Non-subject Antecedent Potential of *Caki* in Korean

CHUNG-HYE HAN  
*Simon Fraser University*

DENNIS RYAN STOROSHENKO  
*Yale University*

1. Introduction

Traditionally, *caki* has been described as subject-oriented, always taking a clausal subject as its antecedent (Lee, 1973; Chang, 1977). This view predicts that in (1), only the matrix subject *Yuli* can be the antecedent of *caki* in the embedded clause.

(1)  

| Yuli-nun | Swuni-lopwuthe  
| caki-ka  
| iki-lke-lako |

Yuli-Top Swuni-from self-Nom win-Fut-Comp  

| tul-ess-ta. |

| hear-Past-Decl |

‘Yuli heard from Swuni that self would win the race.’

However, the view that *caki* has the potential for non-subject antecedents has steadily been gaining ground (Yoon, 1989; Cho, 1994; Kim, 2000; Sohng, 2004; Madigan, 2006). Yoon (1989), for instance, describes *caki* as logophoric, being sensitive to a logophoric centre (Sells, 1987), described in Büring (2005) as the “source of information.” So, according to Yoon, in (1), as the matrix indirect object *swuni* is the source of information, it can be the
antecedent of *caki*. Sohng (2004), on the other hand, without making any reference to logophoricity, observes that though the subject is preferred to be the antecedent of *caki*, the object can also be an antecedent, and takes examples such as (1) to be ambiguous. A moderated view can be found in Kim (2000), proposing that ‘potential antecedents for *caki*’ are ranked according to a thematic prominence hierarchy, where ‘potential antecedents’ are defined as 3rd person nominals that are higher in the thematic hierarchy than *caki*, and that any given instance of *caki* will prefer an antecedent highest on that hierarchy. While subjects are highly-ranked in this system, it leaves open the possibility for non-subject antecedents, and leaves *caki* free to corefer with any nominal when no potential antecedent is present.

In an experiment that combined a forced-choice task and eye-tracking, Han et al. (In press) found that speakers predominately choose the matrix subject over the matrix indirect object as the antecedent for embedded *caki*, even when the discourse context makes the object reference salient. This result re-opens the question of whether a non-subject is a grammatically possible antecedent of *caki* at all. But, what Han et. al. found might be a function of a verb effect and/or a task effect arising from the experimental design rather than from any properties of *caki* itself. In their experiment, all test sentences contained *malha-* (say) as the matrix verb, as in (2). So, their findings could be restricted to *say*-sentences, and may not be representative of the behaviour of *caki* in general.

(2) Jongwu-ka Yuli-eykey chilphan yeph-eyse caki-ka
    Jongwu-Nom Yuli-Dat blackboard beside-at self-Nom
    sihem-ul cal chi-ess-tako malha-n-ta.
    test-Acc well take-Past-Comp say-Pres-Decl
    ‘Jongwu says to Yuli beside the blackboard that self did well on the test.’

Furthermore, in Han et. al.’s experiment, participants had to choose between the subject and the object as the antecedent of *caki*. So, even if the object antecedent interpretation was possible, speakers could have chosen the subject because that is what they prefer; a forced-choice task may not necessarily reveal ambiguity or even the presence of a second potential antecedent.

In this paper, we present two experiments on the antecedent potential of *caki*, a Likert scale experiment and a Truth Value Judgment task (TVJT) experiment, controlling for the potential verb and task effects. In the Likert scale experiment, we found that an object is indeed a grammatically possible antecedent for *caki*, though speakers showed preference for the subject antecedent. In the follow-up TVJT experiment, we found that the preference for the subject interacts with the preference for a logophoric centre, the source of information. Speakers are more likely to accept an object antecedent inter-
interpretation of \textit{caki} if the object is the source of information. But a subject antecedent interpretation of \textit{caki} is available regardless of the logophoric status of the subject.

This paper is organized as follows. In section 2, we describe the Likert scale experiment, and present the findings. The TVJT experiment and its findings are described in section 3. We conclude in section 4 with implications of our findings and a brief discussion of future work.

2. Likert Scale Experiment

It is generally assumed that \textit{caki} is 3rd person, and as such, it can only take a 3rd person nominal as its antecedent (Sohng, 2004; Kim and Madigan, In press).\footnote{Caki also acts as a term of intimate address, with a 2nd person reading. As this most often occurs as a vocative in spoken discourse with no sentential antecedent, we do not consider this usage to interfere with the judgments of \textit{caki} in argument positions requiring an antecedent.} This means that if a bi-clausal sentence contains \textit{caki} in the embedded subject with two nominal arguments in the matrix clause, one 3rd person and the other 1st person, then the 3rd person nominal is the only possible antecedent of \textit{caki}. Exploiting this fact, we designed a scale experiment to address the following research question.

(3) Research question in the Likert scale experiment

How readily do speakers allow a 3rd person object to be the antecedent of \textit{caki} when the subject is 1st person?

2.1. Task

Participants were presented with a short text followed by a test sentence describing the text on a computer screen. They were then asked to rate the appropriateness/naturalness of the test sentence using a 7 point scale, where 1 is not appropriate/natural and 7 is appropriate/natural, by pressing a button on the keyboard.

A screen shot of an example trial is given in Figure 1. In the short text, participants read that Swuni and I were talking in the kitchen, about to make ramyen, and I said that Swuni always made ramyen well. They then read the test sentence (shown in the middle of Figure 1 in a smaller, lighter text), glossed and translated in (4). The last line at the bottom of the display presents the scale.

(4) Na-nun Swuni-eykey caki-ka lamyen-ul hangsang cal
    1-Top Swuni-to self-Nom ramyen-Acc always well
    kkulhi-n-tako malha-yess-ta.
    make-Pres-Comp say-Past-Decl
    ‘I told Swuni that self always made ramyen well.’
If the participants allow the object interpretation of caki in sentences such as (4), they will judge such sentences to be appropriate/natural.

2.2. Design, participants and procedure

The experiment had two within-subjects factors with two levels each: Person (the matrix arguments of test sentences are both 3rd person, or one of them is 1st person) and Antecedent (the text is compatible with the subject or the object antecedent interpretation). Crossing the factors thus gives rise to four different conditions, summarized in Table 1. The example trial described in subsection 2.1 belonged to 1st-person/Object condition where the matrix subject argument of the test sentence is 1st person, and the text is compatible with the object antecedent interpretation.

Twenty native speakers of Korean residing in Vancouver, Canada were tested. All had been educated in Korea until at least the age of 14 and all had lived at most a total of 12 months outside of Korea at the time of testing. Participants were first introduced to the task with four practice trials. They were then given 16 test trials (four trials per condition) along with 16 filler trials.
TABLE 1 Likert scale experimental design

<table>
<thead>
<tr>
<th>1st-person/Subject</th>
<th>Test sentences containing matrix 3rd person subjects and 1st person objects are presented in contexts compatible with the subject antecedent interpretation of embedded caki.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1st-person/Object</td>
<td>Test sentences containing matrix 1st person subjects and 3rd person objects are presented in contexts compatible with the object antecedent interpretation of embedded caki.</td>
</tr>
<tr>
<td>3rd-person/Subject</td>
<td>Test sentences containing matrix 3rd person subjects and 3rd person objects are presented in contexts compatible with the subject antecedent interpretation of embedded caki.</td>
</tr>
<tr>
<td>3rd-person/Object</td>
<td>Test sentences containing matrix 3rd person subjects and 3rd person objects are presented in contexts compatible with the object antecedent interpretation of embedded caki.</td>
</tr>
</tbody>
</table>

trials. Experimental trial ordering was fully randomized, with a unique order generated for each participant.

2.3. Test sentences

We constructed four test sentences for each of the four conditions. In each condition, two test sentences contained malha- (say) as the matrix verb, and the other two contained tut- (hear). Examples from the four conditions are given in (5)-(8) ((6) repeated below from (4)).

(5) 1st-person/Subject condition

Swuni-nun na-eykey caki-ka iki-lke-lako malha-yess-ta.
Swuni-Top I-to self-Nom win-Fut-Comp say-Past-Decl
‘Swuni told me that self would win the race.’

(6) 1st-person/Object condition

Na-nun Swuni-eykey caki-ka lamyen-ul hangsaŋ cal
I-Top Swuni-to self-Nom ramyen-Acc always well
kkulhi-n-tako malha-yess-ta.
make-Pres-Comp say-Past-Decl
‘I told Swuni that self always made ramyen well.’
2.4. Findings and discussion

Mean ratings from all participants on each of the Person/Antecedent combinations are shown in Figure 2. A repeated-measures ANOVA revealed a main effect of Antecedent ($F(1,19) = 56.461, p < .001$) and an interaction between Person and Antecedent ($F(1,19) = 5.865, p = .026$). This means that speakers are significantly more likely to rate sentences with a subject antecedent higher than those with an object antecedent, but speakers’ preference for the subject
The antecedent over the object antecedent is significantly greater in sentences with 3rd person arguments than those that contained a 1st person argument in the matrix clause.

Furthermore, pairwise comparisons of means across conditions with Bonferroni adjustment revealed that while the rating in the 1st-person/Object condition is significantly lower than the rating in the 1st-person/Subject condition \((t(1,19) = 5.294, p < .001)\), it is significantly higher than the rating in the 3rd-person/Object condition \((t(1,19) = 3.363, p = .003)\).

In sum, when a test sentence contains a 1st person subject and a 3rd person object, speakers readily allow the object antecedent interpretation, as evidenced by the higher rating in the 1st-person/Object condition than in the 3rd-person/Object condition \((4.03 \text{ vs. } 2.88)\). But given a choice between the subject and the object antecedent interpretation, when both the matrix subject and the matrix object are 3rd person, the object antecedent interpretation is dispreferred \((5.71 \text{ vs. } 2.88)\). This result supports that an object argument is a grammatically possible antecedent of caki, although there is a preference for the subject antecedent.

It should be noted that the 1st-Person/Object condition of this experiment constructs the exceptional scenario in which, according to Kim’s (2000) definition, caki actually has no potential antecedent: the matrix subject is not 3rd person, and the matrix indirect object is thematically less prominent than the embedded subject. That caki’s interpretation as having a non-subject antecedent is significantly better when there is no other possible antecedent can be interpreted as an expected outcome of this exceptional circumstance. Kim’s approach is not entirely predictive of the results here though, as her system of preferential ranking of potential antecedents predicts a categorical distinction between the two 3rd-person conditions. While we do find a significant difference between these two conditions, a rating of 2.88 may be seen as unexpectedly high for a structure which ought to be completely ruled out.

We have not yet provided any discussion or analysis of the sub-cases within each condition which made use of the matrix verb malha- (say) versus those with tut- (hear). With only two tokens of each per condition, there is insufficient data to make any strong statistical claims, but we do note anecdotally that higher ratings for the object antecedent conditions tended to be more frequent with hear than with say. This ties back to Yoon’s (1989) proposal that caki interpretation may be sensitive to logophoricity. We thus move on to the TVJ experiment which directly tests whether or not caki is sensitive to a logophoric bias.
3. Truth Value Judgment Task (TVJT) Experiment

The purpose of the TVJT experiment is to test whether the subject preference observed in the Likert scale experiment interacts with another factor, the preference for the source of information, or logophoric centre. For this purpose, we manipulated the type of the matrix verb, *malha-* (say) vs. *tut-* (hear), as an independent variable. In *say*-sentences, as in (9), the matrix subject is the source of information, while in *hear*-sentences, as in (10) (repeated below from (1)), the matrix indirect object is the source of information.


Yuli-Top Swuni-to self-Nom win-Fut-Comp say-Past-Decl

‘Yuli said to Swuni that self would win the race.’

(10) *Yuli-nun Swuni-lopwuthe [caki-ka iki-lke-lako]*

Yuli-Top Swuni-from self-Nom win-Fut-Comp

*tul-ess-ta.*

hear-Past-Decl

‘Yuli heard from Swuni that self would win the race.’

Using this fact, we designed a TVJT experiment that addresses the research question in (11).

(11) Research question for the TVJT experiment

Can the subject preference observed in the scale experiment be attributed to the preference for the source of information?

3.1. Task

Participants were presented with a combination of visual and aural stimuli. While watching a screen with two characters (one male and one female) on a computer screen, they heard a scene setting statement by a narrator. They then heard a brief conversation by the two characters, and then a test sentence presented by the narrator. For instance, while watching the picture in Figure 3, participants heard a narrator state that Chelswu (a typical male name) and Swuni (a typical female name) are at a swimming pool and they are going to race. They then heard a male voice stating “You belong to the school swimming team,” a female voice stating “So I practice swimming for an hour every day,” and then the male voice stating “You will win the race.” A test sentence, as in (12), is then presented by the narrator. The participants’ task was to answer whether the test sentence is True or False in the context of the conversation between characters.

(12) *Swuni-ka Chelswu-lopwuthe taiping tay yep-eyse*

Swuni-Nom Chelswu-from diving board beside-at

*caki-ka iki-lke-lako tul-ess-ta.*

self-Nom win-Fut-Comp hear-Past-Decl
‘Swuni heard from Chelswu beside the diving board that self would win.’

If the participants allow the subject antecedent interpretation of caki in hear-sentences, then they will say ‘True’ in this trial. But if they only allow the object antecedent interpretation of caki in hear-sentences, then they will say ‘False’ in this trial.

3.2. Design, participants and procedure

The experiment had two within-subjects factors: Verb (the matrix verb of the test sentence is malha- (say) or tut- (hear)) and Antecedent (the dialogue is compatible with the subject or the object antecedent interpretation of caki). Crossing the factors again gives rise to four different conditions, as summarized in Table 2. The example trial described in subsection 3.1 was presented in the condition in which the matrix verb of the test sentence is tut- (hear) and the dialogue is compatible with the subject antecedent interpretation.

Twenty native speakers of Korean, different from the Likert scale exper-
3.3. Test sentences

We constructed 16 say-sentences, as in (13), and 16 hear-sentences, as in (14) (repeated below from (12)).

    win-Fut-Comp tell-Pres-Decl
    ‘Chelswu tells Swuni beside the diving board that self will win.’

    self-Nom win-Fut-Comp hear-Past-Decl
    ‘Swuni heard from Chelswu beside the diving board that self would win.’

Each sentence was presented twice: once with a dialogue compatible with the subject antecedent interpretation, and once with a dialogue compatible with the object antecedent interpretation. The test sentences were balanced with equal number of male and female subjects.
3.4. Findings and discussion

Mean percent acceptance from all participants on each of the Verb/Antecedent combination are shown in Figure 4. A repeated-measures ANOVA revealed a main effect of Antecedent ($F(1,19) = 40.862, p < .001$) and an interaction between Verb and Antecedent ($F(1,19) = 25.415, p < .001$). That is, speakers are significantly more likely to accept sentences with a subject antecedent than those with an object antecedent, but speakers’ preference for the subject antecedent is significantly higher in *say*-sentences than in *hear*-sentences. Pairwise comparisons of means across conditions with Bonferroni adjustment revealed that the acceptance rate in Say/Subject condition is significantly higher than the acceptance rate in Say/Object condition ($t(1,19) = 12.107, p < .001$), but the difference between the Hear/Subject condition and Hear/Object condition is not significant.

These findings suggest that the subject preference observed in the scale experiment cannot be attributed to the preference for the source of information, and that both the preference for the subject and the preference for the source of information are at work in the speakers’ choice of the antecedent of *caki*. In *say*-sentences, both the source and the subject preference converge on the subject antecedent accounting for the high 98% acceptance rate in Say/Subject condition and the low 15% acceptance rate in Say/Object condition. But in *hear*-sentences, the two preferences diverge: the subject preference points to the matrix subject, and the source preference points to the ma-
The fact that the acceptance rates in both Hear/Subject and Hear/Object conditions hover around 60% strongly suggests that the two preferences are in competition, interacting with each other.

This result is contrary to both the view that caki is subject-oriented, as in Lee (1973) and Chang (1977), and the view that caki is simply logophoric, as in Yoon (1989). It also provides a more robust counterexample to Kim’s (2000) prominence hierarchy, as she would predict there to be no difference between the say and hear cases. In all conditions, the only potential antecedent for caki under her definition should be the matrix subject, as the matrix indirect object is less prominent than the embedded subject. However, the results in the hear condition show that this is not the case, and that a manipulation based on logophoricity can make a thematically less-prominent antecedent more acceptable. The non-categorical result in the hear condition points to a more complex system of competing factors rather than a system with an overarching preferential antecedent ranking according to a single factor such as a thematic prominence hierarchy.

4. Conclusion
Echoing the results in Han et al., and hearkening back to the earliest literature on caki, both the Likert scale and TVJT experiments demonstrate a preference for caki to have a subject antecedent. However, our experiments also demonstrate that this preference is in competition with other factors, leading speakers to accept instances of caki with non-subject antecedents. As a result, we are left with a more nuanced view of how caki must be treated. Any formal analysis of caki must be flexible enough to admit subject and non-subject antecedents; where there is the potential for ambiguity, other factors such as agreement and logophoricity militate against certain readings, though the dis-preferred option must remain available in principle. Rather than aligning with single-factor preferential antecedent ranking treatments of caki, our overall findings are compatible with the multiple-constraints approach to anaphor resolution in that interpretation of anaphors is determined by multiple interacting constraints (Kaiser et. al. 2009, and references therein). Future work in this area will seek to identify further factors influencing the interpretation of caki, using on-line methodology such as eye-tracking. Using experimental methods which allow us to identify the effects of biases as the sentence unfolds, we can find further evidence that multiple such effects exist.

Acknowledgments
We thank Jennifer Hinnell and Samantha Kim for their help with data collection, and Ilkyu Kim for assistance in the piloting stage of the scale experiment. This research is partially funded by SSHRC grant #410-2007-2169 to Han.
References


