

## Clausal pied-piping in Karitiana\*

Karin Vivanco  
University of São Paulo

**Abstract:** The goal of this paper is to discuss the formation of long-distance questions in Karitiana (Tupi). It will be shown that these complex questions involve an operation called clausal-piping, which fronts the whole embedded clause containing the interrogative pronoun. In these constructions, the subordinate clause behaves as the *Wh*- itself, since it receives all the *Wh*-morphology that usually attaches to *Wh*-words in simple questions. Clausal pied-piping also seems to be a strategy to overcome an island violation, as embedded clauses in Karitiana behave like complex noun phrases. Additionally, it will be showed that focus triggers clausal pied-piping. Finally, we argue that unlike other languages for which the existence of clausal-piping has been questioned, the examples depicted here cannot be subsumed under an analysis of clausal pied-piping as a general fronting operation.

**Keywords:** Karitiana, questions, pied-piping, embedded clauses, islands

Complex sentences involving embedded clauses have been a topic of interest in many indigenous languages. Nevertheless, phenomena involving these structures are still under-documented and poorly understood. One of these under-documented phenomena is the so-called long-distance question, in which the interrogative pronoun comes from an embedded clause:

- (1) What<sub>i</sub> did John say [ that Mary bought  $t_i$  ]?

In Karitiana,<sup>1</sup> only simple questions have been described so far. In this paper, we describe the formation of long-distance questions and argue that they involve a strategy known as **clausal pied-piping**, i.e., *Wh*-movement of the whole embedded clause containing the interrogative pronoun. If this analysis is correct, Karitiana would belong to a restricted group of languages that allow *Wh*-movement of larger constituents.

In Section 1, we lay the foundation for the subsequent sections by describing simple *Wh*-questions in Karitiana. Embedded clauses are detailed in section 2 with special attention to the four types involved in our data. The long-distance questions out of these embedded clauses are presented in Section 3. In Section 4, it will be argued that these long-distance questions involve *Wh*-fronting of the embedded clause containing the interrogative pronoun, a process firstly described for languages like Basque and Imbabura Quechua and known as clausal pied-piping. The relation between clausal-piping and islandhood is discussed in Section 5, and we show evidence that embedded clauses in Karitiana behave like complex NPs. Additionally, we show in Section 7 that

---

\* This research was supported by FAPESP (grant numbers: 2014/15141-1 and 2016/07643-2). I also thank Luciana Storto, Norvin Richards, Ana Müller, Thomas Finbow, Jairo Nunes, Mar Bassa, Suzana Fong, the audience of the VIII Conference on Indigenous Languages of Latin America, and all members of the Group of Indigenous Languages of the University of São Paulo for the helpful comments on earlier versions of this work. I am also indebted to the Karitiana community, who was willing to share their language and knowledge with me. Needless to say, any remaining errors are my responsibility.

Contact info: karincv@usp.br

<sup>1</sup> Karitiana is a Tupian language of the Arikém family (Rodrigues 1986), spoken in northwestern Brazil by around 317 people (Storto and Rocha 2017).

just like other languages with this operation, focus of embedded constituents also triggers clausal pied-piping. We also claim that these long-distance questions are truly cases of *Wh*-movement, and not a kind of left-dislocation of the embedded clause in Section 8. Section 9 summarizes our main findings.

## 1 Question formation

Since Landin's (1984) first description of the language, Karitiana has been classified as an ergative language. Question formation is one evidence of its ergative alignment, as *Wh*-questions are sensitive to the case of the extracted argument. Specifically, Storto (2008, 2010) claims that there are two strategies for question formation: one for absolutes (intransitive subjects and direct objects) and another for any non-absolute element (oblique objects, transitive subjects, and adverbs).

In the non-absolute strategy, the *Wh*-word is fronted and the verb only has agreement prefixes:<sup>2,3</sup>

(2) Morā y-sokōĩ?  
 WH- 1-tie.up  
 "Who tied me up?" [Storto 2008]

(3) Morā-ty aj-andyj?  
 WH-OBL 2-laugh  
 "Who did you laugh at?" [Storto 2008]

The absolute strategy is somewhat more complex and involves a process called cleft by Storto. Namely, the *Wh*-word is marked with the suffix *-mon* and the verb has either a prefix *i-* if it is an intransitive subject or *ti-* if it is a direct object. In both cases, the verb is also suffixed by the copular marking *-t/-Ø*, whose allomorphy is conditioned by the last segment of the verbal root:

(4) Mora-mon i-oty-t?  
 WH-INT.COP NMZ-bathe-COP.AGR  
 "Who bathed?" [Storto 2008]

(5) Mora-mon an ti-opi-t?  
 WH-INT.COP 2 INV-cut-COP.AGR  
 "What did you cut?" [Storto 2008]

Since the *Wh*-morphology will be relevant for cases of long-distance extraction, these affixes are summarized in the table below:

<sup>2</sup> We focus on question formation in the non-future, since our long-distance examples are inflected for this tense. For question formation in the future, see Storto (2010, 2008).

<sup>3</sup> The following abbreviations are used in this paper: 1 = 1st person singular; 2 = 2nd person singular; 1PL = 1st person plural; 2PL = 2nd person plural; 3 = 3rd person; ASS = assertive mood; CAUS = causativizer; COP = copula; COP.AGR = copular agreement; DECL = declarative mood; DES = desiderative; FUT = future; INT.COP = interrogative copula; INV = inverse; IMPERF.MOT = imperfective motion; IMPERF.SITTING = imperfective sitting; NFUT = non-future; NMZ = nominalizer; OBL = oblique; PERF.ASP = perfective aspect; WH = interrogative pronoun; <e.v.> = epenthetic vowel. Data from other authors are quoted with the original glosses.

**Table 1:** *Wh*-morphology

ARGUMENT EXTRACTED		<i>Wh</i> -	V
Non-absolutive	Transitive Subject	WH	AGR-V
	Oblique Object	WH- <i>ty</i>	AGR-V
Absolutive	Intransitive Subject	WH- <i>mon</i>	NMZ-V-COP.AGR
	Direct Object	WH- <i>mon</i>	INV-V-COP.AGR

## 2 Embedded clauses

Karitiana exhibits important asymmetries between matrix and embedded clauses. According to Storto (1999), simple matrix clauses are verb-second in the majority of cases,<sup>4</sup> whereas embedded clauses are strictly verb-final:

- (6) Yn a-ta-oky-j an.  
 1 2-DECL-kill-FUT 2  
 “I will kill you.” [Storto 1999]

- (7) Yn  $\emptyset$ -na-aka-t i-sondyp- $\emptyset$  Inácio ’ep opĩ-ty.  
 1 3-DECL-COP-NFUT NMZ-know-COP.AGR Inácio tree cut-OBL  
 “I know that Inácio cuts the tree.” [Storto 2011]

Verbal morphology is also another important difference, since only matrix clauses allow fully inflected verbs.

Four types of embedded clauses were investigated: embedded clauses as direct objects such as (8), as oblique objects (9), as relative clauses (10), and as adjunct clauses (11):

- (8) Ivan  $\emptyset$ -naka-sadn- $\emptyset$  [ Ana obaky oky. ]  
 Ivan 3-DECL-tell-NFUT Ana jaguar kill  
 “Ivan reported that Ana killed a jaguar.”
- (9) Karin  $\emptyset$ -na-aka-t i-koro’op kãra-t [Ana médico mĩ]-ty.  
 Karin 3-DECL-COP-NFUT NMZ-inside suspect-COP.AGR Ana doctor hit-OBL  
 “Karin thinks that Ana hit a doctor.”
- (10) Yn  $\emptyset$ -naka-’y-t [ kinda’o taso ti-amangã tyĵa. ]  
 1 3-DECL-eat-NFUT fruit man INV-plant IMPERF.MOT  
 “I ate (some of the) fruits that the man is planting.” [da Silva 2016]
- (11) [ Ambi-sok ĵonso otam byyk ]  $\emptyset$ -na-oky-t him taso.  
 house-in woman arrive PERF.ASP 3-DECL-kill-NFUT game man  
 “After the woman had arrived, the man killed the game.” [da Silva 2016]

<sup>4</sup> Some sentence types require verb-initial orders, such as intransitive verbs in non-copular constructions and the assertive mood (see Everett 2006; Storto 2014).

### 3 Long-distance questions

Long-distance questions formed out of these embedded clauses are shown below:

- (12) [Morã Karin ti-ko]-mon João ti-sadnan-Ø?  
 WH Karin INV-break-INT.COP João INV-tell-COP.AGR  
 “What did João reported that Karin broke?”
- (13) [Morã pop ’it]-<i>ty i-kāra Karin?  
 WH die do-<e.v.>OBL 3-suspect Karin  
 “Who did Karin think fainted?”
- (14) [Morã ti-m-’a]-mon syke Luciana ti-’y-t?  
 WH INV-CAUS-make-INT.COP porridge Luciana INV-eat-COP.AGR  
 “Who made the porridge that Luciana ate?”  
 (Literally: “[The porridge that who made] Luciana ate?”)
- (15) [ Morã Thiago by-hip byyk ] i-pa’ira Luciana?  
 WH Thiago CAUS-cook PERF.ASP 3-get.angry Luciana  
 “Thiago’s having cooked what got Luciana angry?”  
 (Literally: “[After Thiago had cooked what] Luciana got angry?”)

In the examples above, the whole embedded clause containing the *Wh*-element seems to be fronted. Moreover, the embedded clause seems acts as a *Wh*-word, since it receives all the morphology that *Wh*-elements usually have in questions, such as *-mon* in (12) and *-ty* in (13).

Three pieces of evidence build up this conclusion. First, it seems to be marginal to move the *Wh*- only:

- (16) \*Morã<sub>i</sub> i-kāra Karin [<sub>i</sub> Ana mĩ]-ty?  
 WH 3-suspect Karin Ana hit-OBL  
 (Intended “What did Karin think that Ana killed?”)
- (17) \*Morã<sub>i</sub> João ti-sadnan-Ø [<sub>i</sub> Karin ti-ko]-mon?  
 WH João INV-tell-COP.AGR Karin INV-break-INT.COP  
 (Intended “What did João reported that Karin broke?”)<sup>5</sup>

The second evidence is the *Wh*-morphology that arises in these long-distance data. As seen before, the *Wh*-morphology is sensitive to the case of the argument extracted, i.e., whether it is absolutive or non-absolutive. Crucially, the *Wh*-morphology of the matrix verb makes reference to the syntactic function of the embedded clause. In other words, the case of the *Wh*-word is irrelevant, as the real extraction revolves around the embedded clause. We examine this pattern in more detail with the examples below.

In (18), the matrix verb *kāra* requires an oblique argument marked with *-ty* — a role fulfilled by the whole embedded clause. Crucially, the *Wh*-word is a direct object (an absolutive) in this

<sup>5</sup> Some speakers parse it as two independent clauses, such as “What did João reported and Karin broke?”.

example, but there is no sign of the morphology required for absolutive extraction (namely, *-mon*, *ti-*, and *-t*). In fact, the main verb is only marked with agreement, resembling the oblique argument extraction in (19):

(18) [Morã Ana oky]-ty i-kāra Karin?  
 WH Ana kill-OBL 3-suspect Karin  
 “What does Karin think that Ana killed?”

(19) Morã-ty aj-andyj?  
 WH-OBL 2PL-laugh  
 “Who did you laugh at?”

[Storto 2008]

The long-distance question below has an interrogative pronoun as direct object inside an adjunct clause. Again, no absolutive morphology arises on the matrix verb, and its form is strikingly similar to extraction of adverbs, as in (21):

(20) [ Morã Thiago by-hip byyk ] i-pa'ira Luciana?  
 WH Thiago CAUS-cook PERF.ASP 3-get.angry Luciana  
 “Thiago’s having cooked what got Luciana angry?”  
 (Literally: “[After Thiago had cooked what] Luciana got angry?”)

(21) Tikat i-harãxa aka gooj-<o> Ana?  
 How 3-fix COP car-<e.v.> Ana  
 “How did Ana fix the car?”

Therefore, what the *Wh*-morphology shows is that the element being extracted is the whole embedded clause, not the *Wh*-word. This would also explain why the morphemes that usually suffix the interrogative pronoun in simple question (i.e., *-mon* and *-ty*) appear on the embedded clause.

Finally, the last evidence for movement of embedded clauses is the fact that fully-inflected clauses tend to be verb-second (as shown in Section 2, example 6). If one assumes that embedded clauses occupy the first position in long-distance questions, then the position of the matrix verb follows. If no such operation is posited, its position would hardly be accounted for.

#### 4 Clausal pied-piping

The behavior of long-distance questions in Karitiana resemble other documented cases of a phenomenon known as **clausal pied-piping**. Pied-piping is a construction first described by Ross (1967), in which additional elements are dragged along with the *Wh*-word that triggered movement (see also Horvath 2006):

(22) [ Whose book ]<sub>i</sub> did you read *t<sub>i</sub>*?

In clausal pied-piping, the fronted constituent is the whole embedded clause. Even though it seems to be quite rare, this phenomenon has also been reported for several non-related languages, such as Basque (Arregi 2003; de Urbina 1993), Imbabura Quechua (Cole 1982), Latin (Heck 2008), Bangla (Simpson and Bhattacharya 2000), and Tlingit (Cable 2010). Two examples are provided below:

(23) [ Nor etorriko d-ela bihar ] esan diozu Mireni?  
 who come AUX-that tomorrow said AUX mary-D  
 “That who will come tomorrow have you told Mary?” [Basque, de Urbina 1993]

(24) [ Wáa kwligeyi xáat ] sá i tuwáa sigóo?  
 how it.is.big.REL fish Q your spirit.at it.is.happy  
 “A fish that is how big do you want?”  
 (Literally: “[How big a fish] do you want?”) [Tlingit, Cable 2010]

Therefore, *Wh*-movement of embedded clauses in Karitiana would be similar to these languages. However, there is one important difference: in some languages cited above, clausal pied-piping seems to be optional in most cases. This optionality can be seen in these examples from Imbabura Quechua:

- (25) a. Ima-ta-taj ya-ngui [ Juan randi-shka-ta ]?  
 what-ACC-INT think-2 Juan buy-NMLZ-ACC  
 “What do you think that Juan bought?” [Imbabura Quechua, ?]
- b. [ Ima-ta Juan randi-shka-ta-taj ] ya-ngui?  
 what-ACC Juan buy-NMLZ-ACC-INT think-2  
 “What do you think that Juan bought?” [Imbabura Quechua, ?]

In Karitiana, clausal pied-piping is mandatory. In other words, only structures like (25b) are possible, and (25a) is ruled out (cf. 16 and 17). In the next section, this difference will be discussed in more detail, and it will be claimed that the structure of embedded clauses in Karitiana forces clausal pied-piping in all cases.

## 5 Clausal pied-piping and islands

Even if we accept that Karitiana has clausal pied-piping, one question remains: why is it impossible to move the interrogative pronoun only? This pattern turns out even more mysterious if matrix questions are taken into account, as interrogative pronouns can easily be fronted in these constructions. In this scenario, the conclusion has to be that embedded clauses in Karitiana are somehow opaque for movement.

The nature of this restriction may lie in the behavior of other clausal pied-piping languages: At least in some of them, clausal pied-piping is a strategy to overcome violations of some types of islands. In Imbabura Quechua, for example, it was seen in examples (25a) and (25b) that clausal pied-piping is optional in most cases. However, when the embedded clause is an island, such as a relative clause (see Ross 1967), only clausal pied-piping is possible:

- (26) a. \*Ima-ta-taj riku-rka-ngui [randi-shka runa-ta]?  
 what-ACC-INT see-PAST-2 buy-NMLZ man-ACC  
 “What did you see the man who bought?” [Imbabura Quechua, ?]
- b. [ Ima-ta randi-shka runa-ta-taj ] riku-rka-ngui?  
 what-ACC buy-NMLZ man-ACC-INT see-PAST-2  
 (Literally: “The man that bought what did you see?”) [Imbabura Quechua, ?]

The fact that clausal pied-piping is required in all long-distance questions in Karitiana would imply that embedded clauses are islands of some sort. Obviously, the question now is what type of island they are. As pointed out by Vivanco (2018), there is some evidence that embedded clauses are **complex noun phrases**. Here, we review some of her main arguments.

First of all, it was already mentioned that verbs in embedded clauses appear in a non-finite form (Storto 1999):

(27) Yn a-ta-oky-j an  
 1 2-DECL-kill-FUT 2  
 “I will kill you.” [Storto 1999]

(28) [ Boroja taso oky tykiri ]  $\emptyset$ -naka-hyryp- $\emptyset$  õwa  
 snake man kill PERF.ASP 3-DECL-cry-NFUT child  
 “When the man killed the snake, the child cried.” [Storto 1999]

Typically, erosion of clausal morphology is an indication of nominalization in some of the world’s languages (see Baker 2011; Borsley and Kornfilt 2000; Comrie and Thompson 2007; Koptjevskaja-Tamm 1993).

Person marking on the embedded verb may be another indication of the nominal status of embedded clauses. Storto (1999) showed that agreement is not possible in embedded clauses, as the impossibility of co-occurrence of *yn* and *y-* shows in (29). Nonetheless, it is possible to use a person marking attached to the verb in these environments — a morpheme that she analyzes as a cliticized pronoun.

(29) \* [ Yn y-opiso ] a-taka-kāra-t an.  
 1 1-listen 2-DECL-suspect-NFUT 2  
 “You thought I listened.” [Storto 1999]

(30) [ Y-opiso ] a-taka-kāra-t an.  
 1-listen 2-DECL-suspect-NFUT 2  
 “You thought I listened.” [Storto 1999]

However, this person marking is strikingly similar to the possessive markers found on nouns:

(31) Y-haj  
 1-brother  
 “My (younger) brother.” (male speaker) [Storto 1999]

Again, it is common cross-linguistically to express arguments of a nominalized verb through genitives (Koptjevskaja-Tamm 1993).

Another evidence for the nominal status of embedded clauses is that case-markers are possible on them as on any other noun phrase. In the examples below, one can see that the oblique morpheme *-ty* appears on any internal argument of the verb *pyting* — be they noun phrases or embedded clauses:

(32) Ø-py-pyting-yn    ãjonso    opi-ty.  
 3-ASS-want-NFUT woman earring-OBL  
 “The woman wants the earring.” [da Silva 2011]

(33) Yn Ø-na-aka-t            i-pyting-Ø            [gijo Luciana ti-tak]-<a>ty.  
 1 3-DECL-cop-NFUT NMZ-want-COP.AGR corn Luciana INV-grind-<e.v.>OBL  
 “I want the corn that Luciana ground.”

Finally, the last argument for the nominal analysis is the presence of overt nominalizers on some types of embedded clauses. Karitiana has a nominalizing suffix *-pa* that turns verbs into nouns (instruments or locations) (Everett 2006; Storto 1999):

(34) ambo - pa            (ambopa)  
 go.up - NMZ            (“stairs”) [Storto 1999]

This morpheme can be used in embedded clauses that have a locative meaning, like some indirect questions and relative clauses:

(35) Ana i-sondyp-<y>wak-Ø            Ivan pytim’adn-<i>pa-ty.  
 Ana NMZ-know-<e.v.>DES-COP.AGR Ivan work-<e.v.>NMZ-OBL  
 “Ana wants to know where Ivan works.”

(36) 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.”

If embedded clauses are complex nouns, extraction out of them would be a Complex Noun Phrase violation (see Ross 1967). In this sense, long-distance questions in Karitiana would be prohibited for the very same reasons that extraction out of a relative clause or of a clause headed by a noun is bad:

(37) \*The hat<sub>i</sub> which I believed [ the claim that Otto was wearing t<sub>i</sub> ] is red. [Ross 1967]

In sum, since embedded clauses behave as complex noun phrases, extraction out of them would constitute an island violation. If this is true, then clausal pied-piping would be the strategy to avoid ungrammaticality in long-distance questions. Crucially, noun phrases seem to be islands as well, given that possessors cannot be extracted out of them. In these cases, pied-piping is required:

(38) \*Morã-mon Ivan ti-’y-t            gok?  
 WH-INT.COP Ivan INV-eat-COP.AGR manioc

(39) Morã gok-<o>mon            Ivan ti-’y-t?  
 WH manioc-<e.v.>INT.COP Ivan INV-eat-COP.AGR  
 “Whose manioc did Ivan eat?”

Hence, the occurrence of pied-piping in long-distance questions would parallel extraction out of noun phrases, which in turn provides additional evidence for the nominal status of the former.

## 6 Factive complements

The fact that Karitiana has a strategy to overcome a Complex Noun Phrase Constraint makes important predictions. Specifically, it opens up the possibility of clausal pied-piping being used to avoid island violations of other types of nominal clauses.

One of such clauses is the clausal complement of a verbal class called factive verbs. Crosslinguistically, it was observed that factive verbs do not allow extraction out of their clausal complements (Kiparsky and Kiparsky 1971). Therefore, it is not possible to extract a constituent of clausal complements of verbs like ‘to regret’ or ‘to forget’, whereas extraction is possible with non-factive verbs such as ‘believe’:

(40) \*Who<sub>i</sub> do you regret/forget [ *t<sub>i</sub>* likes this book ]? [Adams 1985]

(41) Who<sub>i</sub> do you believe [ *t<sub>i</sub>* likes this book ]? [Adams 1985]

Crucially, it has been argued that this impossibility is due to the nominal status of these complements — possibly, their structure is something like “the fact of...” (see Adams 1985; Kiparsky and Kiparsky 1971; Zubizarreta 1982). These clausal complements also show several nominal features crosslinguistically, such as determiners (42) and the possibility of gerunds (43):

(42) Lamentamos el que Pedro no haya pasado el exámen.  
 “We regret that Pedro didn’t pass the exam.” [Spanish, Zubizarreta 1982]

(43) a. Sally regrets having agreed to the proposal. [Adams 1985]  
 b. \*Sally believes having agreed to the proposal. [Adams 1985]

In other words, clausal complements of factive verbs seem to be complex NPs. Given that Karitiana already has a strategy to handle extraction out of nominal clauses, we predict both that extraction out of clausal complements of factive verbs is possible and that these constructions will involve clausal pied-piping. These predictions are borne out, as can be seen in the examples below with the factive verbs ‘*koro’op oky*’ (‘to regret’) and ‘*diwyt*’ (‘to forget’):

(44) [Morã harahara ko]-ty i-koro’op oky Pedro?  
 WH mirror break-OBL 3-inside hurt Pedro  
 “Whose breaking the mirror did Pedro regret?”

(45) [Morã bykypa oky]-ty i-diwyt João?  
 WH cockroach kill-OBL 3-forget João  
 (Literally: “[Who killing the cockroach] did João regret?”)<sup>6</sup>

Summing up, clausal pied-piping in Karitiana opens up the possibility of extracting elements out of many opaque domains, such as complements of factive verbs.<sup>7</sup> Ultimately, it would be another indicative of how widespread the phenomenon of pied-piping is in this language.

<sup>6</sup> This sentence can also mean “Which cockroach/whose cockroach did João forget to kill?”.

<sup>7</sup> It is important to stress that our proposal does not imply that all clausal pied-piping languages must neces-

## 7 Focus

An additional argument showing that Karitiana indeed involves clausal pied-piping is the behavior of focus. As shown by Storto (1999), focused constituents occupy the pre-verbal position:

- (46) Mora-mon taso ti-'y-t?  
WH-INT.COP man INV-eat-COP.AGR  
“What did the man eat?” [Storto 1999]

- (47) OHY a-taka-'y-t taso.  
potato INV.DECL-DECL-eat-NFUT man  
“Potatoes, the man ate.” [Storto 1999]

Crucially, a focused element within an embedded clause also triggers clausal pied-piping. In the question-answer pair below, one can see that the only element focused is the object *ombaky*. Nonetheless, the whole embedded clause is fronted:

- (48) [Morã Ana oky]-ty i-kãra Karin?  
WH Ana kill-OBL 3-think Karin  
“What does Karin think that Ana killed?”
- (49) [OMBAKY Ana oky]-ty Ø-naka-kãra-t Karin.  
jaguar Ana kill-OBL 3-DECL-think-NFUT Karin  
“A jaguar is what Karin thinks that Ana killed.”

Similar facts have been reported for other clausal pied-piping languages, such as Basque (de Urbina 1993) and Bangla (Simpson and Bhattacharya 2000):

- (50) JONEK irakurri du liburua.  
Jon read has book  
“Jon has read the book.” [Basque, de Urbina 1993]
- (51) [ JON etorri den ] galdetu du.  
Jon come has.COMP asked have  
“I have asked whether it is Jon that has come.” [Basque, de Urbina 1993]

Therefore, focus constructions would be another evidence for the existence of clausal-piping in Karitiana.

sarily allow long-distance questions with factive verbs. In Basque, for instance, such sentences sound a little degraded according to de Urbina and Etxepare (2003):

- (i)??Nor etorri dela ohartu dira?  
who come AUX.COMP realiza AUX  
“That who has arrived have they realized?” [de Urbina and Etxepare 2003]

## 8 Clausal pied-piping as a general fronting mechanism

The conclusion that long-distance questions in Karitiana involve clausal pied-piping is an important one, as putative cases of CP pied-piping have been questioned for other languages. As a matter of fact, even languages that do not have clausal pied-piping have some sort of fronting mechanism that could move embedded clauses to the sentence-initial position. In English, for example, topicalization promotes CPs to the first position of a clause:

(52) [ Where did he go ], do you think? [Horvath 2006]

Due to the existence of these fronting operations in many languages, one could claim that the clausal pied-piping examples actually involve some sort of fronting operation (e.g., topicalization/focalization) already available in the language. For some paradigmatic cases of clausal pied-piping, the *Wh*-movement analysis has been questioned to the extent that Horvath (2006) claimed that “no uncontroversial evidence is available in favor of the existence of true CP pied-piping under *Wh*-movement.”

Embedded clauses in Karitiana may not be considered CPs due to the lack of clausal morphology. Nonetheless, they retain many clausal features, such as full argument structure, aspectual and evidential markers, and valency morphology (see da Silva 2016; Storto 1999; Vivanco 2018). Our position is then that the Karitiana examples do involve some kind of clausal pied-piping. Moreover, the existence of this clausal pied-piping is much clearer in Karitiana than in other languages.

As in many other languages, Karitiana has a general fronting strategy named **focus construction** by Storto (2008, 2010). In these, an internal argument is fronted and the verb is prefixed with either *a-* (if it is in the declarative mood) or *ti-* (if it is non-declarative):

(53) Tem Tema a-taka-m-tat- $\emptyset$  Botỹj.  
 Tem Tema INV.DECL-DECL-CAUS-go-NFUT Botỹj  
 “Tem Tema, Botỹj took.” [Storto 1999]

(54) Sepa y-ti-m-’a tyã-t.  
 basket 1-INV-CAUS-make IMPERF.SITTING-NFUT  
 “A basket, I am weaving.” [Storto 1999]

Given that this fronting operation already exists in the language, one could argue that the clausal pied-piping examples in (12) to (15) are actually cases of focalization, not of *Wh*-movement exactly. However, the morphology that arises in examples of long-distance questions are evidence that this cannot be the case in Karitiana. We showed in Section 1 that the verbal morphology in long-distance questions reflects the case of the embedded clause. As can be seen in the example below, long-distance questions trigger the occurrence of *Wh*-morphology, such as *-mon*. This is similar to cases of simple questions that involve *Wh*-movement, such as (56):

(55) [Morã Karin ti-kot]-<o>mon João ti-sadnan- $\emptyset$ ?  
 WH Karin INV-break-<e.v.>INT.COP João INV-tell-COP.AGR  
 “What did João reported that Karin broke?”

- (56) Mora-mon taso ti-'y-t?  
 WH-COP.INT man INV-eat-COP.AGR  
 “What did the man eat?” [Storto 1999]

Summing up, the Karitiana examples cannot be reanalyzed as an ordinary fronting operation, such as focalization, because their structure is identical to other questions involving *Wh*-movement of an interrogative pronoun. Hence, the Karitiana data depicted here would be an evidence for the existence of pied-piping of larger constituents, such as clauses.

## 9 Summary

In this paper, the formation of long-distance questions in Karitiana was described in detail. We argued that these involve an operation called clausal pied-piping, which fronts the whole embedded clause containing the *Wh*-word. Additionally, it was also shown that embedded clauses in Karitiana behave as complex NPs, and the occurrence of clausal pied-piping would be a strategy to avoid an island violation in these constructions. In addition, we showed that focus of embedded constituents also induces pied-piping, similarly to other clausal pied-piping languages. Finally, we also argued that clausal pied-piping cannot be accommodated as simple fronting operation, as it triggers all the morphology required in other instances of *Wh*-movement.

## References

- Adams, Marianne. 1985. Government of empty subjects in factive clausal complements. *Linguistic Inquiry* 16:305–313.
- Arregi, Karlos. 2003. Clausal pied-piping. *Natural Language Semantics* 11:115–143.
- Baker, Mark. 2011. Degrees of nominalization: Clause-like constituents in Sakha. *Lingua* 121:1164–1193.
- Borsley, Robert, and Jaklin Kornfilt. 2000. Mixed extended projections. In *Syntax and semantics: The nature and function of syntactic categories*, ed. Robert Borsley, 101–132. Bingley: Emerald.
- Cable, Seth. 2010. Against the existence of pied-piping: Evidence from Tlingit. *Linguistic Inquiry* 41:563–594.
- Cole, Peter. 1982. *Imbabura Quechua*. Amsterdam: North-Holland Publishing Company.
- Comrie, Bernard, and Sandra A. Thompson. 2007. Lexical nominalization. In *Language typology and syntactic description volume iii: Grammatical categories and the lexicon*, ed. Timothy Shopen, 334–381. Cambridge: Cambridge University Press.
- Everett, Caleb. 2006. Patterns in Karitiana: Articulation, perception, and grammar. Doctoral Dissertation, Rice University, Houston, TX.
- Heck, Fabian. 2008. *On pied-piping: Wh-movement and beyond*. Berlin: Mouton de Gruyter.
- Horvath, Julia. 2006. Pied-piping. In *The Blackwell companion to syntax*, ed. Martin Everaert and Henk Van Riemsdijk, 569–630. Oxford: Blackwell Publishing.

- Kiparsky, Paul, and Carol Kiparsky. 1971. Fact. In *Semantics: An interdisciplinary reader in philosophy, linguistics and psychology*, ed. Danny D. Steinberg and Leon A. Jakobovits, 345–369. Cambridge: Cambridge University Press.
- Koptjevskaja-Tamm, Maria. 1993. *Nominalizations*. London: Routledge.
- Landin, David J. 1984. An outline of the syntactic structure of karitiãna sentences. In *Estudos sobre línguas tupi do brasi*, ed. Robert Dooley, 219–254. Brasília: SIL.
- Rodrigues, Aryon Dall’Igna. 1986. *Línguas brasileiras: Para o conhecimento das línguas indígenas*. São Paulo: Edições Loyola.
- Ross, John Robert. 1967. Constraints on variables in syntax. Doctoral Dissertation, MIT, Cambridge, MA.
- da Silva, Ivan Rocha. 2011. A estrutura argumental da língua karitiana: Desafios descritivos e teóricos. Master’s thesis, Universidade de São Paulo, São Paulo.
- da Silva, Ivan Rocha. 2016. Não-finitude em karitiana: Subordinação versus nominalização. Doctoral Dissertation, Universidade de São paulo, São Paulo.
- Simpson, Andrew, and Tanmoy Bhattacharya. 2000. Wh-clausal pied piping in Bangla. In *Proceedings of NELS 30*, volume 30, 1–14.
- Storto, Luciana. 1999. Aspects of a Karitiana grammar. Doctoral Dissertation, MIT, Cambridge, MA.
- Storto, Luciana. 2010. Copular constructions in Karitiana: A case against case movement. In *Proceedings of SULA 5*, 205–226.
- Storto, Luciana. 2011. Subordination in Karitiana. *Ameríndia 35: La structure des langues Amazoniennes* 219–236.
- Storto, Luciana. 2014. Constituent order and information structure in karitiana. In *Information structure and reference tracking in complex sentences*, ed. Rik van Gijn, Jeremy Hannond, Dejan Matić, Saskia van Putten, and Ana Vilcay Galucio, 163–192. Amsterdam: John Benjamins Publishing.
- Storto, Luciana, and Ivan Rocha. 2017. Inventário sociolinguístico da língua karitiana. Inventário Nacional da Diversidade Linguística, IPHAN-MPEG. Unpublished manuscript.
- Storto, Luciana R. 2008. Marcação de concordância absoluta em algumas construções sintáticas em Karitiana. *Ameríndia 32: La structure des langues Amazoniennes*.
- de Urbina, Jon Ortiz. 1993. Feature percolation and clausal pied-piping. In *Generative studies in Basque linguistics*, ed. José Ignacio Hualde and Jon Ortiz de Urbina, 189–219. Amsterdam: John Benjamins Publishing.
- de Urbina, Jon Ortiz, and Ricardo Etxepare. 2003. Focalization. In *A grammar of Basque*, ed. José Ignacio Hualde and Jon Ortiz de Urbina, 459–516. Berlin: Mouton de Gruyter.
- Vivanco, Karin. 2018. Movimento QU-, subordinação e ordem de constituintes em Karitiana. Doctoral Dissertation, Universidade de São Paulo, São Paulo.
- Zubizarreta, Maria Luisa. 1982. Theoretical implications of subject extraction in Portuguese. *The Linguistic Review* 2:79–96.