

A Bound-Variable Account of the Korean Anaphor *caki*¹

Dennis Ryan Storoshenko
Simon Fraser University

1. Introduction

It has been argued that the binding of long-distance anaphors can be reduced to a series of local dependencies by cyclic head movement of the reflexive to a matrix I position (Chomsky 1986; Cole et al. 1990). As a consequence, these reflexives are predicted to be subject-oriented. In Korean, the long distance anaphor *caki* is not strictly subject oriented, but able to be bound by local and non-subject arguments:

- (1) John₄-i Mary₅-eykey Tom₆-i caki_{4,5,6}-lul coaha-n-ta-ko malha-yess-ta.
John-NOM Mary-DAT Tom-NOM SELF-ACC like-PRES-DECL-COMP say-PST-DECL
'John told Mary that Tom likes self.' (Sohng 2003, ex 11a)

The fact that *caki* can be bound locally as well as at a distance, shown in (1), calls into question the need for postulating LF movement; if *caki* can be bound locally, why should LF movement of *caki* be postulated?

Caki can also be pluralised using the *-tul* morpheme:

- (2) Bill₅-kwa-Mary₆-ka caki-tul₅₊₆-ul piphan-ha-yess-ta.
Bill-and-Mary-NOM SELF-PL-ACC criticise-do-PST-DECL
'Bill and Mary criticised selves.'

Here, *caki-tul* can be interpreted as the set comprised of Bill and Mary. However, the interpretation of *caki-tul* can vary, depending on the nature of its antecedent. This introduces a degree of context sensitivity which is at odds with strict notions of binding.

The goal of this paper is two-fold. First, I will propose that *caki* can be treated as an *in-situ* bound variable, avoiding the need to postulate LF movement of *caki*. Secondly, I will examine the variable interpretations of the plural *caki-tul*, establishing a division of labour between bound *caki* and the plural morpheme *-tul*. Ultimately, it will be shown that both singular and plural *caki* can be captured using one analysis to cover all binding domains. The paper will proceed as follows: in section two, all the relevant facts will be presented. Section three will develop the argument for the treatment of *caki* as a bound variable. In section four, a denotation of *-tul* will be proposed and integrated into the proposal from section three. The paper concludes with a brief discussion of future work.

2. The facts

In this section, the full set of data which needs to be accounted for is presented. I begin with singular *caki* and move on to plural *caki-tul*.

As mentioned above, *caki* can be bound locally or long-distance. In fact, *caki* is even more versatile. *Caki* does not require a c-commanding antecedent within the same sentence:

¹ I would like to thank my supervisor Dr. Chung-hye Han for her guidance and advice in pursuing this research. Thanks are also owed to my colleague Lan Kim, for her patience with my questions about *caki*. Finally, a thank you to the audience at WCCFL for their helpful comments and questions. All errors are my own.

- (3) Na-nun Suni₄-eykey chayk-ul pilye cwu-ess-ta. Kulendey sasil ku chayk-un caki₄
 I-TOP Suni-DAT book-ACC lend give-PST-DECL and yet in fact that book-TOP SELF
 oppa-ka ceney nay-key pilye cwun kes ita
 elder brother-NOM before me-DAT lend give thing be
 ‘I lent a book to Suni. But the fact is that self’s brother had lent it to me before.’
 (Kim 2000, ex 2b)

In (3), *caki* is bound by *Suni* from the previous sentence. Assuming that c-command is a sentential relation, *caki* is not c-commanded by *Suni*, and no LF movement is available to resolve this problem. That *caki* can be bound in this manner has led to claims that it is in fact discourse-bound. Furthermore, *caki* can be bound by a non-c-commanding antecedent within the same sentence:

- (4) Suni₅-uy sinpal-un caki₅-uy pal-pota hwelssin kuta.
 Suni-GEN shoes-TOP self-GEN foot-than a lot big
 ‘Suni’s shoes are a lot bigger than self’s feet.’ (Kim 2000, ex 2a)

Examples such as this have called into question the need to treat *caki* in terms of binding. If it does not need a c-commanding antecedent, why then should it be considered bound? I return to this question in the next section.

I now turn my attention to *caki-tul*. As noted above, where it has a plural antecedent, *caki-tul* can be interpreted as having the same reference as that antecedent. However, Madigan and Yamada (2006) note that where *caki-tul* has a singular antecedent, inclusive reference results:

- (5) John₄-i caki-tul_{4+α}-ul sokay-ha-yess-ta.
 John-NOM SELF-PL-ACC introduce-do-PST-DECL
 ‘John introduced selves.’

Here, *caki-tul* refers to John, plus some other contextually salient group, and this inclusive reference reading is readily available to native speakers.

Another possibility emerges where there are two singular antecedents over a plural *caki* as in (6):

- (6) John₄-un Mary₅-eykey [caki-tul₄₊₅-i iki-lke-la-ko] malha-yess-ta.
 John-TOP Mary-DAT SELF-PL-NOM win-FUT-DECL-COMP say-PST-DECL
 ‘John told Mary that selves would win.’ (Huang 2000, ex 2.179)

Huang (2000) uses this example as evidence that *caki* demonstrates split antecedence. That is, *caki-tul* finds its reference from a composite of the matrix subject and dative arguments.

Finally, there are cases where *caki* does not appear to require an antecedent at all:

- (7) Caki-tul₂₊-i chakhay.
 you-PL-NOM good
 ‘You all are good.’

Though plural, this sentence is equally grammatical in its singular form. Sohng (2003) refers to this as inherent reference, wherein a *caki* without an antecedent is inherently second person. This can also be used deictically, and two antecedent-less instances of *caki* can be used to refer to different people given enough explicit pointing:

- (8) Caki₂-ka caki_{2’/2’’}-lul coaha-n-ta.
 SELF-NOM SELF-ACC like-PRES-DECL
 ‘You like yourself.’
 ‘You like you.’ (Different addressees)

The sentence in (8) is actually ambiguous, with one reading being that the object *caki* is bound by the inherently referential subject *caki*. The second reading indicated is the deictic usage, with each *caki* indicating a different second person addressee. Without an antecedent, *caki* is also used as a term of

intimate address. As these uses of *caki* are substantially different from the other cases, antecedent-less *caki* will not be considered further in this paper.

To summarise the facts presented so far, the example of *caki* being bound from a genitive and from a prior sentence calls into question the requirement for a c-command relation between *caki* and its antecedent. The split antecedence and discourse binding examples are actually hallmarks of coreference, not binding. The variability in the plural readings between matched, inclusive reference and split antecedence further suggests that context plays a role in the interpretation of *caki*. And yet, in the face of all this apparently contradictory data, *caki* is generally considered to be an anaphor, bound, and not a pronominal. In the next section, I provide support for a binding analysis, and advance a proposal which can account for all the counterexamples.

3. The case for binding

The data presented in the previous section appear to make a strong case for abandoning a treatment of *caki* as a bound form. However, some of that evidence can be discounted, and there is equally compelling evidence that *caki* must be bound. The first piece of such evidence is presented by Cho (1996):

- (9) John-i caki-lul kwasinhayss-ko, Mary-to kule-ha-yess-ta.
 John-NOM SELF-ACC overtrusted-and Mary-also so-do-PST-DECL
 ‘John overtrusted himself, and Mary did too.’ (Cho 1996, ex 19a)
 =Mary overtrusted Mary. (√ Sloppy)
 ≠Mary overtrusted John. (* Strict)

In this VP ellipsis construction, there is only one possible reading for the elided conjunct. Only the sloppy reading where Mary overtrusted Mary remains, indicating that the elided *caki* is still bound within its conjunct. If *caki* really were merely coreferential, then the strict reading where Mary also overtrusted John should be available.

One of the pieces of evidence which indicated that *caki* might be coreferential was the split antecedent case. The problem with split antecedence is that it is not the only possible reading for sentences where *caki-tul* has two singular antecedents:

- (10) John₄-un Mary₃-eykey [caki-tul₄₊-i iki-lke-la-ko] malha-yess-ta.
 John-TOP Mary-DAT SELF-PL-NOM win-FUT-DECL-COMP say-PST-DECL
 ‘John told Mary that selves would win.’

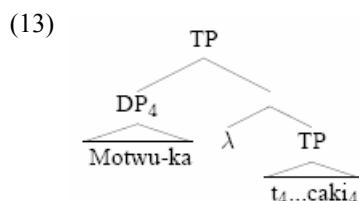
Instead of true split antecedence, what actually results is an inclusive reference reading, wherein *caki tul* refers to John and some other contextually defined group. For some native speakers, it is possible to get a reading where *Mary* is a part of that group, but this is not obligatory, and for other speakers this reading is not possible at all. Purported split antecedence readings are therefore merely cases of inclusive reference where there is accidental overlap of another argument from the sentence in the contextually salient group (Cho 1996; Storoshenko 2007).

Perhaps the strongest argument against treating *caki* as coreferential comes from cases involving quantifiers:

- (11) Motwu₄-ka caki₄-lul salang-ha-n-ta.
 everyone-NOM SELF-ACC love-do-PRES-DECL
 ‘Everyone loves self.’

Here, *caki* is bound by the quantifier *motwu*; the result is unquestionably a bound variable reading. At LF, the underlying structure for (11) will be (12), the result of an application of quantifier raising (QR) and lambda abstraction, illustrated by the tree in (13):

- (12) Every $\lambda x[x \text{ loves } x]$



This bound-variable analysis for quantifier cases applies equally well for local and for long distance cases:

- (14) Motwu₄-ka [John₅-i caki₄-lul salang-ha-n-ta-ko] sayngkak-ha-n-ta.
 everyone-NOM John-NOM SELF-ACC love-do-PRES-DECL-COMP think-do-PRES-DECL
 ‘Everyone thinks John loves self.’ (Moon 1995, ex 3:93)
 Every λx [x thinks [John loves x]]

In discussion of this example, Moon (1995) is quite clear on the point that this is a bound variable reading. As QR and variable binding can account for this long distance relationship without requiring any LF movement of *caki*, I propose that this bound variable analysis can be extended to all instances of *caki*.

If nominals are treated as generalised quantifiers, it then becomes possible to treat all cases of *caki* as bound variables. As generalised quantifiers, nominals would also undergo QR, yielding exactly the same operator variable structure as in (12). In so doing, local and long distance binding of *caki* can be unified under one analysis, as opposed to the long distance case requiring a special mechanism to raise *caki* at LF. The LF semantic forms of parallel local and long distance cases are shown in (15):

- (15) a. John₄-i caki₄-lul sokay-ha-yess-ta.
 ‘John introduced self.’
 John λx [x introduced x]
 b. John₄-i [caki₄-ka iki-ess-ta-ko] mal-ha-yess-ta.
 ‘John said self won.’
 John λx [x said [x won]]

On the surface, it may seem that this analysis is no less costly at LF, in that all antecedents now undergo LF movement, but this movement is independently motivated and is not taking place solely for the benefit of *caki*. Regardless of *caki*, QR is a well-motivated phenomenon, making the current proposal more economical at LF.

The bound variable analysis also allows for an account of the genitive case from earlier (4), repeated below as (16):

- (16) Suni₅-uy sinpal-un caki₅-uy pal-pota hwelssin kuta.
 Suni-GEN shoes-TOP self-GEN foot-than a lot big
 ‘Suni’s shoes are a lot bigger than self’s feet.’ (Kim 2000, ex 2a)

In Korean, it is possible for a quantifier to bind from inside a genitive DP:

- (17) Motwu₅-uy sinpal-un caki₅-uy pal-pota hwelssin kuta.
 everyone-GEN shoes-TOP self-GEN foot-than a lot big
 ‘Everyone’s shoes are a lot bigger than self’s feet.’

In (17), the quantifier *motwu* felicitously binds *caki* from inside the DP. The most reasonable conclusion to draw from this example is that QR allows the quantifier to escape the DP and bind its variable. Thus, the generalised quantifier analysis provides an explanation for the genitive problem: the same instance of QR which accounts for (17) will also account for (16).

At this point, it is worth noting that while QR can, to some extent, allow *caki* to be bound by elements which do not c-command it on the surface, there are restrictions on the power of QR. One

prediction is that *caki* should not be bound by a nominal that is within a subject clause island, a domain typically opaque to QR:

- (18) [Tom₄-i Mary₅-lul salang-ha-n-ta-nun] sasil-i caki₂-lul nollaykh-yess-ta.
Tom-NOM Mary-ACC love-do-PRES-DECL-TOP fact-NOM you-ACC surprise-PST-DECL
'The fact that Tom loves Mary surprised you.'

As shown here, the prediction is borne out, as neither Tom nor Mary may bind *caki*. Instead, the inherent reference of antecedent-less *caki* emerges. Moreover, crossover will block QR from generating spurious readings where an antecedent might QR from a position below *caki* and bind it:

- (19)*caki₅-uy emma-ka Mary₅-lul piphan-ha-yess-ta.
self-GEN mother-NOM Mary-ACC criticise-do-PST-DECL
Self's mother criticised Mary.'

The sentence in (19) is ungrammatical under the reading where *caki* and Mary are coindexed. At best, this sentence could be grammatical if *caki* took the inherent reference reading, but this would not be a case of binding. So while the QR analysis solves some problems through the postulation of LF movement, that movement is constrained enough to avoid overgenerating instances of binding.

To account for discourse binding of *caki*, Gil (1998) proposes that an empty topic operator may serve as an antecedent for *caki*. This operator would itself be coreferential, getting its reference from prior discourse, but it would sit in an A⁰ position similar to that occupied by a quantifier after QR. From this position, the operator would bind *caki*:

- (20) Ani, [*e*₅]_{Top} caki₅-ka kasse.
no SELF-NOM went.
'No, self went.' (where 'self' is from prior context)

In this example, the empty operator, coreferential with something from prior discourse, binds *caki* within the sentence. By adopting this analysis, discourse binding cases can be reduced to sentence local binding by a covert operator.

To summarise the proposal so far: *caki* is to be treated as a bound variable of semantic type *e*, and nominals are to be treated as generalised quantifiers. *Caki* is thus bound in an operator-variable structure formed via QR, where QR is possible. Finally, an empty topic operator can be posited to account for the discourse binding examples. With this proposal in place, local and long distance binding of *caki* have been unified under one analysis, and all the data which appeared to support a coreferential analysis have been addressed.

4. Adding the plural

The next challenge is to unite the bound variable analysis of *caki* with an account of *caki-tul* which will capture the previously observed context sensitivity. The most difficult reading to capture will be the inclusive reference reading, where *caki-tul* refers to a group of people to which its lone singular antecedent belongs. Ideally, *caki-tul* should have the same semantic type as singular *caki*, a simple variable of type *e*. Holding the definition of *caki* constant with the account above, it must then be *-tul* which introduces the context sensitivity.

Nakanishi and Tomioka (2004) discuss a similar type of context sensitivity in the Japanese plural morpheme *-tati*:

- (21) Taro-tati-wa moo kae-ta.
Taro-PL-TOP already go home-PST
'The group of people represented by Taro went home already.'
(Nakanishi and Tomioka 2004, ex 26)

Here, *Taro-tati* refers to Taro and some group with which he has an association. This contextually defined group is ascribed to the semantics of the plural marker *-tati*. Adapting (and simplifying)

Nakanishi and Tomioka's semantics for *-tati*, I propose the following semantic value for *-tul*:²

$$(22) [[-tul]] = \lambda x.Y. |Y| \geq 2 \wedge x \in Y$$

According to this formula, *-tul* is a function of semantic type $\langle e, e \rangle$, which takes an individual as its argument and returns a plurality of type e to which that individual belongs.³

The various readings for *caki-tul* are thus ascribed to the free Y variable, whose cardinality is fixed at two or greater, ensuring that this is a plural set. Binding is maintained in that *caki* remains a simple variable of type e within the semantic value of *-tul*. With this formula, it is now possible to formally illustrate the differences between the matched plural, inclusive reference, and apparent split antecedent.

Looking first at the case which drives the need for context sensitivity, the inclusive reference example is as shown in (23):

$$(23) \text{John}_4\text{-i caki-tul}_4\text{+lul sokay-ha-yess-ta.}$$

‘John introduced selves.’
 John $\lambda x [x \text{ introduced } Y. |Y| \geq 2 \wedge x \in Y]$

In this example, the generalised quantifier *John* binds the x variable, which is a member of the plurality Y introduced by *-tul*.

The apparent split antecedence case looks quite similar:

$$(24) \text{John}_4\text{-i Mary-eykey caki-tul}_4\text{+lul sokay-ha-yess-ta.}$$

‘John introduced selves to Mary.’
 John $\lambda x [x \text{ introduced } Y. |Y| \geq 2 \wedge x \in Y \text{ to Mary}]$

Crucial here is the fact that there is no formal link between Mary and the Y variable. There is nothing blocking Mary from being a part of the contextually defined set, but there is certainly nothing in the formula which forces her to be a part of that set.

Finally, there is the matched plural case:

$$(25) [\text{John-kwa-Bill}]_4\text{-i caki-tul}_4\text{-ul sokay-ha-yess-ta.}$$

‘John and Bill introduced selves.’
 $[\text{John-and-Bill}] \lambda x [x \text{ introduced } Y. |Y| \geq 2 \wedge x \in Y]$

Here, the generalised quantifier resulting from the conjunction of John and Bill binds the x variable. Upon close examination, this looks as though this could be a problem, as there is nothing in the formula which forces the set containing John and Bill to saturate the Y variable. That is, according to the given semantic formula, it is possible for *caki-tul* in (25) to refer to John and Bill, plus some other people. It turns out this is not a problem, as this reading is judged possible by native speakers.

5. Conclusion and future work

In this paper, I have presented two key proposals. First, Korean *caki* can be treated as a bound variable. In conjunction with this proposal, nominals are treated as generalised quantifiers, and the operator-variable structure created through QR allows local and long distance instances of *caki* binding to be united under a single analysis. A covert operator can also be posited to account for the apparent discourse-bound instances of *caki*. The second proposal is that the semantic form of *-tul* introduces a contextually defined plural set. This division of labour between *caki* and *-tul* achieves the goal of

² *-tul* can be applied to other semantic types, such as common nouns or even adverbs, where it has been argued to have a distributive function (Park 2006). The semantics of *-tul* presented here is limited to its application to objects of type e . Other types will naturally call for different denotations of *-tul*.

³ Nakanishi and Tomioka's original formulation is of type $\langle e, \langle e, t \rangle \rangle$, with a lambda binding the Y variable, but a type-shifting function is employed to ultimately reduce the value to e . The simpler form here skips this step. The idea that plural sets can be treated as type e is discussed in Rullmann (2003).

capturing the simultaneously bound, yet context-sensitive, nature of *caki-tul* without having to make any additional claims about the nature of *caki*.

It is left to future work to find additional support for the analysis of *caki* presented here. This future work will comprise a corpus analysis of *caki*. Under investigation in this corpus work will be the question of whether or not instances of *caki* in actual usage are consistent with the bound-variable analysis proposed here.

References

- Cho, Dong-In. 1996. Anaphor or pronominal. *Language Research* 32:621–636.
- Chomsky, Noam. 1986. *Knowledge of Language*. Praeger.
- Cole, Peter, Gabriella Hermon, and Li-May Sung. 1990. Principles and parameters of long-distance reflexives. *Linguistic Inquiry* 21:1–22.
- Gil, Kook-Hee. 1998. Topic and binding in Korean. In *Proceedings of the University of Edinburgh Post-Graduate Conference*, 1–12.
- Huang, Yan. 2000. *Anaphora*. Oxford University Press.
- Madigan, Sean, and Masahiro Yamada. 2006. Asymmetry in anaphoric dependencies: A cross-linguistic study of inclusive reference. In *University of Pennsylvania Working Papers in Linguistics 13-1: Proceedings of Penn Linguistics Colloquium 30*, 183–196.
- Moon, Seung-Chul. 1995. An optimality theory approach to long distance anaphors. Doctoral Dissertation, University of Washington.
- Nakanishi, Kumiko, and Satoshi Tomioka. 2004. Japanese plurals are exceptional. *Journal of East Asian Linguistics* 13:113–140.
- Park, So-Young. 2006. Plural marking in classifier languages: A case study of the so-called plural marking *-tul* in Korean. In *Proceedings of the International Conference on East Asian Linguistics*.
- Rullmann, Hotze. 2003. Bound-variable pronouns and the semantics of number. In *Proceedings of the Western Conference on Linguistics (WECOL 2002)*, ed. Brian Agbayani, Paivi Koskinen, and Vida Samiiian, 243–254. Fresno, CA: Department of Linguistics, California State University.
- Sohng, Hong Ki. 2003. Topics in the syntax of East Asian languages: Long-distance anaphora and adverbial case. Doctoral Dissertation, University of Washington.
- Storoshenko, Dennis Ryan. 2007. Antecedence possibilities for plural *caki* in Korean. In *The Perspectives of Linguistics in the 21st Century: Proceedings of ICLK 2007*, ed. John-Yurl Yoon and Kyoung-Ae Kim.

Proceedings of the 26th West Coast Conference on Formal Linguistics

edited by Charles B. Chang
and Hannah J. Haynie

Cascadilla Proceedings Project Somerville, MA 2008

Copyright information

Proceedings of the 26th West Coast Conference on Formal Linguistics
© 2008 Cascadilla Proceedings Project, Somerville, MA. All rights reserved

ISBN 1-57473-423-2 library binding

A copyright notice for each paper is located at the bottom of the first page of the paper.
Reprints for course packs can be authorized by Cascadilla Proceedings Project.

Ordering information

Orders for the library binding edition are handled by Cascadilla Press.
To place an order, go to www.lingref.com or contact:

Cascadilla Press, P.O. Box 440355, Somerville, MA 02144, USA
phone: 1-617-776-2370, fax: 1-617-776-2271, e-mail: sales@cascadilla.com

Web access and citation information

This entire proceedings can also be viewed on the web at www.lingref.com. Each paper has a unique document # which can be added to citations to facilitate access. The document # should not replace the full citation.

This paper can be cited as:

Storoshenko, Dennis Ryan. 2008. A Bound-Variable Account of the Korean Anaphor *caki*. In *Proceedings of the 26th West Coast Conference on Formal Linguistics*, ed. Charles B. Chang and Hannah J. Haynie, 438-444. Somerville, MA: Cascadilla Proceedings Project.

or:

Storoshenko, Dennis Ryan. 2008. A Bound-Variable Account of the Korean Anaphor *caki*. In *Proceedings of the 26th West Coast Conference on Formal Linguistics*, ed. Charles B. Chang and Hannah J. Haynie, 438-444. Somerville, MA: Cascadilla Proceedings Project. www.lingref.com, document #1700.