Hyunwoo Kim, Kitaek Kim and Haerim Hwang*

Effects of goals and strategies on predictive processing: a visual world eye-tracking study on honorific agreement in Korean

https://doi.org/10.1515/ling-2024-0203
Received October 3, 2024; accepted August 5, 2025; published online January 7, 2026

Abstract: There is ongoing debate about whether prediction is driven solely by bottom-up associative links or is modulated by top-down goals and strategies. The current study attempts to address this issue by investigating the role of top-down factors in Korean speakers' predictive processing of honorific agreement. Two visual-world eye-tracking experiments were conducted, analyzing participants' anticipatory eye movements while manipulating two top-down factors. In Experiment 1, we assigned participants to two groups with different instructions, asking one group to listen to sentences and answer referent-selection questions, and the other group to actively predict the upcoming referent. Experiment 2 manipulated the validity of predictive cues by interspersing experimental items with fillers containing consistent or inconsistent continuations. Results from Experiment 1 showed that participants instructed to actively anticipate the referent used honorific information more quickly to make predictions than the comprehension-only group. In Experiment 2, the group exposed to predictive linguistic stimuli showed an earlier and stronger prediction effect compared to the group exposed to stimuli with no prediction validity. These results suggest that comprehenders engage in different degrees of prediction according to the current demands of task goals and strategies. We discuss these findings in light of recent theories of predictive language processing.

Keywords: prediction; strategies; automatic activation; Korean honorific agreement; visual-world eye-tracking

***Corresponding author: Haerim Hwang**, Department of English, The Chinese University of Hong Kong, Fung King Hey, Room 325, Shatin, NT, Hong Kong, E-mail: haerimhwang@cuhk.edu.hk. https://orcid.org/0000-0003-0888-590X
Hyunwoo Kim, Department of English Language and Literature, Yonsei University, 50, Yonsei-ro, Seodaemun-gu, Seoul, Republic of Korea, E-mail: hyunwoo2@yonsei.ac.kr. https://orcid.org/0000-0003-4810-6333
Kitaek Kim, Department of English Language Education, Seoul National University, 1, Gwanak-ro, Gwanak-gu, Seoul, Republic of Korea, E-mail: kitaek@snu.ac.kr

1 Introduction

One of the major discoveries in the field of language processing over the past decade is the characterization of the human comprehension mechanism as involving a predictive system (Clark 2013; Ferreira and Chantavarin 2018; Kuperberg and Jaeger 2016; Pickering and Gambi 2018). It is widely recognized that language speakers constantly update their mental models of the situation being described through rapid processing, which forms the seeds of proactive activation of upcoming information (Pickering and Gambi 2018). Speakers' preactivation of target linguistic information is commonly indicated by their predictive production of a word biased by linguistic information provided in a discourse context (e.g., Kehler et al. 2008), sensitivity to cloze probability of subsequent words in sentence comprehension where readers show the difficulty of processing unexpected continuations (e.g., Brothers et al. 2017; Thornhill and Van Petten 2012), and, of key interest in this study, predictive eye movements toward a plausible referent visually represented in the visual-world paradigm (e.g., Altmann and Kamide 1999; Garnham et al. 2020; Grüter et al. 2018; Hintz et al. 2017; Kamide et al. 2003; Lau et al. 2006; Lew-Williams and Fernald 2010).

Despite the increasingly influential view of the predictive system as a key feature of the human comprehension mechanism, relatively little is known about specific circumstances under which prediction is generated. A crucial but understudied issue concerns whether prediction is automatically activated by the presence of bottom-up linguistic input, or is constrained by top-down control such as processing goals and strategy (Huettig 2015). While this question has been explored in lexical predictions (e.g., Brothers et al. 2017), it remains unknown how these mechanisms operate beyond processing at the lexical level.

Morphosyntactic information often plays a crucial role in facilitating anticipatory processing, allowing comprehenders to predict upcoming grammatical structures and morphological forms, including number, gender, tense, and agreement features, even before these elements are explicitly encountered in the input. A growing body of research has demonstrated that listeners and readers can rely on various cues to predict morphosyntactic features in languages that require such agreement (e.g., Grüter et al. 2018; Kamide et al. 2003; Lew-Williams and Fernald 2010). These findings suggest that predictive mechanisms are not limited to lexical or semantic cues but are also sensitive to abstract grammatical relations. However, the extent to which this ability is shaped by task demands or contextual factors remains less understood.

To address this gap, the current study investigates whether Korean speakers predictively engage with honorific agreement information under different task manipulation contexts. In Korean, verbs show agreement with honorable subjects through the use of the honorific marker *-si*, as shown in (1).[1] However, this honorific marker is not obligatory; according to corpus studies, it is used only 36.4 % (Song et al. 2019) or 41.5 % of the time (Kim et al. 2005) for honorable subjects. This suggests that in a canonical subject(-object)-verb sentence, an honorable subject is not a strong predictor of the honorific marker on the verb. Conversely, when this marker is given early in a prenominal relative clause, as shown in (2), it provides a strong predictive cue for an honorable subject that follows (see Choe 2004; Choi 2010; Song et al. 2019).

(1) *sensayngnim-kkeyse/*haksayng-i* *cohu-si-ta.*
teacher-NOM$_{\text{HON}}$/student-NOM good-HON-DECL[2]
'The teacher/student is good.'

(2) *cohu-si-n* *sensayngnim/*haksayng*
good-HON-ADN teacher/student
'the teacher/student who is good'

While making use of a structure like (2) in this study, we followed the design adopted by previous research on lexical prediction (Brothers et al. 2017). Specifically, we manipulated task instructions (Experiment 1) and the validity of predictive cues within linguistic stimuli (Experiment 2) to examine whether top-down strategies modulate the extent to which comprehenders engage in predictive processing. Before presenting the current study's findings, we first discuss some key theoretical issues underlying the predictive processing mechanisms. We then provide a detailed overview of Korean honorific agreement and explain its relevance in the context of our study.

2 Research on the predictive processing mechanisms

Researchers have long been interested in how predictive processing is shaped in the human mind (Huettig 2015; Kahneman 2011). Several models propose different types

1 Despite the existence of honorific markers for objects in Korean, as shown in (i), we have chosen to focus on the honorific marker for subjects in this study because it is more productive.

(i) *nay-ka* *sensayngnim-kkey/*haksayng-eykey* *yenlak-tuly-ess-ta.*
I-NOM teacher-DAT$_{\text{HON}}$/student-DAT call-HON-PST-DECL
'I called the teacher/student.'

2 Abbreviations used for the glosses in this study: ACC = Accusative marker; ADN = Adnominal marker; CAU = Causative marker; DAT = Dative marker; DECL = Declarative marker; HON = Honorific form; NOM = Nominative marker; PRES = Present tense marker; PROG = Progressive aspect marker; PST = Past tense marker; TOP = Topic marker.

of processing systems that elucidate how comprehenders take up current linguistic cues to activate information about upcoming materials. Some models address prediction as an associative mechanism operating automatically based on linguistic input (Kahneman 2011; Kuperberg 2007). According to this position, the simple associative mechanism (which Huettig [2015] describes as a "dumb route" to prediction) allows comprehenders to rapidly integrate linguistic representations of incoming information in a bottom-up manner, from which they consult prestored information in their memory to preactivate a set of concepts with associative relationships with the incoming information. The simple spreading activation of items with associative relationships is hypothesized to lead to automatic and unconscious prediction based on the bottom-up input (e.g., Kuperberg 2007). Similarly, prediction-by-association posits prediction as an automatic process (Pickering and Gambi 2018; Pickering and Garrod 2013). According to this account, the parallel activation of concepts serves as a primary driving force underlying prediction, while the linguistic context plays a minimal role in this process.

An alternative position assumes that prediction essentially involves cognitively demanding processes in which comprehenders allocate capacity-limited cognitive resources to generate expectations (Kuperberg and Jaeger 2016). Unlike the view that characterizes prediction as automatic activation, this attentional, resource-draining mechanism of prediction (which Huettig [2015] describes as a "smart route" to prediction) may depend on top-down attentional control that constrains comprehenders to selectively devote their resources to prediction according to the availability of memory resources under various cognitive load conditions. Following from this hypothesis is a prediction that the generation of prediction may be modulated by comprehenders' current goals and strategies. For example, one may expect comprehenders to engage in reduced prediction if they repeatedly encounter linguistic information that proves incompatible with their initial inference, thus rendering prediction less helpful for efficient comprehension (Kuperberg and Jaeger 2016). Conversely, comprehenders may strategically enhance predictive processing when their preactivated cues are congruent with the subsequent information, making prediction highly valid for efficient processing (e.g., Brothers et al. 2017; Huettig 2015; Lupyan and Clark 2015).

Similar to this approach, prediction-by-production highlights the role of nonlinguistic contextual information in predictive processing (Pickering and Gambi; see also "prediction-by-simulation" in Pickering and Garrod 2013). According to this account, comprehenders engage in prediction by using the same system employed for production. Specifically, they generate prediction by deriving the intention behind the current utterance based on various types of contextual information. As linguistic and nonlinguistic knowledge shared between interlocutors is assumed to play a crucial role in accurate prediction and successful communication, this account predicts varying degrees of prediction depending on how comprehenders exploit information provided by both the linguistic and nonlinguistic context.

Despite the potential of these accounts in explaining predictive processing, their empirical investigation remains limited. Two studies have attempted to address this gap. In one study, Brothers et al. (2017) provided evidence supporting the strategic modulation of prediction by examining the influence of top-down comprehension goals on English speakers' lexical prediction. In their event-related potential (ERP) experiment using the rapid serial visual presentation paradigm, participants read passages that either contained a predictable final critical word (e.g., *Thomas didn't like the temperature of his drink. He thought it was much too* hot) or made the final word highly unpredictable (e.g., *Thomas didn't like the look of the water. He thought it was much too* hot). The ERP results revealed that participants exhibited a larger N400 amplitude, associated with processing unexpected continuations, for unpredictable words compared to predictable words. Crucially, this effect was more pronounced when participants were asked to actively predict the sentence-final word compared to when they read passages solely for comprehension. Analogous outcomes were obtained in a follow-up self-paced reading task, where participants spent longer time reading non-predictable target words (e.g., *Alex said he wanted to watch the large spider on the porch*) compared to predictable words (e.g., *The web had been spun by the large spider on the porch*), particularly when presented with highly constraining filler sentences. The predictability effect was attenuated when participants encountered only half of the fillers with predictable final words, and the effect was not significant when all fillers contained unpredictable final words. Brothers et al. (2017) interpreted these results as evidence of strategic lexical prediction in English speakers, where processing efficiency is optimized according to comprehenders' processing goals and in response to "the recent success or failure of predictive processing" (p. 213).

In another study, Hopp (2016, Experiment 2) investigated the modulation of variability of lexical gender assignment on German speakers' predictive processing using gender agreement. In a visual-world eye-tracking experiment, native speakers of German encountered filler items that contained either correctly marked determiner–noun combinations (a No-error condition) or incorrectly marked combinations (an Error condition). The results revealed that participants in the No-error condition proactively used gender information to generate expectations about upcoming objects. However, participants in the Error condition exhibited reduced reliance on gender for prediction. Hopp interpreted these findings as evidence of morphosyntactic adaptation, suggesting "the parser is very flexible in adapting to the input in an attempt to facilitate online language processing in immediate environments" (p. 298).

While these studies contribute to our understanding of how comprehenders adapt to variability in input during predictive processing, further research is needed to explore the broader implications and mechanisms of morphosyntactic adaptation in predictive language processing. Unlike the case of lexical prediction, which requires comprehenders to activate specific lexical forms based on their probabilistic co-occurrence with preceding linguistic materials, the predictive use of morphosyntactic information necessitates a high degree of automaticity in preactivating abstract grammatical features associated with upcoming information, which may demand relatively more processing resources (e.g., Huettig and Guerra 2019). Consequently, the utilization of morphosyntactic information may result in reduced predictive processing compared to the use of lexical information. Alternatively, it is also possible that predictive processing using morphosyntactic information occurs to the same extent as, or even more strongly than, predictive processing based on lexical information. This could particularly hold true when speakers demonstrate heightened sensitivity and automaticity in utilizing morphosyntactic cues.

To our best knowledge, this issue has not been empirically tested in the context of morphosyntactic processing. Against this background, the current study extended Brothers et al.'s (2017) experimental setup to probing the effect of top-down factors in Korean speakers' predictive processing of honorific agreement in Korean. Across two visual-world eye-tracking experiments, we investigated whether processing goals and the validity of prediction cues in the linguistic stimuli would influence the comprehenders' use of an honorific marker to generate predictive looks to an upcoming referent in a visual display while listening to audio-recordings of Korean sentences.

3 Honorific agreement in Korean

Korean uses various linguistic devices to promote an honorific person's social status (Sohn 1999). This study focuses on the progressive aspect marker within the verb, which takes one of two suppletive forms based on the honorable status of the subject (for morpho-syntactic details of Korean honorification, see Choi and Harley 2019; Jou 2024). While the non-honorific progressive marker *-ko iss* is used for any types of referents (either honorific or non-honorific person), the use of the honorific progressive marker *-ko kyeysi* is restricted to a referent with an honorific status, such as *sensayngnim* 'teacher' as in (3a). The use of the honorific progressive marker for a non-honorific character, such as *haksayng* 'student' as in (3b), gives rise to an honorific agreement violation.

(3) a. *Ceki* *wus-ko kyeysi-ess-ten/wus-ko iss-ess-ten* *sensayngnim*
there laugh-PROG$_{\text{HON}}$-PAST-ADN/laugh-PROG-PAST-ADN teacher
'the teacher who was laughing over there'

b. *Ceki* **wus-ko kyeysi-ess-ten/wus-ko iss-ess-ten* *haksayng*
there laugh-PROG$_{\text{HON}}$-PAST-ADN/laugh-PROG-PAST-ADN student
'the student who was laughing over there'

The honorific form within a relative clause thus constitutes a crucial predictive cue for a subsequent referent that it premodifies in terms of its honorific status. The use of the honorific progressive form heralds the coming of an honorific referent, whereas the use of the non-honorific progressive marker does not provide any informative cue with respect to the honorific status of the following referent. The reason this study employed the honorific marker in the progressive form is that its morpho-phonological form (*-iss* or *-kyeysi*) remains consistent regardless of the phonetic environment. Using the simple form instead would have resulted in morpho-phonological variations of verb stems (e.g., *wusu-si-ess-ten* laugh-HON-PAST-ADN for an honorable subject, *wus-ess-ten* laugh-PAST-ADN for a non-honorable subject). To avoid unexpected results due to these changes, we maintained consistent *-iss* or *-kyeysi* markers across our stimuli.

Another motivation for our employment of the Korean honorific agreement system is that it involves both semantic and morphosyntactic features, allowing us to examine whether speakers can integrate these dual information sources to generate predictions. The morphosyntactic aspect of Korean honorific agreement is indicated by the consistent use of the verbal suffix for marking honorific agreement with its associate referent. Like the third-person singular morpheme in English verbs, the Korean honorific verbal suffix functions as a grammatical agreement morpheme, requiring comprehenders to check an honorific feature (i.e., [±HON]) for the subject and the predicate (Koopman 2005; Pak 2006). Concurrently, Korean honorific agreement is also semantically and pragmatically constrained, as the honorific verbal suffix encodes contextual information that the speaker intends to show respect for the subject. Because the honorific agreement between the subject and the predicate has been formed through a semantically driven feature-sharing process, comprehenders should compute their agreement under pragmatically constrained contexts (Choi 2003).

Previous research has provided compelling evidence that Korean speakers exhibit sensitivity to honorific information during sentence comprehension (e.g., Kim and Kaiser 2009; Kwon and Lee 2024; Kwon and Sturt 2016, 2024; Lee and Yoo 2023; Mueller and Jiang 2013). For example, Kim and Kaiser (2009) conducted a study employing both offline acceptability judgment task and an online self-paced reading task. In their study, participants were presented with sentences that consisted of a subordinate clause containing a null subject and a predicate, which was either marked with the honorific suffix -*si* or appeared without it. The subordinate clause was followed by a matrix clause containing a subject and an object, either with an honorable or nonhonorable status, as illustrated in (4).

(4) *Ø isa-lul ka-(si)-ki ceney halapeci/sonca-ka*
move-ACC go-(HON)-COMP before grandfather/grandson-NOM
maywu ttatushakey sonca/halapeci-lul wilohasi-ko, sikan-i
very warmly grandson/grandfather-ACC cheer-and time-NOM
ppalukey huluko i-ss-ta.
quickly running be-PRES-DECL
'Before Ø moving away, the grandfather/the grandson very warmly cheers the grandson/the grandfather, and time is running out quickly.'

Kim and Kaiser found that participants rated sentences containing the honorific marking and a nonhonorable subject as less acceptable in the offline task. In the online task, they showed slower reading times for sentences in this condition compared to sentences in other conditions, indicating that the mismatch in honorific features caused processing difficulty.

In another self-paced reading study, Lee and Yoo (2023) manipulated the honorific status of subjects in both the matrix and embedded clauses, as illustrated in (5). The study found that at the embedded verb region without the honorific marker (e.g., *phaakha-key*), embedded subjects with honorific features resulted in longer reading times compared to those without. This suggests that participants expected the embedded verb to contain the honorific marker, and the mismatch between the honorific feature of the embedded subject and the embedded verb resulted in processing difficulty. This was interpreted by the authors as being consistent with predictive processing.

(5) *{cwuimnim-i/*Cayhwuni-ka} ppalli {paksanim-i/Minho-ka}*
chief-NOM/Cayhwun-NOM quickly doctor-NOM/Minho-NOM
sanghwang-ul phaakha-key] pokose-lul nemkyecwu-si-ess-ta.
situation-ACC figure out-CAU document-ACC hand over-HON-PAST-DECL
'The chief/Chayhwun handed over all documents to him so that the doctor/Minho could figure out the situation quickly.'

In summary, prior research indicates that honorific information significantly influences Korean speakers' sentence comprehension. While Lee and Yoo's (2023) study offers support for predictive processing, the honorific feature mismatch was observed at the verbal region in the integrative processing context. Thus, it remains unclear whether their findings definitively demonstrate predictive processing. To address this issue, the present study uses a visual-world paradigm to examine whether honorific information affects predictive processing. Moreover, considering that honorable subjects do not always necessitate the honorific marker within the verb (Kim et al. 2005; Song et al. 2019; see Section 1), this study utilizes sentences in which the honorific marker within the verb precedes an honorable referent.

4 Current study

Taking advantage of the predictive role of Korean honorific marking and its dual informativity (i.e., morphosyntactic and semantic information), the current study assessed Korean speakers' use of the honorific marker to preactivate honorific information about an upcoming referent. More importantly, we tested whether speakers' predictive use of the honorific information is modulated by top-down goals and strategies. Adopting Brothers et al.'s (2017) experimental design, we manipulated top-down factors that can affect comprehenders' comprehension strategies in two visual-world eye-tracking experiments. In Experiment 1, we provided different types of instructions to two groups of participants, asking one group to listen to a sentence and answer a referent-selection question and another group to actively predict the referent who would be asked about by the following question. (See Kim et al. 2025 for the study with L2 Korean learners employing the same design.) In Experiment 2, we manipulated the validity of predictive cues in the linguistic stimuli by interspersing experimental items with fillers whose continuations were either consistent or inconsistent with preceding linguistic cues. If these top-down factors influence comprehenders' predictive strategies, participants will show facilitated anticipatory processing in the experimental condition that promotes prediction. Otherwise, if prediction is automatically generated without conscious control, we expect to see participants' consistent prediction patterns, irrespective of the manipulation of the top-down factors.

5 Experiment 1

5.1 Methods

5.1.1 Participants

We recruited 32 participants (9 male and 23 female) from the undergraduate and graduate student population at a Korean university. They were native speakers of Korean, ranging in age from 23 to 37 (*M* = 27.9, *SD* = 4.3). All had normal or corrected-to-normal vision with no reported history of auditory deficits. These participants were randomly assigned to two groups, who received different instructions (see Procedure below). The study was approved by the Ethics Board at Seoul National University, and all participants provided written consent and received the Korean equivalent of $10 for their participation.

5.1.2 Materials

The linguistic stimuli were 20 sets of items, each of which consisted of a critical sentence and a question, as illustrated in (6). These items were counterbalanced for the type of the progressive marker attached to the relative clause predicate in the critical sentence, with half of the items containing the honorific *-kyeysi* (honorific condition) and half containing the non-honorific *-iss* (non-honorific condition), yielding 10 tokens per condition (see Appendix A for a list of all experimental items).

(6) Critical sentence
Ecey *tosekwan-eyse* *kongpwu-lul*
yesterday library-at study-ACC
ha-ko kyeysi-ess-ten/ha-ko iss-ess-ten *palo* *ce*
do-PROG$_{\text{HON}}$-PAST-ADN/do-PROG-PAST-ADN very that
sensayngnim-un *sengkyek-i* *coh-ta.*
teacher-TOP character-NOM good-DECL
'The teacher who was studying at the library yesterday has a good character.'
Question
Nwu-ka *sengkyek-i* *cohna-yo?*
who-NOM character-NOM good-Q
'Who has a good character?'

Following the relative clause predicate, we inserted the determiner phrase *palo ce* 'very that' ahead of the target referent (e.g., *sensayngnim* 'teacher'), which functioned as a padding phrase to allow for capturing participants' predictive processing. Any evidence of participants' preactivation of honorific information of the following target referent will be indicated by their increased looks to the target image compared to its competitor in the time window including this padding phrase. We thus analyzed this time frame as the critical region (see Data treatment and analysis below).

For the target referent, we selected characters that can receive honor and thus be modified by an honorific marker, including nouns denoting professions (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 with a certain social standing (*samonim* 'madam', *sacangnim* 'boss or sir', *acwumeni* 'madam or ma'am', *senpaynim* 'senior', *elusin* 'elder'). Because the target referent in the critical sentence was always associated with an honorific character, we included 10 fillers that were structured analogously to the critical sentences in the experimental stimuli but included a non-honorific predicate and a non-honorific referent, such as *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'.

Following the critical sentence, a question appeared asking participants to identify the referent mentioned in the preceding sentence by clicking on one of the characters on the display. Participants' responses to the questions were taken to estimate their understanding of the critical sentence, and thus we excluded trials with incorrect responses (see results).

The experimental items were intermixed with 46 fillers, including those containing a non-honorific referent modified by a non-honorific predicate in contexts like (2) and sentences denoting various events where two characters interact with each other. The linguistic stimuli were recorded by a female native speaker of Korean at a natural speech rate using standard intonation. The audio recordings of the critical sentences in the experimental stimuli were edited using Praat (Boersma and Weenink 2017) to control for the length of the recording frames of interest across the honorific and non-honorific conditions. Consequently, the critical sentences between the two conditions did not significantly differ in the overall sentence length, the offset of the relative clause predicate, the onset of the padding phrase, and the onset of the target referent. Detailed descriptive and inferential statistics for these analyses are presented in Table 1.

Table 1: Information about the acoustic duration (milliseconds) of relevant regions of interest.

	Honorific condition		Non-honorific condition		*t*	*p*	*Cohen's d*
	M	SD	*M*	SD			
Overall sentence length	12,022.0	464.0	11,906.5	503.3	0.755	0.455	0.239
Offset of the relative clause predicate	3,187.9	224.3	3,191.9	217.9	−0.058	0.954	−0.018
Onset of the padding phrase	4,550.3	2,668.9	4,544.5	230.6	0.073	0.942	0.023
Onset of the target referent	5,454.8	300.9	5,466.1	281.5	−0.123	0.902	−0.039

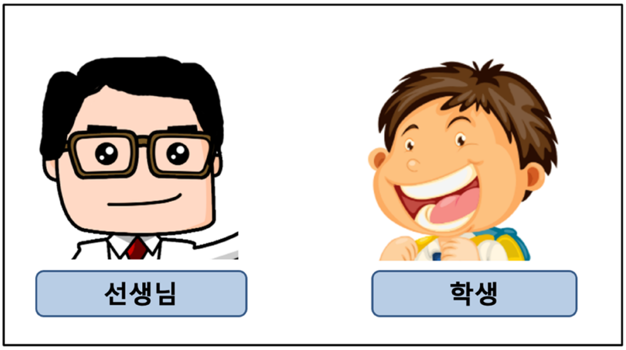

Figure 1: Example of a visual stimulus. The images represent a teacher (left) and a student (right) with the corresponding Korean words printed below.

Each set of experimental stimuli in the honorific and non-honorific conditions was paired with a visual scene that contained two pictures of human faces with the words printed in Korean below (see Figure 1). One of the faces corresponded to the target referent (e.g., teacher) mentioned in the critical sentence, and the other face served as a competitor (e.g., student). The gender of the referents within each visual stimulus was kept identical to forestall any unexpected effects arising from gender stereotypes. The position of the target picture was counterbalanced across items.

5.1.3 Procedure

The eye-tracking experiment was conducted in a quiet lab. We used an EyeLink Portable Duo (SR Research Ltd., Mississauga, Ontario, Canada) attached to a 16-inch monitor to track participants' eye movements from the right eye at a rate of 1000 Hz. Each experimental session began with written and oral instruction, followed by a 13-point calibration and validation procedure. After calibration, participants worked through three practice trials. In order to test the effect of top-down goals on predictive processing, we randomly divided participants into two groups and provided each group with different instructions for the task. Following Brothers et al. (2017), we asked one group to listen to a sentence and try to anticipate which referent the sentence would describe (Prediction group, $N = 16$) and another group to listen to sentences for comprehension without encouraging them to predict a target referent (Comprehension group, $N = 16$). Other than the task instructions, the experimental procedure was identical across the two groups.

At the beginning of each trial, a fixation cross appeared at the center of the screen, prompting participants to fixate on it to initiate the trial. If the tracker failed to capture a fixation for 500 ms within 5 s after the onset of the cross, the screen returned to the calibration page, and participants' eye movements were recalibrated. After successful fixation on the cross, a picture of two human faces, as illustrated in Figure 1, was displayed on the screen for 2,000 ms with no auditory signal. This silent preview was included to ensure that participants had a chance to scan the visual context, including the face images and the printed words (e.g., Ferreira et al. 2013). After the 2,000-ms preview, the critical sentence played, followed by a 1,000-ms pause and a question. The visual scene remained on the screen throughout the duration of the critical sentence and the question. Participants were instructed to click on an image on the display in response to the question. Without feedback, they proceeded to the next trial upon their mouse-click response.

5.1.4 Data treatment and analysis

We first checked participants' responses to the questions for the experimental items. The mouse-click accuracies were at ceiling for both Comprehension (M = 98.3 %, SD = 0.05) and Prediction (M = 97.1 %, SD = 0.06) groups, with no statistical difference between the two groups' accuracies (p = 0.529). We only submitted data from trials into analysis for which participants gave a correct answer to the question (eliminating 0.5 % of the data).

Participants' eye movement data were classified as fixations, saccades, and blinks using the Eyelink DataViewer software. Through this process, participants' fixations on each image were identified and aggregated into 20-ms bins. Based on these fixation data, we calculated a "TargetBias" score as a dependent measure reflecting the difference between the proportion of looks to target and the proportion of looks to competitor for each trial (e.g., McMurray et al. 2009). TargetBias scores were obtained by subtracting the number of 20-ms bins spent looking at the competitor from those spent looking at the target.

Participants' anticipatory looking behavior was operationalized as the increased TargetBias scores in the honorific condition (i.e., sentences with the honorific -*kyeysi*) compared to the non-honorific condition (i.e., sentences with the non-honorific -*iss*) during the relevant time window for each trial. We determined the time windows for analysis in two different ways. First, in light of the definition of predictive processing as "preactivation of aspects of the linguistic representation of a predictable word" before encountering that word (Pickering and Gambi 2018: 1005), we set as the predictive region an entire time frame ranging from the offset of the relative clause predicate including the progressive marker (i.e., -*kyeysi* or -*iss*) to target onset (e.g., *sensayngnim* 'teacher'). In the second analysis, we further segregated this time window into two analysis regions, (a) the relative-clause boundary region, which spans from the offset of the predicate in the relative clause to the onset of the padding

Table 2: Mean duration (milliseconds) and standard deviations of each region of interest across items.

Region	Mean	SD
Predictive region	2,270.5	171.3
Relative-clause boundary region	1,357.5	112.5
Main-clause region	913.0	103.4

phrase (i.e., *palo ce* 'very that') and (b) the main-clause region, which extends from the onset of the padding phrase to target onset. This segmentation allowed for a finer-grained analysis of the temporal dynamics of participants' predictive processing at specific time windows, enabling us to examine at what moment during processing the participants would engage in anticipatory processing. If top-down goals modulate participants' predictive use of the honorific marker, the effect of predictive processing will emerge sooner and/or more prominently for the Prediction group than for the Comprehension group in these time windows. Table 2 presents information of duration in each region.

Following the standard protocol adopted by visual-world eye-tracking studies, we offset 200 ms in our analyses of fixations in each time window, which reflects the saccadic lag taken to initiate a gaze shift (Matin et al. 1993).

For each of the critical windows (i.e., the predictive region, composed of the relative-clause boundary region and the main-clause region), we conducted linear mixed-effects regression (Baayen et al. 2008) on TargetBias scores to statistically compare the two participant groups' processing patterns in the honorific and non-honorific conditions. Each model included fixed effects of *Group* (Comprehension, Prediction) and *Condition* (Honorific, Non-honorific) and random effects of participant and item. Deviation coding was applied to the fixed effects, assigning −0.5 to the Comprehension group and the Honorific condition and 0.5 to the Prediction group and the Non-honorific condition. To control for the imbalanced distribution of data across conditions, we centered the fixed factors around the grand mean (Baayen 2008). Following Barr et al. (2013), we initially created the maximal random-effects structure permitted by the design, but, due to the convergence problem, we progressively simplified the random-effects structure until it converged (Barr et al. 2013). As a result, the random-effects structure contained random intercepts for participant and item and by-item random slope for *Group*. All modelling was conducted in R version 4.2.1 (R Core Team 2022) using the *lme4* package (Bates et al. 2015). To check for the goodness of fit for each model, we inspected the area under the ROC curve (AUC) using the *auc* function (e.g., Perek and Goldberg 2017). The results showed the AUC score of 37.4 % for the model in the predictive region, 50.6 % for the model in the relative-clause boundary region, and 54.4 % for the model in the main-clause region. These results indicate a moderate fit for the data in each model.

5.2 Results

Participants' fixations over the course of the critical time windows are presented in Figure 2 for the Prediction group and in Figure 3 for the Comprehension group. Outcomes of the mixed-effects models conducted for the three critical windows are presented in Table 3.

In the analysis of the predictive region (W1 + W2, from the relative clause predicate offset to target onset), we found a main effect of *Condition*, with a significantly higher TargetBias score (i.e., an increased proportion of looks to the target referent relative to the competitor) in the honorific than in the non-honorific condition. While there was no significant effect of *Group*, we observed a significant interaction between *Group* and *Condition*. To further inspect this interaction, we conducted separate analyses for each group using linear mixed-effects models, with *Condition* (deviation-coded and centered around the mean) as a fixed factor and the random effects of participant and item. We applied an adjusted alpha level of 0.025 (0.05 divided by 2) for multiple comparisons. The results of the by-group analyses showed a significant effect of *Condition* for both the Prediction (b = –0.114, SE = 0.011,

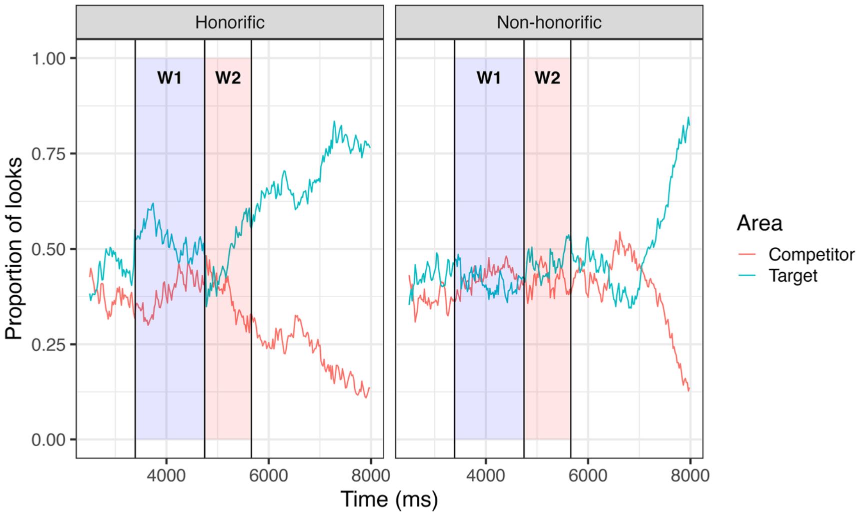

Figure 2: The Prediction group's proportion of looks to the competitor (red line) and the target (green line) in the honorific (left) and non-honorific (right) conditions in Experiment 1. Predictive region = W1 + W2; W1 = relative-clause boundary region (from the offset of the predicate in the relative clause to the onset of the padding phrase); W2 = main-clause region (from the onset of the padding phrase to target onset).

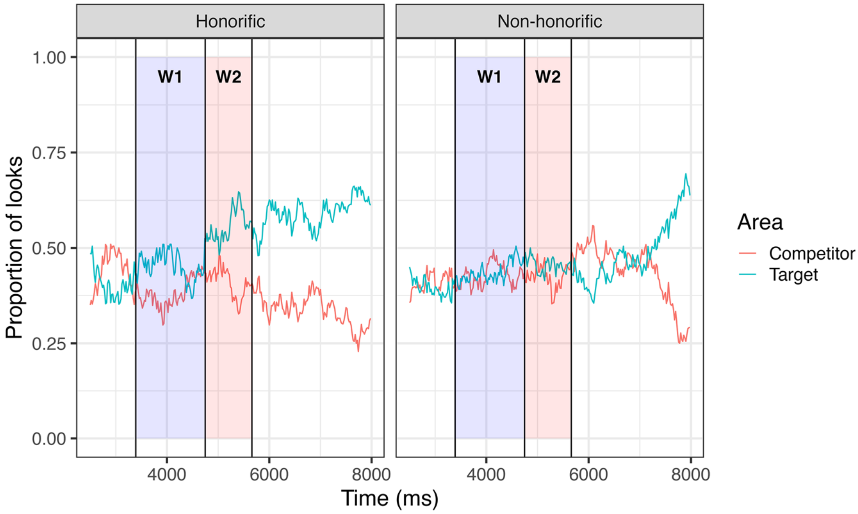

Figure 3: The Comprehension group's proportion of looks to the competitor (red line) and the target (green line) in the honorific (left) and non-honorific (right) conditions in Experiment 1. Predictive region = W1 + W2; W1 = relative-clause boundary region (from the offset of the predicate in the relative clause to the onset of the padding phrase); W2 = main-clause region (from the onset of the padding phrase to target onset).

Table 3: Model outcomes in Experiment 1.

Region	Factor	*b*	SE	*p*	sr^2
Predictive region	Intercept	0.034	0.042	0.425	
	Group	−0.011	0.043	0.810	<0.0001
	Condition	−0.075	0.009	<0.001***	0.0013
	Group × condition	−0.078	0.018	<0.001***	0.0001
Relative-clause boundary region	Intercept	0.041	0.046	0.379	
	Group	0.0001	0.055	0.999	<0.001
	Condition	−0.062	0.011	<0.001***	0.007
	Group × condition	−0.156	0.022	<0.001***	0.0013
Main-clause region	Intercept	0.025	0.054	0.647	
	Group	−0.019	0.061	0.759	0.0001
	Condition	−0.094	0.014	<0.001***	0.0026
	Group × condition	0.057	0.028	0.039*	0.0005

***$p < 0.001$, *$p < 0.05$.

$p < 0.001$) and the Comprehension group ($b = -0.045$, $SE = 0.015$, $p = 0.002$). These results indicate that while both groups were more likely to look to the target referent following the honorific marker compared to the non-honorific marker in the predictive region, the magnitude of this effect was larger for the Prediction group than for the Comprehension group.

We obtained similar findings in the analysis of the relative-clause boundary region (W1, from the relative clause predicate offset to padding phrase onset). There was a significant effect of *Condition*, qualified by its interaction with *Group*. Subsequent by-group analyses showed that the effect of *Condition* was significant only in the Prediction group (b = –0.138, SE = 0.015, p < 0.001), but not in the Comprehension group (b = –0.005, SE = 0.018, p = 0.789). These results suggest that only the Prediction group exhibited predictive looks to target referents in this early time window.

In the following main-clause region (W2, from padding phrase onset to target onset), we observed a main effect of *Condition*, with a significant interaction with *Group*. By-group analyses revealed a significant effect of *Condition* for both the Prediction (b = –0.069, SE = 0.018, p < 0.001) and the Comprehension group (b = –0.111, SE = 0.023, p < 0.001). These results indicate that both groups engaged in predictive processing in this time window.

5.3 Discussion

In Experiment 1, two groups of participants, the Prediction and Comprehension groups, listened to Korean sentences involving honorific agreement under different task instructions. The analysis of the eye movement data in the global predictive region showed that both groups used the honorific marker to generate expectations about the honorific status of the following referent. However, the magnitude of prediction was greater for the Prediction group than for the Comprehension group. Moreover, in the more narrowly defined analysis regions, we observed a group difference in terms of the timing of predictive processing. Participants in the Prediction group demonstrated faster engagement in predictive processing, as they were quicker to detect the honorific information encoded in the honorific marker. These findings suggest that the participants' predictive processing was not solely driven by unconscious automatic associative mechanisms but also modulated by top-down strategic control (Brothers et al. 2017; Huettig 2015).

However, caution is warranted in the interpretation of the current findings due to some potential methodological limitations. The evidence showing the Prediction group's facilitated prediction might be due to the explicit task instruction provided to this group, which may have deliberately induced strategic processing more than is expected under other circumstances without explicit instructions. Although the participants who were asked to predict the target referent during the task showed the early signs of predictive processing, such an explicit instruction is hardly encountered in normal comprehension situations. This issue raises a question as to whether comprehenders' predictive processing is still sensitive to top-down goals in contexts without explicit manipulation of task instructions. We attempted to address this concern in Experiment 2 where we manipulated the validity of predictive cues in linguistic stimuli instead of providing different types of instructions.

6 Experiment 2

6.1 Methods

6.1.1 Participants

A group of 32 Korean speakers (9 male and 23 female), who did not participate in Experiment 1, completed another visual-world eye-tracking task. They were graduate and undergraduate students at a Korean university, aged 19 to 38 ($M = 26.3$, $SD = 4.9$). All had normal or corrected-to-normal vision and reported no history of auditory deficits. They provided written consent and received monetary compensation (Korean equivalent of $10).

6.1.2 Materials

The experimental items, including the auditory and visual stimuli, were identical to those used in Experiment 1, except for the inclusion of specially designed filler items. The decision to use the same critical items in both experiments was made to investigate how participants would process them under different task conditions, thus allowing for a direct comparison of the results from the two experiments. We adopted 48 filler sentences from Kim and Chun (2023) that give rise to verb-induced biases (see Appendix B for a list of fillers). Based on the previous finding that the effect of predictive processing was more evident for object-biased rather than subject-biased verbs, we chose to include only interpersonal verbs that create a bias to imputing the underlying consequence of the event to the object referent. Specifically, we selected verbs such as *hwanakeyha* 'anger', *pwunnohakeyha* 'enrage', *culkepkeyha* 'amuse', and *cilwuhakeyha* 'bore'. These so-called implicit consequentiality verbs (Crinean and Garnham 2006; Stewart et al. 1998) were reported to induce comprehenders' anticipatory looks to an object referent in a visual display prior to the presence of the target referent in the auditory input (e.g., Garnham et al. 2020; Kim et al. 2022). For example, when listening to (7a), Korean speakers were more likely to look at the image of the object referent *Jiho* than the image of the subject referent *Youngsoo* soon after they processed the conjunction *ttaymwuney* 'because of the preceding reason' even before they heard the disambiguating information in the ensuing clause (Kim et al. 2022). This anticipatory processing may be enhanced by the confirmation of the following disambiguating information, which identifies the object referent as the character undergoing the consequence of the event. However, it is conceivable that participants' prediction may be mitigated by following information that disconfirms their initial expectation. For instance, sentence (7b) contains a continuation that disambiguates the null reference to the subject *Yujin*, incongruent with the object bias of the implicit consequentiality verb *pwunnohakeyha* 'enrage'. It is thus expected that under this context, comprehenders should revise their initial prediction upon encountering the bias-inconsistent disambiguating information.

(7) Sample filler items including implicit consequentiality sentences
Bias-consistent continuation

a.

Hakkyo-eyse,	*Youngsoo-ka*	*Jiho-lul*	*hwanakeyha-yss-ki*	
school-at	Youngsoo-NOM	Jiho-ACC	anger-PAST-CONN	
ttaymwuney,	*swuep-i*	*kkuthna-ko*	*Youngsoo-eykey*	*kotpalo*
because	class-NOM	end-after	Youngsoo-DAT	immediately
ttacy-ess-ta.				
complain-PAST-DECL				

'At school, Youngsoo angered Jiho, and so (he = Jiho) immediately complained to Youngsoo after class.'

b. Bias-inconsistent continuation

Swuep-cwung-ey,	*Yujin-i*	*Hyunyoung-ul*	*pwunnohakeyha-yss-ki*	
class-middle-at	Yujin-NOM	Hyunyoung-ACC	enrage-PAST-CONN	
ttaymwuney,	*swuep-i*	*kkuthna-ko*	*Hyunyoung-eykey*	*kotpalo*
Because	class-NOM	end-after	Hyunyoung-DAT	immediately
sakwaha-yss-ta.				
apologize-PAST-DECL				

'During class, Yujin enraged Hyunyoung, and so (she = Yujin) immediately apologized to Hyunyoung after class.'

Taking advantage of these properties of implicit consequentiality sentences, we manipulated the validity of predictive information in our stimuli. We randomly divided the participants into two groups and provided one group with 24 fillers containing bias-consistent endings, such as (7a) (Higher-validity group, N = 16) and another group with 24 fillers containing bias-inconsistent endings, such as (7b) (Lower-validity group, N = 16). This manipulation allowed us to test whether the continuous exposure to highly predictable continuations in the fillers would increase the overall validity of predictive cues in our experimental stimuli and thus facilitate participants' predictive use of the honorific information. In addition to these implicit consequentiality fillers, both groups encountered 22 fillers that described various interpersonal activities (e.g., *Cwumaley, Seongmini Seungholul pangmwunhayssul ttay, sero wusnun elkwullo insahayssta*. ‘On the weekend, when Seongmin visited Seungho, (they) greeted each other with smiling faces.’).

6.1.3 Procedure

Experiment 2 was conducted in the same manner as Experiment 1 (with the same instruction as that given to the Comprehension group in Experiment 1), except for the distribution of the filler items. Following Brothers et al. (2017), we sensitized the participants to the overall validity of predictive cues by presenting 10 fillers involving the implicit consequentiality sentences at the beginning of each experiment prior to the presentation of the experimental stimuli. The other fillers were pseudorandomly intermixed with experimental stimuli, with no more than two experimental items in the same condition presented in a row.

6.1.4 Data treatment and analysis

The data treatment and analysis procedures were identical to those in Experiment 1. The mean accuracy on the comprehension-check questions for the experimental items was high for both the Higher-validity (M = 95.0 %, SD = 0.08) and the Lower-validity group (M = 95.2 %, SD = 0.11), with no significant difference between the two groups (p = 0.950). As in Experiment 1, we only included trials for which participants provided a correct answer to the question (eliminating 1.1 % of the data). We then fit the linear mixed-effects models to fixation data in three time windows (predictive region, relative-clause boundary region, main-clause region) in the same manner as in Experiment 1. The models included fixed effects of *Group* (Higher-validity: −0.5, Lower-validity: 0.5) and *Condition* (Honorific: −0.5, Non-honorific: 0.5) and random effects of participant and item. All fixed effects were deviation-coded and centered. The random-effects structure for these models was simplified by containing random intercepts for participant and item and by-item random slope for *Group*. The AUC scores were 45.7 % for the model in the predictive region, 57.4 % for the model in the relative-clause boundary region, and 52.3 % for the model in the main-clause region. These results indicate the moderate model fit for the data.

6.2 Results

Participants’ fixations over the course of the critical time windows are presented in Figure 4 for the Higher-validity group and in Figure 5 for the Lower-validity group. Model outcomes are presented in Table 4.

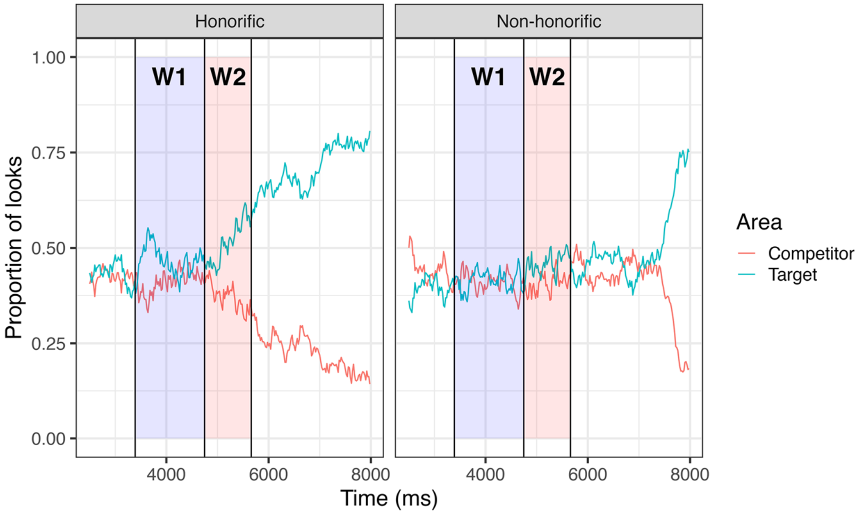

Figure 4: The Higher-validity group's proportion of looks to the competitor (red line) and the target (green line) in the honorific (left) and non-honorific (right) conditions in Experiment 2. Predictive region = W1 + W2; W1 = relative-clause boundary region (from the offset of the predicate in the relative clause to the onset of the padding phrase); W2 = main-clause region (from the onset of the padding phrase to target onset).

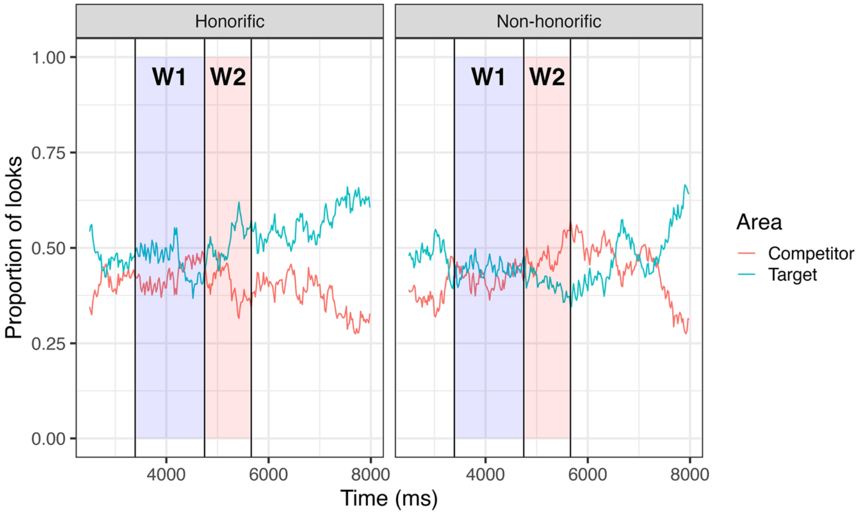

Figure 5: The Lower-validity group's proportion of looks to the competitor (red line) and the target (green line) in the honorific (left) and non-honorific (right) conditions in Experiment 2. Predictive region = W1 + W2; W1 = relative-clause boundary region (from the offset of the predicate in the relative clause to the onset of the padding phrase); W2 = main-clause region (from the onset of the padding phrase to target onset).

Table 4: Model outcomes in Experiment 2.

Region	Factor	*b*	SE	*p*	sr^2
Predictive region	Intercept	0.026	0.055	0.640	
	Group	0.034	0.038	0.387	0.0003
	Condition	−0.062	0.009	<0.001***	0.001
	Group × condition	−0.019	0.017	0.279	<0.0001
Relative-clause boundary region	Intercept	0.016	0.065	0.808	
	Group	0.005	0.051	0.921	<0.0001
	Condition	−0.001	0.011	0.904	<0.0001
	Group × condition	−0.092	0.022	<0.001***	0.0006
Main-clause region	Intercept	0.041	0.054	0.455	
	Group	0.075	0.050	0.142	0.0016
	Condition	−0.147	0.014	<0.001***	0.0065
	Group × condition	0.093	0.027	<0.001***	0.0012

***$p < 0.001$.

In the predictive region (W1 + W2), there was a main effect of *Condition*, with a significantly higher TargetBias score in the honorific than in the non-honorific condition, and a marginal interaction between *Condition* and *Group*. Separate analyses for each group revealed the significant effect of *Condition* in both the Higher-validity group (b = –0.070, SE = 0.011, $p < 0.001$) and the Lower-validity group (b = –0.036, SE = 0.015, p = 0.014) at the adjusted alpha level of 0.025, with a larger effect in the former group. These results suggest that while both groups predictively used the honorific cue, the magnitude of the prediction was greater in the High-validity than in the Lower-validity group in this overall prediction region.

Focusing on the narrowly defined regions, we found a significant interaction between *Group* and *Condition*, without a main effect of each factor, in the relative-clause boundary region (W1). This interaction was driven by the significant effect of *Condition* emerging in both groups yet pulled in the opposite direction. For the High-validity group, the TargetBias score was significantly higher in the honorific than the non-honorific condition (b = –0.047, SE = 0.014, p = 0.001). In contrast, the Lower-validity group had a significantly lower TargetBias score in the honorific than the non-honorific condition (b = 0.054, SE = 0.018, p = 0.003). These results indicate that only the Higher-validity group engaged in predictive processing in this time window. In the following main-clause region (W2), there was a main effect of *Condition* and its significant interaction with *Group*. By-group analyses showed a significant effect of *Condition* both for the High-validity group (b = –0.098, SE = 0.018, $p < 0.001$) and the Lower-validity group (b = –0.164, SE = 0.023, $p < 0.001$), suggesting both groups' predictive processing, with the larger effect size in the Lower-validity group.

6.3 Discussion

In Experiment 2, we manipulated the overall validity of predictive cues in the stimuli by presenting participants with different types of fillers. The analyses of participants' eye movements showed distinct processing patterns between the two groups. The Higher-validity group, who encountered implicit consequentiality sentences with 100 % bias-consistent endings, exhibited early signs of predictive use of honorific cues, whereas the effect of prediction was delayed for the Lower-validity group for whom the prediction effect emerged only in the later time window (i.e., main-clause region). Moreover, the magnitude of prediction in the entire critical time window (i.e., predictive region) was greater for the Higher-validity than the Lower-validity group. Taken together, these results provide confirming evidence of the modulation of comprehenders' top-down goals in the speed and degree of predictive processing.

Unlike in Experiment 1, we manipulated predictive validity through fillers, and despite this implicit strategic manipulation, we obtained outcomes supporting the modulating role of top-down goals. Participants in the Higher-validity group, who were exposed to bias-consistent filler endings, likely recognized the utility of prediction as a resource-optimizing mechanism. The continuous confirmation of their initial predictions through bottom-up input (e.g., honorific cues aligning with referent status) may have enabled them to strategically enhance anticipatory processing. Conversely, the Lower-validity group faced frequent prediction failures due to bias-inconsistent filler endings, where the cognitive cost of revising initial expectations may have outweighed the benefits of prediction. This asymmetry in adaptive strategies highlights how comprehenders dynamically weigh the utility of prediction against its computational demands.

These behavioral patterns align with Kuperberg and Jaeger's (2016) utility function framework, which posits that prediction is dynamically optimized based on the comprehender's ongoing assessment of prediction validity under specific circumstances. This framework integrates the notion that prediction is a flexible cognitive strategy adjusted to maximize efficiency under varying environmental conditions (see also Kuperberg 2013; Lupyan and Clark 2015). More generally, the modulating effect of top-down goals observed here can be interpreted as part of a broader adaptive mechanism whereby listeners continuously update their mental models to "better reflect the broader statistical structure of the current environment" (Kuperberg and Jaeger 2016: 46). This process ensures that predictive mechanisms remain sensitive to changes in linguistic context, thereby optimizing comprehension efficiency.

We further investigated whether participants in the Higher-validity group adapted their predictive processing based on the linguistic cues presented in the experiment. To this end, we conducted additional exploratory analyses that included *Trial number* as an interactive fixed factor in the models for this group across the three time windows. The results revealed significant interactions between *Condition* and *Trial number* in all three regions of interest: W1 + W2 (b = 0.003, SE = 0.001, p < 0.001), W1 (b = 0.002, SE = 0.001, p = 0.004), and W2 (b = 0.003, SE = 0.001, p = 0.005). These findings indicate that as participants in the Higher-validity group encountered a greater number of items, they demonstrated enhanced predictive processing in the honorific condition. These results provide further compelling evidence of adaptive predictive processing in this group.

It is worth noting that the effect of prediction also emerged in the Lower-validity group, albeit delayed until the main-clause region. This suggests that this group did not completely shift away from their predictive strategy for processing the sentences involving honorific agreement despite the challenge of revising their initial expectations when processing the implicit consequentiality cues. One possible account for this outcome may be found in the nature of honorific agreement in Korean. Honorific agreement is a ubiquitous feature in Korean (Brown 2015; Sohn 1999), allowing Korean speakers numerous opportunities to practice honorific agreement in various situations, particularly in the social milieu that emphasizes the expression of social superiority in language use (Kim and Sells 2007). The prevalent and preponderant occurrences of Korean honorification thus may have promoted the learning and entrenchment of this knowledge in the speaker's mind, rendering honorific marking a highly informative precursor to the status of an upcoming referent in the current experiment. However, this account is largely speculative, and it remains to be determined whether the facilitative prediction witnessed among our participants is specific to the processing of honorific agreement or can be generalized to other linguistic features. To address this question, future research needs to test the effect of top-down factors in prediction using diverse linguistic phenomena.

7 General discussion

7.1 Summary of findings

The purpose of this study was to assess the effect of top-down goals and strategies in Korean speakers' predictive processing of Korean honorific agreement. Across two visual-world eye-tracking experiments, we found that top-down strategies and predictive validity modulated the timing and degree of comprehenders' predictive processing. In Experiment 1, participants who received explicit task instruction for prediction were faster in the use of honorific cues for generating prediction than those who processed sentences simply for comprehension. In Experiment 2, an earlier and stronger effect of prediction was observed among the group who constantly encountered linguistic stimuli that promoted prediction, compared to the group who listened to stimuli with low prediction validity. These findings suggest that comprehenders engage in strategic predictive processing that depends on the current goals and the validity of predictive cues in the input (Huettig 2015; Kahneman 2011).

7.2 Theoretical implications for predictive processing

The robust effect of the top-down factors observed in this study is consistent with the findings from Brothers et al. (2017), which demonstrated the modulation of task instruction and prediction validity in English speakers' lexical prediction. The present study extended this research by showing that the strategic manipulations also had measurable impacts on the Korean speakers' predictive eye movements during their processing of honorific agreement in Korean. In addition, our findings go further in demonstrating that the top-down factors influenced comprehenders' processing of honorific agreement not just in the magnitude of prediction but also in terms of the timing of prediction, as indicated by the early effect of prediction that emerged when prediction was promoted. Our study thus provides additional empirical support for the assertion that predictive pre-activation is a general processing mechanism subject to top-down goals and strategies (Huettig 2015; Kuperberg and Jaeger 2016).

As a reviewer pointed out, a question arises as to whether the manipulation employed in Experiment 2 reflects a genuinely context-driven adaptation or whether it imposes a form of top-down guidance. While our use of biased fillers was intended to implicitly encourage participants to adjust their predictive strategies over time, we acknowledge that this manipulation may also be viewed as introducing a structured influence that shapes predictive behavior. However, as Brothers et al. (2017) suggest, such context-sensitive adjustments reflect the flexible deployment of anticipatory strategies without requiring modifications to the underlying architecture of the comprehension system. In other words, as comprehenders repeatedly encounter prediction errors through bottom-up input, they strategically reduce reliance on predictive strategies due to the associated processing costs. Although the experimental setting does not fully replicate naturalistic communication, it plausibly approximates how comprehenders adapt to varying predictive environments, where comprehenders implicitly optimize their predictive behavior in response to the output of a speaker. Consistent with this perspective, we interpret the adaptation observed in our study not as an artificial effect imposed by top-down instruction, but rather as a reflection of the natural sensitivity of the language processing system to distributional regularities in the input. In this sense, participants' predictive behavior in Experiment 2 likely emerged from their ongoing evaluation of the utility of prediction in the given context, closely mirroring the adaptive nature of prediction in everyday language comprehension.

The variable nature of predictive processing depending on the particular processing requirements is consistent with the view of prediction as a utility function, according to which comprehenders adjust the degree of prediction in response to the current need for optimizing cognitive resources for efficient comprehension (Kuperberg and Jaeger 2016). The utility of predictive processing is characterized as the resource-bound comprehender's effortful behavior to "maximise the utility of her predictions and rationally allocate resources" (Kuperberg and Jaeger 2016: 44). This utility function has been identified as a crucial feature of the human processing mechanism. Studies have shown that the speed and accuracy of comprehenders' speech perception and word recognition are influenced by specific task goals and statistical regularities of co-occurrence between linguistic items (e.g., Bicknell and Levy 2012; Hayhoe and Ballard 2005; Kleinschmidt and Jaeger 2015). In keeping with the view that extends this processing mechanism to the domain of prediction (Kuperberg and Jaeger 2016), our participants exhibited substantial variability in the timing and extent of prediction depending on the task goals and strategies, reflecting their flexible adaptation of prediction to the requirements of current processing environments. Our findings thus suggest that predictive processing as a utility function applies to Korean speakers' predictive use of morphosyntactic and semantic features during their processing of honorific agreement (see Fine et al. 2013, for adaption of prediction during syntactic processing).

The present findings also support Huettig's (2015) account of the multiple pathways of prediction whereby comprehenders engage in predictive processing via an interconnected system that incorporates both automatic mechanisms and conscious strategic mechanisms. According to this account, comprehenders unconsciously pre-activate linguistic information of upcoming materials through simple associative mechanisms, but at the same time, when there arise demands for allocation of memory resources to optimize processing efficiency, they often engage in strategic, heuristic prediction through the combinatorial system that includes the active reasoning mechanism sensitive to multiple linguistic and nonlinguistic constraints. Huettig (2015) proposed that the mechanism of heuristic processing is interconnected with other prediction systems such as the production-based and simulation-based prediction, constantly influencing each other. For instance, automatic prediction through the simple spreading activation of items may add to the effect of strategic prediction. Consistent with this account, our participants showed predictive behaviors even without external factors promoting prediction under a normal comprehension situation (i.e., the Comprehension group in Experiment 1) and even with linguistic cues biasing against their initial prediction (i.e., the Lower-validity group in Experiment 2), an indication of the automaticity of their predictive processing. Concurrently, the magnitude and speed of prediction scaled up when participants engaged in strategic prediction in response to the task goal and prediction validity of linguistic cues. While our findings suggest some possible association between these two systems, further research is required to investigate the interconnectivity of the multiple predictive processing mechanisms more thoroughly by examining how different mechanisms operate in concert or in conflict to influence the way that comprehenders engage in prediction under diverse processing contexts (e.g., Hintz et al. 2017).

7.3 Limitations and future directions

We acknowledge several limitations of our study. One of the primary limitations is the relatively small sample sizes. The limited number of participants may have resulted in reduced statistical power, which can impact the generalizability of our findings. Each experiment comprised only 32 participants divided into two groups. Despite the limited number of participants, we were able to observe significant effects resulting from our manipulations (instruction manipulation in Experiment 1 and predictive validity manipulation in Experiment 2). This might be because our study had a larger number of items per condition (k = 10) compared to typical psycholinguistic research. This resulted in an overall sample size per condition comparable to previous studies with larger participant pools but fewer items per experimental condition (e.g., k = 5). Still, a larger sample size would have provided greater statistical power and increased the reliability of our findings. Future research should address this problem by involving larger sample sizes to obtain more reliable and generalizable results.

Another limitation, as pointed out by an anonymous reviewer, is the low proportion of critical items (k = 20) to filler items (k = 70), along with the overall large number of stimuli (k = 90), in Experiment 2. We admit that this design choice might have weakened our experimental manipulation, and that the resulting large number of stimuli may have led to attentional problems or fatigue among participants. In further research, we will take this factor into account in our design.

Furthermore, we only included object-biased implicit consequentiality verbs in Experiment 2. A null referent in sentences like (4) is preferentially associated with the subject in the preceding clause in Korean (Hwang 2023), which introduces a confound to the object bias induced by the verb semantics. In order to disentangle the effects of these factors, it is important to include both subject- and object-biased verbs in future studies.

We also acknowledge that our models did not provide a good fit to the data. In both experiments, the models for each region of interest demonstrated only a marginal fit, suggesting that they may not have accurately captured the processing patterns among participants. This issue could have contributed to some unexpected patterns observed among participants in both experiments. For example, in Experiment 1, the Comprehension group showed reduced fixation on the target even after it was mentioned, which may be attributed to increased processing difficulty with the integration of information towards the end of the sentence. However, it is important to note that this unexpected pattern could also be influenced by the limitations of the model fit. Similarly, in Experiment 2, the Lower-validity group looked at target objects more often in the non-honorific than the honorific condition in W1, which may be due to the limitations of the model fit. To address these limitations, future studies should increase the sample size and the number of items, allowing for a larger number of data points that produce more robust models with a good fit to the observed data.

8 Conclusions

The current study shows that Korean speakers proactively use honorific cues for generating expectations about an upcoming referent, both as a result of the automatic activation of honorific agreement cues and as a function of comprehenders' current processing goals and strategies. Our study adduces additional empirical evidence to the underexplored issue of prediction mechanisms in the domain of morphosyntactic processing by adopting methodologies that afford high ecological validity and targeting the linguistic feature involving morphosyntactic and semantic cues as dual informativity sources in the context of Korean sentence processing. Further research testing the effects of various types of top-down factors in prediction will advance our understanding of specific contexts under which comprehenders engage in enhanced or reduced prediction as well as contributing to current theories of the mechanisms underlying the predictive processing system.

Data availability: Complete resources for this project, including materials, data, and analysis code, are available at https://osf.io/enu5m.

Appendix A: Linguistic stimuli used in Experiments 1 and 2

1. 어제 공원에서 벤치에 앉아 계셨던/있었던, 바로 저 할아버지는 아이스크림을 좋아한다.
2. 어제 마당에서 꽃을 보고 계셨던/있었던, 바로 저 할머니는 강아지를 싫어한다.
3. 어제 도서관에서 공부를 하고 계셨던/있었던, 바로 저 선생님은 성격이 좋다.
4. 어제 카페에서 앉아 졸고 계셨던/있었던, 바로 저 교수님은 취미가 영화감상이다.
5. 어제 방안에서 잠을 자고 계셨던/있었던, 바로 저 아주머니는 코를 잘곤다.
6. 어제 식당에서 밥을 먹고 계셨던/있었던, 바로 저 사장님은 친구가 많다.
7. 어제 극장에서 영화를 보고 계셨던/있었던, 바로 저 어르신은 매우 한가하다.
8. 어제 소파에 누워 잠자고 계셨던/있었던, 바로 저 삼촌은 몸이 허약하다.
9. 어제 화장실에서 혼자 울고 계셨던/있었던, 바로 저 형님은 마음이 여리다.
10. 어제 길에서 가방을 찾고 계셨던/있었던, 바로 저 변호사님은 건망증이 심하다.
11. 어제 운동장에서 소리를 치고 계셨던/있었던, 바로 저 감독님은 화를 잘낸다.
12. 어제 서점에서 책을 읽고 계셨던/있었던, 바로 저 선배님은 머리가 똑똑하다.
13. 어제 시장에서 사과를 사고 계셨던/있었던, 바로 저 어머님은 매우 자상하다.
14. 어제 병원앞에서 혼자 걷고 계셨던/있었던, 바로 저 의사선생님은 매우 친절하다.
15. 어제 연구실에서 실험을 하고 계셨던/있었던, 바로 저 박사님은 매우 부지런하다.
16. 어제 꽃집앞에서 활짝 웃고 계셨던/있었던, 바로 저 사모님은 성격이 밝다.
17. 어제 방안에서 노래를 부르고 계셨던/있었던, 바로 저 아버님은 목소리가 좋다.
18. 어제 버스에서 음악을 듣고 계셨던/있었던, 바로 저 할아버지 성격이 활발하다.
19. 어제 편의점에서 라면을 찾고 계셨던/있었던, 바로 저 할머니는 요리를 잘한다.
20. 어제 약국에서 두통약을 사고 계셨던/있었던, 바로 저 선생님은 스트레스가 많다.

Appendix B: Filler items used in Experiment 2

a. Bias-consistent sentences
   1. 학교에서, 영수가 지호를 화나게했기 때문에, 수업이 끝나고 영수에게 곧바로 따졌다.
   2. 회의중에, 아영이가 채은이를 지루하게했기 때문에, 회의가 끝나고 아영이에게 커피한잔을 얻어마셨다.

3. 운동중에, 태영이가 장훈이를 짜증나게했기 때문에, 운동을 멈추고 태영이에게 꾸지람을 시작했다.
4. 파티에서, 소희가 진미를 매료시켰기 때문에, 파티가 끝나고 소희에게 문자메세지를 보냈다.
5. 며칠전에, 정혜가 찬미를 걱정시켰기 때문에, 전화를 걸어서 정혜에게 안부를 물었다.
6. 복도에서, 혜수가 영희를 놀라게했기 때문에, 걸음을 멈추고 혜수에게 호되게 욕설을 퍼부었다.
7. 시합중에, 준혁이가 우영이를 실망스럽게했기 때문에, 시합이 끝나고 준혁이에게 비난을 퍼부었다.
8. 시험전에, 대호가 민기를 격려했기 때문에, 시험이 끝나고 대호에게 감사인사를 보냈다.
9. 식당에서, 준식이가 기영이를 난처하게했기 때문에, 식사를 마치고 준식이에게 핀잔을 주었다.
10. 여행중에, 윤정이가 미정이를 분노하게했기 때문에, 여행이 끝나고 윤정이에게 절교문자를 보냈다.
11. 수업중에, 동범이가 재혁이를 혼란스럽게했기 때문에, 수업이 끝나고 동범이에게 질문을 했다.
12. 소풍에서, 세영이가 수진이를 즐겁게했기 때문에, 소풍이 끝나고 세영이에게 영화티켓을 보냈다.
13. 극장에서, 철수가 상호를 화나게했기 때문에, 영화가 끝나고 철수에게 절교하자고 말했다.
14. 카페에서, 은하가 채영이를 지루하게했기 때문에, 카페를 나와서 은하에게 잔소리를 했다.
15. 교실에서, 민수가 경식이를 짜증나게했기 때문에, 친구들 앞에서 민수에게 화를 내었다.
16. 공연중에, 정미가 예서를 매료시켰기 때문에, 공연이 끝나고 정미에게 꽃다발을 보냈다.
17. 어젯밤에, 지희가 혜림이를 걱정시켰기 때문에, 아침이 되어서 지희에게 전화달라고 부탁했다.
18. 공원에서 지혜가 희정이를 놀라게했기 때문에, 사람들 앞에서 지혜에게 소리를 질렀다.
19. 경기중에, 현석이가 찬혁이를 실망스럽게했기 때문에, 경기가 끝나고 현석이에게 인상을 찌뿌렸다.
20. 발표전에, 성우가 준기를 격려했기 때문에, 발표가 끝나고 성우에게 고맙다고 말했다.
21. 길가에서, 동호가 민규를 난처하게했기 때문에, 사람들 앞에서 동호에게 핀잔을 주었다.

22. 수업중에, 유진이가 현영이를 분노하게했기 때문에, 수업이 끝나고 유진이에게 소리를 질렀다.
23. 회의중에, 진규가 종석이를 혼란스럽게했기 때문에, 회의가 끝나고 진규에게 다시 물었봤다.
24. 학교에서, 미진이가 은정이를 즐겁게했기 때문에, 수업을 마치고 미진이에게 밥을 사줬다.

b. Bias-inconsistent sentences
1. 학교에서, 영수가 지호를 화나게했기 때문에, 수업이 끝나고 지호에게 곧바로 사과했다.
2. 회의중에, 아영이가 채은이를 지루하게했기 때문에, 회의가 끝나고 채은이에게 커피한잔을 사주었다.
3. 운동중에, 태영이가 장훈이를 짜증나게했기 때문에, 운동을 멈추고 장훈이에게 꾸지람을 들었다.
4. 파티에서, 소희가 진미를 매료시켰기 때문에, 파티가 끝나고 진미에게 문자메세지를 받았다.
5. 며칠전에, 정혜가 찬미를 걱정시켰기 때문에, 전화를 걸어서 찬미에게 안부를 전했다.
6. 복도에서, 혜수가 영희를 놀라게했기 때문에, 걸음을 멈추고 영희에게 호되게 욕설을 들었다.
7. 시합중에, 준혁이가 우영이를 실망스럽게했기 때문에, 시합이 끝나고 우영이에게 비난을 받았다.
8. 시험전에, 대호가 민기를 격려했기 때문에, 시험이 끝나고 민기에게 감사인사를 받았다.
9. 식당에서, 준식이가 기영이를 난처하게했기 때문에, 식사를 마치고 기영이에게 핀잔을 들었다.
10. 여행중에, 윤정이가 미정이를 분노하게했기 때문에, 여행이 끝나고 미정이에게 절교문자를 받았다.
11. 수업중에, 동범이가 재혁이를 혼란스럽게했기 때문에, 수업이 끝나고 재혁이에게 질문을 받았다.
12. 소풍에서, 세영이가 수진이를 즐겁게했기 때문에, 소풍이 끝나고 수진이에게 영화티켓을 받았다.
13. 극장에서, 철수가 상호를 화나게했기 때문에, 영화가 끝나고 상호에게 미안하다고 말했다.
14. 카페에서, 은하가 채영이를 지루하게했기 때문에, 카페를 나와서 채영이에게 사과를 했다.
15. 교실에서, 민수가 경식이를 짜증나게했기 때문에 친구들 앞에서 경식이에게 용서를 구했다.

16. 공연중에, 정미가 예서를 매료시켰기 때문에, 공연이 끝나고 예서에게 꽃다발을 받았다.
17. 어젯밤에, 지희가 혜림이를 걱정시켰기 때문에, 아침이 되어서 혜림이에게 전화를 주었다.
18. 공원에서, 지혜가 희정이를 놀라게했기 때문에, 사람들 앞에서 희정이에게 야단을 맞았다.
19. 경기중에, 현석이가 찬혁이를 실망스럽게했기 때문에, 경기가 끝나고 찬혁이에게 면목없다고 말했다.
20. 발표전에, 성우가 준기를 격려했기 때문에, 발표가 끝나고 준기에게 감사인사를 받았다.
21. 길가에서, 동호가 민규를 난처하게했기 때문에, 사람들 앞에서 민규에게 욕을 먹었다.
22. 수업중에, 유진이가 현영이를 분노하게했기 때문에, 수업이 끝나고 현영이에게 미안하다고 말했다.
23. 회의중에, 진규가 종석이를 혼란스럽게했기 때문에, 회의가 끝나고 종석이에게 자세하게 설명해줬다.
24. 학교에서, 미진이가 은정이를 즐겁게했기 때문에, 수업을 마치고 은정이에게 밥을 얻어먹었다.

References

Altmann, Gerry T. M. & Yuki Kamide. 1999. Incremental interpretation at verbs: Restricting the domain of subsequent reference. *Cognition* 73. 247–264.

Baayen, R. Harald. 2008. *Analyzing linguistic data: A practical introduction to statistics using R*. Cambridge: Cambridge University Press.

Baayen, R. Harald, Douglas J. Davidson & Douglas M. Bates. 2008. Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language* 59. 390–412.

Barr, Dale J., Roger Levy, Christoph Scheepers & Harry J. Tily. 2013. Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language* 68. 255–278.

Bates, Douglas, Martin Mächler, Ben Bolker & Steve Walker. 2015. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software* 67. 1–48.

Bicknell, Klinton & Roger Levy. 2012. The utility of modeling word identification from visual input within models of eye movements in reading. *Visual Cognition* 20. 422–456.

Boersma, Paul & David Weenink. 2017. Praat: Doing phonetics by computer [Computer program]. Version 6.0.36. Available at: http://www.praat.org/.

Brothers, Trevor, Tamara Y. Swaab & Matthew J. Traxler. 2017. Goals and strategies influence lexical prediction during sentence comprehension. *Journal of Memory and Language* 93. 203–216.

Brown, Lucien. 2015. Honorifics and politeness. In Lucien Brown & Jaehoon Yeon (eds.), *The handbook of Korean linguistics*, 303–319. Oxford: Wiley.

Choe, Jae-Woong. 2004. Obligatory honorification and the honorific feature. *Studies in Generative Grammar* 14. 545–559.

Choi, Incheol. 2003. A constraint-based approach to Korean partial honorific agreement. *Texas Linguistic Forum* 53. 157–166.
Choi, Kiyong. 2010. Subject honorification in Korean: In defense of Agr and head-spec agreement. *Language Research* 46. 59–82.
Choi, Jaehoon & Heidi Harley. 2019. Locality domains and morphological rules. *Natural Language & Linguistic Theory* 37. 1319–1365.
Clark, Andy. 2013. Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behavioral and Brain Sciences* 36. 181–204.
Crinean, Marcelle & Alan Garnham. 2006. Implicit causality, implicit consequentiality and semantic roles. *Language and Cognitive Processes* 21. 636–648.
Ferreira, Fernanda & Suphasiree Chantavarin. 2018. Integration and prediction in language processing: A synthesis of old and new. *Current Directions in Psychological Science* 27. 443–448.
Ferreira, Fernanda, Alice Foucart & Paul E. Engelhardt. 2013. Language processing in the visual world: Effects of preview, visual complexity, and prediction. *Journal of Memory and Language* 69. 165–182.
Fine, Alex B., T. Florian Jaeger, Thomas A. Farmer & Ting Qian. 2013. Rapid expectation adaptation during syntactic comprehension. *PLoS One* 8. e77661.
Garnham, Alan, Scarlett Child & Sam Hutton. 2020. Anticipating causes and consequences. *Journal of Memory and Language* 114. 104130.
Grüter, Theres, Aya Takeda, Hannah Rohde & Amy J. Schafer. 2018. Intersentential coreference expectations reflect mental models of events. *Cognition* 177. 172–176.
Hayhoe, Mary & Dana Ballard. 2005. Eye movements in natural behavior. *Trends in Cognitive Sciences* 9. 188–194.
Hintz, Florian, Antje S. Meyer & Falk Huettig. 2017. Predictors of verb-mediated anticipatory eye movements in the visual world. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 43. 1352–1374.
Hopp, Holger. 2016. Learning (not) to predict: Grammatical gender processing in second language acquisition. *Second Language Research* 32. 277–307.
Huettig, Falk. 2015. Four central questions about prediction in language processing. *Brain Research* 1626. 118–135.
Huettig, Falk & Ernesto Guerra. 2019. Effects of speech rate, preview time of visual context, and participant instructions reveal strong limits on prediction in language processing. *Brain Research* 1706. 196–208.
Hwang, Heeju. 2023. The influence of discourse continuity on referential form choice. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 49. 626–641.
Jou, Eunsun. 2024. Honorification as agree in Korean and beyond. *Glossa: A Journal of general linguistics* 9. https://doi.org/10.16995/glossa.9565.
Kahneman, Daniel. 2011. *Thinking, fast and slow*. New York: Farrar, Straus and Giroux.
Kamide, Yuki, Gerry T. M. Altmann & Sarah L. Haywood. 2003. The time-course of prediction in incremental sentence processing: Evidence from anticipatory eye movements. *Journal of Memory and Language* 49. 133–159.
Kehler, Andrew, Laura Kertz, Hannah Rohde & Jeffrey L. Elman. 2008. Coherence and coreference revisited. *Journal of Semantics* 25. 1–44.
Kim, Hyunwoo & Eunjin Chun. 2023. Effects of topicality in the interpretation of implicit consequentiality: Evidence from offline and online referential processing in Korean. *Linguistics* 61(1). 77–105.
Kim, Lucy Kyoungsook & Elsi Kaiser. 2009. Effects of honorific agreement on null subject interpretation in Korean. In Young-Se Kang (ed.), *Proceedings of the 2009 Seoul international conference on linguistic interfaces*, 303–318. Republic of Korea: Hankookmunhwasa.

Kim, Hyunwoo, Kitaek Kim & Joonhee Kim. 2022. Modulation of topicality in the use of implicit consequentiality information in Korean speakers' predictive processing: Evidence from visual world eye-tracking. *Language Facts and Perspectives* 57. 249–279.

Kim, Hyunwoo, Kitaek Kim & Joonhee Kim. 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.

Kim, Jong-Bok & Peter Sells. 2007. Korean honorification: A kind of expressive meaning. *Journal of East Asian Linguistics* 16. 303–336.

Kim, Jong-Bok, Peter Sells & Yang Jaehyung. 2005. Deep processing of honorification phenomena in a typed feature structure grammar. In *Companion Volume to the Proceedings of Conference including Posters/Demos and tutorial abstracts*. Jeju Island, Korea: Asian Federation of Natural Language Processing.

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

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

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

Kuperberg, Gina R. 2013. The proactive comprehender: What event-related potentials tell us about the dynamics of reading comprehension. In Brett Miller, Laurie E. Cutting & Peggy McCardle (eds.), *Unraveling reading comprehension: Behavioral, neurobiological, and genetic components*, 176–192. Baltimore, MD: Paul Brookes.

Kuperberg, Gina R. & T. Florian Jaeger. 2016. What do we mean by prediction in language comprehension? *Language, Cognition and Neuroscience* 31. 32–59.

Kwon, Nayoung & Yeonseob Lee. 2024. When grammaticality is intentionally violated: Inanimate honorification as a politeness strategy. *Journal of Pragmatics* 232. 167–181.

Kwon, Nayoung & Patrick Sturt. 2016. Attraction effects in honorific agreement in Korean. *Frontiers in Psychology* 7. 1302.

Kwon, Nayoung & Patrick Sturt. 2024. When social hierarchy matters grammatically: Investigation of the processing of honorifics in Korean. *Cognition* 251. 105912.

Lau, Ellen, Clare Stroud, Silke Plesch & Colin Phillips. 2006. The role of structural prediction in rapid syntactic analysis. *Brain and Language* 98. 74–88.

Lee, So Young & Myung Hye Yoo. 2023. Honorific agreement in Korean: Retrieval processing and predictive processing. *Linguistic Research* 40. 387–407.

Lew-Williams, Casey & Anne Fernald. 2010. Real-time processing of gender-marked articles by native and non-native Spanish speakers. *Journal of Memory and Language* 63. 447–464.

Lupyan, Gary & Andy Clark. 2015. Words and the world: Predictive coding and the language-perception-cognition interface. *Current Directions in Psychological Science* 24. 279–284.

Matin, Ethel, Kuo-Chih Shao & Kenneth R. Boff. 1993. Saccadic overhead: Information-processing time with and without saccades. *Perception & Psychophysics* 53. 372–380.

McMurray, Bob, Michael K. Tanenhaus & Richard N. Aslin. 2009. Within-category VOT affects recovery from "lexical" garden paths: Evidence against phoneme-level inhibition. *Journal of Memory and Language* 60. 65–91.

Mueller, Jeansue & Nan Jiang. 2013. The acquisition of the Korean honorific affix (u) si by advanced L2 learners. *The Modern Language Journal* 97. 318–339.

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

Perek, Florent & Adele E. Goldberg. 2017. Linguistic generalization on the basis of function and constraints on the basis of statistical preemption. *Cognition* 168. 276–293.
Pickering, Martin J. & Chiara Gambi. 2018. Predicting while comprehending language: A theory and review. *Psychological Bulletin* 144. 1002–1044.
Pickering, Martin J. & Simon Garrod. 2013. An integrated theory of language production and comprehension. *Behavioral and Brain Sciences* 36. 329–347.
R Core Team. 2022. R: A language and environment for statistical computing (Version 4.2.1) [Computer software]. Vienna, Austria: R Foundation for Statistical Computing. http://www.R-project.org.
Sohn, Ho-Min. 1999. *The Korean language*. Cambridge: Cambridge University Press.
Song, Sanghoun, Jae-Woong Choe & Eunjeong Oh. 2019. An empirical study of honorific mismatches in Korean. *Language Sciences* 75. 47–71.
Stewart, Andrew J., Martin J. Pickering & Anthony J. Sanford. 1998. Implicit consequentiality. In Morton Ann Gernsbacher & Sharon J. Derry (eds.), *Proceedings of the 20th annual conference of the cognitive science society*, 1031–1036. Mahwah, NJ: Erlbaum.
Thornhill, Dianne E. & Cyma Van Petten. 2012. Lexical versus conceptual anticipation during sentence processing: Frontal positivity and N400 ERP components. *International Journal of Psychophysiology* 83. 382–392.