The goal of this paper is to investigate the form and structure of information questions in two dialects of Malagasy: Official Malagasy (OM) and Antakarana (Ant). Questions in both of these languages take the form of a cleft, although this is not obvious from the word order. We provide evidence for this claim, building on existing analyses of OM (Paul 2001, Potsdam 2006a,b).

The paper is organized as follows: Section 1 introduces a typology of information-question formation. In section 2, we describe wh-questions in OM and assert that OM is a wh-in-situ language. Wh-questions may be formed using wh-in-situ or a pseudocleft structure. Section 3 turns to the Antakarana dialect of Malagasy, spoken in the northern tip of Madagascar. We describe the form of wh-questions in this dialect and argue that it uses these same two strategies, despite morphosyntactic differences between the two dialects. Section 4 presents our conclusions.

1. Information Questions

Information questions, also known as content questions or wh-questions, contain an interrogative proform and request an answer beyond yes or no. Languages typically use one or more of four strategies for forming such questions. These strategies are illustrated for French in (1a-d). In DISPLACEMENT, the interrogative proform, or wh-phrase, is movement from its canonical, logical position to some dedicated position in the clause, usually clause-initial, (1a). A CLEFT, (1b), is an impersonal biclausal construction in which the wh-phrase is put into focus. A PSEUDOCLEFT, (1c), is a biclausal construction in which the wh-phrase is the predicate and its subject is a complex noun phrase. In SUBSTITUTION, (1d), the wh-phrase is in-situ, in the position where it is interpreted. It simply replaces the constituent that is being questioned.1

---

1 Gloses follow Leipzig glossing conventions (http://www.eva.mpg.de/lingua/resources/glossing-rules.php), with the following additions: ACC-accusative, ACT-active voice, DEM-demonstrative, DET-determiner, EXCL-exclamative, FOC-focus, IRR-irreals, PASS-passive voice.
While these four strategies look very different in French or English, it is often difficult to tell them apart in other languages. Many languages have no copula or a null copula, a null expletive pronoun corresponding to it, a null relative clause head, and/or a null relativizer. All of these conspire to hide the syntactic structure. For example, if the copula, expletive and complementizer were null in the biclausal cleft construction repeated below as (2b), it would be nearly indistinguishable from the monoclusal displacement example repeated as (2a).²

Consequently, more subtle diagnostics beyond surface word order are needed to determine the structure of information questions. Such complications arise in the analysis of information questions in OM and its dialects, to which we turn.

2. Information Questions in Official Malagasy

Malagasy is a Western Austronesian language spoken on the island of Madagascar by approximately 14 million people. The dialect that has been widely discussed in the generative literature is that of standard or Official Malagasy (OM). It is most similar to the Merina dialect spoken in and around the capital city Antananarivo. OM is well known for having rather rigid VOS as its basic word order. More generally, OM can be described as predicate-initial and subject-final, as all predicates, not just verbal ones, precede the subject.

There is a significant body of descriptive and analytical work on wh-questions in Malagasy (e.g. Keenan 1976, Paul 2001, Sabel 2002, 2003,

---
² A difference remains in French because of subject-verb inversion under displacement. Without this inversion, which most languages lack, the two are indistinguishable.
The first strategy for forming Malagasy wh-questions is substitution. Non-subject wh-phrases may generally stay in-situ (Sabel 2003), (3a,b). Example (3c) shows that subjects may not remain in-situ.

(3) a. Nividy inona Rabe? OM
   buy.ACT what Rabe
   ‘Rabe bought what?’ (Sabel 2003:234)

   b. Nividy ny vary taiza Rabe? OM
   buy.ACT DET rice where Rabe
   ‘Rabe bought the rice where?’ (Sabel 2003:234)

   c. *Nividy ny vary iza?
      buy.ACT DET rice who
      (‘Who bought the rice?’) (Sabel 2003:234)

The second strategy for forming wh-questions in Malagasy is to front the wh-phrase and follow it with the particle no, which we gloss FOC(US) because it is also used in the focus construction without a wh-word:

(4) a. Hita-nao ny gidro. OM
   see.PASS-2SG DET lemur
   ‘A lemur was seen by you.’

   b. Inona no hita-nao? OM
   what FOC see.PASS-2SG
   ‘What was seen by you?’

In OM, only subjects and circumstantial (adverbial) phrases can be questioned by fronting accompanied by the focus particle. Questioning of subjects is illustrated in (4b) and (5b). Questioning of adverbial phrases is illustrated in (6). Example (7) shows that questioning other elements, such as a direct object, using this method is ungrammatical.

(5) a. Nihomehy Rasoa. OM
   laugh.ACT Rasoa
   ‘Rasoa laughed.’

   b. Iza no nihomehy? OM
   who FOC laugh.ACT
   ‘Who laughed?’

(6) a. Nahoana no nihomehy ianao? OM
   why FOC laugh.ACT 2SG
   ‘Why did you laugh?’

   b. Taiza no nividy vary Rasoa? OM
   where FOC buy.ACT rice Rasoa
   ‘Where did Rasoa buy rice?’ (Potsdam 2006a:2155)
Several researchers conclude that questions shown in (4b), (5b), and (6) have the structure of a pseudocleft (Dahl 1986, Pearson 1996, Paul 2001, Potsdam 2006a,b). In this pseudocleft, the initial wh-phrase is the predicate and the remaining material is headless relative clause in subject position:

\[(7) \text{Inona no nahita ianao?} \quad \text{OM} \]
\[
\text{what FOC see.ACT 2SG}
\]
\[
\text{‘What did you see?’}
\]

In (7), the initial wh-phrase is the predicate and the remaining material is headless relative clause in subject position.

\[(8) [\text{Iza}] \_\text{predicate} [\text{no nihomehy}] \_\text{subject}? \quad \text{OM} \]
\[
\text{who FOC laugh.ACT}
\]
\[
\text{who the one who laughed}
\]
\[
\text{‘Who laughed?’}
\]

Such an analysis is feasible because Malagasy has no copula, (9), and headless relative clauses are independently available, (10).

\[(9) \text{Mipianatra Rabe.} \quad \text{OM} \]
\[
\text{student Rabe}
\]
\[
\text{‘Rabe is a student.’}
\]

\[(10) \text{ny miasa mafy} \quad \text{OM} \]
\[
\text{DET work.ACT hard}
\]
\[
\text{‘the ones who are working hard’ (Paul 2001:718)}
\]

There are two strands of evidence for this structure, which we briefly present below: i) the initial wh-phrase is a predicate (section 2.1) and ii) the remaining material is the subject (section 2.2). If this result is correct, then Malagasy is essentially a wh-in-situ language: wh-phrases may stay in-situ as arguments or adjuncts, or they may be in-situ as the predicate of a pseudocleft. They never undergo displacement (contra Sabel 2002, 2003).

### 2.1 Evidence for the Predicate

Evidence for the predicate status of the initial wh-phrase in OM questions comes from the placement of particles that accompany predicates. Certain verbal modifiers immediately precede or follow the predicate and cannot appear elsewhere.

Pre-predicate particles include tokony ‘should’, tena ‘indeed’, and the irrealis marker ho ‘IRR’. The placement of tokony ‘should’ is illustrated in (11) for a declarative clause:

\[(11) \text{Tokony hamangy an-dRabe Rasoa.} \quad \text{OM} \]
\[
\text{should visit ACC-Rabe Rasoa}
\]
\[
\text{‘Rasoa should visit Rabe.’ (Potsdam 2006a:2165)}
\]

In wh-questions, these pre-predicate particles can also precede the wh-phrase, supporting its status as predicate:

\[\text{3 Law (2007) disagrees with the pseudocleft analysis and proposes an alternative.}\]
Similarly, there are particles which immediately follow the predicate. These include the universal quantifiers daho ho ‘all’ and aho ‘each’ (Keenan 1995), the exclamative particle anie ‘EXCL’ (Keenan 1995), and VP adverbs like foana ‘always’ (Pearson 1998). For example:

(13) Manapaka (*anie) bozaka (anie) Rasoa (*anie) OM
cut.ACT EXCL grass EXCL Rasoa EXCL
‘Rasoa is really cutting the grass!’ (Potsdam 2006a:2163)

These particles can immediately follow the wh-phrase in information questions:

(14) Iza anie no manapaka bozaka? OM
cut.ACT EXCL grass
‘Who is really cutting the grass?’ (Potsdam 2006a:2164)

Admittedly, additional evidence for the nominal status of the post-wh-phrase material is difficult to find. This has been the major weakness of the pseudocleft analysis (Potsdam 2006b, Law 2007). One of the outstanding issues has to do with the analysis of no. Nonetheless, we adopt the pseudocleft structure for OM wh-questions repeated below.

(17) [Iza] predicate [no nihomely] subject? OM
who FOC laugh.ACT
who the one who laughed
‘Who laughed?’
3. Information Questions in Antakarana

In this section, we turn to the analysis of wh-questions in Antakarana (Ant), a dialect of Malagasy spoken by approximately 330,000 speakers in the northern tip of Madagascar (Lewis et al. 2014). Lewis et al. 2014 indicates 71% lexical commonality with the Merina dialect. The dialect is described by Mbima (no date) and Hanitramalala (2013). We begin by highlighting several salient morphosyntactic differences between OM and Ant in its declaratives and wh-questions (sections 3.1 and 3.2). In sections 3.3 and 3.4 we argue that, despite these differences, which only further hide the syntactic structure of clauses, wh-questions in Ant are also pseudoclefts.

3.1 Antakarana Syntax

Like OM, Ant allows VOS word order. Unlike OM, however, SVO is equally possible:

\[(18)\]
\[
\begin{align*}
\text{VOS} & \\
\text{a. Manana tongatra efatra ny alika rehetra.} & \text{OM have leg four DET dog all} \\
& \text{`All dogs have four legs.'} \\
\text{b. Amboa jiaby manana vity êfatra.} & \text{Ant dog all have leg four} \\
& \text{`All dogs have four legs.'}
\end{align*}
\]

\[(19)\]
\[
\begin{align*}
\text{SVO} & \\
\text{a. *Ny alika rehetra manana tongatra efatra.} & \text{OM DET dog all have leg four} \\
& \text{`All dogs have four legs.'} \\
\text{b. Amboa jiaby manana vity êfatra.} & \text{Ant dog all have leg four} \\
& \text{`All dogs have four legs.'}
\end{align*}
\]

A further difference between OM and Ant is that indefinite/non-specific subjects are allowed in Ant, in both SVO and VOS word orders, (20) and (21). Indefinite/non-specific subjects are not permitted in OM (Keenan 1976, Pearson 1996, Paul 2000, Sabel 2002, but see Law 2006), (22). Instead, an existential construction can be used, (23).

\[(20)\]
\[
\begin{align*}
\text{SV} & \\
\text{a. Olo navy.} & \text{Ant person came} \\
\text{b. VS} & \\
\text{Navy olo.} & \text{Ant came person} \\
& \text{`Someone came.'}
\end{align*}
\]
(21) a. **SVO**
   Tsaiky jiaby tia bonbon. **Ant**
   child all like candy

   b. **VOS**
   Tia bonbon tsaiky jiaby. **Ant**
   like candy child all
   ‘All children like candy.’

(22) *Avy olona.
    come person
    (*Someone came.*)

(23) Misy olona avy.
    exist person come
    ‘There is someone who came.’

Finally, Ant differs from OM in that the focus particle *no* seen above in OM wh-questions is optional, and generally not used:

(24) Ino (no) nivangain-ao?
    what FOC buy.PASS-2SG
    ‘What was bought by you?’

### 3.2 Wh-Questions in Antakarana

This section documents the wh-phrases and the question formation strategies in Ant. Although only some of this is relevant for material that follows, we do this as a contribution to the description of this understudied dialect.

Antakarana wh-phrases are given in Table 1, in comparison to their OM counterparts.

<table>
<thead>
<tr>
<th>OM WH-PHRASE</th>
<th>ANT WH-PHRASE</th>
<th>GLOSS</th>
</tr>
</thead>
<tbody>
<tr>
<td>iza</td>
<td>azôvy</td>
<td>‘who’</td>
</tr>
<tr>
<td>inona</td>
<td>ino</td>
<td>‘what’</td>
</tr>
<tr>
<td>aiza</td>
<td>aia</td>
<td>‘where (non-past)’</td>
</tr>
<tr>
<td>taiza</td>
<td>taia</td>
<td>‘where (past)’</td>
</tr>
<tr>
<td>ho aiza</td>
<td>hanâia</td>
<td>‘where (future)’</td>
</tr>
<tr>
<td>oiviâna</td>
<td>ombiaña</td>
<td>‘when (non-future)’</td>
</tr>
<tr>
<td>rahoviana</td>
<td>ombiaña</td>
<td>‘when (future)’</td>
</tr>
<tr>
<td>ahoanâ</td>
<td>karaha akôry</td>
<td>‘how’</td>
</tr>
<tr>
<td>hoatrinona/ohatrinona</td>
<td>hoatrinô</td>
<td>‘how much’</td>
</tr>
<tr>
<td>fire</td>
<td>fiery</td>
<td>‘how many’</td>
</tr>
<tr>
<td>ahoana</td>
<td>aîîno</td>
<td>‘why (non-past)’</td>
</tr>
<tr>
<td>nahoana</td>
<td>naîîno</td>
<td>‘why (past)’</td>
</tr>
<tr>
<td>N iza</td>
<td>N karaha akôry</td>
<td>‘which N’</td>
</tr>
<tr>
<td>N inona</td>
<td>N azôvy/N ino</td>
<td></td>
</tr>
</tbody>
</table>

Table 1. Wh-phrases in OM and Ant
The same two strategies for forming wh-questions in OM are available in Ant. As in OM, non-subject substitution (wh-in-situ) is permitted, (25), while subject wh-in-situ remains impossible, (26).\footnote{See below for qualification of this statement.}

(25) a. Nivanga \textit{ino} anao? Ant
   \textit{buy.ACT} what 2SG
   ‘You bought what?’

   b. Nivanga vary \textit{taia} anao? Ant
   \textit{buy.ACT} rice where 2SG
   ‘You bought rice where?’

   c. Nivangain’ \textit{azôvy} aombi-ko? Ant
   \textit{buy.PASS} who zebu-1SG
   ‘My zebu was bought by who?’

(26) *Nivanga aombi-ko \textit{azôvy}? Ant
   \textit{buy.ACT} zebu-1SG who
   (‘Who bought my zebu?’)

As in OM, fronting the wh-phrase to initial position is also possible. Illustrative examples are given below along with the corresponding examples in OM. Besides lexical differences, an obvious distinction is the lack of the particle \textit{no} ‘FOC’; as we already mentioned in the previous section, this particle is generally not used in Ant.

(27) a. \textit{Azôvy} namono aomby? Ant
   \textit{who kill.ACT} zebu

   b. Iza \textit{no} namono ilay omby? OM
   \textit{who FOC kill.ACT DEM zebu}
   ‘Who killed the zebu?’

(28) a. \textit{Ino} nivonen’ ny mpitarimy? Ant
   \textit{what kill.PASS DET cattleman}

   b. Inona \textit{no} novonoin’ ny mpiompy? OM
   \textit{what FOC kill.PASS DET cattleman}
   ‘What was killed by the cattleman?’

(29) a. \textit{Taia} namono aomby ny mpitarimy? Ant
   \textit{where kill.ACT zebu DET cattleman}

   b. \textit{Taiza} \textit{no} namono omby ny mpiompy? OM
   \textit{here FOC kill.ACT zebu DET cattleman}
   ‘Where did the cattleman kill the zebu?’

\footnote{See below for qualification of this statement.}
We showed above that the OM (b) sentences have the structure of pseudocLEFTs. The question we address in the remainder of this section is whether the Antakarana wh-questions have the same structure. Given the availability of SVO word order, the possibility of wh-in-situ, and the lack of the focus particle no, a positive answer is far from evident. We will consider two hypotheses. Under Hypothesis 1 (H1), Ant wh-questions involve wh-in-situ of the subject in an SVO word order.\(^5\) We have already seen that Ant allows both SVO word order and wh-in-situ. The analysis of (27), repeated below, would be as in (30a). The initial wh-phrase is the clause-initial subject, and the remaining material is the predicate.

Under Hypothesis 2 (H2), (27) is a hidden pseudocleft. The initial wh-phrase is the predicate and the remaining material is a headless relative clause in subject position, (30b).

\[(27)\quad \text{Azóvy namono aomby\quad \text{Ant}}\]
who kill.ACT zebu
‘Who killed the zebu?’

\[(30)\quad \text{a. H1 analysis: subject in-situ analysis}\]

\[
\begin{array}{ll}
\text{[azóvy]} & \text{[namono aomby]} \\
\text{who} & \text{kill} \\
\text{SUBJECT} & \text{PREDICATE}
\end{array}
\]

\[(30)\quad \text{b. H2 analysis: pseudocleft analysis}\]

\[
\begin{array}{ll}
\text{[azóvy]} & \text{[namono aomby]} \\
\text{who} & \text{the one who killed the zebu} \\
\text{PREDICATE} & \text{SUBJECT}
\end{array}
\]

In the following sections, we present arguments in favor of H2 and conclude that wh-questions in Ant are syntactically the same as OM. The first argument comes from predicate-oriented particles (section 3.3). The second comes from semantic differences between subjects and initial wh-phrases (section 3.4).

### 3.3 Evidence for the Predicate

In section 2.1 we saw that the OM predicate-related particles tokony ‘should’ and anie ‘EXCL’ identify the initial wh-phrase as the predicate. The Ant pre-predicate particles tokony ‘should’ and iseky ‘intend to’ confirm the same claim in Ant. These particles appear preceding the predicate in Ant, and they cannot precede the subject in SVO orders:

\[(31)\quad \text{a. Tokony hamangy Rabe Rasoa.\quad \text{Ant}}\]
should visit.ACT Rabe Rasoa
‘Rasoa should visit Rabe.’

\[\text{This hypothesis clearly cannot extend to (29a), in which the initial wh-phrase taia ‘where’ is not a subject. A second structure would need to be available for such cases. Since we will end up rejecting Hypothesis 1, we will not pursue this.}\]
b. *Tokony Rabe hamangy Rasoa. Ant
    should Rabe visit.ACT Rasoa
    (‘Rasoa should visit Rabe.’)

(32) a. Tseky hivanga tômôbili-nao Rabe. Ant
    almost paint.ACT car-2SG Rabe
    ‘Rabe intends to buy your car.’

b. *Tseky Rabe hivanga tômôbili-nao. Ant
    almost Rabe buy.ACT car-2SG
    (‘Rabe intends to buy your car.’)

The same positioning is found in wh-questions, indicating that the initial wh-phrase is a predicate:

(33) a. Tokony azôvy hamangy Rabe? Ant
    should who visit.ACT Rabe
    ‘Who should visit Rabe?’

b. Tseky azôvy hivanga tômôbili-nao? Ant
    intend who buy.ACT car-2SG
    ‘Who intends to buy your car?’

This observation confirms that the initial wh-phrase is a predicate, as it is preceded by a pre-predicate particle. This fact is unexpected under H1 because these particles do not precede the subject in SVO word orders, as shown by (31b) and (32b).

Regarding post-predicate particles, Ant does not have the exclamative particle anie but it has a question particle ma, whose placement is relevant to our discussion. In Ant, this particle follows the predicate in a yes/no question, (34a). In SVO orders, it cannot follow the subject, (34b):*

(34) a. Nitokiky ma Rasoa? Ant
    laugh.ACT Q Rasoa
    ‘Did Rasoa laugh?’

b. *Rasoa ma nitokiky? Ant
    Rasoa Q laugh.ACT
    (‘Did Rasoa laugh?’)

In wh-questions, the particle ma follows the wh-phrase, confirming that the wh-phrase is a predicate and not a subject, in support of H2.

---

6 Ma corresponds to moa in OM but has a wider distribution.

7 The question in (34b) is grammatical on another interpretation where Rasoa is an initial, clefited constituent: ‘Is it Rasoa who laughed?’ This interpretation is not relevant to the discussion here.
3.4 Evidence that the Initial Wh-Phrase is not a Subject

The distribution of indefinite subjects and wh-phrase subjects provides a second argument in favor of H2. We have already seen that subjects in Ant may be indefinite and precede or follow the verb, (36). Meanwhile, wh-subjects, which are also indefinite, are restricted in only being able to precede the verb; post-verbal wh-subjects in-situ are impossible, (37).

(36) a. SV
    Olo  navy.
    person came
    Ant

b. VS
    Navy olo.
    came person
    ‘Someone came.’
    Ant

(37) a. Azóvy navy?
    who came
    ‘Who came?’
    Ant

b. *Navy azóvy?
    came who
    Ant

If wh-questions were simply instances of SV(O) word order, as H1 claims, the contrast between (37a) and (37b) would be unexpected. We cannot stipulate that subject wh-in-situ is impossible, as that would rule out both (37a) and (37b). One would have to refer specifically to the post-verbal subject position to prevent (37b). Under H2, this difference is more easily accounted for. (37b) is ungrammatical if we stipulate that wh-in-situ is not possible for subjects. We do not need to refer to the pre-verbal post-verbal distinction because the wh-phrase in (37a) is not a subject; it is the predicate of the clause.

Ant data are intriguingly more complex than the corresponding OM data. In Ant, post-verbal wh-phrase subjects are possible if the wh-phrase is d(iscourse)-linked (Pesetsky 1987, 2000). D-linked elements presuppose the existence of a set of contextually determined entities from which the answer is to be chosen. To illustrate, English who/what are generally non-d-linked while wh-phrases such as which student are d-linked:

---

8 Whether this is possible depends upon the structural position of pre-verbal and post-verbal subjects. If the structural position of pre-verbal and post-verbal subjects is the same, for example, if both kinds of subject are in spec,TP, then it is not obvious how one could single out the post-verbal subject in such a generalization. Post-verbal subjects would be in the same place as a pre-verbal subject, and any reference to a DP in subject position would pick out both pre-verbal and post-verbal subjects. If pre-verbal and post-verbal subjects are in different structural positions however, one could refer to the structure of post-verbal subjects alone and impose restrictions just on that position.
Subject wh-in-situ is possible in Ant just in case the wh-phrase is d-linked:

\[(38) \begin{align*}
a. \quad & \text{Who failed?} \quad \text{(non-d-linked)} \\
b. \quad & \text{Which student failed?} \quad \text{(d-linked)} \\
& \text{presupposes a defined set of students from which the answer is to be chosen}
\end{align*}\]

   ('Who hid?')

b. Nisitriky [tsaiky azôvy]? Ant hide.ACT child who D-LINKED WH-PHRASE
   ‘Which child hid?’

(40) Mamaky angano aminao [lobe-nao karaha akôry]? Ant read.ACT story to.you parent-2SG like who D-LINKED WH-PHRASE
    ‘Which of your parents reads tales to you?’

Such data indicate that there are different restrictions on the pre-verbal and post-verbal position of wh-phrases—a phenomenon which is possible under H2 because the two positions are not unified in any way. Under H1, the wh-phrases are all subjects, so it is less clear why restrictions on d-linking or any other contextual properties should exist on the post-verbal but not pre-verbal ones.

We conclude that H2 is supported, and that wh-questions in Ant are also pseudoclefts, as in OM. The initial wh-phrase is a predicate, not a subject. Despite syntactic differences between Ant and OM, they do not differ in the structure of their wh-questions. Both allow wh-in-situ and pseudocleft structures.

4. Conclusion

We have considered the structure of wh-questions in two dialects of Malagasy, Official Malagasy (OM) and Antakaran (Ant). Despite appearances, both languages are wh-in-situ languages. Wh-phrases may be in-situ as predicates in a pseudocleft structure or they may be in-situ as arguments/adjuncts in a substitution structure. Ant further differs from OM in that substitution is permitted for d-linked subjects while it is uniformly banned in OM.

Word order in Ant and OM was of little use in determining the structure of wh-questions. In order to determine the syntactic structure, we appealed to more nuanced diagnostics. This move has both general and language-specific consequences. On a language-specific level, it is clearly necessary to have such
tools available to determine the extent to which the numerous Malagasy dialects are underlingly similar. If they are found to be similar, this needs to be explained; given the lack of obvious syntactic cues and the potential influence of displacement structures in French, such similarity would be striking. On a more general level, our analysis demonstrates that sometimes subtle facts need to be investigated to determine the syntax of the world’s languages. Sometimes surface cues suggest substantive differences, when there really are none.

References


Mbima, Pierre. no date. Grammaire antakarana. MS.


