**Apparent Discontinuous NPs in the Cariban family from the perspective of diachronic typology**

Matter, Gildea, & Cáceres Arandia

(or, if you prefer, Gildea, Matter, & Cáceres Arandia)

Apparent discontinuous NPs are attested in different Cariban languages (Payne 1993; Meira 1999; Meira & Gildea 2009). Since their components are nominal, they may actually not be NPs, but rather consist of two coreferential nouns in apposition (Gildea & Cáceres Arandia 2022). In either case, an important analytical question is how speakers resolve which non-continuous nouns “belong together”. While these constructions readily occur in elicitation, their occurrence in discourse and the factors conditioning it have not been investigated.

The proposed method to investigate said patterns is to examine 1000 clauses each in narrative discourse corpora of 5 Cariban languages: Panare (Cáceres Arandia 2022), Tiriyó (Meira 2022), Hixkaryána (Derbyshire 1965), Ye’kwana (Cáceres 2014), and Yawarana (Cáceres Arandia 2021). The expectation is that discontinuous NPs are exceedingly rare. As for conditioning factors, it is expected that some pragmatic patterns can be identifiable.

As for the analytical question of how referent resolution works, the hypothesis is that it is made possible by discontinuous NPs being restricted to P and S position, not being allowed in A position. While discontinuous obliques can be elicited, they are expected to either not occur in discourse, or to feature the oblique marker on both nouns. The expected pattern is thus that “stray” nouns belong to S in intransitive sentences, and to P in transitive ones.

These expected asymmetries arise out of patterns of language use: A typically features known referents, and needs no further elaboration, whereas S and P are used for introducing new referents (DuBois 1987). This allows for the occurrence of afterthought constructions in said roles, where a clarifying second noun is added. Over time, this pattern may become conventionalized, resulting in discontinuous NPs (or coreferential nouns), no longer being an afterthought.

* Discontinuous NPs appear to be common in Cariban languages
  + Alternative analysis: Not actually NPs, but coreferential nouns in apposition
  + Analytical Question: how do speakers know which Ns are co-referring?
  + Empirical Questions: How common are these things outside of elicitation and what conditions their occurrence?
* Methodology: Examine 1000 clauses of narrative discourse in 5 Cariban languages
  + Panare (Cáceres ELDP corpus)
  + Tiriyó (Meira corpus)
  + Hixkaryana (Derbyshire 1961 corpus)
  + Ye’kwana (Cáceres PhD corpus)
  + Yawarana (Cáceres NSF-DEL corpus)
* Hypotheses for the Empirical Questions (frequency & conditions of occurrence)
  + These are actually pretty damn rare
  + Goal: identify pragmatic conditions that correlate with their occurrence (same conditions as afterthought)
* Hypothesis for the Analytical Question of identifying co-reference: different GRs have different patterns of “discontinuity”
  + We expect most examples to be for S and P (easily elicited)
    - Probably as an afterthought construction?
  + For obliques, rare but attested (allowed in elicitation, but not automatically)
    - Most likely both elements will be marked with the same postposition
  + For A, we expect no discourse examples (not accepted in elicitation)
* Programmatic diachronic hypotheses
  + These asymmetries arise out of patterns of language use
    - A is usually continuing referent, not in need of further qualification
    - Most discontinuous NPs come in initial mentions of S and P, w/ afterthought
      * The afterthought creates a constructional “slot” for co-referring expressions
    - Obliques
      * Probably initial or disambiguating mentions
      * Afterthoughts would require double-marking to clarify (co‑)reference
  + Different patterns of language use are driving the creation of (continuous) NPs (Gildea & Cáceres 2022)
    - Adding Determiners (Always initial — language contact?)
    - Apparently obligatory order of Head-Modifier (varies in different languages)?
* Caceres, Natalia. 2014. “Documentation of Ye’kwana in the Erebato river in Venezuela”. Endangered Languages Archive. Handle: [http://hdl.handle.net/2196/00-0000-0000-0002-59D8-4](https://hdl.handle.net/2196/00-0000-0000-0002-59D8-4”).
* Cáceres Arandia, Natalia. 2022. “A comprehensive documentation of Panare, a Cariban language of Venezuela”. Endangered Languages Archive. Handle: <http://hdl.handle.net/2196/1c9255c2-1f37-40f4-8b6e-1b5c1f8252d9>.
* Derbyshire, Desmond C. 1965: *Textos Hixkaryâna* (Publicações Avulsas). Vol. 3. Belém: MPEG.
* Du Bois, J.W. 1987. “The Discourse Basis of Ergativity”. *Language*, 63(4), pp.805–855.
* Meira, Sérgio. 2022. “Tiriyó shoebox database”.
* Payne, Doris L. 1993: “Nonconfigurationality and discontinuous expressions in Panare”. In David Peterson (Ed.), *Proceedings of the 19th Berkeley Linguistics Society, Special volume on Syntax of Native American Languages*, 122–138.

**Situating Constituency Within Typological Research**

Convenors: Jeff Good & Adam Tallman

The goal of this workshop is to launch a discussion of constituency from a typological perspective, especially in light of recent advances in quantitative approaches to typology and the study of grammar. Questions to be considered include what features of theoretical approaches to constituency can be usefully adapted for typological investigation, what kinds of coding and data analysis strategies are needed to study constituency typologically, and how the tools of typology can be used to shed light on traditional distinctions, such as the contrast between word and phrase, and determine the extent to which such categories can be considered cross-linguistically valid. An additional consideration will be the extent to which observed typological patterns of constituency can be linked to patterns of language use. Contributions exploring concerns emanating from formal linguistic theories, such as so-called bracketing paradoxes or discontinuous constituents, from a typological perspective, are also welcome. While the focus of the workshop is the typology of constituency, papers focusing on specific languages will also be considered as long as they situate their results with respect to broader typological concerns.

Maybe this bit in red is too much to take on in one talk