Large-scale pied-piping and the structure of embedded clauses in Karitiana

**Goal:** The goal of this presentation is to describe and analyze long-distance WH- questions in Karitiana, a Tupian language spoken in Brazil. Specifically, we show how the strategy used for question formation - large-scale pied-piping - reveals important properties of embedded clauses in this language.

**Problem:** In order to build questions, Karitiana fronts the WH- word (Landin 1984, Storto 1999). Nonetheless, when the WH- element comes from an embedded clause (a construction known as ‘long-distance question’), the whole subordinate clause has to be fronted:

1. Morã-ty aj-andyj?
   WH-OBL 2-laugh
   “Who did you laugh at?” [Storto 2008]

2. [Morã Ana oky]-ty i-kãra Karin?
   WH Ana kill-OBL 3-think Karin
   “What does Karin think that Ana killed?”

3. ?*Morã i-kãra Karin Ana mi-ty?
   WH 3-think Karin Ana hit-OBL
   (Intended) “What does Karin think that Ana hit?”

The strategy in 2 - known as **large-scale pied-piping** - is present in languages such as Imbabura Quechua and Basque (Cole 1982, Ortiz de Urbina 1993). The problem posed by 2-3 is the requirement of clausal pied-piping: given that it is possible to move only the WH- word in simple questions, why does Karitiana have to move the whole embedded clause in long-distance questions?

**Proposal:** We claim that the obligatoriness of large-scale pied-piping in long-distance questions is rooted in the structure of embedded clauses. Building on the proposal that pied-piping is a last-resort strategy (Heck 2009), we propose that large-scale pied-piping is a strategy to overcome a Complex Noun Phrase violation (Ross 1967). In our analysis, embedded clauses in Karitiana would be nominalized to a certain degree (in the sense of Baker 2011), since they would be headed by a nominal element that cannot be crossed when a WH- moves from an embedded clause:

4. [CP ... [NP [IP ... WH ... ] N ] ]

Evidence for the proposed nominal status of embedded clauses are: (1) Embedded clauses are always non-finite, and certain verbal morphemes (mood and agreement) are obligatorily absent; (2) Certain subordinate clauses are marked with a nominalizer (the suffix -pa); (3) Person markers in embedded clauses resemble (and may be) possessive markers; (4) Embedded clauses bear case markers as any other noun phrase in the language. These features are exemplified below:

5. [Boroja taso oky tykir] 0-naka-hyrp-0 0wa
   snake man kill ASP.PERF. 3-DECL-cry-NFUT child
   “When the man killed the snake, the child cried.” [Storto 1999]

6. Y-pyr-amynt-yn [[cama taso kat]<a> pa]-ty
   1-ASS-buy-NFUT [bed man sleep]<e.v.>NMZ-OBL
   “I bought the bed where the man slept.”
“You thought I listened.” [Storto 1999]

“I want the corn that Luciana ground.”

**Additional argument:** Our proposal can shed some light onto the presence of a morpheme *a*-that sometimes surfaces in embedded clauses. We claim that even though the nominal head of embedded clauses is usually null, this morpheme *a-* may be the overt realization of this nominal head in some cases. Besides, it would be reminiscent of a nominalizer present in Tupian languages (usually known as “argument case” following Rodrigues 1996), whose allomorphs could be either *a*- or *∅*- in Tupinambá, an ancient Tupian language. In Karitiana, *a* may surface in a noun phrase and also in an embedded clause (and as can be seen in 11, this morpheme is not part of the root):

(9) Ombaky by-'edn-a
jaguar CAUS-son-NMZ

“Dog” (literally: “jaguar turned into son”)

(10) [Jonso húryj- a ] ∅-na-aka-t i-se’a-t
woman beautiful-NMZ 3-DECL-COP-NFUT NMZ-beautiful-COP.AGR.

“The woman who sang is beautiful.” [Rocha 2016]

(11) Atykiri ∅-naka-húryj-∅
taso.
then 3-DECL-sing-NFUT man

‘Then the man sang’ [Rocha 2011]

**In sum:** Pied-piping of the whole embedded clause is required for long-distance WH- questions in Karitiana. We claim that this strategy is used to circumvent a Complex Noun Phrase violation, as embedded clauses are nominalized to a certain degree (as can be detected by their nominal features). Additionally, this nominal element heading embedded clauses may also explain the presence of a morpheme *a*- that sometimes surfaces in embedded clauses.