An experimental investigation of on-line and off-line binding properties of Korean reflexives—evolving grammar of multiple reflexives

[Category: formal syntax, psycholinguistics]

Korean has a rich inventory of reflexives, which raises a question as to how Korean speakers differentiate the different reflexives. We conducted an experimental study in order to investigate the syntactic dimension along which these reflexives are known to vary--the binding distance of reflexives (local vs. long-distance (LD) binding)--focusing on *caki*, *casin* and *caki-casin*.

Few previous studies examined all three reflexives, and those that did were based either on intuitions of researchers (e.g., Yoon 1989) or a corpus search (Kang 1998). The consensus in this literature seems to be that *caki* has a strong LD-binding preference, while *caki-casin* has a strong local binding preference. *Casin* is judged to allow both local and LD-binding, but with a preference for local binding. A similar assessment has been made for the three reflexives of Japanese in Katada (1991). The present study employed both an on-line and off-line measures to investigate the binding distance preference of the reflexives. To the best of our knowledge, our study is the first to systematically investigate how all three of these reflexives are processed on-line (cf. Choi and Kim 2007 examined two reflexives).

Sixteen native speakers of Korean participated in an online eye-tracking task and an offline antecedent identification task. For the eye-tracking task, participants listened to bi-clausal sentential stimuli (appendix 1) while looking at pictures depicting the referents of several NPs that are mentioned in the sentence (appendix 2) (the design was modeled on Clackson, Felser & Clahsen, 2011). In the experimental sentences, the matrix subject and the embedded subject served as potential antecedents of the reflexives (*caki*/ *casin*/ *caki-casin*), which were always the embedded object. A norming test was conducted so that the sentences were equally compatible with the local and long-distance interpretations. Therefore, if a preference for one of the two binding interpretations emerges, it can be attributed to the reflexives. As a way to examine whether such preferences exist, the proportions of fixations to the two subject pictures were compared (matrix subject pictures vs. embedded subject pictures) from the onset of the reflexives. For the antecedent identification task, the experimental sentences of the eye-tracking task were presented to the participants after the completion of the on-line task in written format. Participants were asked to judge whether the reflexive in the sentence can refer to the matrix subject (e.g., for sentences in (1) - "Can *caki/casin/caki-casin* refer to Peterpan?") or to the embedded subject ("Can *caki/casin/caki-casin* refer to Shrek?"). The eye-tracking task was expected to reveal the moment-by-moment process of antecedent search, whereas the antecedent identification task investigated the possible final interpretations speakers assigned to the sentences.

The eye-tracking task results were statistically analyzed within each reflexive condition at two separate time windows − the 'reflexive+adverb' window (reflexive onset - before embedded verb onset) and the 'embedded verb' window (embedded verb onset - before matrix verb onset) (appendix 3). Paired-samples *t*-tests conducted on the fixation proportions averaged in each window revealed that for *caki* and *casin*, the proportions of fixations to the matrix subject and to the embedded subject were not significantly different at reflexive+adverb (*caki*: $t1(15)=-.574$, $p>.05$, $t2(20)=-.812$, $p>.05$; *casin*: $t1(15)=.480$, $p>.05$; $t2(20)=-.168$, $p>.05$) nor at the embedded verb (*caki*: $t1(15)=-1.477$, $p>.05$, $t2(20)=-1.469$, $p>.05$; *casin*: $t1(15)=-1.165$, $p>.05$, $t2(20)=-1.296$, $p>.05$). For *caki-casin*, however, the proportion of fixations to the embedded subject was significantly higher than the proportion of fixations to the matrix subject, at both windows (reflexive+adverb: $t1(15)=-2.999$, $p<.01$, $t2(20)=-4.012$, $p<.01$; embedded verb: $t1(15)=-2.844$, $p<.05$, $t2(20)=-3.092$, $p<.01$). For the antecedent identification task, responses that indicated that the readings were 'possible' were assigned 1 and the 'impossible' responses were assigned 0 (appendix 4). Paired-samples *t*-test comparing the average scores of matrix subject-binding interpretation and embedded subject-binding interpretation revealed that the scores for local and long-distance binding were not significantly different for *caki* ($t1(15)=.000$, $p>.05$, $t2(20)=.438$, $p>.05$). For *casin*, the embedded subject-binding interpretation received a significantly higher score than the matrix subject-binding interpretation ($t1(15)=-2.649$, $p<.05$, $t2(20)=-2.642$, $p<.05$), while for *caki-casin*, the difference was even more highly significant ($t1(15)=-8.676$, $p<.001$, $t2(20)=-11.204$, $p<.001$).

What emerges in online results is a two-way contrast between *caki/casin* and *caki-casin*, and in offline results the significant contrast was between *caki* and *casin*/*caki-casin* (though *caki-casin* showed a stronger local binding bias than *casin*). These patterns are different from what has been assumed thus far. It seems that in moment-by-moment online processing of the LDAs *caki* and *casin*, native Korean speakers postpone assigning a determinate interpretation (but see Choi & Kim, 2007, on embedded verb bias in differentiating *caki* and *casin*). For *caki-casin*, however, they immediately seem to assign the local-binding interpretation. What we failed to find is an early, determinate, LD-binding preference for *caki*, contrary to the claims of the previous literature. In the offline results, a contrast between *caki* and the other reflexives emerged, but again *caki* did not display a marked preference for LD-binding.

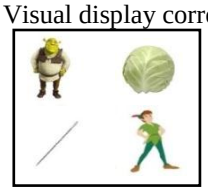
These results seem to reflect the evolving grammar of reflexives in Korean, with *casin* encroaching rapidly on the realm of *caki* (Kim and Yoon 2008). The effect of *caki* being supplanted by *casin* is seen in both online (where the two are not distinguished) and offline (where *caki* does not display a determinate LD-binding preference, contrary to expectations) results.

Appendix

(1) Example sentence (21 experimental items distributed to 3 lists + 59 fillers (Latin Square Design))

Peterpan-un [Shrek-i pyocokhan panul-lo **caki/casin/caki-casin-(l)ul** silswulo cillesstako] malhaysssupnita.
P-top S-nom sharp needle-ins self-acc by.mystake prick-comp said
'Peterpan said Shrek pricked **self** with a sharp needle by mistake.'

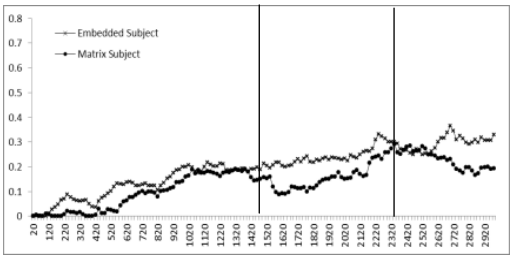
(2) Visual display corresponding to (1)



(3) Eye-tracking task results

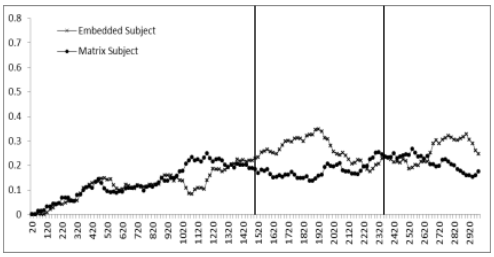
* y-axis: proportion of fixations to the two pictures corresponding to the matrix subject and the embedded subject
x-axis: time in ms from reflexive onset

< *caki* >



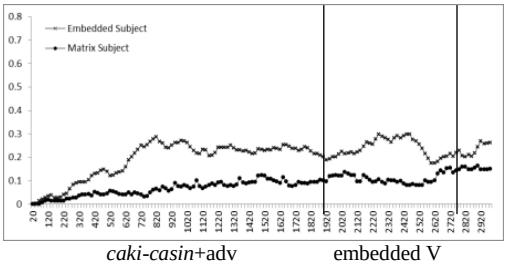
caki+adv embedded V

< *casin* >



casin+adv embedded V

< *caki-casin* >



caki-casin+adv embedded V

(4) Antecedent identification task results (average response scores; 'possible' = 1, 'impossible' = 0)

	caki	*casin*	*caki-casin*
Matrix subject-binding interpretation	0.74	0.58	0.24
Embedded subject-binding interpretation	0.74	0.82	0.96