Extraction from coordination in Khoekhoe

Leland Paul Kusmer NYU Brown Bag • 20 October 2017

1 Introduction

Khoekhoegowab[1] (Central Khoesan; ~200,000 speakers in Namibia; hereafter Khoekhoe) has clausal structure reminiscent of German. In unmarked contexts, clauses contain in order:

- a prefield, containing at most one XP;
- a (set of) second position clitic(s);
- a middlefield with fairly free word-order; and
- a verbal complex.

This general clause structure is shown in (1).[2]

(1) a. taras ge ǁari ǂkhanis -a go mā =te
woman DECL yesterday book -A TAM give 1S.O
"The woman gave me a book yesterday."

b. ǁari =s ge taras -a ǂkhanis -a go mā =te
yesterday 3SF DECL woman -A book -A TAM give 1S.O
"Yesterday the woman gave me a book."

It is possible to coordinate whatever constituent contains middlefield & verb, apparently below the subject:

(2) Khoe-i ge arina |namsa tsi |hôana !khuisa
someone DECL dogs love and cats hate
"Someone loves dogs and hates cats."

It is possible, in this construction, to front the object from the first conjunct into the prefield:

(3) Arina =i ge khoe-e |namsa tsi |hôana !khuisa
dogs 3SN DECL someone love and cats hate
"Dogs someone loves and hates cats."

[1] All Khoekhoe data is from original fieldwork conducted in Windhoek during the summer of 2017. My thanks to Professor Levi Namaseb of UNam for his assistance, and most of all to my wonderful consultants: Markus Kooper, Magdalena Isaak, Michelle Swartbooi, Nadia April, and Irene ǁGaroes.

[2] All Khoekhoe examples are presented in the official orthography. The following abbreviations are used: DECL — declarative clause marker; -A — oblique case marker; PAST — past tense; IMPV — imperfect aspect; NEG.NF — negative non-future tense; (1/2/3) (S/P) (F/M/N) — person, number, & gender marking; .O — object clitic. All Khoekhoe examples are presented in the standard orthography.

There are two surprising (and perhaps contradictory) features of this construction:

1. Something has been extracted out of the first conjunct, apparently in violation of the Coordinate Structure Constraint (Ross, 1967).
2. The subject ('someone') still obligatorily scopes over coordination: The only interpretation available is that one person both loves dogs and hates cats.

This construction is strongly reminiscent of a construction in German that has been called *Subjektlücke in finiten Sätzen* (Subject Gap in Finite clause, hereafter SGF); see Kathol (1995); Höhle (1990); Büring & Hartmann (1998), among others.

(4) a. 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)

b. In den Wald ging der Jäger und fing einen Hasen.
into the forest went the hunter and caught a rabbit
"The hunter went into the forest and caught a rabbit." (Kathol 1995, chp 3 ex. 2, p78)

SGF is problematic because of what Johnson (2002) calls the 'Size Paradox':

- If the two conjuncts are taken to be large (i.e. CPs), the violation of the Coordinate Structure Constraint is avoided, but instead there is no explanation for the obligatory coreference of the subject in the first conjunct with the gap in the second.
- However if the two conjuncts are taken to be small (i.e. vP), then the reverse is true.

In this presentation:

- I will argue for a small-conjuncts analysis, in which the relevant conjuncts are vPs.
- I will follow Lin (2001) and Fox (2000) in arguing that the CSC is a constraint on LFs, not on the syntax — extraction from conjuncts is possible as long as reconstruction is.

2 Background: Khoekhoe clause structure

The Khoekhoe prefield is normally occupied by the (nominal) subject:

(5) Dandagob ge ǁari tarasa ǂkhanisa go mā.
D. DECL yesterday woman book PAST give
"Dandago gave the book to the woman yesterday."

The right edge of the prefield is marked by a cluster of 2nd position clitics. These clitics are:

- *ge* — matrix declaratives
- *kom* — matrix declaratives with "emphasis" (?)
- *kha* — matrix surprise questions
- The subject clitics (see below)

Even in clauses with no clause type marker, we can identify the edge of the prefield by putting something other than the subject there: Any time something other than the subject appears in the prefield, a subject clitic appears at the right edge, preceding any clause type marker.

(6) a. Dandagoba (kha) ǁari tarasa ǂkhanisa go mā?
D. Q yesterday woman book PAST give
"Did Dandago give the book to the woman yesterday?"

b. ǁari =b (kha) Dandagoba tarasa ǂkhanisa go mā?
yesterday 3SM Q D. woman book PAST give
"Yesterday, did D. give the book to the woman?"

Anything can be placed in the prefield, up to and including the verb:

(7) a. ǂkhanisa =b ge Dandagoba ǁari tarasa go mā?
b. mā -b ge go Dandagoba ǁari tarasa ǂkhanisa?

2.1 The nature of the prefield

There is some evidence that the prefield is a prosodic position rather than a syntactic one. For one, CP coordination obligatorily occupies the prefield:

(8) a. ... tsi =b ge Dandagoba ra ǂna.
and 3SM DECL D. IMPV dance
"...and Dandagob is dancing."

b. *... tsi Dandagob ge ra ǂna.
and D. DECL IMPV dance

For another, Brugman (2009) claims (and I am in the process of confirming) that topics fronted before the prefield always form a prosodic unit with the prefield.

The idea that the prefield is somehow post-syntactically defined is supported by a semantic fact: Non-subject prefield elements always reconstruct lower than the subject.

(9) a. ǁîs |gôaba =s ge ma ǁgus hoasa |nam.
her son 3SM DECL which mother all love
"Every mother$_i$ loves her$_{i/j}$ son."

b. ma |gôab hoaba =s ge ǁîb ǁgusa |nam.
which son all 3SM DECL his mother love
"Her$_i$ son loves every mother$_{j/*i}$."

2.2 Tense marking

Khoekhoe packages tense, aspect, and polarity information into enclitic particles. These particles come in two classes:

1. *post-verbal* particles are always final in the verbal complex.
2. *pre-verbal* particles typically occur immediately before the verb, but a) may optionally occur earlier in the prefield, and b) in some syntactic contexts must occur after the verb.

An example of a post-verbal particle is the negative non-future *tama*:

(10) Dandagob ge ǂna tama.
D. DECL dance NEG.NF
"Dandago didn't dance."

An example of a pre-verbal particle is the recent-past *go*:

(11) a. Dandagob ge (go) mâi-e (go) ǂû.
D. DECL (PAST) porridge (PAST) eat
"Dandago ate porridge."

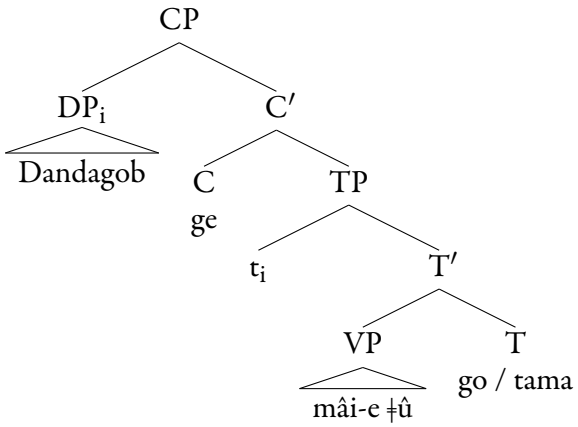
b. ǂû go =b ge mâi-e
eat PAST 3SM DECL porridge
"He ate porridge."

c. ǂû go mâi-e Dandagob ge
eat PAST porridge D. DECL
"As for eating porridge, Dandago did it."

d. Dandagob ge mâi-e ǂû go!
D. DECL porridge eat PAST
"Dandago DID eat porridge!"

In light of these facts, and of the generally-head-final nature of the language, I propose that T is also head-final; pre-verbal particles are derived by some PF displacement mechanism.

(12)



The prosodic displacement analysis is supported by the fact that all post-verbal particles are prosodically heavy, while all pre-verbal particles are light; see Hahn (2013) and eventually my dissertation for more details.

3 Background: The SGF construction

The SGF has been known in German since at least Höhle (1985).

(13) 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)

Recall the key properties of this construction:

1. Some item is fronted into the prefield out of the first conjunct.
2. The subject is apparently inside the middlefield of the first conjunct, but obligatorily scopes over both.
3. (The second conjunct has an apparent V1 order.)[3]

The first two properties are present in the Khoekhoe equivalent:

(14) Bir-e =b ge ǁna khoeba go ā tsi su-e ǂû
beer 3SM DECL that man PAST drink and soup eat
"That man drank beer and ate soup."

The object of the first conjunct has been extracted to the prefield, and the subject (which is apparently in the middlefield of the first verb) is obligatorily the subject of the second verb as well. Compare:

(15) a. *Bir-e =b ge ǁna khoeba go ā tsi ǁna khoesa su-e ǂû
beer 3SM DECL that man PAST drink and that woman soup eat
Intended: "That man drank beer and that woman ate soup."

[3] German-speakers I've consulted tell me that this V1 order is dubious, but it's robustly attested in the literature. Regardless, this property isn't relevant to Khoekhoe.

b. Bir-e =b ge ǁna khoeba go ā tsi -s ge ǁna khoesa su-e go ǂû
beer 3SM DECLthat man PAST drink and 3SF DECL woman soup PAST eat
"That man drank beer and that woman ate soup."

It's worth noting, at this point, that the extraction possibilities are asymmetric: In both languages, extraction from the second conjunct is impossible.

(16) *su-e =b ge ǁna khoeba bir-e go ā tsi ǂû
soup 3SM DECL that man beer PAST drink and eat
Intended: "Soup, the man drank beer and ate it."

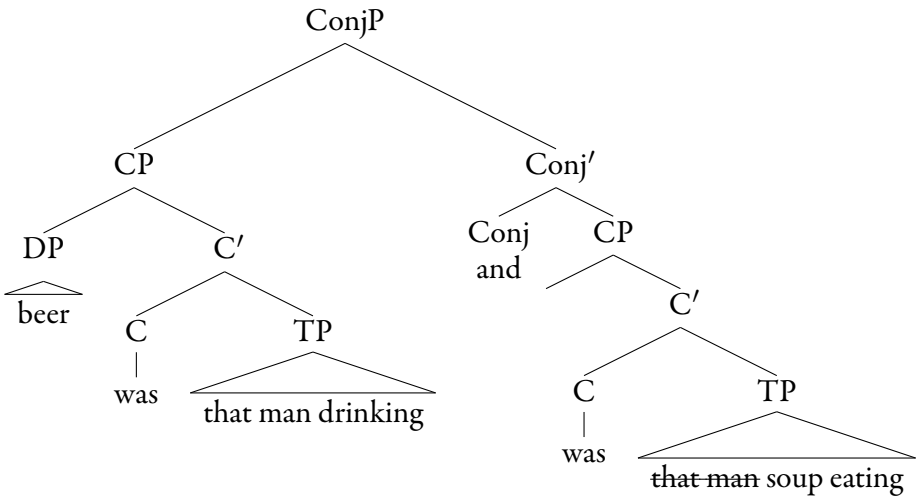
(17) *Den Hund hat einer ihn gefüttert und hat geschlagen.
the dog has someone it fed and has hit
Intended: "Someone has both fed the dog and hit it."

3.1 The Size Paradox

Johnson (2002) notes that analyses of SGF generally fall prey to the 'size paradox':

- Some analyses (e.g. Schwarz 1998) assume the conjuncts are CPs; this allows the prefield to be inside the first conjunct, avoiding violations of the CSC.[4]
- Such analyses need to stipulate that Topic Drop (or something like it) applies to eliminate material in the second conjunct.

(18)



The only problem with this analysis is the Topic Drop: Why is it *obligatory* with subjects and *impossible* with objects in this construction?

(19) *Den Hund hat keiner gefüttert und habe er ___ geschlagen?
The dog has no one fed and he beaten
Intended: "No one has fed the dog and he has beaten (it)."

[4]Schwarz is actually concerned with a similar construction that he reduces to SGF-plus-gapping. For a more fleshed-out 'Big Conjunct' analysis along the lines discussed here, see Büring & Hartmann (1998), which proposes that the 'and' in SGF is not a coordinator but rather a subordinator. The problem with this analysis, as noted by Kathol (1995), is that German SGF constructions allow more than two conjuncts with an overt coordinator present before only the last one.

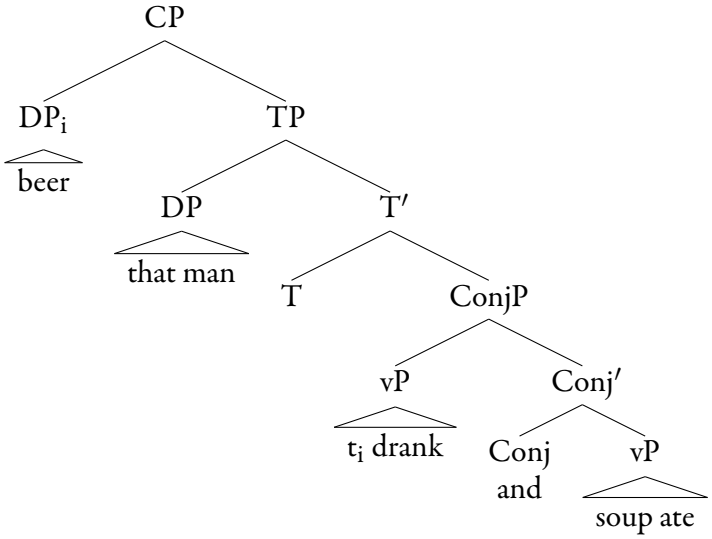
Along similar lines: If the subject gap in the second conjunct is due to Topic Drop, why must indefinite subjects be understood in terms of the whole predicate?

(20) Arina =i ge khoe-i |namsa tsi |hôana !khuisa
dogs 3SN DECL someone love and cats hate
"Dogs someone loves and hates cats."
(Can not mean: "One person loves dogs and another hates cats.")

However, 'Small Conjunct' analyses (e.g. Johnson 2002) fair no better:

- If the conjuncts are smaller than CP, then the prefield item must have been extracted from the leftmost conjunct, in violation of the CSC.

(21)



This analysis has the benefit of keeping the subject above the level of coordination, but we now have to grapple with a CSC violation. What to do?

4 Analysis: Small conjuncts and the CSC

Khoekhoe gives us new evidence that SGF coordination occurs *below TP*:

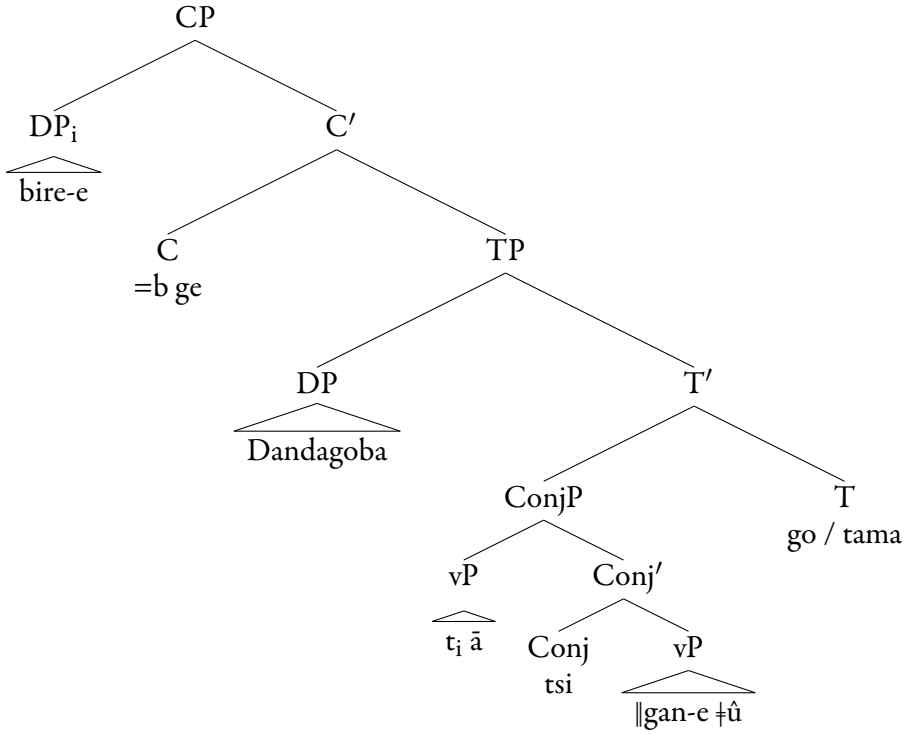
- In SGF constructions (and their non-SGF coordination counterparts), only one tense particle appears.[5]
- Pre-verbal tense particles may appear in *either* conjunct, while still scoping over both.
- Post-verbal particles must appear after the rightmost conjunct in order to scope over both; if they appear earlier, they scope only over the first conjunct.
- This is compatible with coordination **below T**.

[5] In non-SGF coordination, it is certainly possible for two tense particles to appear, especially if the tenses are different. However, my speakers were divided about whether this was grammatical in an SGF context; it needs further investigation.

(22) bir-e =b ge dandagoba (go) ā tsi ǁgan-e (go) ǂû.
beer 3SM DECL d. PAST drink and meat PAST eat
"Dandago drank beer and ate meat."

(23) Bire-e =b ge Dandagoba ā tsi ǁgan-e ǂû tama.
beer 3SM DECL D. drink and meat eat NEG.NF
"Dandago didn't drink beer or eat meat."

(24)



A few notes about this:

- Probably the subject actually originates in spec,vP. That's fine — we can ATB-extract.
- Displacing T into one of the conjuncts is quite surprising! But c.f. Elfner (2012); Bennett et al. (2016).
- Speakers are somewhat divided on whether it is even possible to have two separate T-particles in an SGF context. If this is possible, we lose the explanation for the subject gap — but see Bjorkman (2014) for a possible solution.

This leaves open the question of how to resolve the CSC.

4.1 A constraint on representations

Lin (2001) follows Fox (2000) in proposing that the CSC is a constraint on LF representations, not on syntactic operations:

(25) a. Extraction out of a coordinate structure is possible only when the structure consists of two [or more] independent substructures, each composed of one of the coordinates together with material above it up to the landing site (henceforth, component structures).

b. Grammatical constraints are checked independently in each of the component structures. (Fox 2000: 50)

- Recall that, in Khoekhoe, prefield items obligatorily reconstruct.
- At LF, then, the two structures being evaluated will have the objects in place — no violation occurs.

(26) Two structures for (24):
a. =b ge Dandagoba [$_{vP}$ bir-e ā] go
b. =b ge Dandagoba [$_{vP}$ ǁgan-e go ā] go

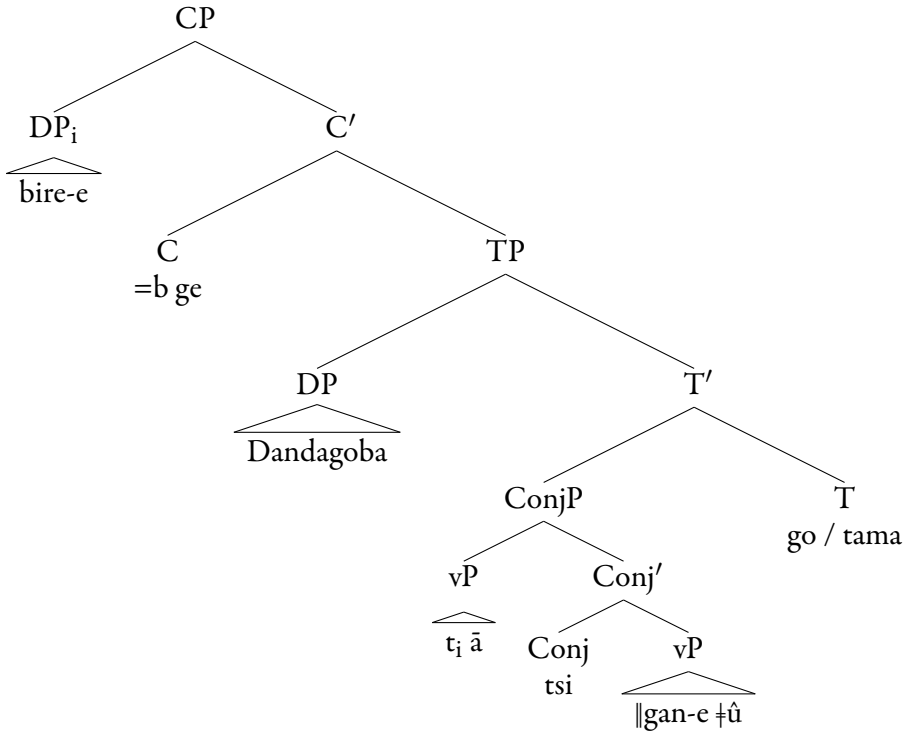
4.2 Asymmetric extraction

One loose end: Why can't we extract from the second conjunct?

(27) *ǁgan-e =b ge Dandagoba bir-e go ā tsi ____ ǂû
meat 3SM DECL D. beer PAST drink and eat
Intended: "Dandago drank beer and ate meat."

An idea: This is a PF linearization constraint.

(28)



- Assume a Kayne (1994) antisymmetric linearization.
- This requires all daughters of vP_1 to precede all daughters of vP_2.
- In licit SGF constructions, this is true: All members of the first conjunct precede all members of the second.
- When something from the second conjunct is extracted, it will violate antisymmetry.

4.3 *But what about German?*

- I believe that it is not generally true that prefield items always reconstruct in German.
- But this analysis makes the clear prediction that prefield items in SGF sentences will always reconstruct.
- I've had trouble confirming this!
- But German has at least one other very similar construction where this holds true.[6]

(29) Bill hat die Bucher$_i$ John$_j$ t_i t_j gegeben und t_j die Magazine geschickt.
Bill has the book to John given and the magazine sent
"Bill has given the book and sent the magazine to John."

- In (29), an indirect object has been ATB-extracted from both conjuncts.
- The direct object from the first conjunct has then been scrambled past it.
- This closely parallels SGF, where a subject is ATB-extracted and an object topicalized past it.
- In (29), the scrambled item obligatorily reconstructs:

(30) Bill hat jedes$_i$ Buch seinem$_{j/*i}$ Autor gegeben und jedes Magazine geschickt
Bill has every book its author given and every magazine sent
"Bill has given it's author every book and sent its magazine." (Binding not possible.)

This seems hopeful — I just need to find some actual SGF speakers!

5 Remaining questions

I've demonstrated that Khoekhoe has the SGF construction; I believe it to be the first non-Germanic language known with this property. How can we go about finding more?

- It's quite striking that both German and Khoekhoe have robust prefield / middlefield clause structure — it's tempting to link SGF to V2 in some way.
- However: Kashmiri is robustly V2 but doesn't allow SGF (Manetta, p.c.).
- Perhaps this is ok: If Kashmiri simply doesn't allow reconstruction, then we predict it not to have SGF. I'm still checking on this!
- I have preliminary evidence that Turkish has SGF (Özyıldız p.c.); Turkish is not V2, but is OV.
- However, some things about Turkish look wrong: Extraction might be possible from the second conjunct, for instance. Perhaps a different construction?

[6] Thanks to Kyle Johnson and NYU's own Haagen Blix for these German examples.

References

Bennett, Ryan, Emily Elfner, & James McCloskey. 2016. Lightest to the right: An apparently anomalous displacement in Irish. *Linguistic Inquiry* 47:169–234.

Bjorkman, Bronwyn. 2014. Accounting for unexpected subject gaps in TP coordination. *The Linguistic Review* 31.

Brugman, Johanna Christina. 2009. Segments, tones and distribution in Khoekhoe prosody. Doctoral Dissertation, Cornell University.

Büring, Daniel, & Katharina Hartmann. 1998. Asymmetrische Koordination. *Linguistische Berichte* 174:172–201.

Elfner, Emily. 2012. Syntax-prosody interactions in Irish. Doctoral Dissertation, University of Massachusetts Amherst.

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, chapter 7.3.2.1.1, 325–340. Routledge.

Hahn, Michael. 2013. Word order variation in Khoekhoe. In *Proceedings of the 20th International Conference on Head-Driven Phrase Structure Grammar*, 48–68.

Höhle, Tilman N. 1985. On composition and derivation: The constituent structure of secondary words in German. In *Studies in German grammar*, ed. Jindrich Toman, 319–376. Dordrecht, The Netherlands: Foris Publications.

Höhle, Tilman N. 1990. Assumptions about asymmetric coordination in German. In *Grammar in progress: GLOW essays for Henk van Riemsdijk*, ed. Joan Mascaró & Marina Nespor, 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.

Lin, Vivian. 2001. A way to undo A movement. In *Proceedings of the 20th West Coast Conference in Formal Linguistics*, ed. Karine Megerdoomian & Leora Anne Bar-el, 358–371. Somerville, Massachusetts: Cascadilla Press.

Ross, John. 1967. Constraints on variables in syntax. Doctoral Dissertation, Massachusetts Institute of Technology.