

An Itsy-Bitsy Puzzle: Asymmetric extraction from coordination in Khoekhoegowab

Leland Paul Kusmer • me@lelandpaul.com

ACAL 49 • 22 MARCH 2018

Khoekhoegowab (Central Khoesan, Namibia, ~200,000 speakers) allows an unusual construction in which a object may be topicalized out of the first of two conjoined VPs.¹

- (1) shows a Khoekhoe clause with two conjoined VPs: two distinct verbs and objects, but only one subject and only one tense.
- In (2), the object of the first verb has been topicalized.
- This construction is surprising because it seems to violate the Coordinate Structure Constraint (Ross, 1967).

(1) Dandagob ge [amsa ||nae] tsi [†naba ra †na.]
D. DECL song sing AND dance IMPV dance
“Dandago is singing a song and dancing a dance.”

(2) Amsa_i=b ge Dandagoba [t_i ||nae] tsi [†naba ra †na.]
song =₃SM DECL D. sing and dance IMPV dance
“As for a song, Dandago is singing it and dancing a dance.”

This construction is also striking because it closely parallels a well-studied but poorly-understood construction in German, called the Subject Gap in Finite Clause (SGF):

(3) **In den Wald**_i [ging der Jäger t_i] und [fing einen Hasen.]
into the forest went the hunter and caught a rabbit
“The hunter went into the forest and caught a rabbit.”

The SGF construction is defined by the following properties:

- Two conjoined predicates share one subject, which nevertheless appears in the middle of the first conjunct.
- Some XP is fronted out of the first conjunct, in apparent violation of the Coordinate Structure Constraint.
- There is an asymmetry: Only the first object may be fronted in this way.

While the SGF construction has been well-studied in German (Kathol, 1995; Höhle, 1990; Johnson, 2002; Schwarz, 1998) and Dutch (Heycock & Kroch, 1993), to the best of my knowledge it has never been found outside of the Germanic family.²

¹All Khoekhoe data comes from original fieldwork carried out in Windhoek during the austral winter of 2017. Thanks to my wonderful consultants Markus Kooper, Magdalena Isaak, Michelle Swartbooi, Nadia April, and Irene ||Garoes for their generous help; any mistakes are my own.

²Bjorkman (2014) points out that the SGF, while not productive in English, exists in frozen form in the nursery rhyme ‘The Itsy-Bitsy Spider’: *Down_i [came the rain t_i] and [washed the spider out].*

Khoekhoe differs from German in a number of key morphosyntactic respects, which shed more light on this unusual construction.

- In particular, the Khoekhoe SGF rules out a number of possible analyses which either argue for hidden structure or for the CSC being outside the syntax proper.
- Khoekhoe also raises a typological puzzle: What exactly is the property that it and Germanic have in common, but other languages lack?

The rest of this talk will proceed as follows:

- §1 reports the basic facts of the SGF construction cross-linguistically and makes the case that the Khoekhoe construction really is the same thing.
- §2 discusses the difficulties with analyzing the SGF and argues that the Khoekhoe rules out approaches which make the conjuncts bigger than VP.
- §3 reports some new facts from Khoekhoe that suggest that we cannot appeal to post-syntactic processes for escape from the CSC: Whatever is allowing the SGF construction is present in the narrow syntax.
- §4 concludes that the asymmetric availability of extraction in SGF sentences must be present in the narrow syntax, and sketches a kind of asymmetric coordination that might do the job. However, this leaves open a larger puzzle: Why do some languages allow SGF while others don't?

I THE SGF, CROSS-LINGUISTICALLY

First, some background on Khoekhoe clause structure:

- Generally verb-final.
- In matrix clauses, second-position clause type clitics, e.g. *ge* 'declarative'.
- By default, the subject appears in first position — as in (4).
- However, any XP may be topicalized into this position, in which case the subject appears after the clause type clitic — as in (5).

(4) Dandagob ge ||ari tarasa †khanisa go mā.
 D. DECL yesterday woman book PAST give
 "Dandago gave the book to the woman yesterday."

(5) †khanisa =b ge Dandagoba ||ari tarasa go mā
 book =₃SM DECL D. yesterday woman PAST give
 "As for the book, Dandago gave it to the woman yesterday."

Descriptively, I'll call the first position the *prefield* and the positions between the clitic and the verb the *middlefield*.

1.1 *The SGF itself*

Recall that the SGF involves topicalizing the argument of the first verb into the prefield, stranding the subject in the middlefield:

- (6) Amsa_i =b ge Dandagoba t_i ||nae tsi †naba ra †na.
 song =₃SM DECL D. sing and dance IMPV dance
 “As for a song, Dandago is singing it and dancing a dance.” (= (2))

In both Khoekhoe and German, the extraction possibilities are asymmetric: only the first object can be extracted this way.

- (7) *†naba_i =b ge Dandagoba amsa ||nae tsi t_i ra †na.
 dance =₃SM DECL D. song sing and IMPV dance
 Intended: “As for the dance, Dandago sang a song and danced it.”
- (8) *Den Hund_i hat einer ihn gefüttert und hat t_i geschlagen.
 the dog has someone it fed and has hit
 Intended: “Someone has both fed the dog and hit it.”

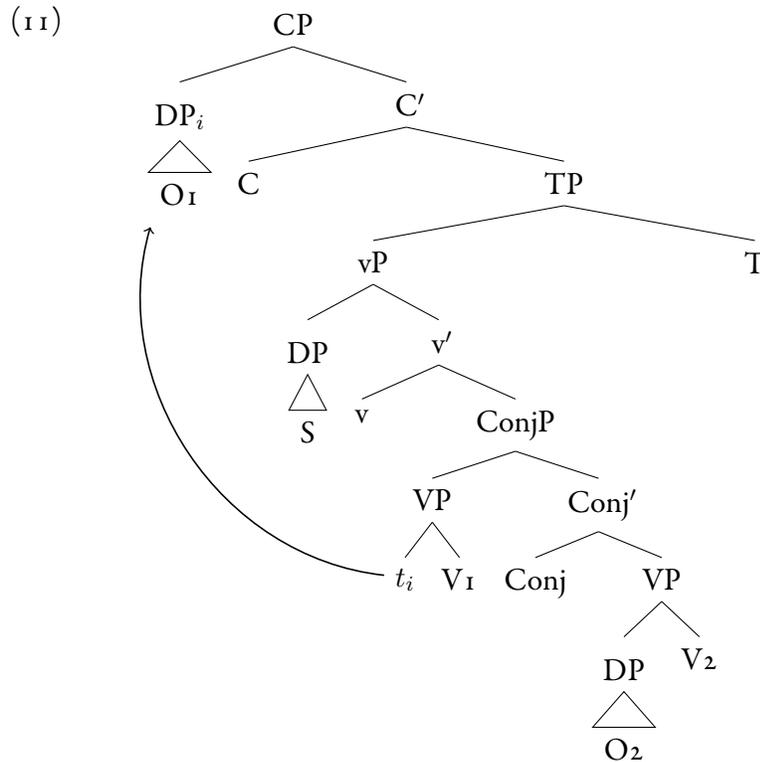
Furthermore, the subject gap in the second conjunct is obligatory:

- As noted by Höhle (1990); Johnson (2002), if the subject is indefinite, it always scopes over both conjuncts (9).
- More directly: if a second subject is placed in the middlefield of the second conjunct, ungrammaticality results (10).

- (9) a. arina =i ge khoe-e |namsa tsi |hōana a !khuisa
 dogs =₃SN DECL someone love and cat STAT.PRES hate
 “Someone (both) loves dogs and hates cats.”
 Not: * “Someone loves dogs and someone hates cats.”
- b. Den Hund hat einer gefüttert und hat ihn geschlagen.
 the dog has someone fed and has it hit
 “Someone has (both) fed the dog and hit it.” (Schwarz 1998, (54b): p213)
- (10) *arina =b ge Dandagoba |namsa tsi Khoedagesa |hōana a !khuisa
 dogs =₃SM DECL D. love and K. cats STAT.PRES hate
 Intended: “As for dogs, D. loves them and K. hates cats.”

This fact — that SGF constructions seem to license only one subject — strongly implies that the coordination takes place below the level at which the subject is merged. This is schematized in (11).

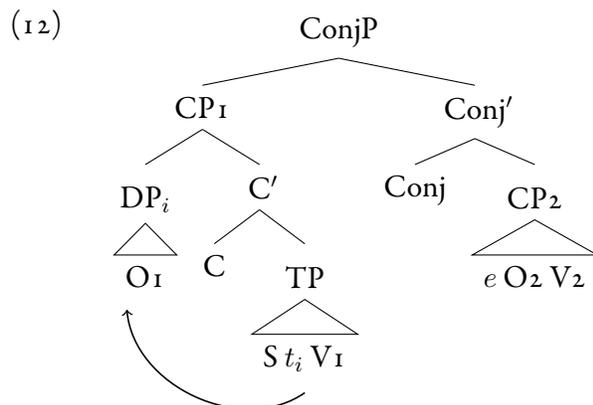
- But in this structure, the topicalization appears to violate the Coordinate Structure Constraint (CSC).



2 AGAINST BIG CONJUNCTS

Any analysis of the SGF must escape the CSC violation somehow; one option pursued by e.g. Schwarz (1998) is to argue that we're conjoining bigger phrases than we thought:

- The first conjunct is taken to be CP-sized, with movement taking place to the specifier of that CP.
- In this model, no CSC violation occurs.



Note the empty category *e* in the second subject position.

- We need some explanation for why this gap is obligatorily coreferential with the subject.

- For example, Buring & Hartmann (1998) propose that the second subject is forced to undergo topic drop.
- This has a number of problems:
 - Why is topic drop obligatory here but optional elsewhere?
 - In German, *only* the subject may be dropped — leaving out the object or an embedded subject results in ungrammaticality.
 - There are other empirical issues arguing against SGF sentences being CP coordination.³

However, I think that in Khoekhoe the evidence against a big-conjuncts solution is even more clear:

- In Khoekhoe SGF constructions, only one clause-type marker appears.
- Worse, in Khoekhoe only one tense marker appears!
- Contrast SGF (13) with a clear case of CP-coordination, as in (14) — multiple subjects are possible, and two clause-type markers and tense markers appear.

(13) Amsa =b ge Dandagoba ||nae tsi †naba †na tama.
 song =₃SM DECL D. sing and dance dance NEG.NF
 “As for the song, Dandago didn’t sing it and dance the dance.”

(14) Amsa =b ge Dandagoba ||nae tama tsi =s ge Khoedagesa †naba †na tama.
 song =₃SM DECL D. song NEG.NF and =₃SM DECL K. dance dance NEG.NF
 “The song, Dandago didn’t sing and the dance Khoedage didn’t dance.”

In sum: the Khoekhoe data seems very difficult to square with a big-conjuncts analysis.

3 ESCAPING THE CSC

Khoekhoe seems to push us towards a small-conjuncts analysis, in which SGF coordination really does take place below the base position of the subject.

- Prior small-conjuncts solutions include Kathol (1995); Johnson (2002); Heycock & Kroch (1993).
- Any such analysis needs a way for the moved item to avoid the CSC.
- In general, prior analyses escape the CSC by appealing to post-syntactic processes.
- I’ll discuss two representative small-conjuncts analyses, one which appeals to PF and one to LF.
- I’ll show that Khoekhoe provides new facts which rule out either of these escape hatches — however we escape the CSC, it must happen in the syntax itself.

³For example: Kathol (1995) presents a problem, noted by Heycock & Kroch (1993), from relative clause extraposition: In German it is not generally possible to extrapose a relative clause from the subject of two conjoined CPs, but it is possible to extrapose an RC from the subject of SGF.

3.1 *Escape at LF*

One road that a small-conjuncts solution can take is to allow escape from the CSC at LF:

- Perhaps the CSC is not a constraint on syntactic processes but on LF representations.
- As long as the syntactic movement is ‘undone’ via reconstruction at LF, no CSC violation obtains.
- This solution is not represented in pure form in the literature; the version I’m presenting is an amalgamation of Johnson (2002); Fox (2000); Lin (2001).

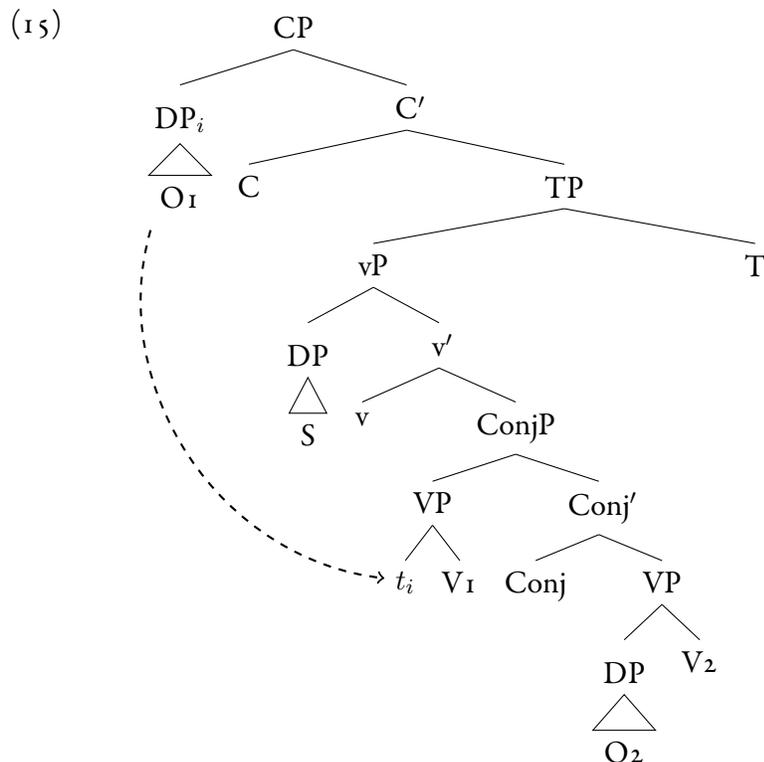
In this model, the CSC will appear violable exactly when the moved item reconstructs at LF:

- CSC-violating movement is permitted in the syntax.
- If the moved item is interpreted in its low (reconstructed) position at LF, neither component sub-structure will be ungrammatical.
- Hence no apparent CSC violation will result.

Lin (2001) does not directly discuss the SGF, but her analysis implies a potential analysis of this construction:

- CSC violations in the SGF are not surprising as long as the moved item reconstructs.⁴

Under this model, we predict that the topicalized object in SGF constructions will always reconstruct **below the subject**:



⁴This analysis doesn't account for the asymmetry between the first and second objects.

We can use Condition C to test this prediction.

- (16) ||ib ge [Dandagob di |hôasa |namsa] tsi [ti ariba a !khuisa.]
 He DECL D. POSS cat love and my dog STAT.PRES hate
 “He_i loves Dandago’s_{j/*i} cat and hates my dog.”
- (17) Dandagob di |hôasa =b ge [(||iba) t |namsa] tsi [ti arina a !khuisa.]
 D. POSS cat =_{3SM} DECL (he) love and my dog STAT.PRES hate
 “As for Dandago’s_{j/i} cat, he_i loves it and hates my dog.”

The prediction is **not borne out**:

- (16) is a control with the subject initial, showing that there is a Condition C disjoint-reference effect between the subject ‘he’ and the possessor ‘Dandago’.
- (17) is the same sentence in SGF order, and there is no longer a disjoint reference effect!
- That is: Topicalization in Khoekhoe ameliorates Condition C effects⁵, even in SGF contexts.⁶

In sum: Lin proposed an escape from the CSC via reconstruction at LF; this new data shows us that this reconstruction does not occur.

3.2 *Escape at PF*

Kathol (1995) proposes an analysis in which the fronted object in SGF sentences escapes the CSC at PF:

- He proposes a wholesale reorganization of our understanding of the linearization algorithm.
- In particular, he proposes that the algorithm places topic-marked XPs leftmost.
- In this system, the fronted item has never actually moved — it just appears in a different place.
- This analysis cannot explain all the Khoekhoe facts.

The evidence for this claim comes not from SGF sentences themselves, but from the properties of WH questions.

- Khoekhoe allows WH items to remain *in situ*.
 - In coordination, WH items can create questions from inside the first conjunct, but not the second.
 - This parallels the asymmetry we see in the SGF.
- (18) Dandagoba (kha) tae-e |namsa tsi |hôana a !khuisa?
 D. INTER what love and cats STAT.PRES hate
 “What does Dandago love while hating cats?”
- (19) *Dandagoba (kha) |hôana |namsa tsi tae-e a !khuisa?
 D. INTER cats love and what STAT.PRES hate
 Intended: “What does Dandago love cats and also hate?”

⁵This is true even in non-SGF contexts — *Dandagob di |hôasab ge (||iba) a |namsa* “As for Dandago’s cat, he hates it.” also allows coreference. This is perhaps surprising?

⁶Tentatively, the same seems to be true for German: One speaker reports that *in Johns Haus ist er gegangen und hat ein Buch genommen* “Into John’s house did he go and a take a book.” allows coreference between *John* and the subject.

This seems to rule out any solution in the vein of Kathol (1995):

- The fact that a WH-word in the second conjunct *cannot* create a question tells us that the CSC is still active.
- But the WH-word in the first conjunct *can* create a question — depending on your theory of WH-*in situ* either by agreeing with a higher head in the syntax or by covertly moving at LF.
- Regardless, no PF reordering will capture these facts.

In addition to ruling out PF mechanisms for escaping the CSC, these facts suggest that the asymmetry of SGF sentences is central to the mystery:

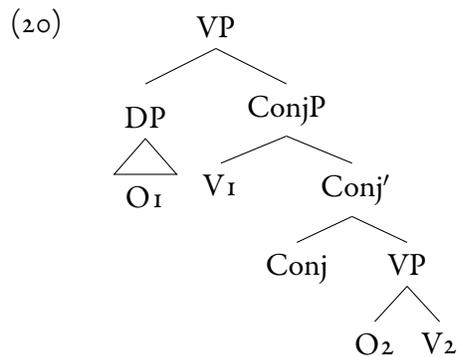
- This asymmetry affects more than just the SGF: Leftmost-conjuncts are transparent, other conjuncts are opaque.
- Whatever the source of this asymmetry, it seems to be present in the narrow syntax.

4 CONCLUSION

Khoekhoe seems to tell us at least two new things about the SGF construction that Germanic morphosyntax obscures:

1. The conjuncts in SGF sentences *must* be small.
2. The asymmetry between the conjuncts must be present in the narrow syntax — the first conjunct is permeable for more than just topicalization.

The simplest possible analysis which captures these two facts results in an atypical syntax:



- This analysis takes seriously the idea that the coordination is asymmetric by coordinating two unlike objects, a phrase and a head.
- The second conjunct VP combines not with the whole first VP, but just with the first verb, creating a complex predicate.
- This way, O₁ simply isn't inside the conjunct, and therefore is available for topicalization, WH-question formation, etc.
- Semantically, this perhaps isn't so strange: VP and V₁ and both predicates of events and could be joined by Event Identification (Kratzer, 1996).

This analysis makes several predictions, which remain to be tested. For example:

- In the structure in (20), O₁ asymmetrically c-commands O₂, which should be testable using a wide variety of phenomena.
- If the conjunction of V and VP is accomplished through event identification, then the two verbs should not be independently modifiable.
- Passivization should only be able to promote O₁ to subject.

4.1 Typology

Regardless of whether the sketch in (20) is the right direction to pursue, the Khoekhoe SGF raises a much more interesting question:

- **Why do Germanic and Khoekhoe have this construction when other languages do not?**
- Put another way: Other than the SGF itself, what property unites Germanic and Khoekhoe?

One place to start: German is underlyingly verb-final (ignoring V₂), as is Khoekhoe.

- **Hypothesis 1:** Perhaps the availability of SGF tracks verb-finality?
- **Not true:** Hindi-Urdu is verb-final and allows scrambling for information-structure, but does not allow SGF extraction.⁷

- (21) a. Raam kuttoN-ko pasand kartaa hai aur billiyoN-se nafrat kartaa hai
 R. dogs-DAT liking do PRES and cats-COM hatred do PRES
 “Raam likes dogs and hates cats.”
- b. *kuttoN-ko Raam pasand kartaa hai aur billiyoN-se nafrat kartaa hai
 dogs-DAT R. liking do PRES and cats-COM hatred do PRES
 Intended: “As for dogs, Raam likes them and hates cats.”

Another hypothesis: While Khoekhoe doesn’t have V₂ word order, it does have 2nd position clitics.

- **Hypothesis 1:** Perhaps the availability of SGF tracks 2nd-position phenomena?
- **Not true:** Kashmiri has V₂ word order with free topicalization into the prefield (Manetta, 2006) but does not allow SGF extraction.⁸

- (22) a. Mohan chu film vuch-aan ta chu su boz-aan.
 M. AUX film see-PART and AUX it hear-PART
 “Mohan is seeing the film and hearing it.”
- b. *Film chu Mohan vuch-aan ta chu su boz-aan.
 Film AUX M. see-PART and AUX it hear-PART
 Intended: “As for the film, Mohan is seeing it and hearing it.”

The asymmetric-coordination analysis sketched above has nothing to say about why this kind of coordination is available in exactly these languages.

- I’m always looking for SGF-like constructions in other languages — if you know of one, please let me know!

⁷Thanks to Jyoti Iyer and Sakshi Bhatia for the Hindi-Urdu data.

⁸Thanks to Professors Emily Manetta and Omkar Koul for the Kashmiri data.

ACKNOWLEDGEMENTS

I am deeply indebted to my consultants Markus Kooper, Magdalena Isaak, Michelle Swartbooï, Nadia April, and Irene ||Garoës. Michelle & Irene in particular have been extremely patient in answering endless WhatsApp questions about Dandago's pet-preferences. Thanks also to Professor Levi Namaseb of UNam for his help.

Professor Chris Collins and all of the NYU Syntax Brown Bag crowd provided insightful comments on an earlier version of this presentation. I'm indebted to Professors Kyle Johnson and Ellen Woolford for considerable help in understanding the SGF literature. I've also benefited greatly from conversations with my colleagues Sakshi Bhatia, Jyoti Iyer, Carolyn Anderson, and I'm sure many others that I've forgotten.

The author is supported by the National Science Foundation Graduate Research Fellowship under Grant No. 1451512. Any opinion, findings, and conclusions or recommendations expressed in this material are those of the author and do not necessarily reflect the views of the National Science Foundation.

SELECTED REFERENCES

- Bjorkman, Bronwyn. 2014. Accounting for unexpected subject gaps in tp coordination. *The Linguistic Review* 31:487–513.
- Büring, Daniel, & Katharina Hartmann. 1998. Asymmetrische Koordination. *Linguistische Berichte* 174:172–201.
- Fox, Danny. 2000. *Economy and semantic interpretation*. Cambridge, Massachusetts: MIT Press.
- Haacke, Wilfred H.G. 2013. Namibian Khoekhoe Syntax. In *The Khoesan Languages*, ed. Rainer Vossen, 325–340. Routledge.
- Heycock, Caroline, & Anthony S Kroch. 1993. Verb movement and the status of subjects: implications for the theory of licensing. *Groninger Arbeiten zur germanistischen Linguistik* 36:75–102.
- Höhle, Tilman N. 1990. Assumptions about asymmetric coordination in German. In *Grammar in progress: GLOW essays for Henk van Riemsdijk*, ed. Joan Mascaró & Marina Nespør, 221–235. Foris Publications.
- Johnson, Kyle. 2002. Restoring exotic coordinations to normalcy. *Linguistic Inquiry* 33:97–156.
- Kathol, Andreas. 1995. Linearization-based German syntax. Doctoral Dissertation, Ohio State University, Columbus, Ohio.
- Kratzer, Angelika. 1996. Severing the external argument from its verb. In *Phrase structure and the lexicon*, ed. Johan Rooryck & Laurie Zaring, 109–137. Dordrecht, The Netherlands: Kluwer Academic Publishers.
- Lin, Vivian. 2001. A way to undo A movement. In *Proceedings of the 20th West Coast Conference in Formal Linguistics*, ed. Karine Megerdooimian & Leora Anne Bar-el, 358–371. Somerville, Massachusetts: Cascadilla Press.
- Manetta, Emily. 2006. Peripheries in Kashmiri and Hindi-Urdu. Doctoral Dissertation, University of California, Santa Cruz.
- Ross, John. 1967. Constraints on variables in syntax. Doctoral Dissertation, Massachusetts Institute of Technology.
- Schwarz, Bernhard. 1998. On odd coordinations in German. *The Journal of Comparative Germanic Linguistics* 2:191–219.