Externally and internally headed relative clauses in Marori

I Wayan Arka
Australian National University / Udayana University

Proceedings of the Joint 2016 Conference on Head-driven Phrase Structure Grammar and Lexical Functional Grammar

Polish Academy of Sciences, Warsaw, Poland

Doug Arnold, Miriam Butt, Berthold Crysmann, Tracy Holloway King, Stefan Müller (Editors)

2016

Stanford, CA: CSLI Publications

pages 23–42

Keywords: relative clauses, information structure, definiteness constraints, Papuan languages, discontinuous expressions

Arka, I Wayan. 2016. Externally and internally headed relative clauses in Marori. In Doug Arnold, Miriam Butt, Berthold Crysmann, Tracy Holloway King & Stefan Müller (eds.), *Proceedings of the Joint 2016 Conference on Head-driven Phrase Structure Grammar and Lexical Functional Grammar, Polish Academy of Sciences, Warsaw, Poland*, 23–42. Stanford, CA: CSLI Publications. DOI: 10.21248/hpsg.2016.2.

Abstract

This paper discusses relative clauses (RCs) in Marori, showing that this language unusually has almost all of relative clause types, from headed/headless, externally/internally headed, single-/double-headed, to pre-/post-head, to attached/detached RCs. Special attention is given to internally headed relative clauses (IHRC). It is argued that Marori IHRCs are of the restrictive or non-maximalising type, which accounts for certain intriguing properties, such as their indefiniteness constraints and the possibility for RC stacking.

1 Introduction

Marori (ISO 639-3: mok; a subgroup-level isolate, TNG/Papuan, highly endangered, around a dozen fluent speakers) is, like many other Papuan languages, predominantly verb final with free word order. It is perhaps unusual as far as its relative clause typology (RCs) is concerned.[1] It has almost all of relative clause types: headed and headless RCs, externally and internally headed RCs, pre- and post- head RCs, as well as detached RCs or co-relatives. In addition, all grammatical relations (subject, objects, obliques and adjuncts) are relativisable.

Internally headed relative clauses (IHRCs), while constrained, may give rise to ambiguity, if out of context. This is discussed in detail in section 4.4. Thus, either the patient 'bench' or the instrument 'club' can be understood as the relativised noun in the following example of IHRC in Marori:[2]

(1) [Keme na **njaj**=i **samagau** ngge terme-ben]$_{IHRC}$
REL 1SG bench=U club with 3SG.U.M.hit-1SGNrPST

tamba keiwei nggu-f
PERF damaged 3SG.M.U.AUX-NrPST
a) 'The **bench** that I hit with the club was damaged.'
b) 'The **club** with which I hit the bench was damaged.'

However, there is an intriguing definiteness constraint, which can disambiguate them. For example, if one were a proper name (further

1 I gratefully acknowledge the support of ARC Discovery Grant (DP10100307). For helpful discussion and feedback, I thank the anonymous reviewers and audience at the HeadLex 2016 Conference and the Departmental Seminar at Sydney University (23/09/2016) in particular Ron Kaplan, Tracy H. King, Doug Arnold, Bill Palmer, Jim Martin, and Sebastian Fedden. I also thank my research assistant, Agus Mahuze, and Marori language consultants in particular Pak Paskalis and Pak Mikael (both deceased), Pak Lukas and Pak Willem, for their help with data and my fieldwork in Merauke.

2 Abbreviations, alphabetically ordered: 1,2,3 (first, second and third person), A (Actor), Acc (accusative), ARG (argument), AUX (auxiliary), F (feminine), FUT (future), Gen (genitive), GF (grammatical function), HAB (habitual), LOC (locative), IRR (irrealis), MID (middle), NML (nominaliser), NOM (nominative), NPL (nonplural), M (masculine), MP (macro present), NrPST (near past), O (object), P (Patient), PL (plural), POSS (possessive), PRES (present), Q (question marker), REAL (realis), REL (relativiser), RmPST (Remote Past), S (intransitive subject), STAT (stative), SG (singular), TOP (topic), U (undergoer).

discussed later in section 4), such ambiguity would not arise. This definiteness constraint of IHRCs in Marori can arguably be accounted for in terms of Grosu's (2012) semantic typology of RCs; that is, IHRCs in Marori are essentially of the restrictive type, having non-specific indefinite intersective force. Proper names are nominals with unique and inherently definite referents, incompatible with the intersective force of the restrictive RC type. The findings on RCs reported in this paper provide a good empirical basis for the typological and theoretical study of RCs.

The paper is organised as follows. After an overview of the clausal structure and RC marking in Marori (section 2), the basic facts and the salient properties of RCs are outlined in section 3. The discussions given in section 4 provide an explicit analysis of Marori RCs in terms of their c-structure properties (4.1), lexical entries and functional annotations involved (section 4.2), and the demonstration of how the analysis works (section 4.3). That IHRCs in Marori belong to the non-maximalising type is discussed in section 4.4. Finally, the conclusion is given in section 5.

2 Clausal structure and marking RCs in Marori

The morphosyntax of the finite clausal structure in Marori is depicted in (2). It captures the following salient properties of Marori morphosyntax: (i) it is non-configurational (i.e. no VP); (ii) clausal word order is typically verb-final, but it allows postverbal arguments; (iii) argument NPs are freely ordered subject to certain information structure constraints (Arka 2016); (iv) the grammar shows a clear verb-noun distinction with the verb being the clausal head and inflected, (v) the predicate can be complex with the NP preceding the inflected (light/auxiliary) functioning as a lexical predicate; (vi) argument marking is semantically transparent, with core NPs flagged for the U role with the clitic *=i;* and (vii), verb inflections show TAM information and agreement with the prefix and suffix indexing U and A arguments respectively. Illustrative examples are given in (3)-(4).

(2) DP*(*=i*) , [PREF:U-V.Root- SUFF:A]$_V$
ARG(=U)/PRED (inflected)

(3) a. *mbe=na kundo-ru*
PART=1SG run-1SG.FUT
'I will run.'

b. *pa=na ter=∅-me-ru*
soon=1SG hit=3-AUX-1SG.FUT
'I will hit him/her.'

(4) a. pa=ka=i kara ku-nggo.
soon=2SG=U sick 2SG-AUX.3NPL.FUT
'You will be sick.'

b. tat, tamba kwon k-imb-ra-f.
grandfather already misquito 2SG-bite-PL-3.NrPST
("you") OBJ SUBJ
'Granpa, you're bitten by misquitos.'

c. Thomas fis mara=i nde-∅-f nggambe.
Thomas yesterday stone=U 3SG.bring-3A-NrPST there
'Thomas brought the stone there yesterday.'

A relative clause (RC) is an adjunct within NP, which itself is part of the nominal structure (DP) whose structure is shown in (5). The D (Determiner) comes before or after the NP. The adjunct RC and other elements are freely ordered within the NP. Sentence (1) exemplifies a postnominal RC in Marori. Note that I assume all nodes in the c-structure are optional.

(5) DP → NP, D. b. NP → PossP, NUM, XP:ADJUNCT, N.

(1) efi moipur ki=kwundo-f tamba soron
that child SG.REL=run-NrPST PERF fall.NrPST
'The child who just ran away already fell off.'

The RC marker in (1) is *ki=*, the shortened form of *kei*/*kefi,* which is also used as a spatial proximal (PROX) deictic in Marori.[3] The full set of RC markers in Marori is given in (2). These markers consist of the PROX formative *k-*, with the stems *efi*/*em(nd)e*, which are actually the third person pronouns in Marori. The same forms are also used as demonstratives. The form *keme* has general number, usable for any number.

(2) RC markers in Marori:

SG	NSG	GENERAL NUMBER
kefi/ *kei*/*ki*	*kemnde*	*keme*

3 Defining RCs and the salient RC properties in Marori

3.1 Definition and challenges

There are three important related aspects in the definition of RCs: syntactic, semantic and pragmatic. Syntactically, a RC is a subordinate clause, functionally an adjunct within a nominal structure. Semantically, it delimits the reference of the nominal by specifying the role of the referent of that NP in the situation described by the RC (Andrews 2007). This defining semantic function of the RC is, however, true only for the RC of the restrictive type. A close scrutiny of the semantics of RCs, however, reveals a complex constellation when other types of RCs are taken into account (further discussed in 4.4). Closely linked to its semantics is the information structure aspect of the RC: a RC introduces a contrast set into the discourse and the referent of the relativised nominal is being focused, in contrast to other referents in the set.[4] For example, in a sentence like (3), the relative clause (within the square brackets) singles out one soldier (implying that there were other soldiers in the discourse). This is done by referring the agentive-subject role of the solder in the event (i.e. firing the shot). The

3 Marori has a complex deictic system, showing a four-way spatial opposition (speaker-proximal, hearer-proximal, semi-distal, and distal) which cross-cuts a three-way number distinction (SG vs. NSG vs. neutral).

4 In this paper, I adopt a traditional LFG analysis, where information structure units (TOPIC and FOCUS) are integrated into f-structure. A precise analysis capturing the semantic/information structure properties of RCs, that allows for (contrastive) set interpretation of FOCUS, would perhaps involve independent information/semantic structures as proposed in Krifka (2008) and Butt (2014), in which the notions of Common Ground (CG) Content and CG Management are important.

relativised nominal is therefore analysed as bearing the discourse function of FOCUS in the proposed LFG-based analysis (further discussed in sections 4.2-4.3).

(3) The soldier [who fired the shot]$_{RC}$ was suspended until completion of the Military Police investigation.

Properties of relative clauses have been of typological and theoretical interest. Typologically, RCs provide a good window to how languages differ in the typological space of complex clause formation. The topics of intense research include, among others, (i) complexity in structure and marking, investigating questions such as headedness in RC (headed vs. headless, externally vs. internally headed); (ii) strategies to encode RC dependency (gapping vs. pronominal copy) and the related restriction showing which grammatical roles are possibly relativised (cf. the accessibility hierarchy (Keenan and Comrie 1977, Comrie and Keenan 1979, Keenan and Comrie 1979), and (iii) RC semantics, investigating the different meanings associated with different RC types (restrictive, vs. non-restrictive/appositive vs. maximalising), which may account the different behaviours of RCs (Arnold 2007, Grosu 2012).

RCs pose a challenge to any theory of grammar, as their complexities involve constraints across all components of the grammar, from semantics to morphosyntax and information structure. This is particularly true for a language that has complex morphology, which also shows different types of RC within same grammatical system. In this paper, it is demonstrated that LFG (Bresnan 1982, Dalrymple 2001, Bresnan et al. 2015, among others) is well equipped to handle the complexities of RCs in such a language, namely Marori. The different types of RC in Marori are outlined in the next subsection. The LFG-based analysis, given in section 4, can be straightforwardly captured in other lexically-based frameworks such as HPSG (Sag, Wasow, and Bender 2003).

3.2 Basic facts: different faces of RCs in Marori

Marori appears to be unusual in that it shows different types of RC. Almost all types are attested in this language: headed/headless, externally/internally headed, pre-/post-head, attached/detached. Clear cases of these types are presented in this section. There is an issue of the identification of double-headed RC, to be addressed later in section 5.

The externally headed relative clause (EHRC) exemplified in (1) is straightforward and needs no further comment. Of particular interest is the internally headed relative clause (IHRC) given in (1). Recall that this IHRC is ambiguous: the patient object and the instrument can be equally relativised. This ambiguity effect is an important issue to be discussed in considerable depth in this paper.

When the referent is clear from the context, the relativised noun is often unexpressed. This gives rise to a headless relative clause, exemplified in (4). The relativiser itself, e.g. *keme* in (5)a can be also dropped, giving rise to a structure shown in (5)b. The syntactic status of this structure is unclear, and ambiguous between IHRC without a relativiser (reading i) or simply two juxtaposed free clauses (reading ii). Further investigation is needed for this.

(4) kefi kwara keswemi tentara=te
REL drum 3SG.M.hit.3SG.PRES REL soldier=BE.3NPL.PRES
'the one who is playing the drum is a soldier'

(5) a. Koro [keme na ife-ben]$_{EHRC}$ tamba kundo-f
dog REL 1SG see-1.NrPST PERF run-NrPST
'The dog that I saw ran off.'

b. [Koro — na ife-ben] tamba kundo-f
dog 1SG see-1.NrPST PERF run-NrPST
i) 'The dog that I saw ran off.'
ii) 'I saw the/a dog (and it) ran off.'

RCs in Marori can appear within or outside their nominal structures. Within the nominal structure, they can be pre- or post-head RCs. Examples given so far are of the post-head RC type. Example (6)a shows a pre-head RC inside its nominal structure whereas examples (6)b shows a detached postverbal RC. This is not unique to RCs in Marori; an ordinary adjunct in Marori can float away from its NP too, as seen in (6)c.

(6) a. [keme na fis ndon-du] ujif]
REL 1SG yesterday 3SG.F.bring.here-1SG.PRES bird

tamba yaba nggwo-f
already dead AUX3SG.F-NrPST

'The female bird that I brought here yesterday already died.'

b. Efi purfam Johni naw te
DET person John named BE.3NPL.PRES

kefi nam bosik yaba ngguf.
RELPOSS pig die AUX-NrPST
'The person whose pig already died is (called) John.'

c. Na mara=i pemje-ben kwebu-wen.
1SG stone=U step.on-1.NrPST sink-SG.NML
'I stepped on the submerged stone.'

Another salient property of Marori grammar is that all grammatical relations, arguments and adjuncts, are relativisable. The previous examples show the relativisation of subject in (1) and (4), patient object in (1), stimulus object in (5), theme object in (6)a, possessor in (6)b and instrumental adjunct i n (1). The following examples illustrate the relativisation of other roles: recipient object (7) and a locative oblique (8).

(7) Maria=i keme njomo-bon bosik, tamba kurye-f.
maria=U REL give.FUT -1PST pig PERF return-NrPST
'Maria who I gave a pig has gone back (home) taking it with her.'

(8) Efi njaj [fis **keme**=na kufamon]
DET bench yesterday REL=1SG sleep-1SG.DUR.NrPST

tamba rafonngin.
PERF broken
'The bench on which I slept yesterday is already broken.'

An important empirical point with a theoretical implication worth mentioning here is the marking and relativisation of obliques and adjuncts. Obliques and adjuncts in Marori must be flagged by their relevant postpositions when they are not relativised. For example, the postpositional clitic *=ku* is obligatory in (9)a, marking the locative-goal. When relativised as in (9)b, *=ku* is not present, either inside or outside the RC. Therefore, this dependency in the RC cannot be accounted for in terms of a filler-gap analysis, as the category of the filler (NP) and that of the gap (PP) are distinct. It must be accounted for in terms of referential identity at the level of functional structure (f-structure), further discussed in the next section.

(9) a. John mara=i sour=ku monjo-f.
John stone=U house=LOC throw-3NPL.NrPST
'John threw a stone to the house.'

b. Efi sour=e [keme John __ mara=i monjo-f].
that house=part REL John stone =U throw-3NPL.NrPST

tamba kewei nggu-f.
PERF damaged AUX-NrPST
'The house at which John threw a stone has been damaged.'

4 Discussion and LFG analysis

In this section, I propose an LFG analyis to account for the RCs in Marori. The analysis consists of the c-structure analysis (4.1) and the functional constraints captured by the lexical entries of the relativisers and the annotations in the c-structure (4.2). The demonstration of how the analysis works is given in 4.3. Finally, intriguing facts about IHRCs with their possible ambiguity and disambiguation are discussed with reference to the semantic constraints of IHRCs 4.4.

4.1 Structural properties

As mentioned in section 2, the RC is an adjunct, part of an NP within DP. Internally there is good evidence to support that the RC is a CP, with the relativiser in C position, and [Spec, CP] filled in by an XP, e.g. adjunct modifying the RC. This is exemplified in (10) below. Thus, the adjunct *fis* 'yesterday', while showing up before the relativiser *keme*, modifies the RC.

The structure in (10) is for EHRC. Likewise, the IHRC has a CP structure but the mother nodes (NP and DP) are not branching (i.e., without their respective heads) as seen in structure (11).[5] .

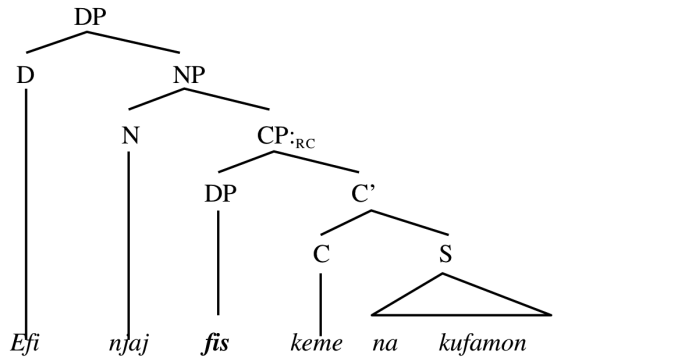
Our c-structure rule can correctly capture the empirical point of the IHRC where the determiner (*efi*) modifying the noun head *ujif* 'bird' shows up outside the relative clause CP, as shown in (12).

5 Note that, since we adopt the conception that all nodes in c-str are optional, the c-structure in (16) can be simplified by not showing the non-branching NP node.

4.2 Lexical entries and CP annotations

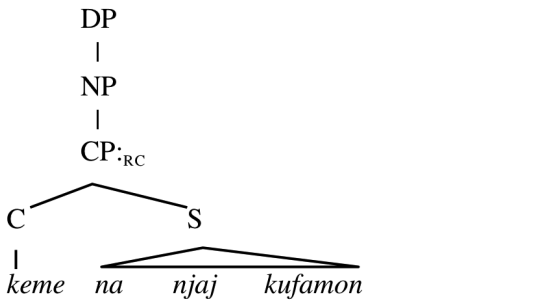
The second part of the analysis deals with the information specified in the lexical entry of the relativiser and the associated annotations on the c-structure of the RC to ensure that both EHRC and IHRC are correctly parsed or generated.

(10)



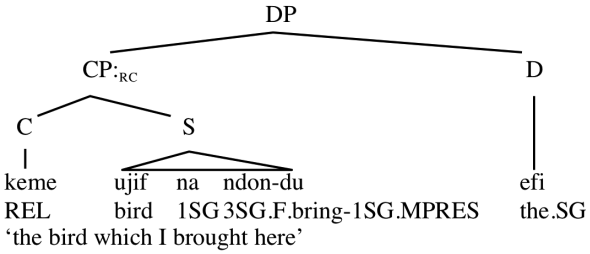
'The bench on which I slept yesterday'

(11)



'The bench on which I slept.'

(12)



'the bird which I brought here'

The entries of the general relativiser *keme* and the singular relativiser *kefi* are given in (13). Each carries a set of equations by which the information associated with the selected relativised NP in the RC, which bears FOCUS and a particular GF (grammatical function), is shared by the

matrix DP. *Keme* and *kefi* differ only in the INDEX specification that regulates the agreement: unspecified for *keme* vs. specified for singular for *kefi*, indicated by [ ]α and [SG]α respectively. The alpha subscript means that the exact value of the index is yet to be determined, depending on the referential features of the nominal in a given RC. This index equation, as part of the agreement constraint, allows for the possibility that more than one noun could be possibly relativised, e.g. as in example (1).

(13) a. *keme*
C
(↑FOCUS INDEX)= [ ]α
((↑FOCUS PRED)='pro')
(↑TYPE)=relative
(↑GF^{+})= (↑FOCUS),

b. *kefi*
C
(↑FOCUS INDEX)= [SG]α
((↑FOCUS PRED)='pro')
(↑TYPE)=relative
(↑GF^{+})= (↑FOCUS),

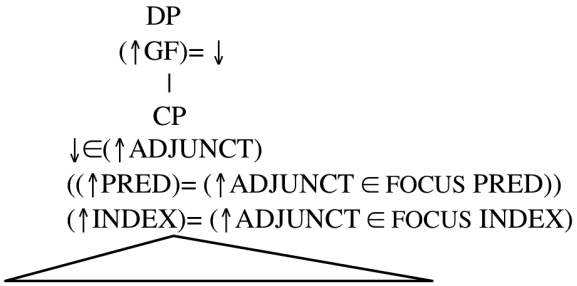
The joint specifications ((↑FOCUS PRED)='pro') and (↑TYPE)=relative mean that the relativiser is a relative pronoun and that its pronominal function is optional. This is to capture the fact that in the absence of any overt noun, the relativiser itself is referential; see the case of headless RCs, example (4). When an overt relativised head noun is present (either inside or outside the RC), however, this pronominal meaning is not used. Instead, the PRED value of the head noun is used.

The last line in (13), (↑GF^{+})= (↑FOCUS), indicates that the relativised GF also bears the discourse function of FOCUS. The notation + means that you can have a path here; e.g. OBJ POSS (i.e. the relativisation of the possessor of OBJ). This equation and the index equation (e.g., (↑FOCUS INDEX)= [ ]α) in the entry (which imposes agreement) determine that the specific GF of a particular relativised NP is identified within the embedded RC. In example (1), for instance, there are three GFs (SUBJ, OBJ and instrumental ADJUNCT) eligible to be picked up by the equations of *keme*. However, the matrix auxiliary (*ngguf*) requires '3SG.M' agreement. This excludes the possibility of picking up the embedded SUBJ '1SG' as the relativised nominal. (Being a pronoun also excludes it from the relativisation; further discussed in the next subsection.) Given their compatible referential features, OBJ and ADJUNCT are then possibly relativisable, giving rise to ambiguity, as expected.

In addition to the functional equation in the entry of the relativiser, we need a set of equations on the CP (relative clause) node, shown in (14). The equations regulate the function of the RC and the flow of information, in particular to ensure that the mother node DP picks up its referential information available from the embedded IHRC. The equation ↓∈(↑ADJUNCT) says that the RC is a member of the ADJUNCT set, associated with the mother's GF. The optional ((↑PRED)=(↑ADJUNCT ∈ FOCUS PRED)) means that the functional head of the DP (i.e. the relativised nominal PRED) is supplied by the RC when the mother DP does not have its head PRED. In this way, we capture the essence of an IHRC, where the relativised PRED comes from an NP internal to the embedded clause. If no NP supplies the PRED internally from the embedded RC, given the entry of the relativiser specified in ((13), then the relativiser itself will supply the functional head of the nominal, namely (↑FOCUS PRED)='pro'. In this way, we capture how the headless RC gains the pronominal interpretation

translatable as ‘the one (who/which)...’ in spite of the missing head. Finally, the selection of the relativised NP is highly constrained by the agreement system. This is imposed by the equation (↑INDEX)=(↑ADJUNCT ∈ FOCUS INDEX). The notation ∈ represents a path into an adjunct set.

(14)



4.3 Accounting for RC properties in Marori

Having outlined the lexical entries and c-structure annotations, we are now ready to account for RC intricacies in Marori. We start in this section with the straightfoward case of an IHRC, exemplified by (15). The partial c-structure of this sentence is shown in (16)a. Its f-structure is shown in (16). Note that the matrix verb *nggwof* requires [3.SG.F] agreement (not shown in the c-structure but shown in the f-structure as index *j*).

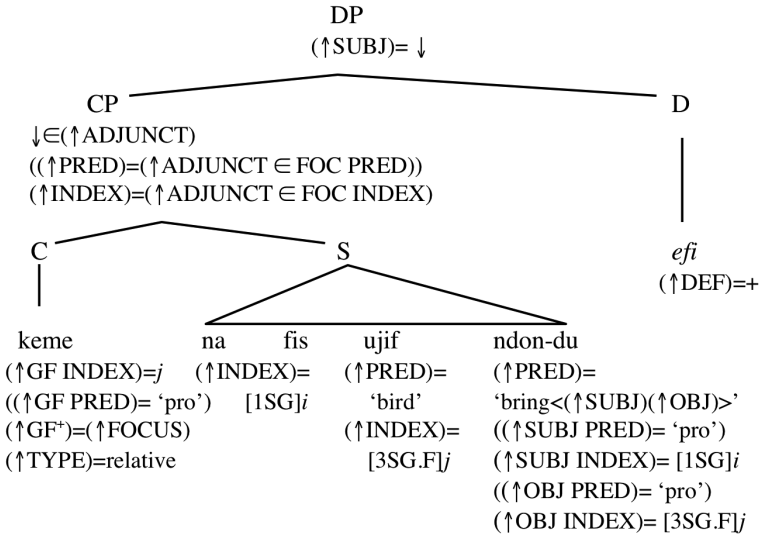
What is special about the IHRC here is that the definite determiner *efi* ‘that.SG’ modifying the noun inside the RC is external to the RC. The proposed analysis, as seen in the f-structure, correctly captures that the [DEF +] information ends up specifying the nominal head PRED ‘bird’, thanks to the equations on the CP node. These equations in effect spread the PRED and INDEX values (tag [1], j) from inside the embedded relative clause (S) to higher nodes in the DP, allowing them to be specified by the determiner carrying [DEF +]. The same mechanism in our analysis accounts for the Marori fact that the quantifier *usindu* ‘all’ or numeral *yanadu* ‘two’ can also quantify a relativised NP inside the RC from a position external to the RC. This is shown in (17).

(15) [[keme na fis **ujif** ndon-du]$_{RC}$]$_{NP}$ **efi**
REL 1SG yesterday bird 3SG.F bring.here-1SG.PRES that.SG

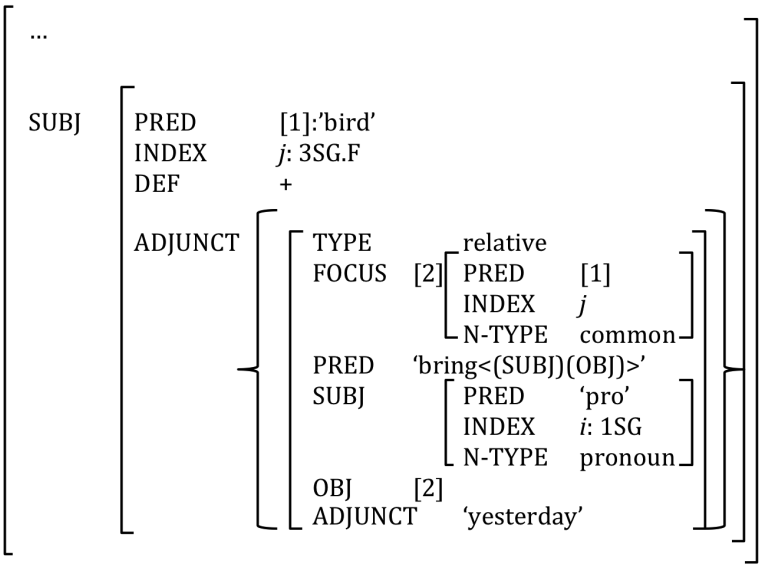
tamba yaba nggwo-f
already dead AUX3SG.F-NrPST

‘The female bird that I brought here yesterday already died.’

(16) a.



b.



(17) [keme na fis **ujif** kein-du]$_{RC}$ **usindu**
REL 1SG yesterday bird 3NSG.bring.here-1SG.PRES all

tamba yaba nggorforo-f
already dead AUX.3PL-NrPST

'All of the birds that I brought here yesterday already died.'

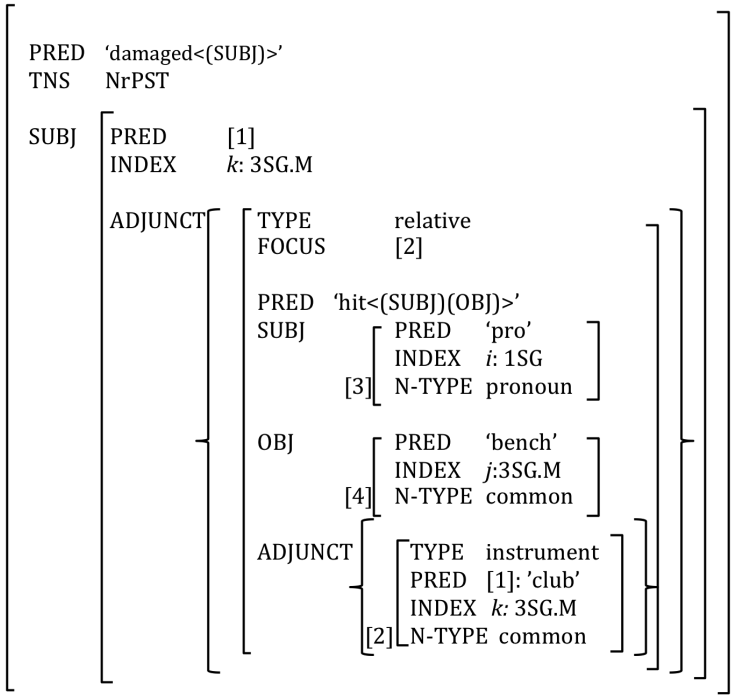
IHRCs lead to ambiguity, a phenomenon also known in IHRCs in other languages. The ambiguity can be now straightforwardly accounted for in our analysis. The relevant example, repeated as (18), is the case where two or more NPs inside the RC whose index features are not in clash with those imposed by the agreement in the matrix predicate. They are therefore eligible to be picked up as the functional head of the nominal. These are 'bench' and 'club' (readings (a)-(b)), but not the pronoun *na* '1SG' (reading (c), which is excluded due to a feature clash).

(18) [Keme na **njaj**=i **samagau** ngge terme-ben]$_{IHRC}$
REL 1SG bench=U club with 3SG.U.M.hit-1SGNrPST

tamba keiwei nggu-f
PERF damaged 3SG.M.U.AUX –NrPST
a) 'The **bench** that I hit with the club was damaged.'
b) 'The **club** with which I hit the bench was damaged.'
c) *'I, who hit the bench with a club, was damaged.'

The f-structure for reading (b) is given in (24). As seen, the nominal PRED 'club' (tag [1], INDEX *k*) is inside the RC as an adjunct. The postpositional *ngge* 'with' is treated like a prepositional case marker flagging the ADJUNCT TYPE of instrument (i.e. carrying no PRED attribute). The whole ADJUNCT value is shared by FOCUS, due to the relativisation enforced by the specification in the entry of *keme* (cf. (24), tag [2]).

(19)



Given the [3SG.M] SUBJ agreement of the matrix structure required by the verbal auxiliary *ngguf*, the value specifications of OBJ 'the bench' (tag

[4], i.e. 'bench', index *j*) also satisfy the requirement. Hence, OBJ [tag [4] is predicted to be possibly selected as FOCUS (i.e. relativised) and matrix SUBJ. The embedded SUBJ, carrying an index of [1SG] is expected to cause a clash in agreement and cannot be selected as the matrix SUBJ, as predicted.

A brief note is needed for detached RCs, as in example (6)b. Detached RCs can be straightforwardly handled in the proposed analysis as floating headless RCs. Thus, the relativiser supplies the PRED (i.e. 'pro') value, taking the detached noun head as its antecedent anaphorically. Sentence (6)b literally means 'The person_i is named John, the one_i whose pig already died', where the relativiser *kefi* 'the one' refers to 'the person' (indicated by index *i*). For reasons of space, no c-structure/f-structure representations are given here, but such representations could be determined with special care given to the annotation of c-structure nodes, to ensure that the right information is passed to the adjunct set of the relative clause.

4.4 Semantic constraints and IHRC typology

This section addresses the definiteness constraint of IHRC in Marori, a phenomenon also observed in other languages such as Lakhota (Williamson 1987, Culy 1990), Dàgáárè (Bodomo and Hiraiwa 2010), Kobon and Wappo (Dixon 2009:331). In what follows, the nature of the constraints is outlined, but no formalisation is proposed. Such formalisation should be handled as part of the semantic structure representation.

We have seen that two NPs are equally relativisable, giving rise to ambiguity. However, when one of NPs in the IHRC is a proper name, no ambiguity arises, as seen in example (20). The proper name cannot be relativised in IHRCs in Marori.

(20) [Keme Markus bosik =i ife -f]$_{IHRC}$
REL Markus pig=U 3SG.M.see-3NPL.NrPST

tamba kundo -f
PERF run-3SG.NrPST

i) 'The pig that Markus saw ran off.'
ii) * 'Markus who saw the pig ran off.'

The effect of the definiteness constraint in Marori is also observed with quantification. Recall that the NP relativised in the IHRC can have its quantifier *usindu* 'all' outside the RC; example (17). The universal quantifier *usindu* presupposes definite referents, and when it quantifies a noun internally within the IHRC it resists relativisation. Thus we have no ambiguity in the following example:

(21) [Kemde usindu meninggon=i purfam paar]$_{IHRC}$
REL allchild.PL=U person money

njemba-b tamba sra-f
3.give-3PL.NrPST already go.PL-NrPST
i)'The people who gave money to all the children already went away.'
ii) *'All the children who were given money by the people already went away.'

I argue that the definiteness constraint, as observed in the preceding examples, is a logical consequence of the typological property of the Marori

IHRC, which is essentially a restrictive type. The restrictive RC is known to have the salient features of non-specific indefinite and intersective force at the level of the RC (Grosu 2012). A highly definite or unique referent like a proper name does not allow the intersective interpretation and is therefore inconsistent with the restrictive RC. Evidence for Marori IHRCs as restrictive is given below, but I will first outline Grosu's (2012) typology of (IH)RC briefly.

Grosu (2012) distinguishes three semantic types of RCs: restrictive, appositive and maximalising. The first two are exemplified from English RCs shown in (22)a-b. In (22)a, there were more than three boys at the party, and only three of them had beards; thus, the RC imposes a further restriction to the denotation of 'boys'. The RC information is essential for the identification of the referent of the head noun. In (22)b, there were only three boys in the party; all of them had beards. The RC imposes no referential restriction, and the RC can therefore elide without affecting the identification of the referent of the relativised noun.

(22) a. At the party, I saw only [[three boys [*who had beards*]$_{RC}$]$_{NP}$ (restr.)

b. At the party, I saw only [[three boys, [*who had beards*]$_{RC}$]$_{NP}$ (apps.)
(Grosu 2012:452, ex. (6))

The maximalising RC is like the appositive RC, in that the denotation is already specific/definite. They are different in the locus of the specificity/definiteness. In the appositive RC, it is fully defined in the matrix NP, as seen in the meaning of (22)b above. In the maximalising RC, the definiteness is fully defined within the RC itself, exemplified by (23). The maximalising RC is, in a way, like the restrictive RC, in that its information is highly essential. It is 'super restrictive', making the referent of the relativised noun maximally definite; e.g. when the noun is referentially plural it gives rise to the totality of plural meaning, as seen in the example from Japanese in (24)b below. In short, maximalising relatives have the characterisation of having strong definite import, presupposing the relativised noun to be maximally definite; hence, dropping *the* in (23) would degrade its acceptability.

(23) I suddenly noticed [the three books [that there were on your desk]$_{RC}$]
(i.e. 'there were exactly three books on your desk and I suddenly noticed them')
(Grosu 2012: 453)

Japanese provides a good illustration involving a maximalising IHRC. It should be noted that Japanese has both EHRC and IHRC, but the IHRC is of the maximalising RC type. The EHRC is exemplified in (24)a and its IHRC counterpart is in (24)b (Grosu's (18) and (20)). Note the difference in meaning as seen in the free translation. The maximalising relative encodes the definite totality of the cookies to be brought by Taro to the party, which is not the case in (24)a.

(24) a. Taro-wa [[Yoko-ga reezoko-ni __ irete-oita]
Taro-TOP Yoko-NOM fridge-Loc put-AUX

hotondo-no kukkii-o] paatii-ni motte itta
almost-all-Gen cookie-Acc party-to brought

'Taro brought to the party almost all the cookies that Yoko had put in the fridge.'

b. Taro-wa [[Yoko-ga reezoko-ni **hotondo-no kukkii-o**
Taro-TOP Yoko-NOM fridge-Loc almost-all-Gen cookie-Acc

irete-oita]-no-o paatii-ni motte itta
put-AUX-NML-Acc party-to brought

'Yuko put almost all the cookies in the fridge and Taro brought {them, *some} to the party.'

Turning to Marori, I argue that IHRCs in this language are restrictive, not maximalising. That is, they do not presuppose definite/specific denotation of the relativised nominal. First evidence for this comes from the fact that the IHRC in Marori can have the full range of quantificational forces, including an existential force, a salient property of the restrictive relative. This is shown in example (25), where the referent of the relativised noun 'boy' is indefinite. It should be noted that, while not presupposing definiteness, the IHRC does provide specifications which make the referent of the relativised noun definite, with the possibility of the (strong) determiner overtly present at the matrix level; see example (15).

(25) Ka einda kefi meipur di nie=fi kuye-den
2SG 3.search.2PL REL child.SG FUT 1NSG=with stay-1DU.PRES
'You look for a child who wants to stay together with me.'

Other evidence that IHRCs are restrictive comes from the fact that they have intersective import, providing a restriction to the denotation of the associated noun in the same way as an ordinary adjunct. To understand this, first consider the restrictive relative clause in English in (26). The denotation of the object bought in (26) is the one in the intersection of sets of 'books', 'cheapest things', and 'things which are not paperbacks'.

(26) I bought the cheapest book which was not a paperback.
(Arnold 2007, ex. (1b))

(27) a. I've never spoken to Kim, who plays poker.

b. *I've never spoken to Kim who plays poker.
(Arnold 2007, ex. (1b))

In contrast to (26), applying a restrictive RC to the proper name Kim in (27) results in downgraded acceptability as seen in (27)b. This is attributed to the referential uniqueness of a proper name which is inconsistent with the intersective interpretation of restrictive RC, which requires a set of referents. Thus, it is not suprising to see why a restrictive RC is not possible with a proper name.

Arguably, the same semantic constraint accounts for why proper names are not relativisable in the IHRC type in Marori, as seen in (20), giving rise to

no ambiguity. The same is true for NPs by the universal quantifier *usindu* 'all', which requires definite referents; see example (21).

Another related effect of this intersective import is the introduction of contrast set into the discourse by the restrictive RC. Then, this contrast set can be accessed anaphorically. In English, this is done by an expression like *others*, as in (28)a. Non-restrictive RCs do not introduce such a set; hence the downgraded acceptability of sentence b. The equivalent example in Marori is (29), where the contrast set introduced by the IHRC in the first clause is accessed by *now* in the second IHRC. The possibility of clauses like in (29) provides supports that Marori IHRC is of the restrictive type.

(28) (a) I like puzzles which require imagination and creativity, and others that just depend on knowledge. [Restrictive]

(b) #I like puzzles, which require imagination and creativity, and others that just depend on knowledge. [Non-restrictive]
(Arnold 2007, examples (4a-b))

(29) Keme=na ujif kein-du, keme Jhon **now** kein
REL=1SG bird bring-1SG.MPRES REL John other bring.NPST

tamba yaba nggorforof
PERF dead AUX.PLURAL-NrPST
'The bird which I brought, (and) the other ones which John brought were dead.'

Given the intersective interpretation, it is therefore possible, and indeed natural, to stack restrictive RCs to provide further intersective specifications to make the denotation more specific. Crucially, both restrictive and appositive RCs can have the stacking, exemplified by (30) and (31). This is not, however, possible with the maximalising RCs (Grosu 2012), example (32). In (30), the depicted (three) referents are on the intersection of sets of individuals who are 'boys', 'bearded', and 'wearing no shoes'. The boldfaced RC further restricts the individuals picked up by the italicised RC, leaving open the possibility that the speaker may have seen additional bearded boys who wore clothes. In the appositive RC (31), the boldfaced RCs provide additional information but the referent of the relativised noun ('the Honourable member') remains unique (without any possibility of others). In (32), the boldfaced RC is unacceptable in its full version. That is, according to Grosu (2012:454), the denotation of the construction is already fully determined by the italicised relative, leaving no interpretation for it.

(30) At the party, I saw only [three boys *who had beards* **who wore no clothes**]. (Grosu 2012: example 8a)

(31) I fear the Honourable Member, *who nobody trusts*, **who nobody believes, who not even his own supporters listen to,** has finally run out of time. (Arnold 2007: example 56; bold and italics added)

(32) I suddenly noticed [the three books *that there were on your desk* **that (*there) had earlier been on my desk**].

In Marori, stacking of IHRCs is possible, as seen in (33). This supports the analysis that IHRCs in Marori are restrictive, not maximalising in nature. Lakhota also has a restrictive IHRC, and is expected to allow RC stacking, as seen in example (34):

(33) Na fis purfam=i eyew=nda-mon
1SG yesterday person=U eye=3.AUX-1SG.NrPST

[*kefi* *koro* *imbirif* **kefi** **kundo-f**]
REL dog bite-NrPST REL run-3NrPST

'Yesterday I was looking for (the/a) person [that dog bit that ran away].'

(34) [Ogle eya šapšapa cha] agli pi wachi ki lena e
shirt some dirty indef take-home Pl 1-want the these be
'These are the shirts that are dirty that I want them to take home.'
(Grosu 2012: 455)

5 Final remarks and further research

In this final section, I will provide a brief summary of the facts and analysis, and then point out further research needed.

This paper has demonstrated the different types of relative clauses in Marori. The paper hopefully contributes to the empirical basis for the typological and theoretical studies of RCs. Marori is unusual in that it has almost all of the relative clause types: headed and headless; if headed, both externally and internally headed; if externally headed, either attached in the nominal structure or detached from it; if attached, either pre- or post-nominal. Of particular interest is the morphosyntax and semantics of the IHRC. It has been demonstrated that LFG formalism is well equipped to capture the structural intricacies and intriguing properties of IHRCs, such as the possibility for ambiguity/non-ambiguity, which relates to nominal types and definiteness. Nominals with unique and definite referents, such as proper names and pronouns, cannot be relativised in the IHRC. It has been argued that this is due to the fact that IHRCs in Marori are of the restrictive type, not the maximalising one, in Grosu's typology. The salient semantic feature of the restrictive RC is that it has non-specific indefinite and intersective import. As such, it provides some specification to a referent of a set, introducing a contrast to the set. This intersective import is therefore semantically incompatible with a unique definite referent where no contrast (set) is possible. As outlined in this paper, salient properties of IHRCs in Marori, such as the possiblility of stacking IHRCs or the inability of proper names and pronouns to be internally relativised, have a good semantic basis.

Most of the data presented in this paper was elicited. Future studies on RCs in Marori must therefore include an investigation of the distribution of different types of RCs in natural texts. The corpus-based study is expected to further reveal and illuminate the semantic-discourse constraint of specificity/definiteness in RCs.[6] It has been mentioned in this paper that a

6 A Marori corpus is being developed as part of the Sothern New Guinea (SNG) and the ELDP projects. The corpus consists of audio-audio recordings with

pronoun and a proper name cannot be relativised in single-headed RCs in Marori (and also other languages such as Lakota (Culy 1990:168)). However, there is a surprising twist, which shows that such a pronoun can be relativised, provided that it is expressed in a double-headed RC. That is, the pronoun is expressed twice, internally within the RC and also externally in the matrix position as shown in (35). This kind of nominal doubling, giving rise to a double-headed RC, is attested in other Papuan languages with IHRCs for disambiguation, e.g. as seen in the contrast in examples in (36) from Kobon (Dixon 2009:331). It is not completely clear at this stage of our research whether structure (35) in Marori and (36)b in Kobon are in fact two juxtaposed free clauses, rather than complex structures with embedded RCs. This could be the case for Kobon, for which the alternative translation would be 'the boy hit the girl (and) I know the girl.' A similar two-free-clause analysis in Marori, however, would have a problem in accounting for the presence of the RC marker *keme*. Dropping *keme*, which is possible in Marori, would indeed make the structure analysable as two juxtaposed free clauses. In the presence of *keme*, the structure in (35) can perhaps be analysable as belonging to a double-headed RC. Further investigation is needed what happens if proper names, and even common nouns, are forced to appear in double-headed RCs in Marori.

(35) keme **na** bosik=i ife-ben tamba=**na**
REL 1SG pig=U 3SG.M.see-1SG.NrPST PERF=1SG

kundo-bon
run-1SG.NrPST

'I, who saw the pig, ran off.'

(36) a. [ña pai pak-öp]$_{RC}$ yad nöŋ-bin
boy girl hit-PERF.3SG 1SG perceive-PERF.1SG
i. 'I know the girl who the boy hit.'
ii. 'I know the boy who hit the girl.'

b. [ña **pai** pak-öp]$_{RC}$ **pai** yad nöŋ-bin
boy girl hit-PERF.3SG girl 1SG perceive-PERF.1SG
i. 'I know the girl who the boy hit.'
ii. *'I know the boy who hit the girl.'

Finally, a RC-related issue which is of particular interest and needs further investigation is the connection between (clausal) nominalisation and the nature of structural embedding with its constraints in head-final (OV) languages like Marori. There is an intriguing behaviour of IHRCs in Marori with a preference for discontinuous post-verbal RCs. This is perhaps what is expected for an OV language, in line with the finding reported in the literature that the reduction of preverbal arguments in SOV languages is a compensatory strategy to reduce the heavy cost in production and comprehension (Hawkins 2004, Ueno and Polinsky 2009, and the references therein). Further research in this area will include an in-depth corpus

transcriptions in ELAN, covering a range of topics. This corpus will be made available to the public in due course through PARADISEC (http://www.paradisec.org.au/) and ELAR (http://www.elar-archive.org/).

investigation of Marori, preferably including comparison with (OV/VO) languages with IHRCs, to gain further empirical evidence for any analysis proposed.

References

Andrews, Avery. 2007. "Relative clauses." In *Language typology and syntactic description*, edited by Timothy Shopen, 206-236. Cambridge: Cambridge University Press.

Arka, I Wayan. 2016. Information structure and its morphosyntactic resources in Marori. Paper presented at The Workshop on Information Structure at TUFS (18-20 February 2016).

Arnold, Doug. 2007. "Non-Restrictive relatives are not orphans." *Journal of Linguistics,* 43 (2):271-309.

Bodomo, Adams, and Ken Hiraiwa. 2010. "Relativization in Dàgáárè and its typological implications: left-headed but internally-headed." *Lingua,* 120:953-983.

Bresnan, Joan. 1982. *The mental representation of grammatical relations*. Cambridge, Massachusetts: the MIT Press

Bresnan, Joan, Ash Asudeh, Ida Toivonen, and Stephen Wechler. 2015. *Lexical-Functional Syntax (second edition)*. Oxford: Wiley-Blackwell.

Butt, Miriam. 2014. "Question and Information Structure in Urdu/Hindi " In *Proceedings of LFG14*, edited by M Butt and T. H King, 158-178. CSLI, http://web.stanford.edu/group/cslipublications/cslipublications/LFG/19/lfg14.html.

Comrie, Bernard, and Edward L. Keenan. 1979. "Noun phrase accessibility revisited." *Language*, 55 (3):649-664.

Culy, Christopher Douglas. 1990. *The syntax and semantics of internally headed relative clauses*, Linguistics Department, Stanford University, Stanford.

Dalrymple, Mary. 2001. *Lexical Functional Grammar*, *Syntax and semantics*. San Diego: Academic Press.

Dixon, R.M.W. 2009. *Basic Linguistic Theory*. Vol. 2. Oxford: Oxford University Press.

Grosu, Alexander. 2012. "Towards a more articulated typology of internally headed relative constructions: the semantic connection." *Language and Linguistics Compass*, 6 (7):447-476.

Hawkins, John A. 2004. *Efficiency and complexity in grammars*. Oxford: Oxford University Press.

Keenan, E, and B. Comrie. 1979. "Data on the Noun Phrase Accessibility Hierarchy." *Language*, 55:333-51.

Keenan, Edward L., and Bernard Comrie. 1977. "Noun phrase accessibility and universal grammar." *Linguistic Inquiry*, 8:63-99.

Krifka, Manfred. 2008. "Basic Notions of Information Structure." *Acta Linguistica Hungarica* 55 (3–4):243–276.

Sag, Ivan A, Thomas Wasow, and Emily M Bender. 2003. *Syntactic theory: a formal introduction (second edition)*. Stanford: CSLI Publications.

Ueno, Mieko, and Maria Polinsky. 2009. “Does headedness affect processing? A new look at the VO-OV contrast.” *Journal of Linguistics*, 45:675-710.
Williamson, J. 1987. “An Indefiniteness Restriction for Relative Clauses in Lakhota.” In *The Representation of (In)definiteness*, edited by E. Reuland and A. ter Meulen. Cambridge, MA.: The MIT Press.