Distinguishing Types of Word Order Complexity: The Shapes of Zippers

Tim Hunter
Department of Linguistics — University of California, Los Angeles

How does the word-order complexity of Australian languages compare with that of other "flexible" word-order languages?

Recent computational research by Wedekind & Kaplan [WK20] builds a bridge between:

- the LFG framework, used in much research on Australian languages, and
- the LCFRS formalism, used in much research on "free word order" in other kinds of languages (e.g. German, Dutch, Turkish, …).

I use this connection to identify **two apparent differences** between the Australian type of word-order complexity and the type that has been the focus of LCFRS work, one where the Australian type is *less complex* and one where it is *more complex*.

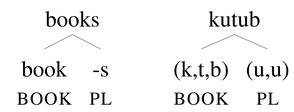
	Depth of zippers	Width of zippers
Australian-style discontinuity	bounded	unbounded
Germanic-style discontinuity	unbounded	bounded

These observations add to the theoretical significance of **specific empirical questions** concerning:
- clausal subordination
- discontinuous nominal expressions

Discontinuous constituents are the key to connecting across frameworks

Discontinuity in LCFRS

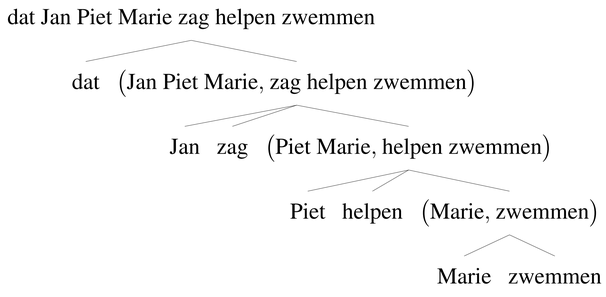
- In a standard phrase-structure grammar, assembling constituents is just concatenating strings, like in simple concatenative morphology.
- In LCFRS, assembling constituents can be a more complex string-building operation, like in non-concatenative morphology.



LCFRS incorporates some **substantive hypothesised limitations on the patterns of discontinuity that can be produced**, motivated by observations of what is and isn't found in certain (non-Australian) languages.

For example it allows the famous Dutch crossing-dependencies pattern [H76], but rules out many conceivable unattested options [K10].

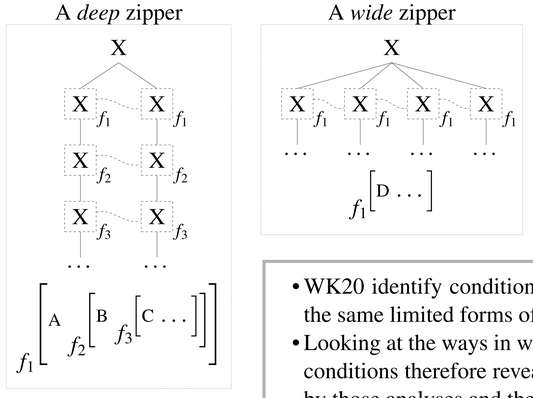
(1) …dat Jan Piet Marie zag helpen zwemmen
…that Jan Piet Marie saw help swim
"…that Jan saw Piet help Marie swim"



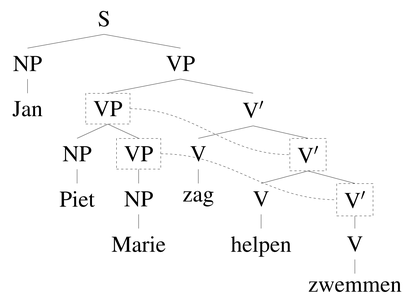
Discontinuity in LFG

LFG implements discontinuous constituency by having distinct c-structure nodes *that are not related by dominance* map to a **shared f-structure**, creating a configuration known as a "zipper". The dashed lines in (3) show the nodes that are linked in this way in the LFG account of the Dutch crossing-dependency construction [B82].

(2)



(3)



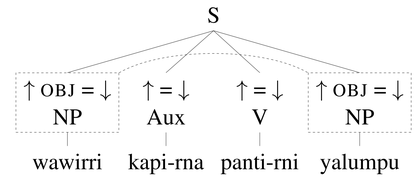
- WK20 identify conditions under which the zippers created by an LFG analysis implement exactly the same limited forms of discontinuity that LCFRS allows.
- Looking at the ways in which LFG analyses of Australian languages do and do not conform to those conditions therefore reveals specific points of difference between the word-order patterns expressed by those analyses and the patterns that motivate the LCFRS formalism's restrictions.

Difference 1: *Depth* of zippers

Languages like Warlpiri show very free word order *within clauses* [H83], to the extent of allowing discontinuous NPs:

(4) **wawirri** kapi-rna panti-rni **yalumpu**
kangaroo AUX spear-NONPAST that
"I will spear that kangaroo"

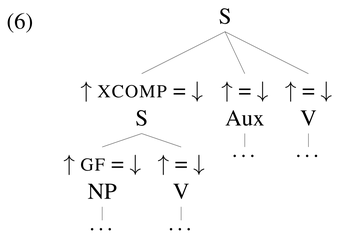
(5)



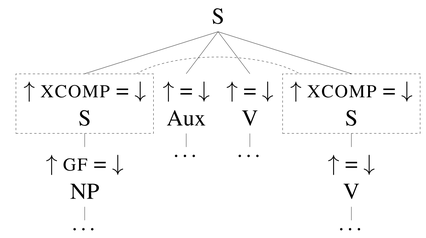
But when a multi-word grammatical element is a subordinate clause, it is (almost?) always contiguous [N06].

Or stated in LFG terms: we do not find pairs of horizontally-related S nodes that map to a shared f-structure: subordinate clauses appear contiguously as in (6), but generally not in a discontinuous form like (7), which would be parallel to (5).

(6)



(7)



- If this generalization is correct and S nodes disallow "open" zippers, then this would amount to a bound on the *depth* of the zippers that produce discontinuities.
- The resulting kind of word-order complexity would be importantly different from the kind that LCFRS produces, which corresponds to zippers of unbounded depth: this is what allows discontinuous noun-verb pairs to accumulate in (3).

Difference 2: *Width* of zippers

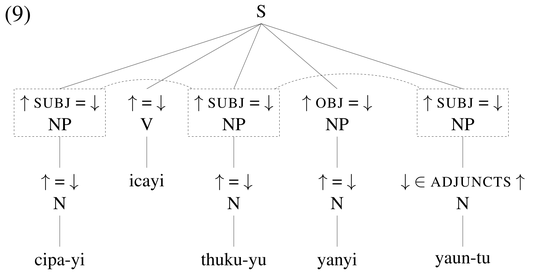
LCFRS allows intermingling between unboundedly many discontinuous elements (deep zippers), but requires that there is a **bound on the number of pieces** into which a discontinuous element is "split" (the width of a zipper).

An important empirical question, therefore, is whether *discontinuous NPs can be split up into more than the two pieces* we see in familiar examples like (4) and (8) (from Kalkatungu [B83, cited in N14]).

LFG analyses generally predict that more than two pieces should be possible: any number of horizontally-related NP nodes might contribute to a unified f-structure (e.g. via a set-valued ADJUNCTS attribute) in a structure like (9).

(8) a. **cipa-yi thuku-yu yaun-tu** yanyi icayi
this-ERG dog-ERG big-ERG white.man bite
"This big dog bit/bites the white man."
b. **cipa-yi thuku-yu** yanyi icayi **yaun-tu**
c. **thuku-yu cipa-yi** icayi yanyi **yaun-tu**
d. **yaun-tu cipa-yi thuku-yu** icayi yanyi
e. **cipa-yi** icayi **yaun-tu thuku-yu** yanyi
f. icayi **yaun-tu cipa-yi thuku-yu** yanyi

(9)



This would mean that analyses of Australian-style discontinuous NPs require zippers of *unbounded width*, unlike the discontinuities modeled by LCFRS.

AB96 Austin & Bresnan 1996, "Nonconfigurationality in Australian Aboriginal languages", *NLLT*. **B82** Bresnan et al. 1982, "Cross-serial dependencies in Dutch", *LI*. **B83** Blake 1983, "Structure and word order in Kalkatungu", *AJL*. **H76** Huybregts 1976, "Overlapping dependencies in Dutch", *Utrecht Working Papers*. **H83** Hale 1983, "Warlpiri and the grammar of non-configurational languages", *NLLT*. **K10** Kallmeyer 2010, *Parsing Beyond Context-Free Grammars*. **N06** Nordlinger 2006, "Spearing the Emu Drinking", *AJL*. **N14** Nordlinger 2014, "Constituency and Grammatical Relations in Australian languages", *The Languages and Linguistics of Australia*. **S91** Simpson 1991, *Warlpiri Morpho-Syntax*. **WK20** Wederkind & Kaplan 2020, "Tractable LFG", *Comp. Ling.*.