Wh-question formation strategies in Guarani

Hunter Johnson (UCLA)
Marisabel Cabrera (UCLA)
05 JAN 2024 • SSILA 2024

1 Preliminaries

- Guarani is spoken by roughly five million people in Paraguay and border regions of Argentina/Brasil with a long prescriptive history.[1] However, there does not exist much *descriptive* work (cf. Estigarribia 2017; 2020 and other works cited below).

- What descriptive work exists does not go into much depth on *wh*-questions or question formation strategies in general.

- The goal of this work is to introduce Guarani to the descriptive literature on *wh*-movement by presenting the first systematic *description* of *wh*-question strategies in Guarani.

Today:

- introduce the following properties of Guarani *wh*-questions:
  - ⇒ *wh*-morphology, *wh*-particles
    - ✽ *wh*-words and their (de)composition
    - ✽ distribution of particles
  - ⇒ *wh*-movement (full) is obligatory (i.e. no partial-*wh*-copying)
    - ✽ intermediate *wh*-elements are not allowed
    - ✽ w/ w/o scope marker
  - ⇒ *wh*-movement (long-distance) leaves no morphological trace
    - ✽ intermediate *wh*-elements are not allowed
    - ✽ w/ w/o scope marker
  - ⇒ *wh*-islands are obeyed
    - ✽ crossing one *wh*-element w/ another banned
  - ⇒ *wh*-elements from embedded clauses exhibit clausal pied-piping
    - ✽ speaker variation (Concepción v. Caaguazu)
    - ✽ some clauses allow for pied-piping, others do not

- We'll conclude by drawing connections between these findings presented here and relative clause (RC) formation in Guarani (another type of $\overline{\text{A}}$-movement) which may have consequences for Guarani linguistics and theoretical linguistics more broadly.

2 Background

2.1 Language background

- We present data primarily from in-situ fieldwork in Coronel Oviedo and Caaguazu, Paraguay with 6 speakers (ages=30–60).

(1) *Map of Paraguay w/ Concepción, Coronel Oviedo, and Caaguazu:*

- Our speakers all speak Guarani as their first language and learned Spanish at a later age.

- Previous work includes: theoretical-oriented Monserrat and Soares (1983); Jensen (1990, 1998, 1999); Velázquez-Castillo (1991, 2002); Rose (2015); Tonhauser (2011a,b, 2020, 2006, 2007); Tonhauser and Colijn (2010); Clopper and Tonhauser (2013); Shain and Tonhauser (2010); Zubizarreta and Pancheva (2017a,b); Pancheva and Zubizarreta (2019); Zubizarreta (2022); Jun and Zubizarreta (2022); Jun et al. (2023) and descriptive/prescriptive (Anchieta 1595 [1979]; Aragona 1625 [1979]; Ruiz de Montoya 1724 [1876]; Restivo 1724 [1892]; Gregores and Suarez 1967; de Guarania 1997, 2008; Krivoshein de Canese and Alcaraz 2001, 2006; Zarratea 2002; Krivoshein de Canese and Alcaraz 2006: a.o.).

- With this strong prescriptive background, in doing fieldwork/public outreach on the language we do our best to bring attention to non-standard dialects.

- One thing we learned through our fieldwork is that, despite its official status, at a national level Guarani is losing government support, but local governments continue to appreciate its value (happy to talk more about this in the discussion period!).[2]

2.2 Linguistic background

- Guarani is highly-agglutinative and is mixed prefixing/suffixing as shown in (2).[3] The verb root is **bolded** in (2). Further, the template in (3) demonstrates that argument structure morphemes (REFL, ANTIP, CAUS, ϕ) appear as prefixes while TAM/Negation (DES, AUG, FUT, CMPL, PROG) appears as suffixes.[4]

[2] The new president of Paraguay has been quoted saying "Do people really still speak Guarani at all?" When, even in the capital Asunción, Guarani is heard on the streets.

[3] Some abbreviations: NEG = negation, ANTIP = antipassive, CAUS = causative, DES = desiderative, EMPH = emphatic, PST = past tense, AUG = augmentive, CMPL = completive, TOT = totalitative

(2) n-o-poro-mbo-guero-**guata**-se-i-terei-rasa-pa-kuri-hikoni
NEG-3-ANTIP-CAUS1-CAUS3-**walk**-DES-NEG-EMPH-EMPH2-TOT-PST-PST
'He didn't want to make him walk too much.'

---

(3) *Template for Guarani verbs:*

NEG - ϕ - ANTIP - CAUS - REFL - RECIP - V - DES - AUG - FUT - NEG - CMPL - PROG

- The *minimal* Guarani verb contains the root and ϕ-agreement marking (person and number) (4).

(4) a-guata
1-walk
'I walk(ed).'

- Person pronouns (5a) for objects and subjects are generally optional, as well as determiners generally (5b) (5c) for Definiteness, for example.

(5) a. (che) ai-pytyvõ (ichupe)
(I) 1-walk (him/her)
'I help(ed) him/her.'

b. (che) a-hecha jagua
(I) 1-see dog
'I saw a/the dog.'

c. (che) a-hecha **pe** jagua
(I) 1-see **the** dog
'I saw the dog.'

- One interesting, but underrecognized, characteristic of Guarani is that it has scrambling of free word order (Tonhauser and Colijn 2010). In a transitive clause, all six orders are possible (6).[5]

(6) a. che ai-pytyvõ Juan-pe
I 1.SUBJ-help Juan-DOM
'I help Juan.' SVO

b. che Juan-pe ai-pytyvõ
I Juan-DOM 1.SUBJ-help SOV

c. Juan-pe che ai-pytyvõ
Juan-DOM I 1.SUBJ-help OSV

d. ai-pytyvõ che Juan-pe
1.SUBJ-help I Juan-DOM VSO

e. ai-pytyvõ Juan-pe che
1.SUBJ-help Juan-DOM I VOS

f. Juan-pe ai-pytyvõ che
Juan-DOM 1.SUBJ-help I OVS

[5] In fact, there is also very free word order in intransitives and ditransitives. While Topichood/Focus is often required to derive these orders in other languages, in Guarani it does not seem to be as closely tied to information structure.

3 *Wh*-morphology and *wh*-particles

- (7) provides a list of *wh*-elements in Guarani with their English translations.

- *Mba'e* 'what' seems to underlie many other *wh*-questions like: *mba'e-re* 'how', *mba'e-rã* 'why', *mba'e-pe* 'with what', and *mba'e-icha* 'how'.

- *Mava* 'who' also underlies the word *mava-ite* 'which'.

(7)

Wh-elements in Guarani	
mba'e	'what'
mba'e-re	'how'
mba'e-rã	'why'
mba'e-pe	'with what'
mba'e-icha	'how'
mava	'who'
mava-ite	'which'
moõ	'where'
araka'é	'when'
mbovy	'how many'

- The markers which tack on to *mba'e* 'what' are found in other corners of the grammar: *re* is the oblique postposition as in *for*, *rã* is a future marker used for nominals, *pe* is the locative/ACC case marker, *icha* is 'like/as' (Estigarribia 2020).

- Curiously, the *ite* used in *mava-ite* 'which' is actually an intensifier/emphatic marker...

- In other words, these markers are not exclusive to *wh*-elements.

3.1 *Wh*-particles

- There are three different *wh*-particles which are used depending on the context—but they all seem to carry the same meaning or interogativity.[6]

- In all cases, the *wh*-element is always *fronted* because it appears at the beginning of the sentence.

[6] The precise discourse contexts which license a particular particle, but do not license others, is unclear for the time being. Therefore, the descriptions on the following page are descriptive.

Wh-particles in Guarani:

✱ *ø* forced question (strongest, considered rude out of the blue)

    (8)   **mba'e** oi-kyti   Romi?
            what   3.SUBJ-cut Romi

            'What did Romi cut?'

✱ *pa* default question (can be used in all contexts)

    (9)   **mba'e-pa** oi-kyti   Romi?
            what-Q   3.SUBJ-cut Romi

            'What did Romi cut?'

✱ *piko* casual question (can be used in casual contexts)

    (10)   **mba'e-piko** oi-kyti   Romi?
             what-Q   3.SUBJ-cut Romi

             'What did Romi cut?'

---

4 *Wh*-movement (full) is obligatory

- With this, we can observe some basic *wh*-questions in Guarani below. A basic *wh*-question is formed by fronting the *wh*-element as in (11).

    (11) a. **máva-pa** oi-kyti   mandi'o?
                 who-Q   3.SUBJ-cut mandioca
                 'Who cut the mandioca?'

         b. **mba'e-pa** oi-kyti   Romi?
                 what-Q   3.SUBJ-cut Romi
                 'What did Romi cut?'

         c. **moõ-pa** oi-kyti   Romi ka'ã?
                 where-Q 3.SUBJ-cut Romi herbs
                 'Where did Romi cut herbs?'

- In general, we will show forms with the *pa* particle because it is the most widely accepted marker in the most contexts.

- Despite its liberal word order, Guarani does not allow for *wh*-items to be left in-situ: *wh*-movement is obligatory and *wh*-objects ((12a), (12b)) and *wh*-subjects ((13a), (13b)) must be fronted.

    (12) a. * Diego oi-kyti   **mba'e-pa**?
               Diego 3.SUBJ-cut **what-Q**

         b. * Diego **mba'e-pa** oi-kyti?
               Diego **what-Q**   3.SUBJ-cut

    c. **mba'e-pa** Diego oi-kyti?
       what-Q Diego 3.SUBJ-cut
       'What did Diego cut?' O$_{wh}$SV

(13) a. * mandi'o **máva-pa** oi-kyti?
        mandioca **who-Q** 3.SUBJ-cut

     b. * mandi'o oi-kyti **máva-pa**?
        mandioca 3.SUBJ-cut **who-Q**

     c. **máva-pa** mandi'o oi-kyti?
       who-Q mandioca 3.SUBJ-cut
       'Who cut the mandioca?' S$_{wh}$OV

- However, as long as the *wh*-element is fronted, other elements in the sentence may freely scramble. Compare the pre-verbal subject in (14a) with the post-verbal object in (14b). The same holds for non-*wh*-objects in (15a) and (15b).

(14) a. mba'e-pa **Diego** oi-kyti?
       what-Q **Diego** 3.SUBJ-cut
       'What did Diego cut?' O$_{wh}$SV

     b. mba'e-pa oi-kyti **Diego**?
       what-Q 3.SUBJ-cut **Diego**
       'What did Diego cut?' O$_{wh}$VS

(15) a. máva-pa **mandi'o** oi-kyti?
       who-Q **mandioca** 3.SUBJ-cut
       'Who cut the mandioca?' S$_{wh}$OV

     b. máva-pa oi-kyti **mandi'o**?
       who-Q 3.SUBJ-cut **mandioca**
       'Who cut the mandioca?' S$_{wh}$VO

- This is not to say that *no* elements may appear to the left of the *wh*-element: adverbs like *kuehe* 'yesterday' may be fronted in (16a).

(16) a. **kuehe** moõ-piko o-ho Huã?
       yesterday where-Q 3-go Juan
       'Where did Juan go yesterday?'

     b. moõ-piko **kuehe** o-ho Huã?
       where-Q **yesterday** 3-go Juan
       'Where did Juan go yesterday?'

     c. moõ-piko o-ho **kuehe** Huã?
       where-Q 3-go **yesterday** Juan
       'Where did Juan go yesterday?'

     d. moõ-piko o-ho Huã **kuehe**?
       where-Q 3-go Juan **yesterday**
       'Where did Juan go yesterday?'

- The same applies for PPs in (17). It may be fronted to the left of the *wh*-element.

(17) a. **mbo'ehao-pe** mba'e ne-mbo'e
    school-LOC what 2OBJ-teach
    'What did they teach you in school?'

b. mba'e **mbo'ehao-pe** ne-mbo'e
    what **school-LOC** 2OBJ-teach
    'What did they teach you in school?'

c. mba'e ne-mbo'e **mbo'ehao-pe**
    what 2OBJ-teach **school-LOC**
    'What did they teach you in school?'

5 Long-distance *wh*-movement

- There does not appear to be any morphological reflex of successive-cyclic movement in $\overline{\text{A}}$-chains in Guarani.

- Long-distance questions are still formed by fronting the *wh*-element to the far left of the sentence when comparing the baseline in (18a) to the embedded object question in (18b) and the embedded subject question in (18c).

(18) a. ne-memby he'i ndéve o-mbo'e-hague tembiasakue ichupe mbo'ehao-pe?
    2-child 3.say to.you 3-teach-PST history them school-LOC
    'Your child told you they were taught history in school.'

b. **mba'e-pa** he'i ndéve ne-memby o-mbo'e-hague ichupe mbo'ehao-pe?
    what-Q 3.say to.you 2-child 3-teach-PST them school-LOC
    'What did your child tell you they were taught in school?'

c. **máva-pa** he'i ndéve ne-memby o-mbo'e ichupe mbo'ehao-pe?
    who-Q 3.say to.you 2-child 3-teach him/her school-LOC
    'Who did you child say teaches them in school?'

- Scrambling does not always appear to be clause bounded in comparing (19a) and (19b).

(19) a. Diego oi-mo'ã Huã o-mbo-'i pe pakova
    Diego 3-think Juan 3-CAUS-DIM the banana
    'Diego thinks that Juan sliced the banana.'

b. Huã Diego oi-mo'ã o-mbo-'i pe pakova
    Juan Diego 3-think 3-CAUS-DIM the banana
    'Diego thinks Juan sliced the banana.'

c. **mba'e-pa** Diego oi-mo'ã Huã 3-mbo-'i ___
    what-Q Diego 3-think Juan 3-CAUS-DIM
    'What does Diego think Juan cut?'

d. **máva-pa** Diego oi-mo'ã ___ 3-mbo-'i pe pakova
    who-Q Diego 3-think Juan 3-CAUS-DIM the
    'Who does Diego think cut the banana?'

- *Wh*-copying is not allowed. In other words, lower traces of *wh*-elements may not be pronounced (20).

(20) a. **mba'e-pa** oi-mo'ã   Arturo (***mba'e-pa**) Diego oi-kyti?
       what-Q 3.SUBJ-think Arturo (***what-Q**) Diego 3.SUBJ-cut
       'What does Arturo think Diego cut?'

   b. **máva-pa** rei-mo'ã   (***máva-pa**) oi-kyti   mandi'o?
      who-Q 2.SUBJ-think (***who-Q**)   3.SUBJ-cut mandioca
      'Who do you think cut the mandioca?

- Partial *wh*-movement is also not allowed (21). In other words, the *wh*-element must be *fully* fronted to the left of the phrase. This applies regardless of whether there is a scope marker (21a) or not (21b).

(21) a. * Isa oi-mõ'ã   **máva** oi-kyti   mandi'o?
        Isa 3.SUBJ-think **who** 3.SUBJcut mandioca
        Intended: 'Who does Isa think cut the mandioca?'

   b. * **mba'e** Isa oi-mo'ã   **máva** oi-kyti   mandi'o?
        what Isa 3.SUBJ-think **who** 3.SUBJ-cut mandioca
        Intended: 'Who does Isa think cut the mandioca?'

6 *Wh*-islands are obeyed

- *Wh*-islands refers to an inability to move one *wh*-element across another and, among other islands, they are a general property of many languages.

(22) a. [che a-ñe-porandu   [$_{CP}$ **mba'e** Huã o-mẽ'ẽ-ta   chéve]]
        I   1.SUBJ-REFL-ask   **what** Juan 3-give-FUT to.me
        'I wonder what Juan will give me.'

   b. [**mba'e-pa**$_1$ nde re-ñe-porandu   [$_{CP}$ o-me'ẽ-ta   ndéve Huã ___$_1$?]]
      what-Q   you 2.SUBJ-REFL-ask   3.SUBJ-give-FUT you   Juan
      'What do you wonder Juan will give you?

   c. * [**mba'é-pa**$_1$ nde re-ñe-porandu   [$_{CP}$ **moõ**$_2$ o-me'ẽ-ta   ndéve Huã
        what-Q   you 2.SUBJ-REFL-ask   **where** 3.SUBJ-give-FUT to.you Juan?
        ___$_1$ ___$_2$]]
        Intended: 'what do you wonder where Juan will give you?'

7 Clausal pied-piping

- In this final section, we provide data on *clausal pied-piping* the process by which extracted *wh*-elements drag along an entire clause with them.

- There is speaker variation w.r.t. clausal pied-piping. The data in

(23) a. [Arturo he'i   chéve [che sy   o-jeroky-hague   kuehe   ]]
        Arturo 3.SUBJ-tell me   my mom 3.SUBJ-dance-PST yesterday
        'Arturo told me my mom danced last night.'

   b. [**máva** he-'i   ndéve Arturo [o-jeroky-va'ekue kuehe?   ]]
      who 3.SUBJ-tell you   Arturo   3.SUBJ-dance-PST yesterday
      'who did Arturo tell you danced last night.'

8

    c.   [[**máva** la o-jeroky-va'ekue] he-'i   ndéve Arturo?]
           who RC 3.SUBJ-dance-PST 3.SUBJ-tell you   Arturo

       'who did Arturo tell you danced last night?'

    d. *[[che sy   o-jeroky-hague   ] he-'i   chéve Arturo]
           my mother 3.SUBJ-dance-PST 3.SUBJ-tell me   Arturo

- However, speakers vary w.r.t. the (un)acceptability of sentences like ([piedpiped]).

- The data above are from a speaker from Concepción (Northern/Central-Paraguay) while the data below, in which the pied-piping is allowed, are from our Caaguazu speaker.

8 Conclusion

- Today we have presented data from Guarani to introduce it to the literature on *wh*-question typology.

Properties of *wh*-questions in Guarani:

⇒ *wh*-morphology, *wh*-particles

⇒ *wh*-movement (full) is obligatory (i.e. no partial-*wh*-copying)

⇒ *wh*-movement (long-distance) leaves no morphological trace

⇒ *wh*-islands are obeyed

⇒ *wh*-elements from embedded clauses exhibit clausal pied-piping

---

References

Anchieta, Joseph de. 1595 [1979]. *Arte de grammatica da lingoa mais usada na costa do Brasil*. Coimbra: António de Mariz.

Aragona, Alonso de. 1625 [1979]. Breve introducción para aprender la lengua guaraní. In *Amerindia. revue d'ethnolinguistique amérindienne*, volume 4, 24–-61. Colegio del Salvador de Buenos Aires.

Krivoshein de Canese, Natalia, and Feliciano Acosta Alcaraz. 2001. *Gramática Guaraní*. Asunción: Ñemitỹ.

Krivoshein de Canese, Natalia, and Feliciano Acosta Alcaraz. 2006. *Guarani dictionary*. Asunción: Instituto Superior de Lenguas.

Clopper, Cynthia G, and Judith Tonhauser. 2013. The prosody of focus in Paraguayan Guaraní. *International Journal of American Linguistics* 79:219–251.

Estigarribia, Bruno. 2020. *A grammar of Paraguayan Guarani*. Series Grammars of World and Minority Languages. London: UCL Press.

Estigarribia, Bruno, and Justin Pintas. 2017. *Guarani linguistics in the 21st Century*. Leiden: Brill.

Gregores, Emma, and Jorge A Suarez. 1967. *A Description of Colloquial Guarani*. The Hague, Netherlands: Mouton.

de Guarania, Félix. 1997. *Diccionario Guarania ilustrado.*. Buenos Aries, Argentina: Colihue.

de Guarania, Félix. 2008. *Table sinóptica para una nueva gramática Guarani: ñe'ẽrekokatu ha ñe'ẽmorangatu*. Asunción, Paraguay: Servilibro.

Jensen, Cheryl. 1990. Cross-referencing changes in some Tupí-Guaraní languages. *Amazonian linguistics: Studies in lowland South American languages* 393–405.

Jensen, Cheryl. 1998. Comparative Tupí-Guaraní morphosyntax. *Handbook of Amazonian languages* 4:489–618.

Jensen, Cheryl. 1999. Tupí-Guaraní. In *The Amazonian languages*, ed. RMW Dixon & Alexandra Y. Aikhenvald, 125–164. Cambridge: Cambridge University Press.

Jun, Sun-Ah, Jian-Leat Siah, Marisabel Cabrera, and Hunter Johnson. 2023. Realizations of the tritonal pitch accent in Paraguayan Guarani. In *Proceedings of the Interational Congress of Phonetic Sciences (ICPhS)*.

Jun, Sun-Ah, and María Luisa Zubizarreta. 2022. Paraguayan Guarani: Tritonal pitch accent and Accentual Phrase. In *Proceedings Interspeech 2022*, 5303–5307. Seoul Korea.

Monserrat, Ruth, and Marília Facó Soares. 1983. Hierarquia referencial em línguas tupi. *Cadernos de Linguística e Teoria da Literatura* 164–187.

Ruiz de Montoya, Antonio. 1724 [1876]. *Tesoro de la lengva gvarani*. Leipzig: B. G. Teubner.

Pancheva, Roumyana, and Maria Luisa Zubizarreta. 2019. On the role of person features in the evidential-temporal connection. *Canadian Journal of Linguistics/Revue canadienne de linguistique* 64:673–708.

Restivo, Pablo. 1724 [1892]. *Arte de la Lengua Guaraní*. Stuttgart: Kohlhammer.

Rose, Françoise. 2015. When "You" and "I" mess around with the hierarchy: a comparative study of Tupi-Guarani hierarchical indexing systems. *Boletim do Museu Paraense Emílio Goeldi. Ciências Humanas* 10:347–369.

Shain, Cory, and Judith Tonhauser. 2010. The synchrony and diachrony of Differential Object Marking in Paraguayan Guaraní. *Language Variation and Change* 22:321–346.

Tonhauser, Judith. 2006. The temporal semantics of noun phrases: Evidence from Guaraní. Doctoral Dissertation.

Tonhauser, Judith. 2007. Nominal tense? The meaning of Guaraní nominal temporal markers. *Language* 831–869.

Tonhauser, Judith. 2011a. The Paraguayan Guaraní future marker–ta: Formal semantics and cross-linguistic comparison. *Tense across languages* 207–231.

Tonhauser, Judith. 2011b. Temporal reference in Paraguayan Guaraní, a tenseless language. *Linguistics and Philosophy* 34:257–303.

Tonhauser, Judith. 2020. Projection variability in Paraguayan Guaraní. *Natural Language & Linguistic Theory* 38:1263–1302.

Tonhauser, Judith, and Erika Colijn. 2010. Word order in Paraguayan Guaraní. *International Journal of American Linguistics* 76:255–288.

Velázquez-Castillo, Maura. 1991. The semantics of Guaraní agreement markers. In *Annual Meeting of the Berkeley Linguistics Society*, volume 17, 324–335.

Velázquez-Castillo, Maura. 2002. Grammatical relations in active systems: The case of Guaraní. *Functions of Language* 9:133–167.

Zarratea, Tadeo. 2002. *Gramática elemental de la lengua Guaraní*. Asunción, Paraguay: Marben.

Zubizarreta, María Luisa. 2022. The verbal phrase in Paraguayan Guarani: A case study on the role of prosody in linearization. *Languages* 7:221.

Zubizarreta, María Luisa, and Roumyana Pancheva. 2017a. A formal characterization of person-based alignment: The case of Paraguayan Guaraní. *Natural Language & Linguistic Theory* 35:1161–1204.

Zubizarreta, María Luisa, and Roumyana Pancheva. 2017b. A note on the syntax of possession in Paraguayan Guaraní. *Boundaries, phases, and interfaces* 85–104.