. Unlike Brothers et al. (2017), who instructed participants to anticipate specific words, the Prediction group in our study was prompted to predict a particular referent between two characters. As a reviewer noted, this specific instruction, along with a limited set of choices in our study, could have enhanced the predictive processing of the Prediction group. To gain a better understanding of the impact of instructions on predictive processing, future studies should explore the various nuances of instructional approaches.

In addition to the observed effect of instruction, there was some indication that our L2 participants used the absence of honorific marking as a predictive cue for anticipating nonhonorable referents. Although not statistically significant, there was a noticeable decline in their looks at the target in the nonhonorific condition, particularly in the main–clause region (W2) for the Prediction group and in the relative–clause boundary region (W1) for the Comprehension group, as illustrated in Figures 3 and 4.[4] This tendency suggests the learners' awareness of the systematic link between neutral predicates and nonhonorable characters. Further research is needed to explore these subtle yet discernible tendencies, examining potential variables influencing participants' predictive processing during the absence of honorific marking.

Our study provides novel evidence that the influence of task instruction on prediction among L1 speakers extends to the domain of L2 predictive processing. In earlier research, Brothers et al. (2017) demonstrated that when English speakers were explicitly instructed to anticipate specific target words during word–by–word

[4]We thank an anonymous reviewer for highlighting this tendency.

https://doi.org/10.1017/S0272263124000639 Published online by Cambridge University Press

comprehension, their anticipatory processing was significantly enhanced. Although Brothers et al.'s study provided crucial insights into the role of top–down goals in predictive processing, their findings were confined to lexical prediction among L1 speakers. The current study attempted to extend this research scope by exploring L2 processing, specifically in the context of honorific agreement in Korean. Our findings show that manipulating task goals had a substantial impact on predictive processing among L2 learners, further underscoring the intricate interplay between task instructions and prediction in L2 contexts.

Furthermore, our findings expand the scope of prior investigations that predominantly focused on comparing predictive abilities between L2 learners and native speakers. While numerous studies have examined L2 predictive processing across diverse learner populations and linguistic phenomena, much of their emphasis has been on evaluating how L2 predictive abilities measure up against those of native speakers. However, as Kaan and Grüter (2021) pointed out, differences in prediction ability between L1 and L2 speakers are inevitable due to their distinct language acquisition experiences and learning backgrounds. Consequently, there is a growing need for research to shift its focus toward understanding "when, how, and why L2 (and first language!) comprehenders engage in prediction, and what the relation is between prediction and learning" (Kaan & Grüter, 2021, p. 1). In this context, our study contributes to this line of inquiry by demonstrating that L2 learners can effectively engage in prediction, even to a degree comparable to L1 speakers, particularly when they recognize the utility of prediction as a strategic tool for accomplishing task–related objectives. The observed variations in the extent of prediction, contingent upon the specific task goals, indicate the flexible and adaptable characteristics of L2 prediction mechanisms. These mechanisms are subject to various factors, including the current circumstances of language use (Grüter & Rohde, 2021), underscoring the dynamic interplay between prediction, utility function, and task demands in L2 processing.

Finally, we acknowledge limitations that need to be addressed in future research. First, our investigation of the task effect employed a between–group design. As highlighted by a reviewer, this design choice may not entirely control for variations that could exist between the two groups, potentially influencing the outcomes. To address this concern, further research should explore a within–group design, wherein a single learner group engages in different types of tasks (e.g., Ito, 2023). Another limitation concerns the inclusion of printed words below each image during the eye–tracking task. We made this design choice to aid participants in recognizing the honorific status of each character. However, as noted by a reviewer, this addition may compromise ecological validity, potentially hampering participants' naturalistic processing. Hence, further research should explore whether similar findings are observed when presenting target images without printed words. Lastly, our mixed–effects models do not fully meet the normality assumption, as indicated by deviations in the residuals observed in the Q–Q plots. While strict normality is not always required for mixed–effects models, these deviations suggest that some caution is needed when interpreting the results. Future studies should address this issue by increasing sample sizes and considering alternative modeling approaches that better accommodate nonnormal data.

Despite these limitations, this study shows that L2 learners of Korean can proactively utilize honorific cues to generate expectations, particularly when provided with explicit task instructions that emphasize anticipation. The increased validity of prediction among our L2 learners may be attributed to their perceived benefits of generating predictions to fulfill ongoing processing goals, which outweigh the cognitive costs associated with processing target information without relying on their L1. Further

research is needed to explore how the effects of task goals might interact with other factors that enhance the utility of prediction. This exploration would broaden our understanding of specific contexts that can boost L2 learners' predictive abilities, while also contributing to the characterization of the mechanisms underlying the L2 predictive processing system.

Competing interest. The authors have no competing interests.

References

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

Altmann, G. T., & Mirkovic, J. (2009). Incrementality and prediction in human sentence processing. *Cognitive Science*, *33*, 583–609.

Andringa, S. (2020). The emergence of awareness in uninstructed L2 learning: A visual world eye tracking study. *Second Language Research*, *36*, 335–357.

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

Boersma, P., & Weenink, D. (2017). Praat: doing phonetics by computer [Computer program]. Version 6.0.36. http://www.praat.org/

Bovolenta, G., & Marsden, E. (2022). Prediction and error-based processing in L2 processing and acquisition. *Studies in Second Language Acquisition*, *44*, 1384–1409.

Brothers, T., Swaab, T. Y., & Traxler, M. J. (2017). Goals and strategies influence lexical prediction during sentence comprehension. *Journal of Memory and Language*, *93*, 203–216.

Choi, I. (2003). A constraint-based approach to Korean partial honorific agreement. *Texas Linguistic Forum*, *53*, 157–166.

Curcic, M., Andringa, S., & Kuiken, F. (2019). The role of awareness and cognitive aptitudes in L2 predictive language processing. *Language Learning*, *69*, 42–71.

Dijkgraaf, A., Hartsuiker, R. J., & Duyck, W. (2017). Predicting upcoming information in native-language and nonnative-language auditory word recognition. *Bilingualism: Language and Cognition*, *20*, 917–930.

Ferreira, F., Foucart, A., & Engelhardt, P. E. (2013). Language processing in the visual world: Effects of preview, visual complexity, and prediction. *Journal of Memory and Language*, *69*, 165–182.

Fine, A. B., Jaeger, T. F., Farmer, T. A., & Qian, T. (2013). Rapid expectation adaptation during syntactic comprehension. *PLoS ONE*, *8*, e77661.

Foucart, A., Romero-Rivas, C., Gort, B. L., & Costa, A. (2016). Discourse comprehension in L2: Making sense of what is not explicitly said. *Brain and Language*, *163*, 32–41.

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

Hopp, H., & Grüter, T. (2023). The time-course of competition from the L1 grammar in L2 sentence processing: evidence from cross-linguistic structural priming. *Second Language Research*, *39*, 133–159.

Hopp, H., & Lemmerth, N. (2018). Lexical and syntactic congruency in L2 predictive gender processing. *Studies in Second Language Acquisition*, *40*, 171–199.

Huettig, F. (2015). Four central questions about prediction in language processing. *Brain Research*, *1626*, 118–135.

Ito, A. (2023). *Encouraging prediction enhances predictive eye-movements in L2 speakers*. [Poster]. AMLaP Conference, San Sebastian, Spain.

Ito, A., Corley, M., & Pickering, M. J. (2018). A cognitive load delays predictive eye movements similarly during L1 and L2 comprehension. *Bilingualism: Language and Cognition*, *21*, 251–264.

Ito, A., & Knoeferle, P. (2023). Analysing data from the psycholinguistic visual-world paradigm: Comparison of different analysis methods. *Behavior Research Methods*, *55*, 3461–3493.

Jackson, C. N., & Hopp, H. (2020). Prediction error and implicit learning in L1 and L2 syntactic priming. *International Journal of Bilingualism*, *24*, 895–911.

Jeong, J. (2017). *A Study on the L2 Korean acquisition of tense, aspect, and modality pre-final endings by L1 Chinese learners*. (Unpublished doctoral dissertation). Ewha Womans University, Seoul, South Korea.

Kaan, E. (2014). Predictive sentence processing in L2 and L1: What is different? *Linguistic Approaches to Bilingualism*, *4*, 257–282.

Kaan, E. & Grüter, T. (2021). Prediction in second language processing and learning: Advances and directions. In E. Kaan and T. Grüter (Eds.) *Prediction in second language processing and learning* (pp. 2–24). Amsterdam: John Benjamins.

Kamide, Y., Altmann, G. T. M., & Haywood, S. L. (2003). The time-course of prediction in incremental sentence processing: Evidence from anticipatory eye-movements. *Journal of Memory and Language*, *49*, 133–159.

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*, 133–154.

Kim, J. B., & Sells, P. (2007). Korean honorification: A kind of expressive meaning. *Journal of East Asian Linguistics*, *16*, 303–336.

Kleinschmidt, D. F., & Jaeger, F. T. (2015). Robust speech perception: Recognize the familiar, generalize to the similar, and adapt to the novel. *Psychological Review*, *122*, 148–203.

Koch, E., Bulté, B., Housen, A., & Godfroid, A. (2021). Using verb morphology to predict subject number in L1 and L2 sentence processing: A visual-world eye-tracking experiment. *Journal of the European Second Language Association*, *5*, 115–132.

Koopman, H. (2005). Korean (and Japanese) morphology from a syntactic perspective. *Linguistic Inquiry*, *36*, 601–633.

Kuperberg, G. R. (2007). Neural mechanisms of language comprehension: Challenges to syntax. *Brain Research*, *1146*, 23–49.

Kuperberg, G. R. (2013). The proactive comprehender: What event-related potentials tell us about the dynamics of reading comprehension. In B. Miller, L. Cutting, & P. McCardle (Eds.), *Unraveling reading comprehension: Behavioral, neurobiological, and genetic components* (pp. 176–192). Baltimore, MD: Paul Brookes.

Kuperberg, G. R., & Jaeger, T. F. (2016). What do we mean by prediction in language comprehension? *Language, Cognition & Neuroscience*, *31*, 32–59.

Kutas, M., DeLong, K. A., & Smith, N. J. (2011). A look around at what lies ahead: Prediction and predictability in language processing. In M. Bar (Ed.), *Predictions in the brain: Using our past to generate a future* (pp. 190–207). Oxford University Press.

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

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*, 323–349.

Pak, M. (2006). Jussive clauses and agreement of sentence final particles in Korean. *Japanese/Korean Linguistics*, *14*, 295–306.

Pickering, M. J., & Gambi, C. (2018). Predicting while comprehending language: A theory and review. *Psychological Bulletin*, *144*, 1002–1044.

R Core Team (2024). *R: A language and environment for statistical computing (Version 4.3.3) [Computer software]*. Vienna, Austria: R Foundation for Statistical Computing. http://www.R-project.org

Rayner, K. (2009). The 35th Sir Frederick Bartlett Lecture: Eye movements and attention in reading, scene perception, and visual search. *Quarterly Journal of Experimental Psychology*, *62*, 1457–1506.

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*, 253–267.

Sohn, H.–M. (1999). *The Korean language*. Cambridge University Press.

Tanenhaus, M. K., Spivey-Knowlton, M. J., Eberhard, K. M., & Sedivy, J. C. (1995). Integration of visual and linguistic information in spoken language comprehension. *Science*, *268*, 1632–1634.

Cite this article: Kim, H., Kim, K., & Kim, J. (2025). The role of processing goals in second language predictive processing: A visual–world eye–tracking study of Korean honorific agreement. *Studies in Second Language Acquisition,* 47: 136–156. https://doi.org/10.1017/S0272263124000639