The Predictive Processing of Korean Control Sentences

Jina Song\textsuperscript{1} and Hongoak Yun\textsuperscript{2*}

\textsuperscript{1}Seoul National University
\textsuperscript{2*}Gachon University

joenjina@gmail.com, oaktreepark@gmail.com

The goal of this study is to investigate whether an antecedent for pro in Korean control sentences is predictively determined and if so, which constituent is preferred to be selected as an antecedent for pro. We manipulated the type of preverbal markers (i.e., control-creating markers, – keys or –la, and a control-neutral marker, -kes) and the type of control verbs (subject control verbs, object control verbs). Using a self-paced moving window reading task with a secondary judgment, we found that 1) reading times at control verbs were slower when preverbal markers (i.e., control-neutral markers) did not provide any particular semantic information on pro than when they (i.e., control-creating markers) cued semantic information on pro. 2) The words prior to control verbs took longer to read when the antecedent for pro corresponded to matrix subjects than matrix objects. 3) The rejection rates of sentences were higher from the position of control verbs when control-neutral markers were used and pro had to be co-indexed with matrix subjects than when control-creating markers were used and pro had to be co-indexed with matrix objects. Our results suggest that licensing of pro identity was predictively determined and matrix objects were preferred the antecedent for pro. Taken together, we claim that Korean readers actively and immediately use information from preverbal markers and control verbs associated with pro resolution.

\textbf{Keywords:} predictive processing, Korean control sentences, control-creating markers
1. Introduction

It has been well documented that readers incrementally integrate words or phrases into sentence fragments without waiting for necessary evidence in the processing of reduced relative clauses (Ferreira & Clifton, 1986; Trueswell, Tanenhaus, & Garnsey, 1994; Altmann, 1999; Clifton, Traxler, Mohamed, Williams, Morris, & Rayner, 2003), filler-gap sentences (Crain & Fodor, 1985; Stowe, 1986; Clifton & Frazier, 1989; Traxler & Pickering, 1996), syntactically complex sentences (Mazuka, Itoh, & Kondo, 2002; Miyamoto & Takahashi, 2002, 2004), and so on. Many of these studies have revealed that diverse cues from semantic, syntactic, or statistical information simultaneously contribute to facilitating readers’ expectation of what is coming up for a given context during online comprehension (Altmann & Kamide, 1999; Hale, 2001; Rayner, Binder, Ashby, & Pollatsek, 2001; Kamide, Altmann, & Haywood, 2003; McDonald & Shillcock, 2003; Boland, 2005; Levy, 2008, Roland, Yun, Mauner, & Koenig, 2012).

Evidence supporting such incremental and predictive processing has also been observed in the processing of head-final languages like Korean and Japanese (Inoue & Fodor, 1995; Kamide & Mitchell, 1999; Miyamoto, 2002; Kamide et al., 2003; Kamide, 2006; Aoshima, Yoshida, & Phillips, 2009; Yun & Hong, 2014; Yun, Nam, Yoo, & Hong, 2015). However, in comparison to studies of head-initial languages in which verb information (e.g., verb argument information, verbs’ syntactic regularities, verb-associated event information, and so on) plays a significant role in driving predictive comprehension, little is known about which sources of information lead to the predictive comprehension in head-final languages. We extended this issue into the predictive use of morphological information encoded by preverbal markers in the processing of empty subjects in control constructions in Korean.¹

Linguistically, the referentiality of an embedded empty subject (pro)²

¹ We do not attempt to distinguish anticipation from integration in our use of terms like prediction or expectation. In either case, processing difficulty is reduced either by the fact that a particular word or phrase is anticipated or that it is easy to integrate into a sentence due to the predictive use of given information.
² In this paper, the empty embedded subject of Korean control sentences (with a control-
should be the function of intrinsic semantic properties of control verbs (Pollard & Sag, 1991; Jackendoff & Culicover, 2001; 2003; Comrak & Smith, 2004; Gamerschlag, 2007). Aspectual, implicative, or desirative verbs trigger subjects to control the interpretation of pro, whereas directive or manipulative verbs trigger objects to control the referent for pro. For example, PRO in (1a), in which a desirative verb, promise, is used, refers to the matrix subject, Mary, whereas PRO in (1b), in which a directive verb, persuade, appears, is co-indexed with the matrix object, Bill.

(1) a. Mary promised Bill [PRO to study hard].
   b. Mary persuaded Bill [PRO to study hard].

Note that in processing sentences like (1a-b), the matrix verb appears prior to PRO. This means that upon the recognition of matrix verbs, the semantic information of those verbs is already available as a cue for pro referentiality and that readers are able to have PRO co-indexed either to Mary or Bill provisionally before they actually encounter to-infinitives. In fact, Boland, Tanenchaus, Garnsey, and Carlson (1995) demonstrated that the argument role associated with PRO was provisionally assigned upon the recognition of main verbs. By using a stop-making-sense task in the processing of wh-filler-gap sentences like (2), they found that readers started rejecting sentences from the pronoun, them. Such an early rejection indicated that readers provisionally assigned the wh-phrase which child as an antecedent for an upcoming PRO at the verb, remind. Thus, when the pronoun them appeared instead, they had to revise the provisional assignment; eliciting immediate rejections at them, although there might be another gap (i.e., the object gap of the verb watch) that the filler, which child, could have been filled into.

creating / a control-neutral marker) is represented as pro, not as trace or PRO (Yang, 1984; Borer, 1989; Lee, 2009). First, since the embedded clauses of control sentences can contain TopP or FocP, they should be considered CPs. Because of the Phase-Impenetrability Condition (Chomsky, 2000), the empty subject of embedded clauses cannot be a trace of A-movement. Second, embedded empty subjects can bear different Case from their antecedents. So, the empty subject cannot be PRO (See Lee, 2009 for further discussion).
(2) [Which child]i did Markj remind themk [PROi*/j/k to watch ____ this evening]? 

Now, suppose that the verb’s information functions as the locus of predictive processing for pro resolution in head-final languages in a similar way that it does in head-initial languages. On the one hand, assigning the referent for pro should be delayed at the end of sentences until control verbs are encountered. On the other hand, although control verbs appear sentence finally, pro resolution might occur predictively by using another source of information that would be available early. In this study, we take the second account and attempt to obtain evidence indicating the predictive processing of pro in Korean control sentences. In particular, we examined the role of preverbal markers (e.g., control-creating/control-neutral marker) in pro referentiality that could lead to the early resolution for pro before the recognition of control verbs (c.f., Yang 1984, 1985; Gamerschlag, 2007; Madigan, 2008). By testing the processing of Korean control sentences either with preverbal control-neutral or preverbal control-creating markers, our goal is to investigate if the antecedents of pro could be provisionally assigned as a result of the immediate use of the information encoded by the preverbal markers before control verbs are actually encountered.

2. The interpretation of Korean control constructions: control verbs, preverbal markers and their relationship

We focus on testing the function of two subtypes of Korean control structures: control constructions with control-neutral markers (i.e., the nominalizer –kes) and control constructions with control-creating markers (i.e., the modal marker –keyss and the mood marker -la). First, control sentences with the nominalizer –kes have a control interpretation only when the lexical properties of control verbs hint at which type of control interpretation (subject-control interpretation vs. object-control interpretation) it should have. Gamerschlag (2007) classified control verbs into Subject Control (SC) verbs and Object Control (OC) verbs based on the semantic characteristics of Korean control verbs. All of the SC verbs listed in (3) share common semantic properties such that the referents co-indexed
with matrix subjects execute the action denoted by the embedded verbs volitionally and intentionally. On the other hand, all of the OC verbs listed in (4) belong to a manipulative/directive verb class. They share the semantic properties that referents corresponding to matrix objects are manipulated to make him/her bring about the action/event denoted by the embedded verb.

(3) Subject Control (SC) verbs
- aspectual/phasal verb class: memchwuta (stop); kkuthmachita (finish)
- desiderative verbs classes: kepwuhatta (refuse), yaksokhata (promise)
- implicative verb class: samkata (refrain), soholhihata (neglect)
- factive/commentative verb class: hwuhoyhata (regret)

(4) Object Control (OC) Verbs
- manipulative/directive verb class: yokwuhata (demand), kangyohata (force), seltukhata (persuade), myenglyenghata (order), yochenghata (request), pwuthakhata (ask (as a favor)), tokchokhata (press), pwuchwukita (encourage), kwenyuhata (induce), chwungkohata (advise)...

When the nominalizer -kes occurs with control verbs, as in the sentences like (5a-b), a control relation is elicited. For example, 1) in example sentence (5a), where a SC verb hwuhoyhata (meaning regret in Korean) is used, pro within the nominalizer complement is obligatorily co-indexed with the matrix subject, Mary, because one can only promise actions one has done on its own (Gamerschlag, 2007). 2) In Example (5b), where an OC verb kangyohata (meaning force in English) is used, pro is obligatorily co-indexed with the matrix object, because the verb means act upon a person in order to make him do the action expressed by the control verb (Gamerschlag, 2007). However, when a nominalizer occurs with non-control verbs like alta (meaning know in English) as in (5c), a control interpretation is not elicited and the referential element atul (meaning son in English) is taken as an embedded subject. Likewise, since the nominalizer

3 Note that in many literatures of Korean control constructions, dative marked NPs in control sentences are regarded as matrix objects and they serve as controllers in object control sentences (Gamerschlag, 2007; Madigan, 2008; Lee, 2009; Park, 2011).
-kes is control-neutral, referent licensing to pro tends to be fixed via the lexical semantics of control verbs.

(5)a. Mary-ka, Bill-eykey \textsubscript{j} [pro\textsubscript{i,j,k} swul-ul masi-l kes-ul] \textit{yaksokhay-ss-ta}. Mary-NOM Bill-DAT alcohol-ACC drink-ADN NML-ACC promise-PAST-DECL \textit{Mary promised Bill to drink alcohol.}
b. Mary-ka\textsubscript{j} Bill-eykey\textsubscript{j} [pro\textsubscript{i,j,k} swul-ul masi-l kes-ul] \textit{kangyohay-ss-ta}. Mary-NOM Bill-DAT alcohol-ACC drink-ADN NML-ACC force-PAST-DECL \textit{Mary forced Bill to drink alcohol.}

Second, unlike the nominalizer (i.e., -kes), modal and mood preverbal markers in embedded complementized clauses contribute to determining the interpretation of Korean control sentences. While the volitional modal marker \textit{–keys} makes pro be construed as a subject of matrix clause, the imperative mood marker \textit{-la} forces the referent of matrix objects to be identical with pro (Yang, 1984, 1985; Gamerschlag, 2007; Madigan, 2008). To illustrate, in examples (6a-c), a non-obligatory control verb \textit{malhata} (meaning \textit{tell} in English) is used and it encodes no particular control interpretation. When the volitional marker \textit{-keyss} is used in (6a), pro is co-referenced with the matrix subject, \textit{Mary}, whereas, when the imperative marker \textit{-la} is used in (6b), pro is co-referenced with the matrix object, \textit{Bill}. However, when neither verbs nor makers contribute to elicit a particular control interpretation as in (6c), in which the declarative mood marker \textit{-ta} is used, no obligatory control interpretation is triggered, resulting in its control interpretation being completely ambiguous.

(6)a. Mary-ka\textsubscript{i} Bill-eykey\textsubscript{j} [pro\textsubscript{i,j} swul-ul masi-keyss-ta-ko] malhay-ss-ta. Mary-NOM Bill-DAT alcohol-ACC drink-VOL-DECL-COMP say-PAST-DECL \textit{Mary told Bill that she (Mary) would drink alcohol.}
b. Mary-ka\textsubscript{i} Bill-eykey\textsubscript{j} [pro\textsubscript{i,j} swul-ul masi-la-ko] malhay-ss-ta. Mary-NOM Bill-DAT alcohol-ACC drink-IMP-COMP say-PAST-DECL \textit{Mary told Bill that he (Bill) would drink alcohol.}
Third, the semantic relationship between control verbs and their preverbal markers should also be taken into account in the resolution of pro identity. When obligatory control verbs co-occur together with embedded clauses attached with control-creating markers, the meaning of the markers must align with the meaning of control verbs. Otherwise, the whole sentence becomes implausible. In other words, the lexical meaning of control-creating markers must “fit” well with those of control verbs (Madigan, 2008).

(7) a. ?? Inho-ka Chelswu-eykey hakkyo-ey ka-la-ko yaksok-ha-yess-ta  
   I-NOM C-DAT school-LOC go-IMP-C promise-do-PST-DC
b. * Inho-ka Chelswu-eykey hakkyo-ey ka-keyss-ta-ko myenglyeng-ha-yess-ta  
   I-NOM C-DAT school-LOC go-VOL-DC-C order-do-PST-DC
c. Inho-ka Chelswu-eykey hakkyo-ey ka-keyss-ta-ko yaksok-ha-yess-ta  
   I-NOM C-DAT school-LOC go-VOL-DC-C promise-do-PST-DC
d. Inho-ka Chelswu-eykey hakkyo-ey ka-la-ko myenglyeng-ha-yess-ta  
   I-NOM C-DAT school-LOC go-IMP-C order-do-PST-DC

The goodness of semantic-fit between control verbs and preverbal markers accounts for why the examples (7a-b) are implausible but the examples (7c-d) are plausible. The imperative marker, –la, in (7a) requires the referent of the matrix object to be co-indexed with the embedded pro, but the SC verb, yaksokhata (meaning promise in English) requires the matrix subject to be co-referenced with pro. The mismatch between the control verb and the marker makes the sentence (7a) implausible. Similarly, the volitional marker, –keyss, in (7b) makes the matrix subject become co-referenced with pro; but it is followed by the OC verb, myenglyenghata, (meaning order in English) which requires the matrix object to be co-referenced with pro. In both of the cases, the extremely poor goodness of the semantic fit between preverbal markers and control verbs results in the implausibility of these sentences. In contrast, there is no mismatch relationship found in the
examples of (7c-d); the preferred type of pro by the preverbal markers is well matched with the type of control verbs. The volitional marker, -keyss, is aligned with the SC verb, yaksokhata (promise in English), and the imperative marker, -la, is accompanied with the OC verb, myenglyenghata, (order in English). To account for the plausibility associated with the goodness of semantic fit between control verbs and preverbal markers, Madigan (2008) proposed the subcategorization frames for control verbs, as shown in (8). SC verbs must select for volitional complements, as listed in (8a), whereas OC verbs must select imperative complements, as noted in (8b). Sentences like (7a-b) are implausible because the sectional restrictions of these control verbs are violated.

(8) a. promise-type (SC verbs): +vol, *imp, *exh 
    b. order-type (OC verbs): *vol, +imp, *exh

Under this theoretical linguistic approach, the referent of pro is assigned only when control verbs appear and pro’s licensing information becomes available. Moreover, controller assignment is completely determined depending on the lexical meaning of control verbs and thus control verbs have the necessary information for pro licensing (Pollard & Sag, 1991; Jackendoff & Culicover, 2001, 2003; Gamerschlag, 2007). However, from the processing perspective, given the fact that not only control verbs but also preverbal markers are involved in the interpretation of control constructions, it is possible that readers’ early use of preverbal marker information would prevent relatively increased processing costs related to keeping an unassigned pro in working memory until appropriate control verbs arrive at the end of the sentence (c.f., Gibson, 2000). The early resolution of pro referentiality might also occur in other head-final languages like Japanese. We discuss this issue further in the following section.

---

4 Exhortative complement is selected by the proposal-type verbs (Madigan, 2008).
3. Incremental processing of pro in control constructions: evidence in Japanese

Previous studies on control processing in Japanese showed that Japanese readers provisionally assigned an antecedent to pro before control verb information was available (Sakamoto, 1995, 1996; Oda et al., 1997; Ninose et al., 1998; Witzel & Witzel, 2011, 2016). Specifically, these studies suggested that readers had bias on which constituent should be an antecedent for pro before the point when sentence-final verbs appeared. Using an end-of-sentence antecedent retrieval task, Sakamoto (1995, 1996) observed that Japanese readers had an OC bias in the interpretation of Japanese control sentences, regardless of whether sentences were canonical or scrambled. On the contrary, when a recognition task was used, Oda et al. (1997) and Ninose et al. (1998) found that Japanese readers had a SC bias. However, due to the use of offline behaviors, these results have some limitations in that they might not reflect online processing behaviors (Witzel & Witzel, 2011).

Using an online reading comprehension task, Witzel and Witzel (2011) compared control sentences with a null embedded subject (pro) to those with an overt embedded subject. As shown in (9a-b), the gender-marked reflexive (i.e., karezisin or kanozyozisin) was co-referenced with either the matrix subject (i.e., Daisuke) or the matrix object (i.e., Kyoko). The researchers observed that control verbs in sentences with overt embedded subjects were read faster than those in sentences with empty pro. Importantly, the reading time differences emerged only in the SC condition (i.e., 9a) but not in the OC condition (i.e., 9b). Overall, SC sentences took longer to read than OC sentences. These results indicated that matrix objects were strongly preferred as antecedents for pro.

(9) a. Daisuke]-ga  Kyōko]-ni  [(karezisin]-ga  kōnyūsuru koto]-o
  Daisuke-NOM Kyoko-DAT (himself]-NOM) car-ACC purchase fact]-ACC
  kyūtōsitu-de zimansita  yō desu.
  office kitchen-in showed off  seems
  It seems that Daisuke showed off to Kyoko in the office kitchen
  that he would purchase a car.
To sum up, several studies in Japanese demonstrated that control sentences in head-final languages as well as those of head-initial languages are processed without delay. For such an immediate processing, certain matrix arguments (objects in case of Japanese) are provisionally selected as a highly-likely antecedent for pro. By doing so, readers try to make the most complete representation possible even before they meet necessary control information later at control verbs (Witzel & Witzel, 2011). Thus, if any processing difference appears at control verbs due to the function of readers’ pre-determined object bias, this could support, in part, the predictive processing of pro in Japanese. However, the current results from the Japanese studies simply show readers’ bias for pro interpretation and do not provide direct evidence supporting the early determination to pro-identity.

4. Our study

There are two sources of information that contribute to determining the referent of pro in Korean control constructions. One is the semantic information denoted by control verbs and the other is the semantic information of preverbal control-creating markers. Also, for control sentences to be plausible, the semantic information of control-creating markers that licenses the referent of pro has to be matched with the semantic information denoted by following control verbs.

We have a couple of things to clarify. First, if Korean control sentences are processed predictively, pro resolution should be initiated before a head (i.e. a control verb) appears and the representational features of the head can be postulated before the head arrives. Thus, in the cases that control sentences have preverbal licensing information, readers are likely to assign
a role to pro immediately using this preverbal information. Consequently, readers are highly likely to expect to encounter relevant control verbs whose semantic information is well matched to the semantic properties of pro’s referential identity. Second, as found in the Japanese studies, it is also important for us to obtain evidence that Korean readers have a provisional bias on which constituent is preferred as an antecedent for pro before they encounter control verbs. If readers predictively use the preverbal marker information, they would develop their bias on who is more likely to be an antecedent for pro. We think that readers might have different degree of bias depending on how strongly potential cues combinatorily indicate which constituent should be selected. Namely, when both control verbs and their preverbal markers refer to the same constituent, the strength of readers’ bias on which one has to be an antecedent for pro may be very strong. Whereas, if neither control verbs (e.g., non-obligatory control verbs) nor their preverbal markers (e.g., control-neutral marker) refer to a certain constituent, the degree of readers’ bias on which one has to be an antecedent for pro may be very weak. The strength of readers’ bias toward pro might vary by how much available information is accessible in determining an antecedent for pro.

We have two hypotheses to test.

1) The null hypothesis is to support no predictive resolution of pro in the processing of Korean control constructions. If this hypothesis is correct, there would be no particular bias for which constituent should be assigned as an antecedent for pro, prior to the occurrence of control verbs. For the results to support this hypothesis, it is necessary to observe no processing differences between when the preverbal marker information is control-neutral and when the preverbal marker information is control-creating. The null hypothesis is in line with the theoretical linguistic point of view positing that controllers are determined by the information specified in control verbs (Cormack & Smith, 2004; Choe, 2006; Madigan, 2008). In other words, pro’s controller cannot be assigned at the position of the preverbal control-creating marker. Thus, pro resolution should be delayed until control verbs appear, regardless of whether preverbal markers create a control interpretation or not.
2) Our alternative hypothesis is to support the predictive resolution of pro in the processing of Korean control constructions. If this hypothesis is correct, there would be particular bias for which constituent should be assigned as an antecedent for pro depending on the type of preverbal marker information. For the results to support this hypothesis, it is necessary to observe significant processing differences at control verbs between when the preverbal marker information is control-neutral and when the preverbal marker information is control-creating. Moreover, if the referent of pro is predictively determined, a further question emerges about whether there is a constituent that is strongly preferred as an antecedent for pro. If Korean readers behave similarly to Japanese readers, we expect to observe object-bias as an antecedent for pro.

5. Experiment

Our goal was to investigate if pro in control sentences in Korean would be predictively processed. We also asked whether there would be a bias on which constituent would be preferred as an antecedent for pro. In order to do this, we manipulated the types of preverbal marker and control verbs, resulting in an experiment with a 3 preverbal markers (Neutral, SC-creating, OC-creating) x 2 Plausibility (Match, Mismatch) within-subjects design. For the test of our hypotheses, we employed a region-by-region self-paced moving window task with a secondary judgment.

5.1 Method

5.1.1 Participants

Forty-six native speakers of Korean (22 females, 17 males) took part in the online reading experiment. They were graduate or undergraduate students

---

5 According to our hypothesis that Korean readers might have object bias in the processing of control sentences as similarly as Japanese readers, the mismatch condition of the neutral marker condition holds sentences in which subject-control verbs are used. The idea was that if Korean readers have object-bias, they are likely to determine objects as antecedents for pro. Thus, they might expect to encounter object-control verbs rather than subject-control verbs, although both types of verbs are quite plausible.
of Seoul National University and paid 4,000 won for their participation.

5.1.2 Materials

Thirty-six sets of experimental sentences were used, as indicated in Table 1. All NPs at W1 were subjects marked with nominative case and all NPs at W2 were objects marked with dative case. The dative marked NPs are considered as matrix objects (c.f., see footnote 3). The nouns at W1 and W2 were of either proper names or common nouns. (The position of the names and the common nouns were balanced in W1 and W2.) The NPs associated with accusative markers (embedded objects) appeared at W3 and embedded verbs attached with preverbal markers followed at W4. Note that empty subjects (pro) for the embedded verbs were implicitly included in the embedded clause regions (W3-W4). Next, at W5, control verbs occurred. The control verbs at (10a) and (10c) were SC verbs, meaning that embedded pro in these sentences were controlled by sentential subjects rather than datives (matrix objects). On the other hand, in sentences with OC verbs in (10b) and (10e), the embedded pro referred to datives (matrix objects) rather than sentential subjects. To avoid wrap-up effects by critical control verbs being at the end of sentences, we added some words at W6 and W7. They were the evidential predicate, kes-ulo poi-, ‘seem’. Finally, words at W8 and W9 contained utterance verbs and their subjects.

Each set of experimental stimuli differed in two ways. First, the morphological information associated with pro encoded by preverbal markers was manipulated into three levels by using the control neutral marker -kes as in (10a-b), the SC-creating marker -keyss as in (10c-d), and the OC-creating marker -la as in (10e-f). The control neutral marker -kes in (10a-b) does not control which constituent should be determined as pro’s referent at the position of embedded verbs (W4), and thus, the interpretation for pro tends to be temporarily ambiguous until control verbs (W5) are encountered. However, control sentences combined with a volitional modal marker -keyss like (10c-d) strongly control pro to be co-indexed with the matrix subject, while control sentences combined with an imperative mood marker -la like (10e-f) clearly cue objects (or datives) to be antecedents of pro. The interpretation of pro is likely to be provisionally determined upon
<table>
<thead>
<tr>
<th>VERB</th>
<th>MARKER</th>
<th>PLAUSIBILITY</th>
<th>W1</th>
<th>W2</th>
<th>W3</th>
<th>W4 (Preverbal marker)</th>
<th>W5 (Control Verb)</th>
<th>W6</th>
<th>W7</th>
<th>W8</th>
<th>W9</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Jee-NOM</td>
<td>friend-DAT</td>
<td>tabacco-ACC</td>
<td>quit-ADN-thing-ACC</td>
<td>swear-ADN</td>
<td>thing-DR</td>
<td>seem:PRS-DC-C</td>
<td>Sue-NOM</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Sue said that Jee seemed to swear to (his) friend [pro, to stop smoking]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10b</td>
<td>OC</td>
<td>Match</td>
<td></td>
<td>Jee-ka</td>
<td>chinkwiw-eykey</td>
<td>tampay-lui</td>
<td>kkumul-keu-ul</td>
<td>seliukha-n</td>
<td>kesi-ulo</td>
<td>po-in-ta-ko</td>
<td>Sue-ka</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Jee-NOM</td>
<td>friend-DAT</td>
<td>tabacco-ACC</td>
<td>quit-ADN-thing-ACC</td>
<td>persuade-ADN</td>
<td>thing-DR</td>
<td>seem:PRS-DC-C</td>
<td>Sue-NOM</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Sue said that Jee seemed to persuade (his) friend [pro, to stop smoking]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Jee-NOM</td>
<td>friend-DAT</td>
<td>tabacco-ACC</td>
<td>quit-VOL-DC-C</td>
<td>swear-ADN</td>
<td>thing-DR</td>
<td>seem:PRS-DC-C</td>
<td>Sue-NOM</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Sue said that Jee seemed to swear to (his) friend [pro, to stop smoking]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10d</td>
<td>OC</td>
<td>Match</td>
<td></td>
<td>Jee-ka</td>
<td>chinkwiw-eykey</td>
<td>tampay-lui</td>
<td>kkumul-keuul-tako</td>
<td>mayngyesan</td>
<td>kesi-ulo</td>
<td>po-in-ta-ko</td>
<td>Sue-ka</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Jee-NOM</td>
<td>friend-DAT</td>
<td>tabacco-ACC</td>
<td>quit-VOL-DC-C</td>
<td>persuade-ADN</td>
<td>thing-DR</td>
<td>seem:PRS-DC-C</td>
<td>Sue-NOM</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>* Sue said that Jee seemed to persuade (his) friend [pro, to stop smoking]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10e</td>
<td>OC</td>
<td>OC-creating</td>
<td>Match</td>
<td>Jee-ka</td>
<td>chinkwiw-eykey</td>
<td>tampay-lui</td>
<td>kkumul-la-ko</td>
<td>seliukha-n</td>
<td>kesi-ulo</td>
<td>po-in-ta-ko</td>
<td>Sue-ka</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Jee-NOM</td>
<td>friend-DAT</td>
<td>tabacco-ACC</td>
<td>quit-IMP-C</td>
<td>persuade-ADN</td>
<td>thing-DR</td>
<td>seem:PRS-DC-C</td>
<td>Sue-NOM</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Sue said that Jee seemed to persuade (his) friend [pro, to stop smoking]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10f</td>
<td>SC</td>
<td>Mismatch</td>
<td></td>
<td>Jee-ka</td>
<td>chinkwiw-eykey</td>
<td>tampay-lui</td>
<td>kkumul-la-ko</td>
<td>mayngyesan</td>
<td>kesi-ulo</td>
<td>po-in-ta-ko</td>
<td>Sue-ka</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Jee-NOM</td>
<td>friend-DAT</td>
<td>tabacco-ACC</td>
<td>quit-IMP-C</td>
<td>swear-ADN</td>
<td>thing-DR</td>
<td>seem:PRS-DC-C</td>
<td>Sue-NOM</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>* Sue said that Jee seemed to swear to (his) friend [pro, to stop smoking]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
the recognition of embedded verbs with mood markers like –keyss and -la, prior to the occurrence of control verbs.

Second, control verbs occurring at W5 were either SC verbs (10a, 10c, and 10f) or OC verbs (10b, 10d, and 10e) in each condition of preverbal markers. Note that the role of modal/mood markers in the interpretation of pro fit well with that of control verbs in plausible sentences like (10c) and (10e), whereas the role of modal/mood markers in interpreting who should be co-indexed with pro does not fit with that of control verbs in implausible sentences like (10d) and (10f). Consequently, the plausibility of sentences emerged as a function of the goodness of thematic fit match between preverbal markers and control verbs. Note that we marked sentences with subject-control verbs in neutral marker condition to belong to the mismatch condition. This was done given our hypothesis that Korean readers might have object bias in the processing of control sentences as similarly as Japanese readers. The idea was that if Korean readers have object-bias, they are likely to determine objects as antecedents for pro. Thus, they might expect to encounter object-control verbs rather than subject-control verbs, although both types of verbs are quite plausible.

All experimental sentences were intermixed with 72 filler sentences which consisted of 48 plausible sentences and 24 implausible sentences and presented in counterbalanced lists. Implausible sentences could be rejected for semantic reasons because the semantic fit between subjects and verbs was poor (i.e. an inanimate subject was combined with an action verb). The filler sentences had various forms of syntactic structures (e.g., a raising structure, a relativization clause, and a declarative complementized structure).

5.1.3 Norming study

In order to confirm that our experimental materials were manipulated as they should be without having other potential concerns, we conducted a preliminary norming study. Forty-five native speakers of Korean (32

---

6 10 control verbs (4 SC verbs and 6 OC verbs) were drawn from Gamerschlag (2007); remaining 2 verbs’ status as SC verbs was confirmed by the result of the preliminary experiment.
females and 13 males) participated in this study. They were undergraduate or graduate students in universities in Seoul (mean age = 25.9). In this study, we examined 1) how natural the experimental sentences would be and 2) which constituent would be a likely actor of the events in embedded clauses, respectively; that is, which matrix argument should be co-referenced with pro. For these purposes, we asked the participants to rate the naturalness of each sentence on the scale from 1 (completely unnatural) to 7 (completely natural). Then, we also required them to select who would be the actor of the event in the embedded clause in plausible sentences like (10a), (10b), (10c), and (10e). All sentences were presented in a pseudorandom order using Google Form.

First, the results of naturalness rating are displayed in Table 2. All plausible sentences were judged to be fairly natural, having their means of naturalness over 5 out of 7, and their naturalness was not different from each other. Plausible sentences with SC control verbs and those with OC control verbs were equally natural ($t(35) = -1.36, p = .18$). The naturalness ratings of these sentences was not different between the sentences in which control-neutral markers and SC-biased control verbs were used, ($t(35) = .49, p = .63, t(35) = -.74, p = .47$) and between those in which control-neutral markers and OC-biased control verbs were used, ($t(35) = .01, p = .98, t(35) = -1.58, p = .12$). The naturalness ratings between matched and mismatched sentences significantly differed not only when SC-creating

<table>
<thead>
<tr>
<th>Condition</th>
<th>Mean</th>
<th>SD</th>
<th>Mean</th>
<th>SD</th>
<th>Mean</th>
<th>SD</th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Neutral-SCV (e.g., 10a)</td>
<td>5.35</td>
<td>.75</td>
<td>5.24</td>
<td>.56</td>
<td>5.24</td>
<td>.96</td>
<td>2.31</td>
<td>.45</td>
</tr>
<tr>
<td>Neutral-OCV (e.g., 10b)</td>
<td>5.24</td>
<td>.96</td>
<td>5.24</td>
<td>.45</td>
<td>5.46</td>
<td>.62</td>
<td>2.23</td>
<td>.40</td>
</tr>
<tr>
<td>SCM-Match (e.g., 10c)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SCM-Mismatch (e.g., 10d)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>OCM-Match (e.g., 10e)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>OCM-Mismatch (e.g., 10f)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Note.** Naturalness ratings ranged from 1 to 7, 1 corresponding to completely unnatural and 7 corresponding to completely natural. ‘Neutral-SCV’ refers to the condition with the control-neutral marker and SC verbs. ‘Neutral-OCV’ refers to the condition with the control-neutral marker and OC verbs.

Table 2. Results of naturalness ratings
control markers were used \( (t (35) = 16.52, p < .00) \) but also when OC-creating control markers were used \( (t (35) = 25.56, p < .00) \). Sentences of control-neutral markers were equally natural, regardless of whether they were combined with SC-biased control verbs or OC-biased control verbs \( (t (35) = .64, p = .53) \), indicating that these sentences should be equally acceptable and thus control-neutral markers would not elicit any significant bias to which constituent should be marked as pro, at least in offline. The results of naturalness norming confirmed us that the effect that we would observe from an online study would not be influenced by poorly controlling the goodness of naturalness.

Second, the results about selecting who would be the actor of embedded verbs are illustrated in Figure 1. The patterns of selection preference were obvious. Subjects were predominantly selected as the antecedent for pro when SC verbs were used in sentences in which control markers were either neutral (96%) or SC-creating (99%). Similarly, dative marked objects were highly preferred as the controller for pro when OC verbs were used in sentences in which control markers were either neutral (99%) or object-creating (98%). These results again confirmed that the interpretation of pro would be clear at the position of control verbs, supporting previous studies that the role assignment for pro should be determined by control verbs.
Crucially, the strength of pro preference in the condition of control-neutral marker was equally extremely high in both SC verbs and OC verbs. If any significant bias for who should be an antecedent for pro would NOT emerge online, there should be no significant difference in the processing of control verbs between SC verb condition and OC verb condition. However, if any bias to ‘which’ constituent as an antecedent would occur during online comprehension in a way that previous Japanese studies showed (Sakamoto, 1995, 1996; Oda et al., 1997; Ninose et al., 1998; Witzel & Witzel, 2011), there should be significant difference in the processing of control verbs. Our norming results showing the equated preference of selection in the control-neutral marker condition allowed us to test our hypothesis in a fair manner.

5.1.4 Procedure

A self-paced moving-window procedure with an incremental judgment task (a.k.a., stop-making-sense task) was conducted using Linger (http://tedlab.mit.edu/~dr/Linger/ by Doug Rohde). We used the same task as in Boland et al.’s (1995) study to increase the sensitivity of the methodology to subtle effects that might not be observed in a straight reading paradigm (Mauner, Tanenhaus, & Carlson, 1995). In the beginning, a row of dashes and white spaces were shown on a computer monitor. Stimulus sentences were presented on one line. To reveal the first region, participants had to press a spacebar, resulting in the dashes corresponding to this region being replaced by words. To reveal the next region, participants had to press the spacebar again. Then, the first region reverted to dashes and the second region was revealed. Participants continued pressing the spacebar to read each subsequent region as long as this region made sense given the sentence fragment that they were reading. If a sentence turned out to be insensible at any time, participants had to press a Q-key corresponding to “No” response. The “No” response ended the current trial immediately and initiated the next trial. “Yes” Reading times and “No” judgments were obtained as dependent variables for each region, but No judgements were primarily
considered for our analysis. Before the experiment began, participants were asked to read the instructions that described the task with some examples. After reading the instructions, they completed 7 sensible trials and 6 nonsensical practice trials to familiarize themselves with the task and the response keys.

5.2 Analysis and results

The participants who had accuracy rates\(^7\) less than 80% were excluded from the data. So, the data of 39 participants (22 females, 17 males) was included for the analysis.

5.2.1 No rejections

The adjusted percentage of “No” judgments were tabulated at each region per each participant, using the procedure outlined in Boland et al. (1990). Adjusted percentages for each sentence trial were computed by dividing the number of “No” judgments at a given region by the number of remaining opportunities that a participant had for responding “No” in that sentence. Mean adjusted percentages were then computed for each condition and region for each participant. Figure 2 displayed the results of “No” rejections across target regions.

For our materials, the earliest moment that sentences started insensible was from W5. Most rejections appeared in the Mismatch condition of sentences with control-creating markers (i.e., SCM and OCM). Crucially, “No” rejections were elicited even in the conditions with the control-neutral marker. Furthermore, for sentences with control-neutral markers, rejections occurred more in the SC condition than in the OC condition, suggesting a strong bias to select dative marked objects as an antecedent for pro. These results indicating that “No” rejections appeared differently depending on the types of control markers supported the predictions of the

---

\(^7\) The accuracy rates were measured based on participants’ responses of filler sentences. For example, when a participant pressed “No” buttons for sensible filler sentences or did not press “No” buttons for insensible filler sentences, the participant was considered not to comprehend the sentences accurately.
predictive-processing view rather than the theoretic linguistic view. For further statistical analyses, we conducted a series of mixed-effect logistic regressions.

For a statistical scrutiny of rejection rates across conditions and regions, we conducted a linear mixed logit model (Baayen, 2008), using lme4 (Bates, Maechler, Bolker, & Walker, 2013) and the languageR library (Baayen, 2014) for the R statistics program (R Development Core Team, 2014). In the model, a binary code was assigned to the dependent variables (i.e., judgment responses): 1 corresponding to “No (rejection)” and 0 corresponding to “Yes (acceptance)”. As for fixed factors, we had Match, Marker, Word, and Trial. 1) The matchedness of control verbs was coded as 1 for the match condition and 0 for the mismatch condition. 2) The variable of Marker was coded into three: the neutral marker condition as 0, the subject-favoring control marker (SCM) condition as 1, and the object-favoring control marker (OCM) condition as 2. 3) Measurements observed from Word 4 to Word 7 were included. 4) In order to control the effect of presentation order,
we also included an additional factor called Trial number in all analyses across all regions. Both participants and items were submitted as random variables. For the structure of random effects, fully crossed and specified random effects were reduced step by step until the model converged. Only the effects which contributed to the significant improvement for the model were included in the final model. Table 3 displays the results from the final logistic model in which the fixed factors (Marker, Match, Word, and Trial) and their interactions were included.

The main effect of Match indicated that sentences in the mismatch condition were rejected more than sentences in the match condition. The main effect of Word meant that rejections occurred more in later positions of sentences (i.e., W5 and W6) than earlier position of sentences (i.e., W4). Three significant two-way interactions appeared. The three-way interaction of Marker*Match*Word was also observed, indicating that the interaction of Marker*Match differed as a function of Word. To investigate the nature of this three-way interaction, we conducted a mixed-effect logistic regression analysis at each word. The two-way interaction of Marker * Match emerged at W5 and W6 but not at W4 and W7. The results at W5
Table 4. The results of the linear mixed effect logistic regression at W5 (i.e., control verbs) and at W6 (after control verbs)

<table>
<thead>
<tr>
<th></th>
<th>W5</th>
<th></th>
<th>W6</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Estimate</td>
<td>S.E.</td>
<td>z-value</td>
<td>p-value</td>
</tr>
<tr>
<td>Intercept</td>
<td>-1.02</td>
<td>.13</td>
<td>-7.69</td>
<td>.00***</td>
</tr>
<tr>
<td>Marker</td>
<td>.26</td>
<td>.09</td>
<td>3.00</td>
<td>.00**</td>
</tr>
<tr>
<td>Match</td>
<td>2.03</td>
<td>.15</td>
<td>13.78</td>
<td>.00***</td>
</tr>
<tr>
<td>Log (verb)</td>
<td>-.20</td>
<td>.14</td>
<td>-1.43</td>
<td>.15</td>
</tr>
<tr>
<td>Trial</td>
<td>.00</td>
<td>.00</td>
<td>1.55</td>
<td>.12</td>
</tr>
<tr>
<td>Marker*Match</td>
<td>1.10</td>
<td>.18</td>
<td>6.25</td>
<td>.00***</td>
</tr>
</tbody>
</table>

Note. All variables were centered. As for the condition of control-neutral markers, when control verbs were SC-biased, it was coded as being mismatched, whereas when control verbs were OC-biased, it was coded as being matched.

and W6 are reported in Table 4.

For the analysis of rejections at W5, we included the log-transformed frequency of target verbs to control for the effect of verb frequency. Also, to control for the effect of presentation order, we also included Trial as a factor. As shown in Table 4, the main effect of Match occurred, indicating that sentences in the mismatch condition were rejected more than sentences in the match condition. The main effect of Marker meant that rejections occurred more in the condition of control-creating markers than in that of control-neutral condition. Importantly, the two-way interaction of Marker*Match emerged, indicating that rejection rates were modulated by the types of control markers. After splitting the data by the factor of control markers, we found that the effect size for rejections was much stronger in the condition of SCM (coefficient = 2.44, S.E. = .27, p < .000) and the OCM (coefficient = 3.14, S.E. = .32, p < .000) than in that of the Neutral marker (coefficient = .79, S.E. = .26, p < .002).

The very same patterns of results were yielded in the analysis of rejections at W6 (see Table 5). The main effect of Match occurred, indicating that sentences in the mismatch condition were rejected more than sentences in the match condition. The main effect of Marker meant that
rejections occurred more in the condition of control-creating marker than in that of control-neutral condition. Importantly, the two-way interaction of Marker*Match emerged, indicating that rejection rates were modulated by the types of control markers. After splitting the data by the factor of control markers, we found that the effect size for rejections was much stronger in the condition of SCM (coefficient = 3.05, S.E. = .42, \( p < .000 \)) and the OCM (coefficient = 6.03, S.E. = 1.11, \( p < .000 \)) than in that of the Neutral marker (coefficient = .96, S.E. = .48, \( p < .04 \)).

5.2.2 “Yes” reading times

“Yes” reading times were analyzed only when sentences were plausible. Prior to the analyses of “Yes” reading times for target words in sensible sentences, reading times were trimmed in two steps. First, reading times faster than 200 milliseconds and greater than 5,000 milliseconds were removed leading to the removal of 55 data points out of the total number of reading times (2% of the data). We removed these extreme reading times because including them might have led to an inflated estimation of the data.

As in the analysis of the “No” rejections, a linear mixed effects regression was conducted. The same version of the R program was used as in the analysis of the “No” responses. A set of fixed factors was submitted into the model: 1) based on Control verb type, 2 object bias in which 0 corresponded to the condition that the antecedent of pro was an object and 1 corresponded to the condition that the antecedent of pro was a subject, 2) 2 Marker in which 0 referred to the condition of control-neutral marker and 1 referred to the condition of control-creating marker, and 3) the interaction of object bias and Marker. The effects of these factors were tested in four regions of Word (i.e., W4, W5, W6, and W7). In addition to these factors, log-transformed verb frequency was added in the analyses of W5 and W6 to control for the frequency effect of verbs on reading times. The factor of Trial was not included because it had no effect on results. As random factors, items and participants were included in the models. For the structure of random effects, fully crossed and specified random effects were reduced step by step until the model converged. Only the effects which contributed to the significant improvement for the model
Unexpectedly, at W4, we observed the effect of object bias, indicating that the words at W4 were read faster when the antecedents for pro were objects (OC verbs) than when they were subjects (SC verbs) regardless of whether preverbal markers were control-neutral or control-creating. The interaction of object bias and Marker revealed that the OC-bias was statistically significant when the preverbal marker was control-creating (Estimate = 123.68, S.E. = 31.53, t-value = 3.92) but numerically trending when the preverbal marker was control-neutral (Estimate = 47.12, S.E. = 31.61, t-value = 1.49). The differences of reading times at W4 need further exploration.

Table 5. The results of the linear mixed effect regression at Word 4-6

<table>
<thead>
<tr>
<th></th>
<th>W4</th>
<th></th>
<th>W5</th>
<th></th>
<th>W6</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Estimate</td>
<td>S.E.</td>
<td>t-value</td>
<td>Estimate</td>
<td>S.E.</td>
<td>t-value</td>
</tr>
<tr>
<td>Intercept</td>
<td>850.72</td>
<td>26.32</td>
<td>32.32*</td>
<td>986.70</td>
<td>38.72</td>
<td>25.48*</td>
</tr>
<tr>
<td>Marker</td>
<td>-3.34</td>
<td>22.84</td>
<td>-0.15</td>
<td>-208.24</td>
<td>40.14</td>
<td>-5.19*</td>
</tr>
<tr>
<td>OC-bias</td>
<td>94.65</td>
<td>22.83</td>
<td>4.15*</td>
<td>44.25</td>
<td>40.42</td>
<td>1.10</td>
</tr>
<tr>
<td>OC-bias*Marker</td>
<td>95.77</td>
<td>45.66</td>
<td>2.10*</td>
<td>-144.58</td>
<td>80.64</td>
<td>-1.79</td>
</tr>
<tr>
<td>Log (verb)</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>-93.14</td>
<td>42.01</td>
<td>-2.22*</td>
</tr>
</tbody>
</table>

Note. All predictors were centered. If the absolute t-value of a fixed factor was over 2, the effect of the factor was considered to be significant at α < .05, marked with * (Gelman & Hill, 2007). N/A meant that log-transformed verb frequency was not applied in the analysis of W4.
At W5, not surprisingly, a significant effect of verb frequency was observed such that higher frequency verbs were read faster than lower frequency words. Most importantly, the main effect of Marker was significant. The control verbs in the control-creating condition were read faster than the verbs in the control-neutral condition. The processing of control verbs was facilitated in the control-creating condition, due to the already available semantic information of preverbal markers. When the semantic information of the marker did not contribute to the resolution of pro identity, the processing of control verbs was slowed down presumably because readers spent extra time in the integration of verbs while they were determining who was an antecedent for pro. The main effect of Marker supported our hypothesis.

At W6, there were significant main effects of OC bias and Marker. The words after control verbs were read faster when they followed sentence fragments in which antecedents for pro were objects and preverbal marker information was control-creating than when they followed sentence fragments in which antecedents for pro were subjects and preverbal marker information was control-neutral. No significant effects occurred at W7. 

Figure 3. The means of reading times of target words across conditions from W4 to W6

The Predictive Processing of Korean Control Sentences
summary, the analyses of “Yes” reading times revealed that the processing of pro in Korean constructions was predictive and the antecedent of pro was preferred to be objects rather than subjects.

6. General Discussion

An offline preliminary experiment and an online stop-making-sense task were conducted to investigate whether the processing of Control sentences in Korean would be predictive, with the hypothesis that the resolution of pro would be provisionally determined prior to the recognition of control verbs. Our hypothesis was supported. A significant reading-time difference was detected depending on the presence of the semantic information encoded by preverbal markers (control-neutral marker condition vs. control-creating marker condition). We also had an additional question about whether Korean readers would have a bias on which constituent would be preferred as an antecedent for pro. We observed more rejections and longer reading times emerged when pro was co-indexed with subjects than when it was done with objects. In the processing of Korean control sentences, objects, rather than subjects, were strongly preferred as an antecedent for pro. Taken together, our results showed that the processing of Korean control sentences was incremental and predictive and that matrix objects were strongly biased to be antecedents for pro. These findings provided the first behavioral evidence indicating that Korean readers actively and immediately use the preverbal marker information as well as control verb information in the processing of control sentences.

6.1 Pro licensing by the use of preverbal markers

We claim that the antecedent of pro was immediately assigned without delay in two ways. First, the reading times at control verbs were faster in the control-creating marker conditions than in the control-neutral marker conditions. We think that readers had already started assigning the referent for pro upon the recognition of control-creating markers, leading to the facilitation of the integration of upcoming control verbs. In contrast, when control preverbal markers were neutral, its semantic information did not
elicit enough bias to determine the particular referent for pro. Thus, when control verbs appeared, readers had more processing cost in the integration of the verbs into sentences because readers also needed to determine pro’s referential identity. Also, Korean’s strong bias to an object as an antecedent for pro might lead to the immediate integration of control verbs, like Japanese’s strong bias to objects led to provisional reading of control sentences (Witzel & Witzel, 2011; 2016). Similar to the claim by Witzel and Witzel, Korean readers might have assigned objects as antecedent for pro before readers encountered control verbs, and such a bias became clear only when control markers were neutral, that is, when there was no provisional information from control markers. These results supported the predictive view that the semantic information of control verbs as well as that of preverbal markers play an important role in determining pro’s controller and completing control interpretation.

Our data was not in line with the verb-centered view in theoretic linguistics. From the linguistic point of view, the antecedent of pro should be determined only by the semantic properties of control verbs. Under this view, since readers select the complement of control verbs and assign the antecedent for pro in the complement clause when encountered with the control verbs regardless of either when control-neutral or control-creating markers were used, there should be no difference in processing load across these conditions. That being said, under this view, the processing cost (e.g., reading times) at control verbs should not be modulated by the presence of the licensing information of the preverbal markers. The claim based on the theoretic linguistic view was not supported by our online behavioral results. Second, in addition to the reading times results, rejection rates were higher at control verbs in the control-neutral marker condition (20%) than in the control-creating marker matched condition (13%) (coefficient = -.28, S.E. = .002, z = -144.4, p < .001). The higher rejections at control-neutral marker condition indicated that readers felt more implausibility in the integration of control verbs into sentences. Importantly, the rejection differences were not due to the differences of sentence naturalness because our offline norming results revealed that the naturalness of the control-neutral marker conditions was equated with that of the matched control-creating marker conditions. Our online results suggested that assigning the
antecedent of pro might have been initiated even at the moment when the preverbal markers appeared in the control-neutral marker condition. Under the predictive processing point of view, the control-neutral marker provided the information on the embedded clause boundary in which an unassigned pro would be present. Thus, although the preverbal marker did not have the semantic information to determine ‘which’ controller, readers might try to create the most complete representation by provisionally assigning some matrix argument as an antecedent of pro. By doing so, they would be able to reduce associated processing cost. Given the assumption that the provisional assignment could be made without reference to the semantic information of forthcoming control verbs, there were higher possibilities that the provisional assignment needed being revised when control verbs actually occurred in the cases that preverbal markers were neutral than in the cases that preverbal markers were control-creating. These rejection rates at control verbs also supported our claim that the processing of the Korean control sentences is predictive.

To be brief, our results showed that readers actively and immediately used available local information to determine the antecedent of pro. Given the fact that control markers appear before control verbs arrive, it is economic and rational decision to initiate the early assignment of an antecedent for pro.

6.2 Object bias in the predictive processing in Korean control sentences

Recall that Japanese readers had object bias in their processing of Japanese control sentences (Sakamoto, 1995, 1996; Witzel & Witzel, 2011). Consistent with the findings observed in the Japanese studies, we also found object bias in the processing of Korean control sentences both with the control creating and with the control neutral markers. Yes reading times and rejection rates showed that matrix objects were much more preferred as an antecedent for pro than matrix subjects. Like Japanese readers, Korean readers strongly preferred assigning matrix objects as antecedents of pro and then developed their expectation on what kind of information was likely to be coming up. When the verb information turned out to violate their expectation, readers might reject sentences more even though sentences were grammatical
and plausible. What fundamentally underlies object bias in Korean? What would make Korean readers favor assigning a matrix object as an antecedent of pro relative to a matrix subject?

Two existing models are suitable to account for our results of object bias. One approach is Most Recent Filler Strategy (Frazier, Clifton, & Rall, 1983; Clifton and Frazier 1986). According to this strategy, pro resolution is completed in a way that less distant fillers are initially selected as antecedents of pro and then their assignment is checked when control verbs arrive. For us, since matrix objects are the less distant filler for pro than matrix subjects, readers initially assign the linearly closer filler, matrix objects, as potential antecedents of pro. This linear distance model (i.e., Most Recent Filler strategy) could be one way to explain the data of Korean control sentences.

An alternative approach to explain readers’ object bias is to take the claim by the Structural Distance Model (i.e., phrase structure hypothesis by O’Grady (1997). In this model, the degree of structural complexity increases as the structural distance between a filler and a gap gets longer. The degree of structural distance proportional to that of structure’s complexity is computed by counting the number of intervening labelled brackets of XP categories between a filler and a gap, as it is generalized in (11) (O’Grady, 1997:137).

\[(11) \text{A structure’s complexity increases with the number of XP categories (S, VP, etc.) between a gap and the element with which it is associated.}\]

(O’Grady, 1997:136)

For example, when a control-neutral marker is used in (12a), the complexity rating associated with the subject filler (i.e., *Joe*) is greater than that associated with the object filler (i.e., *Mary*). The same is also true even when a control-creating marker is used in (12b). Likewise, object bias could emerge because the hierarchical distance between the object filler and the gap is shorter than the distance between the subject filler and the gap, which in turn would yield lower degree of processing load.
Joe swore/persuaded to Mary [pro to stop smoking] \\
\textbf{Complexity rating} = 4 (subject filler), 2 (object filler)

Joe swore/persuaded Mary [pro to stop smoking] \\
\textbf{Complexity rating} = 3 (subject filler), 1 (object filler)

Not only the linear distance model but also the structural distance model provide underlying mechanism on why matrix-objects, rather than matrix-subjects, should be favored on pro referentiality.

7. Conclusion

This study aimed to examine whether an antecedent for pro in Korean control sentence is predictively determined prior to the recognition of control verbs and which constituent is favored as an antecedent for pro. We observed that 1) control verbs were read more slowly when control-neutral markers were used than when control-creating markers were used. 2) The words prior to control verbs were taken longer to read when the antecedent for pro corresponded to matrix subjects than matrix objects. 3) Control verbs were rejected more when control-neutral markers were used and pro had to be co-indexed with matrix subjects than when control-creating markers were used and pro had to be co-indexed with matrix objects. Taken together, we claim that Korean readers immediately used the information of preverbal markers and control verbs associated with pro resolution.
Acknowledgements

We appreciate anonymous reviewers for their precious comments. Our acknowledgment also goes to Sang-Hee Kim for her helping this research. With regard to any inquiries or requests on this paper, please correspond to the corresponding author.α

References

Clifton, C., Traxler, M. J., Mohamed, M. T., Williams, R. S., Morris, R. K., & Rayner, K. 2003. The use of thematic role information in parsing: Syntactic

α Corresponding author. Gachon Liberal Arts Colleague, Gachon University, 1342 Sungnam-ro, Sujeong-gu, Gyeonggi-do, 13120, Korea, oaktreepark@gmail.com (Hongoak Yun).


Nam, Yunju, Hong, Upong, & Yun, Hongoak. 2016. Toward a unified account of the length effects in sentence production and sentence comprehension. Language Research 33(1), 135-175.


Rayner, Keith, Kathy S. Binder, Jane Ashby, & Alex&er Pollatsek. 2001. Eye
movement control in reading: word predictability has little influence on initial
introduction (Vol. 92). Stanford, CA: Center for the Study of Language and
Information.
the transparency hypothesis. Fukuoka: Kyushu University Press.
of empty subject sentences. In M. Nakayama (Ed.), Sentence processing in
East Asian languages, 189–221. Stanford: CSLI.
Language and Cognitive Processes 1, 227–245.
dependencies: An eye-tracking study. Journal of Memory and Language 35,
454–475.
Trueswell, J. C., Tanenhaus, M. K., & Garnsey, S. M. 1994. Semantic influences on
parsing: Use of thematic role information in syntactic ambiguity resolution.
In H. Yamashita, J. Packard, & Y. Hirose (Eds.), Processing & producing head-
final structures, 23–47. London: Springer.
Witzel, J. D., & Witzel, N. O. 2016. Incremental sentence processing in Japanese:
A maze investigation into scrambled and control sentences. Journal of
Psycholinguistic Research, 45, 475-505.
Yang, Dong-Whee 1985. On the integrity of control theory. In Berman, S., J.-
W. Choe, & J. McDonough (eds.): Proceedings of the North East Linguistic
Society. Amherst, Mass.: GLSA, University of Massachusetts 15, 389-408.
Yun, Hongoak, & Hong, Upyong.2014. Role predictability and role filler
Yun, Hongoak, Nam, Yunju, Yoo, Duckgeun, & Hong, Uyong. 2015. The effect of
role shifting and expectation in the processing of center-embedded relative
clauses in Korean. Language Research, 32(2), 313-353. DOI: 10.17250/
khisli.32.2.201508.002