

On the topic of pseudoclefts*

Ileana Paul

University of Western Ontario

This paper presents arguments in favor of a pseudocleft analysis of a certain class of sentences in Malagasy, despite the lack of an overt *wh*-element. It is shown that voice morphology on the verb creates an operator-variable relationship much like the one created by *wh*-movement in free relatives in English and other languages. The bulk of the paper argues in favor of an inversion analysis of specificational pseudoclefts in Malagasy: a predicate DP is fronted to a topic position from within a small clause constituent. Moreover, it is shown that the same inversion occurs in equative and specificational sentences in Malagasy, suggesting that these types of sentences share the same syntactic structure. The proposed analysis also provides support for the view that specificational pseudoclefts have a topic>focus structure, where the *wh*-clause has been overtly topicalized.

keywords: pseudoclefts, information structure, topicalization, Malagasy

1. Introduction

Pseudoclefts in English have been the topic of much linguistic interest over the past few decades (Akmajian 1970, Higgins 1979, Schlenker 2003). A typical example is given in (1).

(1) What Jessie is is important.

Most researchers focus on how to derive certain interpretive properties of pseudoclefts (e.g. connectivity). In this paper, however, I address the question of the syntax of pseudoclefts in one particular language, Malagasy, with a focus on their information structure. The examples in (2)

are illustrative: there is a nominalized predicate in initial position, followed by the topic particle *dia*, which is in turn followed by an XP (e.g. DP).¹

- (2) a. Ny mahafinaritra dia izany vaovao izany.
DET AT.happy TOP that news that

What is pleasing is that news.

- b. Ny nahatongavany dia omaly.
DET PST.CT.arrive.3(GEN) TOP yesterday

When he arrived was yesterday.

[Rajaona 1972]

This construction, to the best of my knowledge, has not been studied in detail in the literature, though Rajaona (1972) provides some initial important observations.

The central goal of this paper is to provide an analysis of Malagasy data like those in (2). In particular, I argue that despite the lack of a *wh*-element, the sentences in (2) are parallel to English pseudoclefts. In particular, *ny mahafinaritra* (lit.) ‘the pleasing thing’ in (2a) functions like a free relative. In my analysis, this free relative is generated as the predicate of a small clause and undergoes movement to the clause-initial topic position. I therefore argue in favor of an inversion derivation of these pseudoclefts (Williams 1983, Moro 1997, Mikkelsen 2004, den Dikken 2006a). Moreover, like Mikkelsen (2004), Heycock and Kroch (1999) and den Dikken (2006a), I explore the parallels between specificational copular sentences, specificational pseudoclefts and equatives. I show that all of these types of clauses involve inversion in Malagasy.

Another important aspect of this paper is the analysis of topicalization in pseudoclefts. The topic status of the *wh*-clause in English pseudoclefts has been argued for by den Dikken et al. (2000).² Malagasy provides overt evidence in favor of topicalization (i.e. the presence of the

topic particle *dia*). I show that topicalization is forced in Malagasy to avoid syntactic constraints on the form and position of predicates and subjects. In particular, in Malagasy the *wh*-clause is not a possible subject and must be topicalized. One effect of this topicalization is to put the *wh*-clause in a sentence initial position, such that both English and Malagasy pseudoclefts share the same word order despite Malagasy being a subject-final language. In fact, specificational pseudoclefts in both language share similar underlying and surface structures.

The remainder of this paper is organized as follows. In section 2, I provide an overview of some of the issues surrounding pseudoclefts and some analyses that have been proposed. Section 3 provides background information on Malagasy clause structure. In section 4, I discuss nominal predication in Malagasy and introduce pseudoclefts in this language. Section 5 is dedicated to pseudoclefts in Malagasy and their particular properties. I illustrate the syntactic structure of pseudoclefts in section 6 and rule out alternative analyses. In section 7, I turn to broader considerations of information structure and cross-linguistic implications and section 8 concludes.

2. Issues in pseudoclefts

Despite the large body of literature on pseudoclefts (see references below), there remain many unanswered questions surrounding their syntactic structure. Before we begin, let us consider an example of a pseudocleft and the terminology associated with it.

(3) What Jessie is is important.

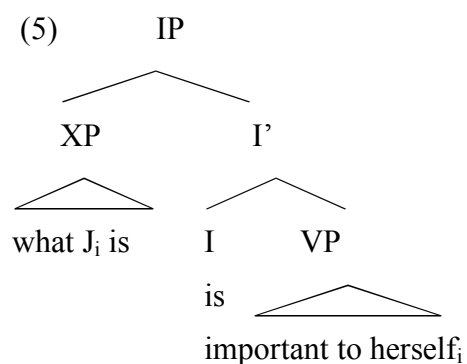
I will refer to the constituent *what Jessie is* as the *wh*-clause. Following Heycock (1994), I call the post-copular constituent (*important* in (3)) the “counterweight”. As noted by Akmajian (1970), (3) is ambiguous between a predicational and a specificational reading. On the predicational reading, (3) means: Jessie is an X (e.g. a firefighter); being an X is important. On

the specificational reading, (3) simply means that Jessie is important. We will see these same two readings are available for pseudoclefts in Malagasy.

One of the major sources of linguistic interest in pseudoclefts is the so-called “connectivity” (or “connectedness”) effect found in specificational pseudoclefts.³

- (4) a. What Jessie is is important to herself.
 b. What everyone proved was his own theory.
 c. What Jessie didn't buy was any pictures of Alex.

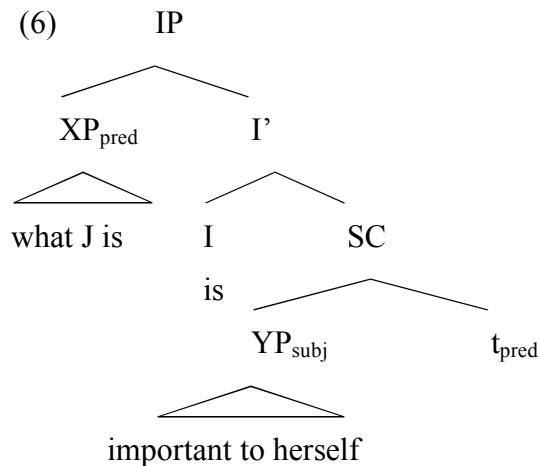
For the purposes of licensing anaphors, bound pronouns and negative polarity items, a specificational pseudocleft behaves as if the counterweight were somehow “connected” to the *wh*-clause. This connectivity is mysterious given that elements in the *wh*-clause do not (apparently) c-command the counterweight, as illustrated schematically in (5), violating the ordinary licensing conditions for anaphors, bound pronouns, etc.



Moreover, predicational pseudoclefts lack connectivity effects: (4a) is unambiguously specificational. For this reason, specificational pseudoclefts have been the focus of more attention than their predicational counterparts. Although connectivity is not the focus of this paper, it is a useful test to distinguish between the two types of pseudocleft and will be relevant to the discussion of the Malagasy data. For analyses of connectivity, I refer the reader to

Jacobson (1995), Sharvit (1999) and Schlenker (2003), among many others. I now describe two issues in specification pseudoclefts that are crucial for this paper: their syntactic structure and their information structure.

Turning first to the syntactic structure of pseudoclefts, there are roughly two types of analysis in the literature.⁴ The first, supported by Higgins (1979), Bošković (1997) and Schlenker (2003), is what I will call the non-movement analysis. Although the details of the above analyses differ greatly, they all assume that the free relative is a subject in the underlying representation (it is either base generated in subject position or raises from a VP-internal subject position). The second type of analysis, the one that I will argue for in this paper, assumes movement (see e.g. Williams 1983 and Moro 1997). For these researchers, in a specificational pseudocleft the *wh*-clause originates as a small clause predicate and moves into the subject position. The counterweight is the small clause subject. Abstracting away from the details of the particular analyses proposed by the above authors, the structure of a specificational pseudocleft is as in (6).



Because it is the predicate rather than the subject that raises, this analysis is often called “inversion”.

The information structure of pseudoclefts has also attracted a certain amount of attention in the literature and is an aspect that I will focus on in this paper. As documented by Prince (1978), the *wh*-clause is old information while the counterweight is new information. In order to account for this pattern, the analyses of Heycock and Kroch (1999, 2002) and den Dikken et al. (2000) invoke the notions of topic and focus. While Heycock and Kroch (2002) claim that the counterweight moves into a focus projection at LF, den Dikken et al. (2000) place the *wh*-clause in [Spec, TopicP] in the overt syntax. In Malagasy, we will see direct evidence for overt topicalization of the *wh*-clause.

Stepping back from pseudoclefts for a moment, I would like to point out some closely connected sentence types that are relevant for discussion of the Malagasy data: specificational copular sentences and equatives. Higgins (1979) notes the parallel between specificational pseudoclefts and other kinds of specificational sentences. Thus the sentences in (7) are roughly synonymous.

- (7) a. What I am pointing at is a kangaroo.
b. The animal I am pointing at is a kangaroo. [Higgins 1979: (68)]

Pursuing this parallel, den Dikken (2006a) argues that all specificational sentences such as (7) are instances of inversion and he extends this analysis to equatives, such as (8).

- (8) Cicero is Tully.

Mikkelsen (2004) also applies the inversion analysis to specificational copular sentences, such as (7) and (9).⁵

- (9) [_{PREDICATE} The lead actress in that movie]_i is [_{SUBJECT} Ingrid Bergman *t*_i].

Mikkelsen argues for three central points: first, that the two DPs in (9) are of different semantic types: the first is a predicate and the second is an argument; second, the predicate DP is the

syntactic subject; third, the predicate DP is interpreted as a topic (discourse-old).⁶ What distinguishes Mikkelsen’s analysis from other approaches involving inversion is her focus on information structure, another important aspect of specificational sentences as discussed above. As we will see, the Malagasy data support treating specificational pseudoclefts, specificational copular sentences and equatives on a par, supporting den Dikken’s and Mikkelsen’s analyses.⁷

Pseudoclefts clearly raise issues that touch on syntax, semantics and pragmatics. The goal of this paper is to provide some insight into the nature of the interactions between these components of the grammar.

3. Background on Malagasy

Malagasy is a western Austronesian language spoken in Madagascar by approximately 13 million people. The word order, which is fairly rigid, is VOS, as illustrated in (10) (the subject is marked with a dotted underline). Another striking feature of Malagasy is the verbal voice morphology. As will be important later in the paper, the voice morphology indicates the “role” of the subject. (10) illustrates the standard voice paradigm for the root *tapaka* ‘cut’.⁸ Actor Topic morphology (usually the prefix *an-* or *i-*), as in (10a), indicates an agent subject; Theme Topic morphology (here the suffix *-ina*) in (10b) indicates a theme subject; Circumstantial Topic morphology (a circumfix consisting of AT morphology and a suffix *-ana*) in (10c) indicates that some other role is in subject position (e.g. instrument, time, location).

(10) a. Actor Topic (AT)

Nanapaka ity hazo ity tamin’ny antsy i Sahondra.

PST.AT.cut this tree this PST.P’DET knife Sahondra

Sahondra cut this tree with the knife.

b. Theme Topic (TT)

Notapahin'i Sahondra tamin'ny antsy ity hazo ity.

PST.TT.cut.GEN.Sahondra PST.P'DET knife this tree this

This tree was cut by Sahondra with the knife.

c. Circumstantial Topic (CT)

Nanapahan'i Sahondra ity hazo ity ny antsy.

PST.CT.cut.GEN.Sahondra this tree this DET knife

The knife was used by Sahondra to cut the tree.

The precise nature of the so-called voice system is the subject of much debate in the literature (Pearson 2005 analyzes Malagasy voice as *wh*-agreement; Rackowski and Richards 2005 treat Tagalog voice as case agreement; Aldridge 2004 claims that voice marks transitivity in both languages). For present purposes, I will continue to call this verbal morphology “voice”.

Similarly, the status of the clause-final argument is hotly disputed in current research on Malagasy and western Austronesian in general (see Schachter 1976 for a seminal article on this issue). Some refer to this argument as a subject (an A position) (Guilfoyle, Hung and Travis 1992), others as a topic (an A-bar position) (Pearson 2005) and some (Aldridge 2004) claim that these languages are ergative and there is no unified notion of subject. The data discussed in this paper may in fact provide evidence in favor of the A-bar analysis. As we will see, however, while the clause-final argument may be an A-bar position, it is a highly restricted position, reserved for referential arguments, much like subjects. In addition, Malagasy has a special clause-initial position reserved for topics (discourse familiar elements). I will discuss this topic position in 4.2.1. Finally, there are several different proposals concerning the structure of VOS word order in Malagasy. In this paper, I will assume that the clause-final subject is in a rightward

specifier position, following Guilfoyle, Hung and Travis (1992). But see Pearson (2001) and Rackowski and Travis (2000) for more Kaynean (Kayne 1994) alternatives involving predicate fronting.

Turning now to non-verbal clauses, we see that Malagasy lacks a copular verb and can be described as predicate initial, where the predicate can be an NP, AP or PP, as illustrated below.

(11) a. [Vorona ratsy feo]_{NP} ny goaika

bird bad voice DET crow

The crow is a bird with an ugly voice.

b. [Faly amin'ny zanany]_{AP} Rasoa

proud P'DET child.3SG(GEN) Rasoa

Rasoa is proud of her children.

c. [Any an-tsenana]_{PP} Rakoto

P ACC-market Rakoto

Rakoto is at the market.

Nominal predication is one of the core topics of this paper, so I will discuss it in more detail in the next section.

Finally, although Malagasy is predicate initial, it is possible for subjects to appear in the pre-predicate position. These apparently fronted subjects are marked by either the topic particle *dia* as in (12a) or by the focus particle *no*, as in (12b).

(12) a. Ny mpianatra dia mamaky teny.

DET student TOP AT.read word

The students, they are reading.

- b. Ny mpianatra no mamaky teny.
 DET student FOC AT.read word

It is the students who are reading. [Rajemisa-Raolison 1971: 30]

If both *dia* and *no* are present, *dia* precedes *no*.

- (13) [Ity radara ity] dia [ny Rosiana] no nanao azy.

this radar this TOP DET Russian FOC PST.AT.do 3(ACC)

As for this radar, it was the Russians who made it. [Keenan 1976: (69)]

I discuss topicalization in more detail in section 4.2.1. The structure of focus constructions is discussed in Dahl (1986), Paul (2001), Potsdam (in press), who argue that although these are often called clefts, the underlying structure is similar to pseudoclefts.⁹ In (12b), for example, *ny mpianatra* ‘the students’ is the predicate and *no mamaky boky* (lit.) ‘the ones who read books’ is a headless relative in subject position, as schematized below.

- (14) [[_{PREDICATE} Ny mpianatra] [_{SUBJECT} no mamaky teny]].
 DET student FOC AT.read word

‘The ones who are reading are the students.’

If the above analysis and the one presented in this paper are correct, Malagasy has two different kinds of pseudoclefts: one where the relative-like constituent is in the subject position (as in (14)) and one where the relative-like constituent has been topicalized – the pseudoclefts under analysis in this paper. Moreover, both kinds of pseudoclefts are “concealed”: neither has an overt *wh*-element.¹⁰ Because the structure of the focus construction is tangential to this paper, I do not discuss it in any more detail, but refer the reader to the above articles.

4. Nominal predication

Pseudoclefts in Malagasy crucially involve nominal predication – that is, a predication relationship between what appear to be two DPs. Because nominal predication is central to the analysis of pseudoclefts, I provide an overview of this topic in this section.

4.1 Indefinite predicates

As illustrated in (11a) above and (15a) below, indefinites (nominals without a determiner) can be predicates in Malagasy. Crucially, however, definites cannot (but see (19) for some exceptions). Definite DPs are those with a determiner or a demonstrative, as well as names and pronouns.¹¹ So while examples like (15a,c) are grammatical, (15b,d) are not.

(15) a. [predicate Mpanjaka] [subject Rakoto].

king Rakoto

Rakoto is/was (the) king.

b. * Ny mpanjaka Rakoto.

DET king Rakoto

c. Vadiko izy.

spouse.1SG(GEN) 3(NOM)

S/he is my spouse.

d. * Ny vadiko izy.

DET spouse.1SG(GEN) 3(NOM)

S/he is my spouse.

It is important to note that these predicates can be interpreted as definite, in spite of being formally indefinite (lacking a determiner). In other words, (15c) doesn't mean 'He is one of my

spouses’. Moreover, when used as an argument, a possessed DP must have a determiner (unlike in English).

- (16) Namangy *(ny) vadiko Rabe omaly.
PST.AT.meet DET spouse.1SG(GEN) Rabe yesterday
‘Rabe met my spouse yesterday.’

I take the contrast between (15c) and (15d) as key evidence that the “definiteness constraint” on the predicate position is a formal one, not a semantic restriction. A semantic definiteness restriction would erroneously rule out (15c). Instead, it appears that the predicate position excludes DPs that have some formal marking of definiteness.¹² If a predicate can move (e.g. to a topic position) it can escape the restriction. We will see this is precisely what happens in pseudoclefts.

The definiteness restriction on the predicate position means that simple identity claims, of the type ‘Cicero is Tully’, are not possible.

- (17)* Rabe Rakoto.
Rabe Rakoto
Rakoto is Rabe.

In order to express the equivalent of (17), one of the names must be topicalized.

- (18) Rakoto dia Rabe.
Rakoto TOP Rabe
Rakoto is Rabe.

This restriction on the predicate position and topicalization as a means to escape this restriction are two key elements of pseudoclefts in Malagasy, as I will show in section 6.

Although predicates cannot generally be definite, as shown by (15d) repeated in (19a), there are exceptions, such as (19b,c), originally pointed out by Rajaona (1972).

- (19) a. * Ny vadiko izy.
 DET spouse.1SG(GEN) 3(NOM)
 She/he is my spouse.
- b. Ny vadiko iny.
 DET spouse.1SG(GEN) that
 That is my spouse.
- c. Ny vadiko ilay olona teto omaly.
 DET spouse.1SG(GEN) DEF person PST.here yesterday
 The aforementioned person who was here yesterday is my spouse.

Note in particular that the crucial difference between (19a) and (19b) is that in the former the subject is a personal pronoun while in the latter the subject is a demonstrative pronoun. Similarly, in (19c) the subject appears with the definite determiner *ilay* and not the specific determiner *ny*. I do not have an explanation for the above examples, but I believe they show that there seems to be a requirement that predicates be “less definite” than subjects.¹³ We can violate the definiteness restriction on the predicate position just in case the subject is “highly definite” (e.g. a demonstrative). Obviously this account of the distinction between (19a) and (19b,c) is only a description. Moreover, the terms “less definite” and “highly definite” remain to be defined.¹⁴

4.2 Definite predicates

We now turn to the construction under investigation in this paper. As we just saw, in order to circumvent the definiteness restriction on predicates, Malagasy has recourse to what I will for the

moment call the *dia* construction. Further examples are given in (20). (20a) illustrates the unmarked word order, with an indefinite nominal predicate. (20b,c) illustrate topicalized definite nominal predicates.

- (20) a. Filoha Ravalomanana
 president Ravalomanana
Ravalomanana is the president.
- b. Ny filoha dia Ravalomanana.
 DET president TOP Ravalomanana
The president is Ravalomanana.
- c. Spiderman dia i Tobey Maguire ao amin'ily sary mihetsika.
 Spiderman TOP Tobey Maguire there P'DEF picture AT.move
Spiderman is (played by) Tobey Maguire in this film.

In (20b,c), two definite DPs flank the *dia* particle. Note that unlike in English, the word order is typically fixed. The “role” or pseudonym must precede *dia*, while the true name or identity must follow, as shown by the contrast between (20) and (21).

- (21) a. * Ravalomanana dia ny filoha.
 Ravalomanana TOP DET president
Ravalomanana is the president.
- b. * I Tobey Maguire ao amin'ily sary mihetsika dia Spiderman.
 Tobey Maguire there P'DEF picture AT.move TOP Spiderman
Tobey Maguire is (plays) Spiderman in this film.

The only examples that can be reversed are ones where the two names have the same status (neither is more basic than the other), as in (22) where one city has two different names.

(22) a. Diego dia Antsiranana.

Diego TOP Antsiranana

Diego is Antsiranana.

b. Antsiranana dia Diego.

Antsiranana TOP Diego

Antsiranana is Diego.

As we will see, specificational pseudoclefts in Malagasy are also not reversible. I take these facts as initial evidence that *dia* is not a copula (see section 6.5 for further arguments). The lack of reversibility points to a fixed underlying structure for equatives and specificational sentences: the role or pseudonym is predicated of the name and not vice versa.

Of interest for this paper are examples where the first DP is a definite nominalized predicate, as seen in (2), repeated in (23). Note that this is an instance of bare nominalization: the determiner *ny* precedes a predicate that is not otherwise morphologically marked as being nominal.

(23) a. Ny mahafinaritra dia izany vaovao izany.

DET AT.happy TOP that news that

What is pleasing is that news.

b. Ny nahatongavany dia omaly.

DET PST.CT.arrive.3(GEN) TOP yesterday

When he arrived was yesterday.

[Rajaona 1972]

In section 5, I will argue that the *dia* construction is the Malagasy equivalent of a pseudocleft, but before discussing these examples in more detail, it is necessary to discuss the *dia* particle and nominalizations.

4.2.1 *Dia*

In the literature, *dia* has been identified as a topic marker (Keenan 1976).

(24) Rakoto dia manasa lamba.

Rakoto TOP AT.wash cloth

Rakoto, he is washing clothes.

Keenan calls this kind of example “weak topicalization” and contrasts it with “strong topicalization”, as in (25).

(25) Raha Rakoto dia manasa lamba.

if Rakoto TOP AT.wash cloth

As for Rakoto, he is washing clothes.

He notes that the “weak topic” does not have to correspond to a topic of conversation, and is often an adverbial. There is very little research on topicalization in Malagasy (but see Flegg 2003), and there are no studies on how *dia* is used in connected discourse. Nevertheless, typical examples can be found in texts, such as the following sequence from a story (Rajohanesa 1963), where the weak topic is used to talk about a recently introduced entity:

(26) [...] manana andevo maromaro ihany izy. Ary izany andevony izany dia

AT.have slave many EMPH 3(NOM) and those slave.3(GEN) those TOP

manompo azy telo mianaka irery

AT.serve 3(ACC) three together-child alone

... he has many slaves. And those slaves of his serve them three (parents and one child) alone.

I assume that weak topics correspond to discourse-familiar information (“links” in the terminology of Vallduví 1992).¹⁵ That is, topics have either been previously mentioned, as in

(26), or they can be inferred from other material (“bridging”). In what follows, I argue that *dia* is the same lexical item in the various sentences under consideration; that is, it is consistently a topic particle.¹⁶

4.2.2 Nominalizations

Although Malagasy has productive nominalization morphology (Rajemisa-Raolison 1971: 26-27; Paul 1996; Ntelitheos 2005), it is also possible to nominalize a predicate by simply adding a determiner (Rahajarizafy 1960: 101).

- (27) a. ny ratsy
 DET bad
 evil (e.g. good vs. evil)
- b. ny anatin’ny vata
 DET inside’DET suitcase
 what is inside the suitcase
- c. ny nataony
 DET PST.TT.do.3(GEN)
 what he did
- d. ny mandainga
 DET AT.lie
 lying

These examples show that the interpretation of such bare nominalizations varies – some are event nominals, some are abstract nouns, some are object nouns. An object reading is possible with verbs marked with Actor Topic morphology, but requires a demonstrative rather than the regular determiner *ny*:

- (28) ireto miasa ireo
 these AT.work these
these who are working

The reasons for this restriction on the interpretation of bare nominalizations are not yet understood; for the purposes of this paper, however, what is important is that bare nominalizations are possible and productive and they can denote entities as well as events.

4.4 Past analyses

As mentioned in the introduction, the *dia* construction has received little attention in the literature, with the notable exception of Rajaona (1972). Rajaona claims that the pre-*dia* XP is the predicate and the post-*dia* XP is the subject. I will discuss Rajaona’s approach in more detail in section 6.4, but at first blush his analysis does not account for the fact that the *dia* construction has the reverse information structure of regular predicate-subject sentences. With standard word order, the predicate is focused (new information), as shown in (29a). In the *dia* construction, however, the initial XP is a topic (old information), as seen in (29b).

- (29) a. Tonga ny ankizy.
 arrive DET child
The children arrived. (rather than left)
- b. Ny tonga dia ny ankizy.
 DET arrive TOP DET child
The children arrived. (rather than the adults)

Due to this difference in information structure, (29a) cannot be the answer to the question ‘Who arrived?’, while (29b) can be. The change in information structure supports the hypothesis that *dia* is a topic marker in (29b). Despite this problem, we will see that Rajaona’s insight will be

crucial to the present analysis; see section 6.4 for further discussion. I further hypothesize that the *dia* construction is parallel to English pseudoclefts. I explore this hypothesis in more detail in the next section.

5. The *dia* construction = pseudocleft

Let us now look at the *dia* construction in more detail. As noted above, the *dia* construction involves a nominalized predicate followed by the *dia* particle, followed by another XP. Some examples are given in (30), which show that the post-*dia* XP can be of any category: DP, PP, VP, AP, CP.

- (30) a. Ny milalao dia [ny ankizy]_{DP}.
 DET AT.play TOP DET child
The ones who are playing are the children.
- b. Ny nahatongavany dia [tamin'ny Talata]_{PP}.
 DET PST.CT.arrive.3(GEN) TOP PST.P'DET Tuesday
When he arrived was on Tuesday.
- c. Ny ataon-dRabe dia [manasa lamba]_{VP}.
 DET TT.do.GEN-Rabe TOP AT.wash cloth
What Rabe is doing is washing clothes.
- d. Ny nariny dia [lafa]_{AP}.
 DET PST.TT.lost.3(GEN) TOP expensive
What he lost was expensive.
- e. Ny anontanian-dRasoana tena dia [raha misy Andriamanitra]_{CP}.
 DET CT.ask.GEN-Rasoana self TOP if AT.exist God
What Rasoana wonders is if God exists.

Based on their interpretation, I claim that these examples are similar to pseudoclefts in English, despite the lack of a *wh*-word.¹⁷ As a more robust point of similarity, consider the two possible interpretations of English pseudoclefts. As mentioned in section 2, pseudoclefts are typically ambiguous between a predicational reading and a specificational reading.¹⁸ On the predicational reading of (31), the interpretation is roughly “Something about Jessie is interesting”. On the specificational reading, the interpretation is simply “Jessie is interesting”.

(31) What Jessie is is interesting.

Although it is not possible to construct a single Malagasy sentence with the two readings, it is possible to have both predicational and specificational readings of the *dia* construction.¹⁹ (32a) is predicational and (32b) is specificational.

(32) a. Ny nomeko azy dia lafo.
 DET PST.give.1SG(GEN) 3(ACC) TOP expensive
 What I gave him was expensive.

 b. Ny nomeko azy dia ity peratra ity.
 DET PST.give.1SG(GEN) 3(ACC) TOP this ring this
 What I gave him was this ring.

Thus as in English, what I am calling pseudoclefts in Malagasy can be either specificational or predicational. In this way, we see the initial parallels between English pseudoclefts and the *dia* construction. I pursue these parallels in the following sections.

5.1 Specification

The difference between the specificational and predicational readings has been described in detail by Higgins (1976). Turning first to specification, we can draw a parallel between

specification and lists.²⁰ In the following Malagasy example, the pre *dia* XP acts like the heading of a list and the post-*dia* XP supplies the elements on the list.

- (33) a. Ny ilaiko dia fiara sy trano.
DET TT.need.1SG(GEN) TOP car and house
What I need is a car and a house.
- b. I need the following things: a car and a house.

Specificational readings also pattern with question-answer pairs: “What do you need?” “I need a car and a house”.

One way of formalizing the specificational interpretation is by appealing to the notion of variables. Consider the English example in (34):

- (34) a. What John is is important to himself.
- b. John is x, x=important to himself.

The *wh*-clause sets up a variable and the counterweight supplies the value for this variable. In English, the variable is created by *wh*-movement. As noted by many authors, the *wh*-clause in a specificational pseudocleft is non-referential (but see Heycock and Kroch 1999 for a dissenting view).

In Malagasy, on the other hand, there is no *wh*-phrase in the *dia* construction. Instead, the variable is set up by voice morphology. In other words, the voice morphology on the verb indicates the “role” of the variable. Thus with an actor topic verb, the role of the variable is agent (35a); with a theme topic verb, the variable corresponds to the theme (35b); and with circumstantial topic, the variable is any circumstance related to the verb, in (35c) an instrument.²¹

- (35) a. Ny manasa lamba dia Rabe.
 DET AT.wash cloth TOP Rabe
Who is washing clothes is Rabe.
- b. Ny sasan-dRabe dia ny lambany.
 DET TT.wash.GEN.Rabe TOP DET cloth.3(GEN)
What Rabe is washing are his clothes.
- c. Ny anasan-dRabe lamba dia ny savony.
 DET CT.wash.GEN.Rabe cloth TOP DET soap
What Rabe is washing clothes with is the soap.

This relation between the voice on the verb and the role of the specified element can account for the following contrast (initially noted by Rajaona 1972):

- (36) a. Ny tsy tiako dia ny tsy nahafahany fanadinana.
 DET NEG TT.like.1SG(GEN) TOP DET NEG PST.CT.pass.3(GEN) exam
What I don't like is his not passing the exam.
- b. * Ny tsy nahafahany fanadinana dia ny tsy tiako.
 DET NEG PST.CT.pass.3(GEN) exam TOP DET NEG TT.like.1SG(GEN)

Rajaona claims that (36b) is excluded because the predicate (the pre-*dia* XP) must have greater “extension” than the subject.²² Translating his terms into the ones we are using, we can say that the value of the variable in the pre-*dia* XP must be supplied by an element with the correct role. In (36a), the variable is the theme of *tiana* ‘like’ as marked by the Theme Topic morphology. The value is specified as the event of him not passing the exam. In (36b), on the other hand, the value of the variable must be some circumstance related to the event of him not passing the exam

(e.g. a time, a place, a reason), as marked by Circumstantial Topic morphology. My not liking cannot fill this role. Malagasy thus illustrates what we could call “voice connectivity”.

The examples in (36) illustrate one important difference between Malagasy and English. As is well known, specificational pseudoclefts in English are reversible, as shown in (37), unlike the Malagasy examples in (36):

- (37) a. What Jessie is is important to herself.
 b. Important to herself is what Jessie is.

An explanation of the lack of reversibility in Malagasy is provided in section 6.2.

5.2 *Predication*

The predicational reading of pseudoclefts has received the least attention in the literature, probably because it is fairly straightforward. Consider the English example:

- (38) What John is is worthwhile.

In this example, *what John is* simply functions as the referential subject of the sentence and *worthwhile* is predicated of the subject. In other words, *what John is* is an argument DP – it can appear in other argument positions.

- (39) I don’t like what John is.

The Malagasy examples of predicational pseudoclefts are similar. In (40a), the nominalized predicate *ny nomeko azy* ‘what I gave him’ is the subject of the predicate *lafo* ‘expensive’.²³ As a regular argument DP, it can also appear in the clause-final subject position, as in (40b).

- (40) a. Ny nomeko azy dia lafo.
 DET PST.TT.give.1SG(GEN) 3(ACC) TOP expensive
What I gave him was expensive.

- b. Nariny ny nomeko azy.
 PST.TT.lose.3(GEN) DET PST.TT.give.1SG(GEN) 3(ACC)

He lost what I gave him.

Thus the topicalization in predicational *dia* constructions can always be “undone”, as seen in (41).

- (41) a. Ny nomeko azy dia lafo.
 DET PST.TT.give.1SG(GEN) 3(ACC) TOP expensive

What I gave him was expensive.

- b. Lafo ny nomeko azy.
 expensive DET PST.TT.give.1SG(GEN) 3(ACC)

What I gave him was expensive.

I return to the issue of reversibility in section 6.2.

5.3 Connectivity

Most studies of pseudoclefts focus on connectivity effects. It has long been noted that specificational, but not predicational pseudoclefts show connectivity effects.

- (42) a. What John is is important to himself. [specificational only]
 b. What John is is important to him. [predicational only]

Malagasy specificational pseudoclefts pattern with English. (43) illustrates connectivity for Condition C: it is not possible to interpret the pronoun in the pre-*dia* XP as coreferent with the name (*Rabe*) in the post-*dia* XP.

- (43) Ny namidiny*_{i/j} dia ny tranon-dRabe_i.
 DET PST.TT.sell.3(GEN) TOP DET house.GEN.Rabe

What he sold was Rabe's house.

In a predicational pseudocleft, however, Condition C effects are absent: the pronoun and the name may be coreferential.

- (44) Ny novidiany_{i/j} dia lafo be loatra hoan-dRakoto_i.
 DET PST.TT.sell.3(GEN) TOP expensive very too for-Rakoto
What he bought was too expensive for Rakoto.

Similarly, Condition A connectivity is seen in the specificational pseudocleft in (45); in this example, the verb is marked with reciprocal morphology:²⁴

- (45) Ny ataon-dRabe sy Rakoto dia mifanasa.
 DET TT.do.GEN.Rabe and Rakoto TOP AT.RECIP.invite
What Rabe_i and Rakoto_j do is invite each other_{i+j}.

Binding of the reciprocal is blocked in a predicational pseudocleft:

- (46) * Ny novangian-dRabe sy Rakoto dia nifandainga.
 DET PST.TT.invite.GEN.Rabe and Rakoto TOP PST.AT.RECIP.lie
The ones who Rabe_i and Rakoto_j invited lied to each other_{i+j}.

Finally, as mentioned in 5.1, Malagasy exhibits “voice connectivity”. In other words, the voice on the pre-*dia* XP sets up a variable that the post-*dia* XP must fill. The semantic role of the post-*dia* XP is thus determined by voice. In this way, the post-*dia* XP acts like it is interpreted within the pre-*dia* XP. “Voice connectivity” is absent in predicational pseudoclefts: the post-*dia* XP is interpreted as a predicate, not as an argument of the pre-*dia* XP.

5.4 Other semantic effects

Many authors suggest that pseudoclefts are associated with certain semantic or pragmatic effects: first, the *wh*-clause must be old information, not new; second, the *wh*-clause is interpreted as

exhaustive. Rejecting the new versus old distinction, Declerck (1988) claims that it is possible to put new information in the *wh*-clause:

(47) A: John washed the dishes.

B: No, the one who *broke* the dishes was John.

Whatever the status of this exchange in English, such examples are not possible in Malagasy.

(48) shows that it is not possible to put new information in the nominalized predicate in a Malagasy pseudocleft. Thus in the context of A's statement in (48), B's reply is inappropriate, even if we attempt to emphasize the verb *namaky* 'to break'.

(48) A: Nanasa lovia Rabe.

PST.AT.wash dish Rabe

Rabe washed dishes.

B: #Tsia, ny namaky lovia dia Rabe.

no DET PST.AT.break dish TOP Rabe

No, the one who broke dishes is Rabe.

Turning now to exhaustivity, as in English, the results seem to vary. Consider first the following context: a group of friends (Rabe, Rakoto and Rasoa) go to see a comedy, but no one laughs. In this context (49) would be inappropriate because it presupposes that everyone else laughed and that the only one who didn't was Rabe.

(49) Ny tsy nihomehy dia Rabe.

DET NEG PST.AT.laugh TOP Rabe

The one who didn't laugh was Rabe.

(49) thus appears to illustrate the exhaustivity of pseudoclefts. In other cases, however, exhaustivity disappears. Thus the following conjunction is perfectly natural, allowing for my car to need more than just a new battery.

- (50) Ny ilain'ny fiarako dia batery vaovao
 DET TT.need.GEN.DET car.1SG(GEN) TOP battery new
 ary mila zavatra hafa koa ilay izy.
 and AT.need thing different too DEF 3(NOM)

What my car needs is a new battery and it also needs something else.

In this way, exhaustivity is an implicature rather than an entailment and is defeasible in the correct context. Declerck (1988) comes to the same conclusion about English pseudoclefts based on similar examples.²⁵

6. Structure

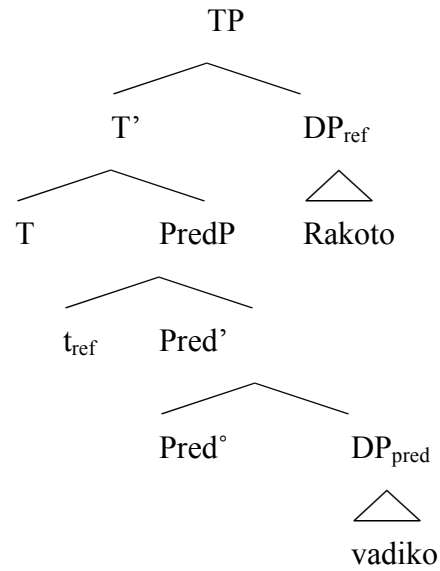
In this section, I propose a structure for pseudoclefts in Malagasy that involves overt topicalization to [Spec, TopP]. The head of Top° is spelled out as *dia*.

6.1 Nominal predicates

Following recent work on nominal predication (Adger and Ramchand 2003, Baker 2003), I assume that in order to act as a predicate, a noun must appear as the complement of a predication head (Pred°).²⁶ The subject of predication is merged in the specifier of the PredP small clause and raises to the matrix subject position. A simple example is given below: the predicational noun *vadiko* ‘my spouse’ is DP_{pred} and the subject of predication *Rakoto* is DP_{ref}.

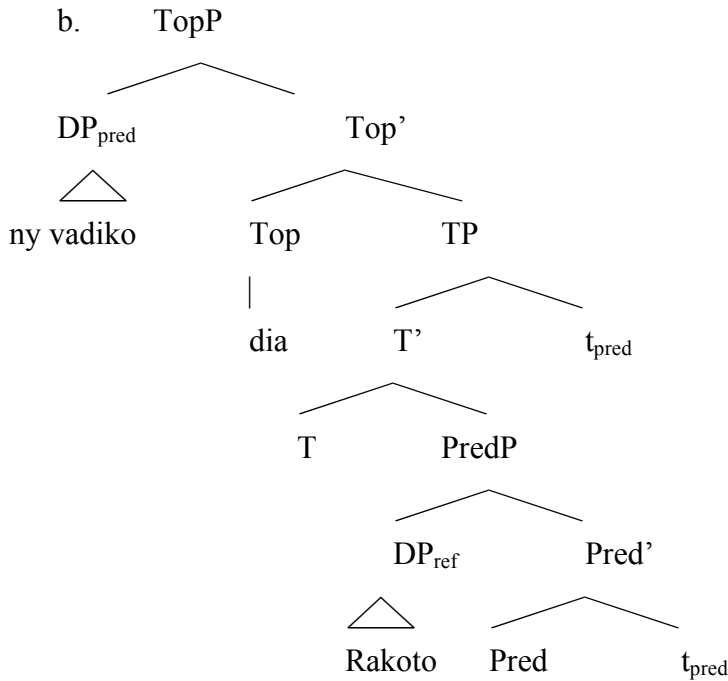
- (51) a. Vadiko Rakoto
 spouse.1SG(GEN) Rakoto
 Rakoto is my spouse.

b.



If the nominal predicate is definite, the underlying structure is the same, but it is the predicational DP that raises to subject position and is then topicalized.²⁷

- (52) a. Ny vadiko dia Rakoto.
 DET spouse.1SG(GEN) TOP Rakoto
My spouse is Rakoto.

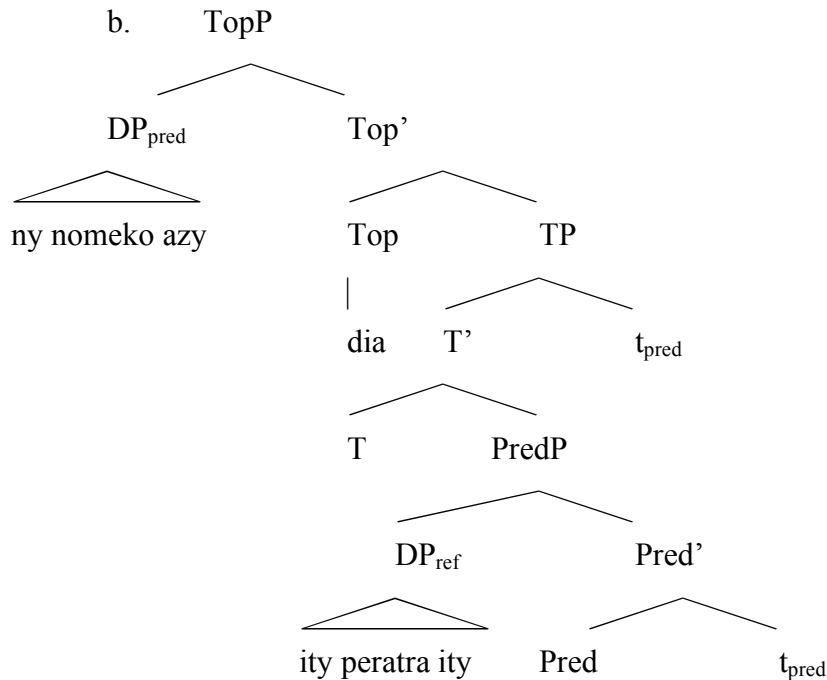


Recall from the discussion in 4.1 that due to a formal constraint, definite predicates are ungrammatical in situ in Malagasy. If the predicate DP remains in situ, the derivation crashes. To escape this definiteness restriction, movement of the predicate is obligatory in (52). As shown in the structure above, the predicate is topicalized, first passing through the subject position. To formalize this movement, I propose that the definite predicate can have a topic feature that checks the uninterpretable topic feature on Top° via movement to Spec, TopP. Two questions arise here. First, why doesn't the predicate move directly to the topic position? Why does it move via the subject position? I assume this two-step derivation because, as with all A-bar movement in Malagasy, topicalization is restricted to subjects.²⁸ The second question is why the predicate topicalizes. Why can't it remain in the subject position? I answer this question in section 7.

6.2 Specificational pseudoclefts

The structure of specificational pseudoclefts is the same as in (52). The nominalized predicate is merged as the DP_{pred} and the subject of predication is the small clause subject (DP_{ref}). The DP_{pred} is raised to subject and subsequently to the topic position.

- (53) a. Ny nomeko azy dia ity peratra ity.
 DET PST.give.1SG(GEN) 3(ACC) TOP this ring this
 What I gave him was this ring.



The proposed structure draws a direct parallel between specificational copular constructions (e.g. *Ny vadiko dia Rakoto* ‘My spouse is Rakoto’) and specificational pseudoclefts, as suggested by Mikkelsen (2004) and den Dikken (2006a). As with (52), topicalization is obligatory in (53). According to the proposed analysis, specificational pseudoclefts are non-reversible due to the definiteness restriction on the predicate position in Malagasy. Note that the subject of predication

can in fact be of any category, not just DP. In (30), we saw examples of PP, AP, and CP small clause subjects.²⁹

Although connectivity is not the focus of this paper, I follow Bošković (1997) in assuming that there is LF movement of the referential DP to a position within the predicational DP.³⁰ Connectivity (e.g. binding) is therefore calculated at LF. Let us consider this movement in more detail. In particular, I assume that the predicate DP contains a gap, a variable created by A-bar movement. In English pseudoclefts, this A-bar movement is made visible by the *wh*-word. In Malagasy, however, there is no *wh*-word; A-bar movement is signaled instead by the voice morphology on the verb, marking movement of an empty operator (Pearson 2005).³¹ Consider the relevant material from (53), as schematized in (54):

- (54) [Op_i ny nomeko t_i azy]
 DET PST.give.1SG(GEN) 3(ACC)
what I gave him

At LF, the referential DP (*ity peratra ity* ‘this ring’) replaces the trace.

- (55) [Op_i ny nomeko ity peratra ity azy]
 DET PST.give.1SG(GEN) this ring this 3(ACC)

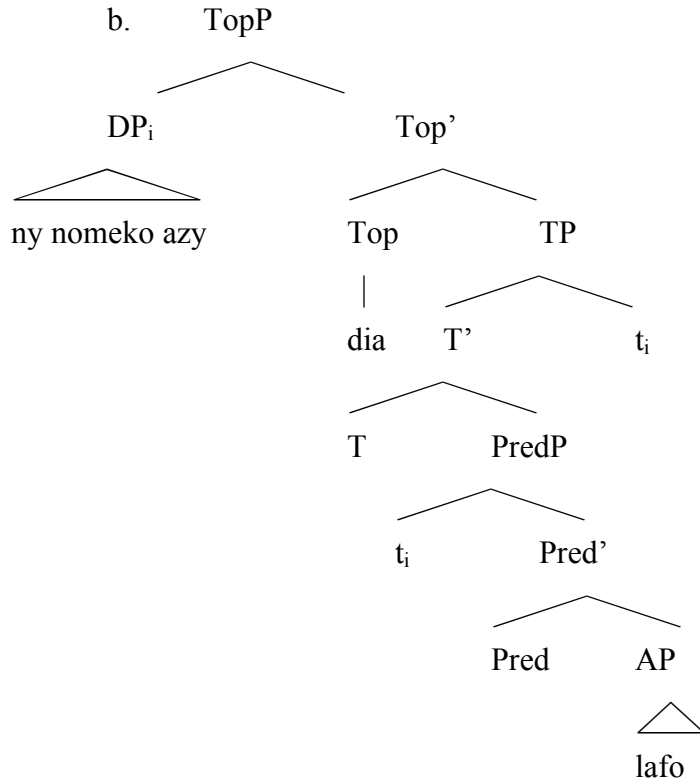
This LF replacement allows the post-*dia* XP to act like it occurs within the pre-*dia* DP for the purposes of binding and other dependencies.

6.3 *Predicational pseudoclefts*

Turning finally to predicational pseudoclefts, I suggest that these do not involve a predicate DP. Instead, the nominalized predicate is a referential DP. It is merged in the subject position of whatever predicate it occurs with. This difference between specificational and predicational pseudoclefts can be seen by comparing the tree in (53b) with the one below.³²

- (56) a. Ny nomeko azy dia lafo.
 DET PST.give.1SG(GEN) 3(ACC) TOP expensive

What I gave him was expensive.



Because the DP *ny nomeko azy* ‘what I gave him’ in (56) is a subject, not a predicate, it does not have to be topicalized. Instead, as in a typical case of predication, it is the subject of predication that raises to [Spec, TP]. Topicalization of this DP is therefore optional.

What explains the lack of connectivity in predicational pseudoclefts? Note that in both kinds of pseudocleft, the pre-*dia* DP contains a verb that has voice marking and therefore both have operator movement. Recall, however, that I have claimed an important difference between specificational and predicational pseudoclefts: in the former, the pre-*dia* XP is a predicate (semantic type <e,t>), while in the latter, it is an argument (semantic type <e>). I suggest that only in predicate DPs is the gap created by operator movement available to be filled at LF. The

DP in a predicational pseudocleft is an argument, not a predicate and therefore connectivity fails. This description corresponds to the standard distinction between arguments and predicates: predicates have a gap that is satisfied by an argument.

6.4 Tests for structure

We can now examine Rajaona's (1972) claim concerning pseudoclefts. According to him, in a sentence such as (57), the nominalized predicate *ny milalao baolina* 'the ones who are playing ball' is the predicate and the post-*dia* XP *ny ankizy* 'the children' is the subject.

(57) Ny milalao baolina dia ny ankizy.

DET AT.play ball TOP DET child

The ones who are playing ball are children.

As I pointed out in section 4.4, the information structure of (57) does not fit with Rajaona's characterization. There is also syntactic evidence that the post-*dia* XP is not in the matrix subject position, but surfaces in a structurally "lower" position. First, Malagasy has adverbs that appear on either side of VP. Pre-VP adverbs, such as *tokony* 'should' precede the predicate, as shown in (58a). Post-VP adverbs such as *foana* 'always' surface between the VP and the subject, as seen in (58b).

(58) a. Tokony hilalao baolina ny ankizy.

should FUT.AT.play ball DET child

The children should be playing ball.

b. Milalao baolina foana ny ankizy.

AT.play ball always DET child

The children are always playing ball.

A simple prediction follows: if *ny ankizy* ‘the children’ in (57) is the subject, both pre- and post-VP adverbs should precede it. This prediction is incorrect, as shown in (59). Pre-VP adverbs precede *ny ankizy* (59a), but post-VP adverbs follow it (59b).

- (59) a. Ny milalao baolina dia tokony ny ankizy.
 DET AT.play ball TOP should DET child
The ones who are playing ball should be the children.
- b. Ny milalao baolina dia ny ankizy foana.
 DET AT.play ball TOP DET child always
The ones who are playing ball are always the children.

Similarly, negation may precede *ny ankizy* ‘the children’ in a pseudocleft, but not when *ny ankizy* is a subject.

- (60) a. Ny milalao baolina dia tsy ny ankizy.
 DET AT.play ball TOP NEG DET child
The ones who are playing ball are not the children.
- b. * Milalao baolina tsy ny ankizy.
 AT.play ball NEG DET child
 (lit.) *Not the children are playing ball.*

A second argument against treating *ny ankizy* as a subject comes from pronouns. Subject pronouns in Malagasy have a particular form, often called nominative (Pearson 2005 calls them the ‘default’ forms). These pronouns can also appear in clefted and topicalized positions, with the exception of the first person singular pronoun. When this pronoun appears in a non-subject position (e.g. focus or topic), it must be realized as *izaho*, not *aho* (data adapted from Rajemisa-Raolison 1971: 60).

- (61) a. Tsy mahalala izany aho.
 NEG AT.know that 1SG(NOM)
I don't know that.
- b. Izaho no tsy mahalala izany.
 1SG FOC NEG AT.know that
It's me who doesn't know that.
- c. * Aho no tsy mahalala izany.
 1SG(NOM) FOC NEG AT.know that

If the post-*dia* XP is a first person singular pronoun, it must be *izaho*, not *aho*, suggesting it is not in the regular subject position.

- (62) a. Ny mihira dia izaho.
 DET AT.sing TOP 1SG
The one who is singing is me.
- b. * Ny mihira dia aho.
 DET AT.sing TOP 1SG(NOM)

The data in (59), (60) and (62) are difficult to account for if *ny ankizy* is the (matrix) subject. The proposed structure in (53b), however, accounts for the position of adverbs and negation as well as for the form of pronouns. In the structure, *ny ankizy* is not the matrix subject, but the subject of the small clause. At the same time, the proposed structure captures Rajaona's original insight: *ny ankizy* is indeed a subject, but not the matrix subject.

6.5 *Dia* ≠ *be*

Given the parallels I have drawn between the *dia* construction and pseudoclefts, one might ask whether *dia* is in fact a copula. In fact, Malagasy grammars often refer to *no* (the focus particle)

and *dia* as copulas (e.g. Rajemisa-Raolison 1971: 159, 161). In other words, the structure of a typical *dia* construction would be as below, where *dia* heads the VP:

- (63) [TP [DP Ny milalao baolina] [VP dia [DP ny ankizy]]].
 DET AT.play ball BE DET child

The ones who are playing ball are children.

There are several reasons, however, to reject this approach.³³ First, as we have already seen, Malagasy allows various categories as the main predicate without an overt copula.

- (64) a. [Vorona ratsy feo]_{NP} ny goaika
 bird bad voice the crow

The crow is a bird with an ugly voice.

- b. [Faly amin'ny zanany]_{AP} Rasoa
 proud P'DET child.3SG(GEN) Rasoa

Rasoa is proud of her children.

- c. [Any an-tsena]_{PP} Rakoto
 P ACC-market Rakoto

Rakoto is at the market.

The first puzzle is why the copula is required in (63) but banned from (64). The second puzzle involves word order. In order to successfully account for the position of *dia*, we would have to stipulate that it is the only predicate that must occur non-initially (a strange restriction given the overwhelmingly head-initial structure of Malagasy). Third, if *dia* were a predicate, we would expect the position of adverbs to reflect this. In other words, pre-predicate adverbs, such as

tokony ‘should’, would precede *dia* and post-predicate adverbs, such as *foana* ‘always’, would follow it. These predictions are not borne out, as shown by the contrasts below.

(65) a. Ny milalao baolina dia tokony ny ankizy.

DET AT.play ball TOP should DET child

The ones who are playing ball should be the children.

b. * Ny milalao baolina tokony dia ny ankizy.

DET AT.play ball should TOP DET child

The ones who are playing ball should be the children.

c. Ny milalao baolina dia ny ankizy foana.

DET AT.play ball TOP DET child always

The ones who are playing ball are always the children.

d. * Ny milalao baolina dia foana ny ankizy .

DET AT.play ball TOP always DET child

The ones who are playing ball are always the children.

Third, the lack of reversibility of sentences with *dia* between two DPs is surprising if *dia* is simply a copula. I repeat the relevant examples in (66):

(66) a. Ny filoha dia Ravalomanana.

DET president TOP Ravalomanana

The president is Ravalomanana.

b. * Ravalomanana dia ny filoha.

Ravalomanana TOP DET president

Ravalomanana is the president.

I therefore conclude that *dia* is not a copular verb. Given that *dia* otherwise consistently marks topicalization, as discussed in section 4.2.1, I treat it as the head of [TopicP].

7. Why topicalization?

In the previous section, I have proposed that specificational pseudoclefts involve overt topicalization of a predicational DP. Part of the motivation for movement in the above analysis is the definiteness restriction on the predicate position in Malagasy: the predicate must move out of its base position if it is formally definite. I left unexplained, however, why topicalization obtains; I therefore now discuss topicalization in pseudoclefts.

7.1 *Subjects vs. topics*

As we have seen, a definite predicate in Malagasy occurs in a clause-initial topic position (67a). Partial motivation for the movement of the predicate comes from the definiteness restriction on the predicate position. But this restriction does not explain why the predicate can't simply raise to the subject position. As shown by (67b), this is not possible in Malagasy.

(67) a. Ny mahafinaritra dia izany vaovao izany.

DET AT.happy TOP that news that

What is pleasing is that news.

b. * Izany vaovao izany ny mahafinaritra.

that news that DET AT.happy

What is pleasing is that news.

Nothing in the proposed analysis rules out (67b). I suggest that (67b) is ungrammatical because the subject position in Malagasy is restricted to argument DPs – DPs that are assigned a referential theta-role (Rizzi 1990).³⁴ Pearson (2001) shows that the subject position in Malagasy

can only host phrases of the category DP. Moreover, non-argument DPs such as measure phrases are blocked from surfacing in the subject position, as illustrated by the ungrammaticality of (68).

(68) * Lanjain'ity voankazo ity ny iray kilao.

TT.weigh.GEN'this fruit this DET one kilo

(lit.) *One kilo is weighed by this fruit.*

The same restriction that rules out (68) will also exclude (67b) given that the *wh*-clause is a predicate DP and therefore not an argument. Therefore the only converging derivation is one where the *wh*-clause is topicalized.

This account of (67b) requires two points of clarification. First, note that the restriction on the subject position is relevant to specificational sentences in general in Malagasy. Therefore even sentences with two proper names are subject to overt topicalization, as we have already seen and as illustrated in (69).

(69) a. * Rabe Rakoto.

Rabe Rakoto

Rakoto is Rabe.

b. Rakoto dia Rabe.

Rakoto TOP Rabe

Rakoto is Rabe.

I take these examples to show that specificational clauses are inherently asymmetric: one of the DPs is generated as a predicate and one is a subject. The predicate DP must topicalize for precisely the same reasons as the *wh*-clause in pseudoclefts.

Second, the proposed account raises the question of what counts as a topic. Linguistic research on western Austronesian languages has long puzzled over the status of the “subject”

(see e.g. Schachter's 1976 paper on Tagalog). Some claim that the subject is in fact a topic, an A-bar element (Richards 2000 on Tagalog, Pearson 2005 on Malagasy). It has long been noted that the subject position is highly topical: subject DPs typically are old information and must have a determiner. Keenan (1976) and Schachter (1976) therefore claim that subject in Malagasy and Tagalog must be definite (but see Law 2006 for arguments that some subject DPs are interpreted as indefinite). Although the arguments in favor of an A-bar position are strong, I think it would be incorrect to label it as a topic.³⁵ For example, it is possible to have negative polarity items in this position.

(70) Tsy mahatakatra izany na iza na iza.
 NEG AT.reach that or who or who

No one can afford that.

(lit.) *Anyone can't afford that.* [Dez 1990]

Moreover, as we have seen, the subject position is highly restricted, only hosting argument DPs, while the topic position is open to other categories (e.g. PPs, predicate DPs). I take the data presented in this paper as further evidence that the clause-final DP is not a topic, per se. Given that *dia* topicalization is obligatory in pseudoclefts, the clause-final DP is clearly not a possible position for at least certain kinds of topics. One solution is to claim that the two positions (clause-initial and clause-final) host distinct types of topics. Some kinds of topics are permitted clause-finally and others must front. A simpler solution is to conclude that the clause-final position is not a topic position at all. Clearly more research is needed on topicalization in Malagasy, but the present paper provides one argument in favor of the one-topic analysis.

7.2 *Inversion and information structure*

As described above, Malagasy pseudoclefts involve the fronting of new information. In this way, Malagasy pseudoclefts bear a certain resemblance to inversion, as discussed by Birner (1994).

(71) Sitting in the garden was an old man. [Birner 1994: (4)]

Birner shows that the preposed XP in inversion structures must not be less familiar than the post-verbal XP. In particular, she argues in favor of the notion of discourse familiarity as the relevant distinction for inversion. In the conclusion to her article, Birner argues that discourse familiarity correlates with relative linear position in the sentence, rather than with the subject position. In other words, discourse familiar elements typically appear clause-initially but not necessarily in the subject position. This ordering corresponds to a cross-linguistic preference to place old information before new information (this ordering is a preference and not an inviolable constraint, see e.g. Tomlin and Rhodes 1992 on Ojibwa). In English it can sometimes be difficult to distinguish the subject from the topic position, but the data from Malagasy support Birner's claim. As has been clear from the data presented in this paper, the discourse familiar information in an inversion sentence appears clause-initially, not clause-finally (i.e. not in the subject position).

The Malagasy facts also pattern with English pseudoclefts. Drawing on a corpus of naturally occurring discourse, Prince (1978) argues that the *wh*-clause in English pseudoclefts must contain information inferable from the discourse. I take this to mean that the *wh*-clause is a discourse topic. Pseudoclefts in English thus involve a kind of inversion that places the *wh*-clause in initial position. Den Dikken et al. (2000) argue that the *wh*-clause is in a topic position, much like I have argued for Malagasy. Once again, since English is a subject-initial language, these notions (subject vs. topic) are often difficult to tease apart. The data from Malagasy,

however, show clearly that the discourse topic in an inversion structure is preposed and is not in the subject position. Thus we have further evidence that information structure is typically sensitive to word order rather than to hierarchical structure.

8. Conclusion

This paper has explored the structure of pseudoclefts in Malagasy. In particular, I have argued that specificational pseudoclefts are derived via topicalization of the predicational DP from a small clause. Although the paper has focused on Malagasy, the data are relevant to current discussions of the structure of pseudoclefts cross-linguistically and to research on specificational predication in general. The data support so-called inversion analyses of specificational pseudoclefts and equatives and also underline the importance of information structure in specificational clauses. Moreover, this paper has touched on issues that are central to the syntax of Malagasy: first, I have argued that voice morphology acts much like *wh*-movement in setting up a variable in pseudoclefts, which suggests that voice morphology is indeed *wh*-agreement, as argued by Pearson (2005); second I have addressed the issue of what is a topic in Malagasy and I have concluded that the clause-final subject is not a topic position, despite the fact that it typically hosts discourse-old information. In order to fully understand the status of the subject position, what remains to be determined is the precise role played by determiners in Malagasy and whether they indicate definiteness, specificity or something else entirely.

* I would first like to thank the Malagasy speakers who very patiently helped me with the data: Emma Mamifarananahary, Hasina Mihaingosa, Dina Rakoto-Ramambason, Hanta Rakotoarivony, Georges Ralaisoa, Vololona Rasolofoson, Francine Razafimboaka. Unless otherwise indicated, all data are from my own fieldwork. I would also like to thank three anonymous reviewers, whose careful comments and insightful questions greatly improved this

paper. Finally, I would also like to acknowledge the feedback from Diane Massam, Matt Pearson, Eric Potsdam and Lisa Travis, as well as from participants at the Stanford Austrofest 2005 and audiences at McGill, University of Toronto and University of Western Ontario. Any errors are my own.

¹ Abbreviations used in this paper:

ACC – accusative	AT – actor topic	CT – circumstantial topic	DEF – definite determiner
DET – specific determiner	FOC – focus particle	FUT – future	GEN – genitive
NEG – negation	NOM – nominative	P – preposition	PRT – particle
PST – past	RECIP – reciprocal	SUPER – superlative	TOP – topic particle
TT – theme topic			

² Note that den Dikken et al. (2000) follow Ross (1972) and argue that the *wh*-clause is a full IP, while I will argue that the *wh*-clause in Malagasy is a predicate DP. The structure I propose for Malagasy is in fact more like the structure they argue underlies “reverse” or “Type B” pseudoclefts (*important to herself is what Jessie is*).

³ The examples in (4) are not an exhaustive list of connectivity effects. For a more complete discussion, I refer the reader to the survey in den Dikken (2006b).

⁴ Den Dikken et al. (2000) and den Dikken (2006a) argue that pseudoclefts fall into two types, A and B. Type A pseudoclefts are base-generated, while Type B involve inversion.

⁵ Mikkelsen does not focus on pseudoclefts, but her thesis inspired the analysis presented in this paper.

⁶ Mikkelsen crucially argues against overt topicalization in specificational sentences.

⁷ Den Dikken (2006a) notes that Scots Gaelic lacks both specificational pseudoclefts and equatives, further confirming the close connection between the two sentence types. Mikkelsen (2004), however, explicitly denies that equatives are inversion structures.

⁸ Here I provide a “traditional” description of voice morphology (e.g. Rajemisa-Raolison 1971). Pearson (2005), however, has a different analysis of Actor Topic. Moreover, I have simplified the description to only include three different voices. There are in fact several others, as well described in Rajemisa-Raolison (1971) and others. I believe, however, that these details are tangential to the present paper.

⁹ Law (2005), however, argues for a cleft analysis.

¹⁰ As noted by an anonymous reviewer, one important difference between the two kinds of pseudoclefts is the class of elements that can be focused. In examples such as (14), only DPs and PPs can be focused (much like English clefts). As we will see, in the pseudoclefts discussed in this paper, all categories can be focused (much like English pseudoclefts).

¹¹ Note that names and pronouns arguably contain a determiner. The determiner for adult proper names is *Ra* (always written as a prefix) and the determiner for children is *i* (written separately). The determiner *i* also shows up in pronouns and the demonstratives.

¹² Interestingly, this constraint on the predicate is the mirror image of the constraint on the subject position in Malagasy. It is traditionally said that subjects in Malagasy must be definite (Keenan 1976) or specific (Paul 2000). But as noted by Law (2006), this is a purely formal constraint and subjects marked with a determiner are not always interpreted as definite or even specific. And as a reviewer points out, this is in general true in Malagasy – the presence of a

determiner does not always correlate with definiteness/specificity. In the following example, the genitive agent *ny ankizy* ‘the children’ can be interpreted as indefinite.

(i) Takatry ny ankizy ny baolina.

reach.GEN DET child DET ball

The/some children reached the ball.

Although much more careful research is required on this topic, all of these observations point to the definiteness constraint on the predicate being a formal device, rather than a semantic one.

¹³ Givón (1973: 118) makes the same claim about language in general: “A predicate may never be **more** referential than its subject.”

¹⁴ This notion of definiteness may provide an explanation for why definite predicates are always possible in the focus construction if we assume that the headless relative in the subject position is “highly definite”.

(i) Ny mpianatra no mamaky teny.

DET student FOC read word

The ones who are reading are the students. [Rajemisa-Raolison 1971: 30]

Law (2005), however, takes data such as (i) to show that the focused element is not in fact a predicate. I leave this debate to future research.

¹⁵ Topics (including weak topics) must be definite (Keenan 1976; Paul 2000). But as we have seen, formal marking of definiteness doesn’t always track semantic definiteness, so this restriction is not necessarily an indication of topics being discourse-old.

¹⁶ There are other uses of *dia*, but they have a very distinct interpretation, clearly not a part of the meaning of pseudoclefts. For example, *dia* can be used to mean ‘and then’ and it is also used to form one kind of superlative (Rajemisa-Raolison 1971: 159).

¹⁷ The examples show that the range of possible *dia* constructions is larger than English pseudoclefts.

¹⁸ But see Iatridou and Varlokosta (1998) for arguments that some language only have predicational pseudoclefts. See also Alexiadou and Giannakidou (1998) for counterarguments.

¹⁹ The best English examples to illustrate specificational versus predicational readings use *be*. Since Malagasy doesn't have a copula, it is impossible to create parallel examples. Other examples that are ambiguous in English are apparently unambiguously specificational in Malagasy:

(i) Ny tsy hanin-dRabe dia ny sakafon'ny alika.

DET NEG TT.eat.GEN.Rabe TOP DET food.GEN.DET dog

What Rabe doesn't eat is food for the dog.

The only interpretation of (i) is that Rabe eats all sorts of things but never dog food. This sentence can't mean that Rabe's leftovers are given to the dog. There seems to be a preference to interpret DPs as arguments rather than predicates which blocks the predicational reading here.

²⁰ In fact, *dia* is used at the beginning of lists (example from Jedele and Randrianarivelo 1998).

(i) Ny tanjon'ity fikambanana ity moa dia: voalohany indrindra,

DET goal.GEN.this organization this PRT TOP first SUPER

ny fitsinjovana ny ho aviny, ary faharoa...

DET taking-care-of DET FUT come.3(GEN) and second

The goals of the organization are first of all, to provide for its own future and second...

²¹ The pattern in (35) fits with Pearson's (2005) analysis of voice morphology as *wh*-agreement.

These data could therefore be taken as evidence in favor of the A-bar analysis of the clause-final subject position in Malagasy. See section 6.2.

²² Once again, this is reminiscent of Givón (1973: 119): “A predicate may never be **less general** than its subject.”

²³ For reasons which are not yet clear to me, speakers have much more difficulty judging predicational pseudoclefts reading, often hesitating over examples such as (40a). This difficulty may also explain why it has been impossible to construct truly ambiguous examples. See footnote 19.

²⁴ Due to the restricted distribution of anaphors in Malagasy, combined with the lack of a copular verb, it is impossible to create sentences parallel to (42).

²⁵ Eric Potsdam (p.c.) points out that exhaustivity may be relativized to the particular context. If we’re talking about why my car won’t run and it’s because someone stole the tires and the battery, *What my car needs is a new battery* is infelicitous. On the other hand, *What my car needs is a new battery and tires and it also needs a new radio come to think of it*, is not infelicitous because the radio is incidental to getting my car running.

²⁶ Unlike Adger and Ramchand, however, I allow Pred° to select DP.

²⁷ In what follows, I assume that topicalization involves movement rather than base-generation. There are obvious differences between the two approaches but these differences, to my best knowledge, are not relevant to the analysis proposed in this paper.

²⁸ As will be shown in section 6.4, the small clause subject is not in the matrix subject position. I take the movement of the predicate DP through the subject position to block raising of the small clause subject.

²⁹ A reviewer points out that this class of elements can never be the grammatical subject and asks if there is evidence that the AP, PP, CP are small clause subjects (and not predicates). As far as I have been able to determine, there are no syntactic arguments for this structure. My motivation

for placing the counterweight in the specifier position is instead essentially conceptual. In specificational pseudoclefts, the AP or PP or CP is not acting like a predicate. Instead, it acts like an argument of the nominalized predicate: it fills in the value for the variable inside the predicate. I take this to be the role of the small clause subject. I would like to thank the anonymous reviewer for making me think through this point in more detail.

³⁰ As noted by a reviewer, this movement is “sideways” movement: the landing site of the raised DP does not c-command its trace. For Bošković (1997), however, this movement does not leave a trace and hence is not ruled out as improper movement.

³¹ The exact position of the empty operator is not important for present purposes; see Pearson (2005) for discussion.

³² For the sake of parallelism, I have posited a PredP dominating AP. If APs can be predicates without PredP, the analysis remains the same: the DP subject would be projected in the specifier of AP.

³³ Unlike verbs in Malagasy, *dia* does not inflect for tense or mood (e.g. imperative) and it lacks voice alternations. Given the unusual morphological properties of copulas cross-linguistically, I do not take this as strong evidence against *dia* being a copular verb.

³⁴ As pointed out by a reviewer, traces seem to be immune to this restriction, which is potentially a problem given the Minimalist assumption that traces are copies. I suggest here that the restriction to argument DPs holds for overt subjects.

³⁵ Matt Pearson (p.c.) agrees that the subject position is not a true topic position.

References

- Adger, D. and G. Ramchand. 2003. Predication and equation. *Linguistic Inquiry* 34: 325-359.
- Akmajian, A. 1970. *Aspects of the grammar of focus in English*. PhD thesis, MIT.

- Aldridge, E. 2004. *Ergativity and word order in Austronesian languages*. PhD thesis, Cornell University.
- Alexiadou, A. and A. Giannakidou. 1998. Specificational pseudoclefts as lists. In *Proceedings of WCCFL 17*, ed. K. Shahin et al., 1-15, Stanford: CSLI Publications.
- Baker, M. 2003. *Lexical categories. Verbs, nouns and adjectives*. Cambridge: Cambridge University Press.
- Birner, B. 1994. Information status and word order: An analysis of English inversion. *Language* 70: 233-259.
- Bošković, Z. 1997. Pseudoclefts. *Studia Linguistica* 51: 235-277.
- Dahl, O.C. 1986. Focus in Malagasy and Proto-Austronesian. In *FOCAL I: Papers from the Fourth International Conference on Austronesian Linguistics*, ed. P. Geraghty, L. Carrington, S.A. Wurm, 21-42. Canberra: Australian National University.
- Declerck, R. 1988. *Studies on copular sentences, clefts and pseudo-clefts*. Leuven: Leuven University Press/Foris Publications.
- Dez, J. 1990. *Cheminements linguistiques malgaches*. Paris: PEETERS/SELAF.
- Dikken, M. den. 2006a. *Relators and linkers. The syntax of predication, predicate inversion and copulas*. Cambridge, MA: MIT Press.
- Dikken, M. den. 2006b. Specificational copular sentences and pseudoclefts. In *The Blackwell companion to syntax*, ed. M. Everaert and H. van Riemsdijk, 292-409. Oxford: Blackwell.
- Dikken, M. den, A. Meinunger and C. Wilder. 2000. Pseudoclefts and ellipsis. *Studia Linguistica* 54: 41-89.
- Flegg, J. H. 2003. Topics and clitic left dislocation in Malagasy. Paper presented at AFLA X, University of Hawai'i.

- Givón, Talmy. 1973. Opacity and reference in language: An inquiry into the role of modalities. In *Syntax and Semantics Volume 2*, ed. J. Kimball, 95-122. New York: Seminar Press.
- Guilfoyle, E., H. Hung and L. Travis. 1992. Spec of IP and Spec of VP: Two subjects in Austronesian languages. *Natural Language and Linguistic Theory* 10: 375-414.
- Heycock, C. 1994. *Layers of predication: The non-lexical syntax of clauses*. New York: Garland.
- Heycock, C. and A. Kroch. 2002. Topic, focus and syntactic representations. In *WCCFL 21 Proceedings*, ed. L. Mikkelsen and C. Potts, 141-165. Somerville, MA: Cascadilla Press.
- Heycock, C. and A. Kroch. 1999. Pseudocleft connectivity: Implications for the LF interface. *Linguistic Inquiry* 30: 365-397.
- Higgins, F. 1979. *The pseudocleft construction in English*. New York: Garland.
- Iatridou, S. and S. Varlokosta. 1998. Pseudoclefts crosslinguistically. *Natural Language Semantics* 6: 3-28.
- Jacobson, P. 1995. On the quantificational force of English free relatives. In *Quantification in natural language*, ed. E. Bach, E. Jelinek, A. Kratzer and B. Partee, 451-486. Dordrecht: Kluwer.
- Jedele, T.P. and L. Em. Randrianarivelo. 1998. *Malagasy newspaper reader*. Kensington, MD: Dunwoody Press.
- Kayne, R. 1994. *The antisymmetry of syntax*. Cambridge, MA: MIT Press.
- Keenan, E. 1976. Remarkable subjects in Malagasy. In *Subject and topic*, ed. C. Li, 249-301. New York: Academic Press.
- Law, P. 2005. Questions and clefts in Malagasy. In *Proceedings of AFLA XII (UCLA Working Papers in Linguistics, no. 12)*, ed. J. Heinz and D. Ntelitheos, 195-209. UCLA.

- Law, P. 2006. Argument-marking and the distribution of wh-phrases in Malagasy, Tagalog and Tsou. *Oceanic Linguistics* 45: 153-190.
- Mikkelsen, L. 2004. *Specifying who: On the structure, meaning and use of specificational copular clauses*. PhD thesis, University of California at Santa Cruz.
- Moro, A. 1997. *The raising of predicates: Predicative noun phrases and the theory of clause structure*. Cambridge: CUP.
- Ntelitheos, D. 2005. Malagasy participant nominalizations: A structural account. In *Proceedings of AFLA XII (UCLA Working Papers in Linguistics, no. 12)*, ed. J. Heinz and D. Ntelitheos, 313-327. UCLA
- Paul, I. 1996. F-words and other (Malagasy) nominals. In *Proceedings of Canadian Linguistics Association*. Calgary Working Papers in Linguistics: 321-332.
- Paul, I. 2000. *Malagasy clause structure*. PhD thesis, McGill University.
- Paul, I. 2001. Concealed pseudo-clefts. *Lingua* 111: 707-727.
- Pearson, M. 2001. *The clause structure of Malagasy: A minimalist approach*. PhD thesis, UCLA.
- Pearson, M. 2005. The Malagasy subject/topic as an A-bar element. *Natural Language and Linguistic Theory* 23: 381-457.
- Potsdam, E. To appear. More concealed pseudo-clefts. *Lingua*.
- Prince, E. 1978. A comparison of wh-clefts and it-clefts in discourse. *Language* 54: 883-906.
- Rackowski, A. and N. Richards. 2005. Phase edge and extraction: A Tagalog case study. *Linguistic Inquiry* 36: 565-599.
- Rackowski, A. and L. Travis. 2000. V-initial languages: X or XP movement and Adverbial Placement. In *The Syntax of Verb Initial Languages*, ed. A. Carnie and E. Guilfoyle, 117-141. Oxford: Oxford University Press.

- Rahajarizafy, A. de Padoue. 1960. *Essai de grammaire malgache*. Antanimena Tananarive: Imprimerie Catholique.
- Rajaona, S. 1972. *Structure du malgache: Étude des formes prédicatives*. Fianarantsoa: Librairie Ambozontany.
- Rajemisa-Raolison, R. 1971. *Grammaire malgache*. Fianarantsoa: Librairie ambozontany.
- Rajohanesa, P. 1963. *Ilay kintana mamirapiratra*. Antananarivo: Imprimerie lutherienne.
- Rizzi, L. 1990. *Relativized minimality*. Cambridge, MA: MIT Press.
- Ross, J. 1972. Act. In *Semantics of natural language*, ed. D. Davidson and G. Harman, 70-126. Dordrecht: Reidel.
- Schachter, Paul. 1976. The subject in Philippine languages: Topic, actor, actor-topic or none of the above? In *Subject and topic*, ed. Charles N. Li, 491-518. New York: Academic Press.
- Schlenker, P. 2003. Clausal equations. *Natural Language and Linguistic Theory* 21: 157-214.
- Sharvit, Y. 1999. Connectivity in specificational sentences. *Natural Language Semantics* 7: 299-341.
- Tomlin, R. and R. Rhodes. 1992. Information distribution in Ojibwa. In *Pragmatics of word order flexibility*, ed. D. Payne, 117-135. Amsterdam: John Benjamins.
- Vallduví, E. 1992. *The informational component*. New York: Garland.
- Williams, E. 1983. Semantic vs. syntactic categories. *Linguistics and Philosophy* 6: 423-446.