
Towards understanding information structure in Vera'a

Author(s): Schnell, Stefan

Source: Proceedings of the International Workshop on Information Structure of Austronesian Languages, 10 April 2014, pp.47-89.

Published by: ILCAA, TUFSS

Permanent URL: <http://hdl.handle.net/10108/75987>

The Prometheus-Academic Collections are a repository of academic research. In them are found the research and educational achievements of the Tokyo University of Foreign Studies and also historical materials which have been preserved. The University shares such academic achievements on the Internet to ensure an open university. Copyright is retained by each author, academic society, association, publisher, and/or other rights owners. The whole or parts of the materials can be used (referred to, reproduced, printed), with copyright acknowledged, for academic and personal use only.

Towards understanding information structure in Vera'a

Stefan Schnell (S.Schnell@latrobe.edu.au)

1. Introduction

In this working paper¹ I present preliminary findings from an ongoing investigation into information structure (IS) in the Oceanic language Vera'a. The study is exclusively based on corpus data of mainly narrative texts, plus some descriptive texts. However, the current study is still less systematically corpus-based than would be desirable, hence what I have to present here is clearly a very preliminary outline, meant to lay the ground for future more detailed and systematic investigation of the topic. Nevertheless, even the selective data examined so far seem to suggest some interesting findings that I would like to share at an early stage of investigation and understanding. I shall also mention a further desideratum at this stage: I have so far not undertaken a systematic analysis of prosodic features of the Vera'a language. As considerations of prosodic marking are, however, indispensable for a study in information structure, I shall give comments on the prosodic marking, be they preliminary and impressionistic, where relevant.

The paper is organised as follows: in section 2, I outline the basic structure of the clause in Vera'a. In section 3 I deal with information status. An important point to make in this area is that the choice of referential form is probably not exhaustively accountable in terms of 'accessibility' or 'activation', concepts that are grounded basically prior (shared) knowledge or linguistic context. Instead, the Vera'a data suggest that 'forward' planning (e.g. in cataphoric reference, or the establishment of discourse topics) is equally important. Section 4 then discusses information structuring strategies as attested in the texts. While IS has canonically been associated with 'topic' and 'focus' as (quasi-)universal functional categories, my findings corroborate Matic & Wedgwood's (2013) critical assessment of their universality in so far at least two morphosyntactic devices are employed for marking of what might be identified as 'topic' and 'focus' respectively, but have essentially other basic functions. Hence, their "topic-marking" and "focus-marking" properties appear to be mere by-effects of other functions. Section 5 concludes the paper, mainly by outlining directions of future research into the topic.

¹ The research reported in this paper was made possible through two grants within the VWStiftung-funded DoBeS programme which enabled the bulk of fieldwork in the Vera'a community, conducted between 2007 and 2012; a further fieldtrip in 2013 and ongoing work on corpus linguistic work is currently being funded by a ARC DECRA grant (DE120102017, and a grant from the DAAD/Go8. I am furthermore grateful to the organisers of the workshop *Information structure in Austronesian languages* for inviting me to contribute to this meeting, and to the participants of the workshop for critical comments. Most of all, though, I am grateful to the members of the Vera'a community for their engagement in our collaborative language project, their kindness, friendship and patience. All shortcomings of this work are of course my own responsibility.

2. Vera'a basic clause structure

Vera'a is a largely isolating language that marks core arguments by their order relative to the verb complex (VC henceforth), which in turn functions as the predicate of a verbal clause. The alignment of core argument encoding is accusative: S and A arguments (i.e. subjects) immediately precede the VC, and P arguments (objects) follow it. Flagging by prepositions is employed to mark oblique arguments and adjuncts. The latter can occur in the clause-initial or clause-final periphery.

(1) a. [ote mu-n e ni'i rēne anē]^{SUBJ} [ne ma]^{VC2} |
 mum POSS.GEN-CS PERS.ART small woman DEM1 TAM2 dead

e [=n 'ama-gi]^{SUBJ} [ne 'ēn gōr]^{VC} [ēn ni'i rēne anē]^{OBJ3} vaavan
 DISC =ART father-3SG TAM2:3SG see secure ART small woman DEM1 RED-go

'And since the little girl's mum died, her father then looked after the little girl, until ...'

ANV.005-006

b. [misiin wal ros] [di]^{SUBJ} [ne van]^{VC} [lē =n qarān 'alēn ōlōl]^{OBL} |
 lengthy once NEG 3SG TAM2 go LOC =ART hole ASSOC RED:conjure

[di]^{SUBJ} [ne mom]^{VC} [ēn 'enge 'alēn wede]^{OBJ}
 3SG TAM2 put ART plant ASSOC rain

'Not long (after that), he went to the conjuring hole and installed some rain herbs.'

JJQ.101

There is no structural requirement for either subject or object to be overtly expressed, and a clause can consist solely of a VC, minimally consisting in turn of an initial TAM marker and at least one verb:

(2) ne maran | ne qōn
 TAM2 daylight TAM2 night

² Morpheme glossing of examples follows the *Leipzig Glossing Rules*. A list of non-standard abbreviations is provided preceding the references section. Certain elements, like TAM markers and demonstratives, are not glossed by labels for functional categories, but simply by a number, as functional labels necessarily invoke functions from those languages that a reader may happen to be familiar with, but not necessarily enlighten the situation in Vera'a.

³ The structural position of the demonstrative in this example is debatable, and alternatively it can be analysed as a clause-level constituent, as will be discussed in section 4 below.

‘Day broke, and noight broke, [and then day broke again].’

ISAM.031

In addition to non-argument clauses, we find clauses where not all or even none of the arguments are overtly expressed, but are instead “realised” by zero form:

(3) ___ *ne* *rēv* ___
 TAM2 drag

‘(He) pulled (it).’

AS.1.051

(cf. example (6b) below for preceding context). Zero forms will be discussed in more detail in 3.1 below.

In non-verbal clauses the predicative phrase is not a VC but either a NP, PP or some other type of phrase. The following is an example with a NP predicate:

(4) [*duru*]^{SUBJ} [= *m* *'ēn*]^{VC} | *so* [*dī*]^{SUBJ} [= *n* *mēr* *ga* *wēe*]^{PRED}
 3DL =TAM1 see CPL 3SG =ART kid TAM3 good

‘They saw that he was a fine kid.’

ISWM.034

As mentioned above, at the current stage of investigation, intonation has not been studied in any significant detail. It can nevertheless be stated impressionistically that sentential accent is situated sentence-finally as the default. Other types of accent may occur in the leftward area of a sentence, but as a principle rule, individual constituents do not usually receive individual accent, so that for instance subject NPs or verbs cannot receive special prosodic marking⁴. I do mark perceived prosodic marking by capitalisation of object language material in the transcription line, where this is relevant. Obviously, these observations are very preliminary and require thorough substantiation by more thorough prosodic studies. In what follows, I will therefore largely confine myself to morphosyntactic phenomena.

3. Information status

In this section I outline the formal realisation of discourse referents with different information statuses as attested in the text corpus. I will first outline the referential forms

⁴ Needless to say that a more systematic and detailed phonetic analysis of prosody in Vera’s might in fact reveal some marking of individual constituents that went un-noticed as to now.

available. I then describe the formal realisation of two general information statuses, namely discourse-new (in 3.2) and discourse-given (in 3.3). According to Prince (1981) and Lambrecht (1994:109) – among many others – further subtypes can be distinguished, for example discourse-new, inferable or anchored referents; also, for anaphoric reference to discourse-given referents, different types of antecedent relationships can be distinguished (cf. e.g. Baumann & Riester ms). I will allude to such finer-grained distinctions only where they seem to be justified by contextual evidence⁵, and only where they appear to account for a given formal choice.

3.1 Referential forms and their distribution

Vera'a has three main types of referential forms, namely NP, pronoun and zero. Among the first ones, I distinguish common and personal NPs, the latter may have a personal pronoun as its head (cf. Schnell 2011:53):

- (5) a. [*E ote*] *ne sursur* [*ēn NES*]
 PERS.ART mum TAM2 RED:sing ART song

'Mum is singing a song.'

HHAK.077

- b. [*n vē vē -m*] *SA* [*e no*]
 ART mother-2SG FOC? PERS.ART 1SG

'But I am your mother!'

1.TNU.023

Personal NPs headed by personal pronouns are restricted to left-dislocated position and predicative phrases in non-verbal clauses, as in (5b) above. Common and other personal NPs can occur in any syntactic position and function. The same goes for pronouns and zeroes, which both may "fill" the subject and object positions:

- (6) a. [*kimi*] =*k* *le no*
 2PL =TAM2 take 1SG

'Can you take me?.'

HHAK.023

- b. *alē* ___ *ne gis sa* [= *n wōbin mu-gi*]

⁵ I do not believe that categories like 'inferable' are straightforwardly applicable, especially not to data from a hitherto understudied language like Vera'a, and they should be understood as very preliminary, hypothetical characterisations of some usage contexts.

alright TAM2 hold FOC? =ART penis POSS.GEN-3SG

___ *ne rēv* ___
 TAM drag

‘Alright, (he) grabbed his penis and (he) pulled (it).’

AS.1.051

Also, both pronoun and zero forms may be oblique arguments:

(7) a. *e* [=n *lasbon*] *ne dam kel mulō* [*sir diē*]^{OBL}
 DISC =ART lastborn TAM2 hang back <MUL>wards for 3SG

‘And then the lastborn was going to swing over to her.’

ANV.061

b. *e* [*di*] =*m* *surga no* [*sir* ___]^{OBL}
 DISC 3SG =TAM1 send 1SG for

‘[The almonds uncle has sealed ...], and he sent he for (them).’

JSU.136

There are a few further types of referential form: numeral phrases (NumPs) are headed by a numeral and introduced by a special numeral article, cf. (8a). Pronominal NPs consist of a personal pronoun and another word or phrase to form a complex phrase – as opposed to simple bare pronouns –, cf. (8b) (cf. Schnell 2011:85f.):

(8) a. [*ne* ___ *vōwal*]^{SUBJ} *ne van ma*
 NUM.ART one TAM2 go hither

‘Then one (of the spirits) came here.’

MVBW.111

b. [...] *raksag* [*mē kamam ‘a* ___ *Vera’a*]
 especially DAT 1PL.EX SPEC.ASSOC *place.name*

‘... especially for us, (the people) of Vera’a.’

JWR.019

Despite the possibility for NPs, pronouns and zeroes to occur in all syntactic functions, actual language usage constrains their distribution in fairly systematic ways. The proportion of the three referential forms over the core argument functions S, A and P as attested in a sub-corpus of narrative texts – comprising approx.. 3500 clause units – is

given in Table 1. Common and personal NPs, and NumPs are subsumed here under “NP”, and pronominal NPs under “pro(nouns)”.

Table 1: Formal realisation of subjects (S and A) and objects in narrative texts

	S	A	P	TOTAL
NP	565 (23%)	135 (12%)	687 (60%)	1387 (30%)
PRONOUN	1399 (58%)	669 (61%)	229 (20%)	2297 (49%)
ZERO	470 (19%)	288 (27%)	224 (20%)	982 (21%)
TOTAL	2434 (100%)	1092 (100%)	1140 (100%)	4666 (100%)

Though I will not go into greater detail here, a few brief remarks seem useful for the following discussion: firstly, Table 1 shows the familiar low proportion of NPs in subject function, thus confirming what Chafe (1994) has termed the “Light Subject Constraint” (cf. also Prince 1981 on data from English). The constraint is much more pronounced for ‘transitive subjects’, i.e. the A function, a tendency known as the “Avoid lexical A” constraint, as part of DuBois’ (1987, 2003) “Preferred Argument Structure”. Moreover, non-lexical subjects are indeed more often pronouns than zeroes, a tendency that is probably best accounted for in terms of an ongoing grammaticalization process that turns subject pronouns into bound subject markers, as proposed by Schnell (2012, 2013, in prep).

Objects on the other hand have predominantly lexical form, in fact half of all NPs appear in object function. The proportion of pronouns is fairly low compared to that in subjects, corroborating roughly the “Avoid pronominal P” constraint determined by Genetti & Crain (2003) for Nepali and by Haig et al. (2011) for corpora from four typologically diverse languages, including Vera’a. I will occasionally come back to these distributional patterns of referential forms in the following discussion.

3.2 Introducing and establishing discourse referents

First-mention forms of discourse-new referents are generally lexical in Vera’a narratives, i.e. a NP is used to introduce a referent into discourse. This confirms what is proposed in the literature with regards to the interaction of information status and referential form (e.g. Ariel 1990; Chafe 1976, 1994; Lambrecht 1994:108): non-lexical forms are confined to discourse-given referents (sufficient, not necessary condition on givenness, activeness); lexical NPs can have both new or given (inactive or active) referents, but new referents are (usually) introduced (or given referents re-activated) by lexical NPs (lexical form is a necessary, not sufficient condition on newness / inactiveness). One exception to this

general rule is the use of 3rd person plural pronouns for first mentions attested in the Vera'a narratives:

- (9) a. *alē* =*n* *qōn ne vōwal dir* =*ēm gis ēn lavet vōwal*
 alright =ART night NUM.ART one 3PL =TAM1 hold ART feast one

'Alright, one day they hold a party (in a neighbouring village).'

1.PALA.021

- b. *n vono-n e Qo' dir ga ul so Tamlinlin*
 ART home-CS PERS.art Qo' 3PL TAM3 call QUOT place.name

'Qo's home (village), they call it Tamlinlin.'

JJQ.159

Such first-mention pronouns may either refer to a specific group of people associated with a particular place, as in (9a), or have generic reference, as in (9b). They are extremely rare in the corpus of narrative text where only 20 cases are attested, all in subject function (of a total of 1399 subject pronouns in the corpus; cf. Table 1 above). Also note that they are not the topic in both examples: (9a) represents athetic articulation where the entire event of these people holding a feast is introduced; and in (9b), the topic is Qo's home village.

More commonly thus referents are introduced into discourse by means of a lexical or personal NP. They are then resumed by either a pronoun or another lexical form. The following are typical examples from the corpus:

- (10) a. *qōn vōwal e ruwa mē =n gunu-ruō*
 night one PERS.ART two.people DAT =ART spouse-3DL

duru =m 'ogo 'ogo vaa-van
 3DL =TAM1 stay stay RED-go

'Once upon a time, there was a couple. And as time went by [and the two stayed and stayed, on and on]'

ANV.001-002

- b. *no me kaka =n kaka ne vōwal*
 1SG TAM4 tell =ART story NUM.ART one

kaka anē dī [...]'amē =n qono wo =n gusōwō
 story DEM1 3SG ASSOC =ART seagull and =ART rat

‘I will tell a story. This story, it is about a seagull and a rat.’

GAQG.001-

002

The first clause in (10a) represents the canonical beginning of custom story in Vera’a, introducing initially the parents of the actual protagonists. This pattern is attested in 21 of 26 stories examined in greater detail so far. The referent is introduced via a NP, and then immediately taken up by a personal pronoun. In (10b), the story that the speaker is about to tell is referred to by a common NP marked as ‘indefinite’ by the numeral phrase *ne vōwal* ‘one’, which basically has quantifying function. The referent is subsequently taken up by a (further) left-dislocated lexical NP and a resumptive subject pronoun in the following clause, the left-dislocated NP containing the demonstrative *anē* ‘that, this’. I will turn to the function of this particular demonstrative below.

Hence, two strategies for the introduction and establishment of discourse-new referents can be observed in these examples: 1. the use of the numeral phrase in first mention NPs, and 2. the use of a repeated lexical form in second mentions. These appear to serve particular discourse-structural functions that are probably not exhaustively accountable in terms of ‘accessibility’ or ‘activation’, but are instead more relevant in terms of the role of the referent in subsequent discourse. Let us first compare first mentions with and without a numeral phrase:

(11) a. *duru =k van rōw ___ 'ēn ēn 'ili tōo ne vōwal ba*

3DL =TAM2 go seawards see ART egg fowl NUM.ART one but

[...] *ba =n 'ili tōo ēn varaba*

but =ART egg fowl ART twin

duru =k van 'ō ___ sar = ma

3DL =TAM2 go carry bushwards = hither

‘And then the two turned seawards, and spotted a fowl’s egg. But the fowl’s egg was twin.’

ANV.035-036

b. *duru =m van rōw vovōnodo*

3DL =TAM1 go seawards fishing

buskat di 'ō =n vus wo =n gōsuwō di 'ō =n qe-
go-go'

cat 3SG with =ART bow and =ART rat 3SG with =ART LONG-RED-hook

‘The two went down to the sea, fishing. [...] Cat had bow and arrow, and Rat had the fishing rod.’

GABG.036-038

There are two factors that seem to determine whether a discourse-new referent is introduced by a numeral-marked NP or not: firstly, “brand-new” referents that are somewhat “unexpected” from the discourse context and not inferable from world or cultural knowledge are more likely to be introduced by a numeral-marked NP. Thus, in (11a) the two protagonists unexpectedly come across two eggs in the middle of the bush; in (11b) on the other hand, the bow and arrow and fishing rod are inferable from the context via the frame-semantic structure of the event of ‘going fishing’. Secondly, in (11a) the egg will be of some concern for the two protagonists, while in the subsequent of (11b) the bow and arrow and fishing rod will not be mentioned ever again (which seems unnecessary, given that they are understood to be present anyway, hence the connection between these two factors). I interpret the deployment of the numeral phrase in these examples as a strategy to introduce a specific discourse-new referent via the operation of individuation. Individuation, in turn, is employed only to those referents that are not expected and are of some relevance for what follows.

Turning now to the repeated lexical mention of discourse-new referents, let us consider the continuation of the story ANV, the beginning of which was presented in (9a) above:

(12) a. *n= rēne anē ne wotoqtoqo ___ ne visis ēn ni'i rēne*
 ART= woman DEM1 TAM2 pregnant TAM2 lay ART small woman

duru =m 'ēn gōr ēn ni'i rēne anē vaavan
 3DL =TAM1 see secure ART small woman DEM1 RED-go

‘The woman got pregnant, and gave birth to a little girl. The two [parents] looked after the small girl until [the mother of the little girl died. and then her father looked after the small girl; cf. (1a)].’

ANV.003-004

b. *duru =k lele =n nar*
 3DL =TAM2 RED:take =ART almond

duru =k gegen ___
 3DL =TAM2 RED-eat

In (12a) (and its subsequent context), the ‘little girl’, who is the central character in the first part of the story, is mentioned four times with lexical form, and only then it is taken up again by pronominal reference. In contrast, the ‘almonds’ in (12b) are non-specific and are not of major concern in the story (though the almond tree will be of some relevance), and they are referred to by zero form just after having been introduced. The pattern of ‘subsequent’ (i.e. after introduction) lexical mention versus subsequent non-lexical mention is thus related not so much to the givenness, accessibility or activation of the referent in question, but rather to its topicality in the global (subsequent) discourse. In my view, such examples do pose a serious challenge to models like Ariel’s (1990:74ff.) Accessibility scale, where only the ‘state’ of the referent at a given point of mention is taken as determinant of the referential form employed. Both referents in (10a) and (10b) are quite certainly equivalent in this regard at their subsequent mention, however, they do differ in their form, which is determined by subsequent (rather than preceding) context⁶. Lichtenberk (1996) presents similar findings of narratives from the Oceanic language To’aba’ita where protagonists are established by multiple lexical mention, as opposed to typically non-human, non-discourse-topical referents the second-mention forms of which are more often non-lexical.

As pointed out by Himmelmann (1996:222), discourse referents are not firmly established by an (indefinite) NP alone, and instead a sequence of two lexical mentions is often required, hence indefinite NP – definite NP (- pronoun/proform). The Vera’a pattern seems to serve exactly the function of establishing discourse referents; however, those discourse referents that are of primary concern will not get properly established, hence they do not show multiple lexical mention. However, again more systematic (and quantitative) corpus investigations of discourse-new more or less topical referents are necessary to substantiated or falsify this claim.

Also, Himmelmann (1996:222f.) mentions what Wald (1983:93) (apud Himmelmann ibd.) labels ‘*this-new*’ function as a possibility to establish a discourse-new referent through the use of a cataphoric proximal demonstrative, but does not find this strategy attested in his Indonesian or Tagalog data. Vera’a, however, does seem to exhibit this strategy, albeit in a somewhat restricted manner. Consider the following examples showing the use of the proximal demonstrative *agēnē* in first-mention NPs:

⁶ Note that Ariel (1990:78) does acknowledge that lexical forms are occasionally used for accessible discourse referents; however, she interprets such usage as “over-use” or “mistakes”, found for instance in the speech of little children. However, the instances of lexical reference reported here are quite certainly systematic and follow the pattern described.

(13) a. <ah> *gēnē =n kaka agēnē di 'alē =n biēg*
 here =ART story DEM2 3SG ASSOC =ART breadfruit

‘Here, this story here is about breadfruit.’

MVB.001

b. [...] *e sē ga lama'i =n dōdōmia agēnē*
 PERS.ART who TAM3 know =ART thought DEM2

di ne as ēn binigi di =m van [...]
 3SG TAM2 rub ART hand-3SG 3SG =TAM1 go

‘Who knows this (following) idea, (s)he will rub their hands, (s)he will go, ...’

ISWM.195-

196

As a rule, it appears, the use of *agēnē* is restricted to the so-called ‘discourse deictic’ use (cf. Himmelmann 1997:224ff.), referring here to the content of a story (13a) or an idea about a procedure (13b). The use of Vera’s *agēnē* thus does have a *new-this* function, hence allowing for the establishment of a new referent by a single demonstrative-marked NP, but at the same time seems to restrict this function to the discourse deictic use. It remains to be seen whether entity referents may be introduced by the same marking.

3.3 Resuming discourse referents

Pronouns and zero forms are specialised in discourse-given and activated information status, usually not being used for inactive or even discourse-new referents. Also, as pronouns and zeroes together make up the majority of all referential forms found overall in the corpus (cf. Table 1 above), and NPs may also have non-given inactive referents, these non-lexical forms are clearly the most typical referential forms for the resumption of discourse-given referents.

The greater prevalence of non-lexical form is in fact found with subjects rather than with objects, which show a majority of lexical mentions, cf. Table 1 above. Subjects have thus most often discourse-given referents, and these are also mostly topics (cf. 4.1.1 below). Nevertheless, lexical subjects often seem to have discourse-given referents as well, being either non-continuous topics or non-topics. Objects, on the other hand, do indeed seem to be the typical function for introducing new referents into discourse via lexical NPs.⁷

⁷ Again, it needs to be said that these preference claims require substantiation by more systematic corpus investigation: what is required here is a systematic investigation of the information status of lexical subjects and objects, assuming that non-lexical ones always have discourse-give, activated status.

The following example shows typical realisation patterns of subjects and objects with mostly discourse-given referents:

- (14) *di* = *m* *mom ēn mu-gi* = *n* *ga si-sidīn*
 3SG = TAM1 put ART POSS.GEN-3SG = ART rope RED:bird.catch
- di* = *m* *momom qē'* ___ ___ *so ne 'isiwē*
 3SG = TAM1 RED:put finish ?? TAM2 descend
- wo e raga anē* = *k* *sur ēn nes so*
 and PERS.ART people DEM1 = TAM2 sing ART song QUOT

‘[They all climbed on different nutmeg trees, but Qo’, he climbed on the big nutmeg tree. He climbed up to the tree top.] Then he set his bird caching rope, set it all ready, and was about to climbed down. And they [i.e. his brothers] started to sing this song:’

JJQ.145-146

In the first three clauses in (14), the subject has either pronominal or zero form. In all cases, the subject refers to Qo’, the mythical hero of the narrative JJQ, who is also the continuous topic in all these clauses. The topic shifts in the last clause to Qo’s brothers who were last mentioned six clauses before, and hence a lexical NP is employed to signal the topic shift, and presumably to re-activate the referent. The lexical object in the first, and its zero resumption in the subsequent, clause represent the typical pattern of introducing non-topical discourse-new referents, as outlined above.

The choice between a pronoun and a zero form for discourse-given referents is apparently not determined by considerations of accessibility or activation, as suggested by the set-up of Ariel’s (1990:75) Accessibility scale (and similar scales established in the literature). Hence, we find the usage of both forms equally attested for the continuous topical subjects in (14) above with no apparent difference in information status. Also, for subjects animacy does not seem to determine the choice, as can be seen from the figures in Table 2: both pronominal and zero forms are used for both human and non-human subject referents, with a general preference for pronouns. Schnell (2012, 2013, in prep) suggests that the high proportion of pronominal realisation among subjects in fact points to an ongoing grammaticalization process of bound subject indexes from originally free pronouns. At the current stage of this development, then, the choice between overt and zero form for subjects seems to be determined mainly by the clitic versus non-clitic form of the following TAM marker, compare the first two clauses to the third one in (14).

Table 2: Formal realisation of human and non-human subjects in narrative texts

	+hum	-hum	TOTAL
NP	455	245	700
PRONOUN	1936	132	2068
ZERO	670	88	758
TOTAL	3061	465	3526

With objects, a quite different pattern is attested: the choice between pronominal and zero forms is ultimately determined by the animacy of the object referent. Human objects are most commonly expressed by a pronoun, while non-human ones take zero form, cf. Table 3 for the relevant figures.

Table 3: Formal realisation of human and non-human objects in narrative texts

	+hum	-hum	TOTAL
NP	117	570	687
PRONOUN	224	5	229
ZERO	18	206	224
TOTAL	359	781	1140

The following two examples illustrate the pattern found with non-lexical objects:

(15) a. *e ruwa m̄ ē rm̄ ē re =k van e ___ =k le go-r'ō l*

PERS.ART two.people kid =TAM2 go CC =TAM2 take POSS.EAT-3TL

=n gengen dir'ō l 'ē qē l 'ō ' kel ___ ma
 =ART food 3TL descend carry back hither

‘The kids fetched some food for the three of them, and then they (the three) went back down with (it).’ HHAK.147-148

b. *dir =ēk qērē ba'a di sar lē =n mō-gi =n niṁē*
 3PL =TAM2 push into 3SG in LOC =ART POSS.HOUSE-3SG =ART house

‘They would then pish her into her house..’ ISWM.171

In both (15a) and (15b), the referent of the object is discourse-given and activated, and hence may be referred to by either a pronoun or a zero form, the choice being determined by the animacy of the referent. Consequently, these data suggest that in both subjects and objects the choice between pronoun and zero is obviously determined by

factors other than degree of accessibility, as suggested by Ariel's (1990:73) Accessibility scale.⁸

Despite the examples of second-mention lexical forms discussed above, the choice between lexical (NP) and non-lexical form is in many cases obviously determined by the activation of the referent – or distance to its antecedent⁹ –, for instance in the following example:

- (16) *wede di =m luwo*
rain 3SG =TAM1 big
- di ne le =n nak susu ē*
3SG TAM2 take =ART canoe RED-paddle DEM3

'When the rain had become big, it took that canoe [so that it floated down to the sea].'

JJQ.104-105

The antecedent of the lexical object in (16) occurs 20 clause units and 8 intonation units away, and the passage between these two mentions of the canoe has a different theme, namely the making of rain that is required to take the canoe out towards the sea. Hence, the referent is clearly not activated here, and a non-lexical form would not be appropriate. This is further underscored by the employment of the 'recognitional' demonstrative (glossed DEM3) in this object NP. Similarly, in (14) above the subject personal NP *e raga*

anē in the final clause is apparently deployed because its antecedent is a few clause units away, and also because of the competing antecedent Qo'.

However, recall that lexical forms also often occur immediately following on a preceding lexical mention. Thus, it seems that in contexts like (16), non-lexical form may not be appropriate for reference to the brothers, as it would be misunderstood as referring to Qo'. But this does not mean that lexical forms cannot have functions other than retrieving less accessible referent by being lexically more explicit, for instance the establishment of a hitherto discourse-new referent, cf. above.

Note that (impressionistically) quite many of these NPs with discourse-given referents are marked with a demonstrative, either the recognitional *ē* or the demonstrative *anē* which is

⁸ The cross-linguistic issue of availability and usage of High Accessibility Markers, in particular that of pronouns versus zero, is an open research question. Where both forms are available – though their usage may often be somewhat restricted, as in English; cf. Kibrik (2011:103ff.) for discussion – the Accessibility scale clearly predicts a cline in usage according to accessibility, a claim that is not supported by our Vera'a data.

⁹ Ariel (1990) considers 'distance of antecedent' and 'presence of competitors' as the two main components of accessibility.

likely to have (at least partly) topicalising or switch-topic marking functions, a point that I will return to in 4.2.1 below. It appears that NPs with discourse-given referents may also be unmarked:

- (17) a. [...] ___ *ne vrig 'a-'ag ma =n qōrō lie va'anē*
 TAM2 rush RED-follow hither =ART hole cave now?
 ___ *ne dīn ēn lie ___ ne wak*
 TAM2 flick ART cave TAM2 open

‘ (The devil) rushed along the (tunnel of) caves, flicked the caves (which would) open.’

1.PALA.085

- b. *di ne tēk so 'ooo kamam mi'ir lē =n qānris [...]*
 3SG TAM2 say QUOT no 1PL.EX:TAM1 sleep LOC =ART oven
'ama' ē ne lama'i so
 devil DEM3 TAM2 know QUOT
o qiri dir me mi'ir lik lē =n qānris
 INTERJ tonight 3PL TAM4 sleep more LOC =ART oven

‘He said: “No no, we slept in the oven.” [...] So now the devil understood that that night they would again be sleeping in the oven.’

JJQ.310-312

In (17a), the ‘cave(s)’ is in fact an already established discourse referent. It is taken up again in the form of an unmarked NP, and then mentioned again as such. In (17b), the ‘oven’ is an inferable (from architecture of a house) referent, and referred to by the same unmarked NP form when activated. Hence, givenness and activation are both necessary but not sufficient conditions on the deployment of demonstratives, and therefore demonstratives cannot be analysed merely as markers of givenness or activation (or ‘definiteness’), cf. 4.2.1 below.

Impressionistically though, it seems that such unmarked NPs are fairly rare and examples are rather hard to come by with. Quite often, it seems, lexical NPs contain a demonstrative. The three forms usually encountered in recorded texts are summarized in Table 4, giving their extra-textual deictic and their intra-textual phoric functions.

Table 4: Functions of demonstratives

	<i>agēnē</i>	<i>anē</i> (‘ <i>ē</i>)	<i>ē</i>
deictic	proximal near speaker	medial ? perceivable by speaker and hearer	distal ? not perceivable?
phoric	cataphoric	anaphoric	recognitional

The analysis of demonstratives in Vera’a needs much more detail than can be devoted to them here; in particular, the possible spatial and speaker/hearer-related distinctions in their deictic usage is understood only minimally yet. As for their phoric functions, the cataphoric usage of the proximal demonstrative was already demonstrated above.

The forms labelled ‘anaphoric’ and ‘recognitional’ here are indeed both exclusively used for anaphoric retrievable of discourse-given referents; no examples with discourse-new referents are attested in the corpus. The first two clauses of the following example illustrate their usage (cf. (16) above for function of the recognitional form):

- (18) [...] *ne* ‘*ēn* *e* *ruwa* *mal̄mala* *ē* =*s* *sag* ‘*i*
TAM2 see PERS.ART two.people girl DEM3 =TAM6 sit LIM
- e* *ruwa* *mal̄mala* *anē* *so*
PERS.ART two.people girl DEM1 QUOT
- ‘*ei* *kamadu* *anē* =*m* *van* *ma* *sir* *nik* *anē*
INTERJ 1DL.EX DEM1 =TAM1 go hither for 2SG DEM1

‘[After the young man had finished dancing, he was about to come over and sit down,] and then he would spot the two girls sitting there. And the two girls said: “Hey, we came here because of you, [because we want to marry you.]”

1.PALA.059-061

In the first clause, ‘the two girls’ are taken up again after a longer passage about the young man dancing, and then heading towards them; thus, the recognitional demonstrative may be said to signal the ‘reactivation’ of this referent. In the second clause, the NP referring to the very same referent is marked by the anaphoric demonstrative, and this demonstrative seems to underscore that it is co-referential with the lexically identical NP in the immediately preceding context – “and **these (same)** two girls ...”. This seems to be confirmed by the following example, where the second mention of the young man contains the anaphoric demonstrative:

- (19) *duru* =*k* ‘*ēn* *ma* =*n* *lumgav* *ne* *vōwal*
3DL =TAM2 see hither =ART young.man NUM.ART one

lumgav ne vōwal anē di ne laa-laka sēnē
 young.man NUM.ART one DEM1 3SG TAM2 RED-dance like.this

wo ___ ne virig ma [...]
 and TAM2 rush hither

‘Then the two spotted a young man. And the young man, he danced, and then (he) came over,
 [and sat down under the wild kava plant.]’

1.PALA.044-045

In the first clause, the referent is just introduced into discourse by stating that he is being seen by the already familiar two girls. In the second clause, he is taken up again by repetition of the same expression plus the anaphoric demonstrative. The example is particularly interesting because it shows that the anaphoric demonstrative is not a ‘givenness’ or ‘activeness’ marker, as it co-occurs with the numeral phrase here, i.e. the marker for ‘non-identifiability’ and ‘newness’. It seems that the anaphoric demonstrative is indeed strictly anaphoric (“the just mentioned”, “the latter(?)”), regardless of the information status of the referent in question. The second mention in this example is obviously part of the establishing strategy as applied to a topical discourse referent.

Recall, however, that this demonstrative is by no means obligatory for NPs with activated referents, as demonstrated by examples like (17) above. It apparently does not resemble definite articles in languages like English or other similar forms employed as “definiteness’ markers. But what motivates the use of *anē* then?

The same question about the motivation for marking with *anē* obviously comes from examples like the last clause in (18) above, where a speech-act participant (SAP) pronoun is thus marked. Note that this clause is part of direct speech within the narrative, and hence the usage of *anē* relates to the spatial environment of the depicted world within that narrative. Why would one need to employ a form that seems to aid the correct identification of antecedents (or referents, with deictic use) with a personal pronoun, in particular one referring to a speech-act participant? The answer probably is that *anē* has functions other than this referential aid function, and this will be discussed further in 4.2.1.

3.4 Summary: information status

In this section, I have outlined the patterns of formal realisation of discourse referents with different information status. Discourse-new referents are typically introduced by lexical form, and discourse-given ones mostly expressed by non-lexical form. Where

discourse-given referents are taken up by a lexical expression, this is either due to its low accessibility, as predicated by the Accessibility scale (and similar models), or it serves a particular ‘discourse-referent establishing’ function that is reserved for more topical, here human, referents. The choice between pronoun and zero form is largely independent of accessibility. Demonstratives presumably play a role in identifying particular referents, but this does apparently not account for their usage sufficiently. Information structure considerations are likely to play a role, and these will be discussed in the following section.

4 Information structure

In this section I will first present the basic information structure (IS) patterns that appear to be identifiable on the grounds of discourse contexts from the corpus. These findings are to be taken as fairly preliminary at this stage, especially as prosodic means of marking IS functions could not be taken into account yet. Also, as mentioned above with regards to Matic & Wedgwood’s (2013) critical assessment of focus structure analyses, it seems to be doubtful that the functions of topic and focus are attested in Vera’a in the “same way” as they are in English or other languages. On the other hand, we need to start out at some point to arrive at an idea about how information is packaged in this language, and the patterns to be discussed seem to be a good point of departure.

After introducing some basic structures that appear relevant for IS purposes, I turn to the marking of IS functions. Two morphological devices seem to be involved in the marking of IS functions, but neither of them exclusively so. Moreover, the syntactic construction of left-dislocation has IS functions, but none of these seems to neatly coincide with “topicalisation” or “focalisation”. I will suggest that these structures have functions that have effects similar to that of topic or focus marking, but they clearly do not coincide with these.

4.1 Basic IS patterns

In this section I will give a brief overview of how basic information structure patterns correspond to morphosyntactic structures of clauses in Vera’a. As a working basis, I adopt Lambrecht’s (1994) conception of focus, but will refine his understanding of topic as mere ‘aboutness’. The definition of topic as understood here is closer to that of Reinhart (1981), and subsequently Krifka (2008).

Under “focus” I basically understand what Lambrecht (1994:213) defines as “[t]he semantic component of a pragmatically structured proposition whereby the assertion differs from the presupposition.” In this sense, focus describes a relationship to a

proposition whereby the latter expands the Common Ground (CG; adopted from Krifka 2008:245-261). This definition implies an exhaustive partitioning of a proposition into a presupposed component and the focus component. Lambrecht (1994:222ff.) distinguishes three basic types of focus structures, where a particular type of focus complements a type of presupposition: PREDICATE-FOCUS, ARGUMENT-FOCUS, and SENTENCE-FOCUS structures. I will adopt these three types of structure as “basic IS patterns” here.

As Lambrecht (1994:210) points out, the idea of a “focus relation” (on pragmatic and/or semantic level) has to be strictly discerned from that of “focus marking”; below I will discuss the possibility of focus marking in Vera’a. However, in the light of Matic & Wedgwood’s (2013) critical assessment of a universal (grammatical) category “focus”, “focus marking” is not understood here in the sense of Lambrecht (1994), where it seems to allude to a one-to-one form-function relation. Instead, it is assumed here that certain structures may produce an *effect* of signalling a focus relation of an expression’s referent to the proposition expressed by the clause, while possibly expressing a somewhat different (set of) function(s).

The notion of “topic” is probably even more problematic to define than that of focus. I basically follow Lambrecht (1994) in assuming that topic has to do with what a proposition is basically about (rather than equating ‘topic’ with ‘old’ information), being complementary to what is asserted, i.e. the focus element. Thus, focus types thus correspond to particular topic-structure types, namely “topic-comment”, “identification”, and “event-reporting” articulations, respectively (cf. Lambrecht 1994:222). Topics are required to be accessible or activated to a certain degree in order to fulfil their role, cf. Lambrecht’s (1994:165f.) Topic Acceptability Scale. A referent can be promoted on the scale, and languages may have morphosyntactic constructions that serve this purpose, Lambrecht (1994:176-184) discusses “presentational” and “detachment” (“left-dislocation”) constructions, but other types of morphosyntactic devices seem possible.

However, there Lambrecht’s conception of “topic” as global “aboutness” seems very vague, and therefore does not seem to always strictly distinguish a putative topic from other components of a proposition, for instance why isn’t Lambrecht’s (1994:149) example 4.24 (*As for Rosa, John didn’t reallt love her.*) also *about* “love”?¹⁰ A more confined

¹⁰ In fact, Lambrecht (1994:150) argues for his conception of multiple “aboutness” topics on structural grounds, claiming that *John* and *her* both show deaccentuation, hence being likewise “marked” as topic expressions. This seems to be an unnecessary conflation of levels of representation that Lambrecht himself points out need to be avoided. Cf. Reinhart’s (1981:58f.) comment that a proposition is in principle *about* all denotata it contains.

understanding of topic in the sense of Reinhart’s (1981) “file card metaphor” seems intuitively appealing. Both conceptions might in fact work together, so that a proposition can in principle “about” a number of entities, concepts, etc, but may *primarily* be functioning to expand the CG with regards to a particular entity or set of entities (cf. Krifka 2008:265 for a definition of topic along these lines). Secondly, both Lambrecht’s and Reinhart’s (and accordingly Krifka’s) approaches rely entirely on the notion of topic as a category corresponding to a single proposition expressed in a single sentence (or clause). It seems that neither the “aboutness” approach nor the “file card metaphor” require that they apply to single propositions; however, it seems that in some languages information “packaging” works on different levels (cf. Ozerov 2014 for an analysis of information structure in Burmese that essentially dispenses with the traditional notions of “topic” (and to a lesser degree also that of “focus”). The ‘anaphoric’ demonstrative in Vera’a seems to mark a category that clearly

4.1.1 Predicate-focus structure

The PREDICATE-FOCUS structure is considered here the unmarked IS structure, for example in the following clauses (prosodic accent represented by capitalisation):

- (20) *dir* =*k* VAN ___ *k rev lu ma =n KO-re*
 3PL = TAM2 go TAM2 drag out hither = ART POSS.VES-3PL
- ___ *dīn rōw lē =n mē’ērsa a REV*
 reach seawards LOC = ART harbour LOC.SPEC *place.name*
- dir* =*k vilvil ēn nak mu-RE*
 3PL = TAM2 RED:tie ART canoe POSS.GEN-3PL

‘(Then) they went to drag out their vessels, reached down at the harbour of Arep and tied up their canoes.’

JJQ.030-032

In (20) we find a chain of clauses with a continuous topic being expressed as the subject of each clause. All clauses are primarily ‘about’ a group of people (the brothers of the hero) and report on their consecutive actions, hence answering the question ‘what happened *to them?*’, cf. Lambrecht (1994:223ff.), and these are presupposed (by context). Hence, the form of reference is either a 3rd person plural pronoun or zero. The remainder of each clause expresses what is asserted about the group of brothers, and this assertion is in focus relation to the proposition as a whole. Hence, everything except the subject expression is the focus domain in each clause. The focus domain is not marked by

morphosyntactic means. Prosodic stress – as determined on the grounds of preliminary auditive analysis – falls on the last word of each clause-final constituent, interacting with principles of lexical accent assignment. Where this is not the VC, as in the first clause, it is either the object argument (as in clause 2 and 4) or an oblique argument (as in clause 3) or even an adjunct.

As for the information status of referents in topic and focus in (20), the continuous topics are given activated. The referents in focus have differing information status: the object referent in clause 2 is given, but inactive, while in clause 4, this same referent is given activated and bears a focus relation to the proposition. In clause 3, the goal participant in focus is given, but inactive. Note that all these referents are only “part” of the focus denotatum in these examples, as we are dealing with “wider” predicate focus here.

While typically in a predicate-focus structure it is the referent of the subject that bears a topic relation, object referents may be topics in a predicate-focus structure as well:

- (21) *di =m sik ē n menre-gi vō WAL|m SIKĒ* — [...]
 3sg =TAM1 search ART piece-3SG one TAM1 search
m siksik LĒ GĒ [...] | Qo =m mul 'ō ' —
 TAM1 RED:search in.vain Qo' =TAM1 go.home carry

‘He searched for one piece, searched and searched, but in vain: Qo’ had taken it home.’
 JJQ.057-058

In the last clause of (21), it is presumably the object that bears a topic relationship, with the remainder of the clause expressing the assertion that is the comment on this topic: it is essentially explained here why Spider cannot find the last piece (of a tree), and the answer is that Qo’ had taken it. In other words: this clause expresses what has happened to that piece of wood, hence its classification as a predicate-focus (topic-comment) structure. Note in passing that due to the topic status of the object referent and its zero realisation, it is the final constituent word of the VC that appears in clause-final position and receives prosodic marking; additionally, the subject is expressed by a personal NP rather than a pronoun or zero that would be used for a topical subject. It seems that the combination of these two properties contributes to the predicate-focus reading of the clause.

In other cases of presumed predicate-focus structures with object topics the role of the subject referent seems less clear:

- (22) *riar mu-gi [...] di =m mom 'i* — *suWEI*
 k.o.bow POSS.GEN-3SG 3SG(SUBJ) =REAL put LIM down

di e rem 'ō Rō s
 3SG NEG1 climb carry NEG2

‘His bow, he put (it) down, (he didn’t climb with (it)).’

ISWM.133-134

The referent of the dislocated NP (corresponding to the objects of the two clauses) refers to the protagonist’s bow, which is discourse-given but inactive. It has been mentioned 43 intonation units before: “He (the hero) took his bow and arrow, carried it in his hands, and went off.”, and after a longer walk he is now about to climb up a steep edge of a gorge to reach his beloved. He will, however, be ambushed by his rivals, and hence the fact that he does not take his weapon is relevant for the subsequent discourse. In this sense, it seems that the passage represented in (22) is *primarily* about ‘what happened to his bow?’. As the ‘bow’ is inactive here, it needs to be “promoted” to be eligible as topic, which is done by left-dislocation, or “topicalisation” (cf. 4.2.3 below). Again, the referent is realised as zero in object position, leaving the VC-final element stressed, presumably adding to predicate-focus reading.

What makes (22) different from (21) is that the subject relation is also expressed by non-lexical form, i.e. a personal pronoun, which is the typical form for a topic referent. It would indeed seem superficial here to propose that the passage is *not* about the protagonist as well, quite in accordance with Lambrecht’s idea of multiple topics. But, again, I cannot see why under such a loose understanding of “aboutness” the clauses are not also “about” leaving the bow behind? What seems more relevant for the development of the story (and of the CG, by the same token) here is to gain information about the bow. I thus propose that indeed everything except for the object is the focus domain here¹¹, the latter thus containing a non-focal element, quite in the sense of Lambrecht (1994:228), who mentions non-focal possessive pronouns with focal NPs (*My CAR broke down.*).¹²

In this section I have mainly dealt with the canonical cases of established continuous topics that do not require special marking (the object referent in (22) being the exception). I will turn to other types of “marked topics” in 4.2.1 below when discussing the question whether Vera’a possesses any means of special morphosyntactic marking for topics at all.

4.1.2 Argument-focus structure

¹¹ Note that this ‘domain’ does not coincide with the structural domain of a “VP” in the sense of a predicative phrase plus an object.

¹² Note that a significant difference between Lambrecht’s (1994:228) *my CAR* and the Vera’a example is that in the later case, the focus domain does not necessarily correspond to “a constituent”, depending on the PS analysis of the language.

ARGUMENT-FOCUS structures appear to be the ones that are in some of the literature treated as focus par excellence. In this focus structure, it is a particular referent that bears a focus relation to the proposition and therefore represents the assertion, the remainder of the proposition being presupposed (cf. Lambrecht 1994:228). In Vera’a, where the referent in argument focus is that of an object, it may be unmarked¹³:

(23) *ba =n (0.3) gengen GO-RĒ wunva =n (0.3) [...]*
 but =ART food POSS.EAT-3PL probably =ART

[...] so =n wōmō’ō’ wunva di ga gengen vuva =n ‘erē ()
 CPL=ART fish.sp probably 3SG TAM3 RED:eat only =ART PL

gengen ‘alē =n .. NŌN [...]
 food ASSOC =ART sand

‘As for its food, [when we gut it, we see that] the *wōmō’ō’* fish, probably what it eats is all the food found in the sand.’

fish04_CH.012-014

The topic of the passage reproduced in (23) is what the *wōmō’ō’* fish eats, considering what is found in its guts when slaughtering it. After ‘its food’ is set as the “frame” (cf. 4.1.4) for this passage, the following two clauses (represented in free translation only) have the fish expressed as the non-topical object. It is then “topicalised” by means of left-dislocation. The last clause has the fish as its (primary?) topic, and this and its eating are presupposed, with the food being the missing piece of information that is added to the CG, thus bearing the focus relation. Further examples of unmarked clause-final constituent with their referents in focus are the following, both (impressionistically) bearing prosodic stress:

(24) a. *di =m da mamas ēn ‘irbē-gi dē e lama’i RŌS*
 3SG =TAM1 do dry ART body-3SG 1PL.IN NEG1 know NEG2

di =m meñ mamas ēn ‘irbē-gi lē =n SAVA
 3SG =TAM1 towel dry ART body-3SG LOC =ART what

‘He made his body dry, but we don’t know with what he towelled his body.’

ISWM.111

¹³ Recall that prosodic sentence accent, at least impressionistically, is in clause-final position, and predicate-focus structures are thus not distinguished simply by the presence of prosodic accent. Note, however, that prosodic marking awaits further finer-grained investigation which may reveal subtle differences in the marking of object argument focus and predicate focus.

In (24), we find an indirect constituent question expressed in the last clause. Here, the final NP, complement of a preposition in an oblique PP argument, is headed by the interrogative-indefinite noun *sava* ‘what, something’ and expresses the focussed referent. The idea *that* the hero is towelling his body is presupposed, but what he uses as a towel is what the narrator cannot retrieve, hence being in focus here. The PP expressing the oblique argument in focus here appears in clause-final position and the complement NP carries prosodic stress, but is not further marked. In another example involving an indirect constituent question, however, the argument focus domain seems to be a subject constituent, a personal NP headed by an interrogative noun *sē* ‘who’:

(25) *'ei ba e sē =m 'aram mē nikē =n ge ANĒ*
 INTERJ but PERS.ART who =TAM1 tell DAT 2SG =ART DEICT TOP?

‘Hey, but who told you about this thing [i.e the big kind of breadfruit]?’

MVB.174

The clause in (25) represents a passage of direct speech within a narrative. The presupposition is that the addressee was told about the breadfruit, and the speaker asks *who* told him. Again, we can at this stage not rule out the possibility that there is some prosodic marking of the subject NP¹⁴. What might play a role here is the marking of the clause-final constituent with the demonstrative *anē* which may serve as a topic-marking device, as will be discussed in 4.2 below; this may have the effect of leaving the subject the only possible domain for focus, given that the ‘telling’ and the addressee are presupposed.

Another example exemplifies the analytical problems with the identification of focussed object referents on the grounds of discourse contexts as attested in corpus data:

(26) *no lama'i ros so duru =m da =n SAVA*
 1SG know NE2 CPL 3DL =TAM1 do =ART what

‘I don’t know what they did.’

MVBW.137

The trouble with deciding whether we are dealing here with predicate or (object) argument focus seems to be tied to the relative semantic under-specification, so that the indirect question constituent here seems to embrace VC plus object.

¹⁴ Impressionistically, a rise in pitch may be perceivable.

All examples of argument-focus structures presented in this section involved focus constituents that are morphosyntactically unmarked for this IS function. The only marking impressionistically determined is that of prosodic stress, falling by default on the constituents in questions that appear in clause-final position. Recall that this stress does not uniquely identify argument focus though. In 4.2.2 below I will take up the issue of morphosyntactic focus marking again.

4.1.3 Sentence-focus structure

Sentence-focus structures seem to be quite rare in the texts recorded. It will suffice here to present a single example:

(27) *alē =n qōn ne vōwal [..]. dir =ēm gis ēn lavet vōwal*
 alright =ART night NUM.ART one 3PL =TAM1 hold ART feast one

‘Alright, one day they held a feast..’

1.PALA.021

The clause in this example expresses an entirely new proposition, and none of the referents of subject and object can be regarded as topic. Hence, the entire clause (sentence) is the focus domain (cf. Lamrecht 1994:233).

4.1.4 Clause-initial frame-setting expressions

Frame-setting “topics” have the function of delimiting the validity of the following (set of) proposition(s) to a particular domain (cf. Krifka 2008, Chafe 1976). Different types of expressions with varying semantics occur as frame-setters in clause-initial position. The left-dislocated NP in the first line of example (23) above is an example of such a frame-setting topic expression: the referent ‘food’ is not a topic of the proposition expressed by the following clause, but rather sets the frame in which the following clauses are relevant, including the gutting and the observations made. The entire passage is given below as (28):

(28) *ba =n (0.3) gengen go-rē wunva =n (0.3)*
 but =ART food POSS.EAT-3PL probably =ART

masō-gi dē =k vena =n mes
 time-3SG 1PL.IN =TAM2 catch =ART fish

gidē me 'ēn so =n wōmō'ō' wunva
 1PL.IN FUT see CPL =ART fish.sp probably

di ga gengen vuva =n 'erē (0.2) gengen 'alē =n nōn

3SG TAM3 RED:eat only =ART PL food ASSOC =ART sand

‘As for their food, probably ... when we gut the fish, we would see that the *w.* fish, probably it eats only all the food that can be found in the sand.’ fish04_CH.013-015

As is evident from this example, the position of the extra-posed phrase is outside of the clause or even the sentence, and the dislocated NP apparently precedes the entire sequence of clauses presented here. A more detailed analysis of extra-core / -clausal / -sentential slots still needs to be done for Vera’a, and I will leave it with these parsimonious comments. More typically, it seems, temporal or locational expressions are found as frame-setters in clause-initial position. These refer to a point in time or period of time, or a place for which the proposition expressed by the clause is valid, e.g. when a particular event took place:

(29) *lē* =*n* *masō-gi anē* =*n* *mōg di* =*m* *rān ēn vunuō*
 LOC = ART time-3SG DEM1 =ART quake 3SG =TAM1 shake ART island

‘But during this time, the earthquake shook the island.’ BSVH.008

As will be discussed in 4.2.3 below, the clause-initial, ‘left-dislocated’ position has a wider range of functions and is not restricted to that of ‘topicalisation’ or ‘frame-setting’.

4.2 Marking of topic and focus?

I now turn to the question whether Vera’a possesses morphosyntactic means of marking IS functions like topic and focus, and how these spell out in the structure of the language. The main result of this discussion is that Vera’a does not possess specialised topic or focus markers. Instead, a demonstrative that has “usually” deictic functions is employed as a marker of topic in some contexts, and otherwise as a marker of presupposition. A particle that seems to have “emphatic” and “assertive” function can in some contexts have the effect of focus marking. Hence, “marking “ of topic and focus relations in Vera’a is a mere by-effect of the marking of other categories (cf. Matic & Wedgwood 2013). Again, I need to make it clear that possible prosodic marking of IS relations still needs to be investigated systematically, hence I will be discussing morphosyntactic means of marking only here.

4.2.1 Demonstratives as topic markers?

Demonstratives have already been discussed in 3 above as a means of marking the ‘discourse status’ of discourse referents. Specifically, it was shown that in texts the

‘proximal’ demonstrative can have cataphoric function, and therefore be used an aid to introduce discourse-new referents. The ‘medial’ and ‘distal’ demonstratives are used with given referents, and at least have reference-tracking functions. However, as was indicated above, this does not seem to be the whole story, and anaphoric demonstratives seem to have a wider function than just signalling the givenness or activation of a referent. It will be suggested here that their functions comes close to that of topic marking, but that it does not perfectly seem to match this function either.

As noted above, *anē* is used with referents that are clearly given-activated and hence do not seem to ‘require’ such marking. A blunt example is the use of *anē* with 1st and 2nd person pronouns in (30a), but the same applies basically to (30b):

- (30) a. [...] *ne 'ēn e ruwa mālmāla ē =s sag 'i*
 TAM2 see PERS.ART two.people girl DEM3 =TAM6 sit ANT

e ruwa mālmāla anē so
 PERS.ART two.people girl DEM1 QUOT

'ei kamadu anē =m van ma sir nik anē
 INTERJ 1DL.EX DEM1 =TAM1 go hither for 2SG DEM1

‘[After the young man had finished dancing, he was about to come over and sit down,] and then he would spot the two girls sitting there. And the two girls said: “Hey, we came here because of you, [because we want to marry you.]”’

1.PALA.059-061

- b. *e ruwa a =n ni'i-gi anē duru =k van*
 PERS.ART two.people REL =ART child-3SG DEM1 3DL =TAM2 go

‘[Those spirits who were her friends said: “Oh yes, you were right, it’s your children’s voices that the wind is carrying over here.”] And so the two who were her children, they went.’

MVBW.073

In these examples, *anē* does not seem to fulfil functions typically associated with demonstratives, namely activating new referents, reactivating given ones, or picking out a particular referent from a number of competitors (cf. Diessel 2006:470). Hence, *anē* does not contribute to identifying the correct referent in these contexts, or retrieving the

correct antecedent (cf. Ariel 1990:73ff.).¹⁵ Moreover, *anē* is used with discourse referents that are at the same time marked as being non-identifiable, namely via the use of the numeral *vōwal*:

- (31) *duru =k 'ēn ma =n lumgav ne vōwal*
 3DL =TAM2 see hither =ART young.man NUM.ART one
- lumgav ne vōwal anē di ne laa-laka sēnē*
 young.man NUM.ART one DEM1 3SG TAM2 RED-dance like.this
- wo ___ ne virig ma [...]*
 and TAM2 rush hither

‘Then the two spotted a young man. And the young man, he danced, and then (he) came over, [and sat down under the wild kava plant.]’

1.PALA.044-045

On the other hand, as was made clear above, many unmarked lexical NPs have clearly identifiable antecedents in the preceding discourse; thus accessibility is a necessary but not a sufficient condition on marking with *anē*: Therefore, *anē* is apparently not grammaticalising into a default identifiability (or ‘definiteness’) marker, which is – at least to a large extent – what definite articles are in languages like English.

As with the use of numeral phrases or the repeated lexical mention of not yet firmly established discourse referents, it seems that the function of demonstratives is – at least partly - to be accounted for in terms of their referents’ salience in discourse in general (including their role in *subsequent* discourse) rather than just their information status at a given point in discourse. Specifically, I think that the examples mentioned so far lend themselves to the hypothesis that *anē* has topic-marking effects, and this will be explained in the following paragraphs.

A particularly blunt example for the topic-marking effect of *anē* is (30a): here, the two girls have not been mentioned for a longer stretch in discourse, and are being re-activated by the object NP in the first clause. The preceding clauses had described the young man’s dancing and coming back to where he was sitting, only to find the two. The recognitional demonstrative is employed in this re-activating object NP. In the following clause, the

¹⁵ Note, however, that this does seem to be the major function of these demonstratives in their deictic use, namely bringing entities in the extra-linguistic environment to the interlocutors’ mutual attention, quite in the sense of Diessel’s (2006) ‘joint attention’ account.

same re-activated referent is taken up again in a *anē*-marked NP, and is the topic of this clause, performing a direct speech act. Similarly in (30b), the ‘two children of hers (their dead mother)’ are active but not the topic in the preceding discourse. Specifically, some spirits argue with the two kids’ dead mother whether some noise is the kids’ voices or just some other noise, and the spirits think that it is just some birds’ noise. Their discussion takes place on the top of a mountain where the spirits live. In (30b), then, the topic switches to these two kids who are walking in search of her deceased mother through the bush. The switch in topic entails a switch in scene, i.e. location and action, referred to by a left-dislocated NP that takes the demonstrative, and this signals that the two kids are the topic of what follows. Example (31) is parallel to (30a), with the difference that the referent is not yet quite established in discourse here, but yet can be marked as the topic of the second clause after he has been introduced by a numeral-marked NP.

In sum, it seems, the demonstrative *anē* does indeed not have accessibility-marking – more clearly the function of the recognitional demonstrative – functions, but topic-establishing ones. Specifically, it seems to mark pre-VC NPs, either subject or left-dislocated, for expressing a new or switch topic referent, and is in this sense ‘topic-establishing’. This function is corroborated in examples (30b) and (31) by the left-dislocated position of the NP. There also exist examples of *anē* marking NPs in subject position for new / switch topic reference, as in (32) below. This example, however, also shows an object NP that seems to be marked with *anē*:

(32) *masōgi* [...] *dir* = *m* *van* *mē* = *n* *sisidīn* [...] *di* *ne* *rem* *ēn* *daraga*
time 3PL = TAM1 go DAT = ART RED-bird.catch 3SG TAM2 climb ART nutmeg

daraga *ga* *ul* [...] *so* = *n* *darag* *towla* *di* *ga* *lu-luwo*
nutmeg TAM3 call QUOT = ART nutmeg ?? 3SG TAM3 RED-big

so *e* *raga* *anē* = *m* *remrem* *lē* = *n* *'erē* *darag* *ga* *sēsēe*
?? PERS.ART people DEM1 = TAM1 RED:climb LOC = ART PL nutmeg TAM3

different

ba *e* *Qo'* *di* = *m* *rem*
but PERS.ART Qo' 3SG = TAM1 climb

'a = *n* *darag* *anē* *mē* = *n* *sisidīn*
SPEC.LOC = ART nutmeg DEM1 DAT = ART RED-bir

‘When they had arrived where they wanted to go bird catching, [they climbed,] he climbed a nutmeg tree. That nutmeg tree is called [in the Vera’a language (it) is called] *towla* nutmeg. [It is very big.] So, everyone climbed all different (kinds of) nutmeg trees, but Qo’ he climbed that (bespoken) nutmeg tree to bird catch.’

JJQ.140-143

The topic-establishing function of *anē* can be seen from the third line of (32): here, Qo’s brothers are taken up again as the topic of the clause, after the preceding passage was about Qo’s climbing up that particular nutmeg tree, which is then further elaborated on. The marking of the *object* NP, however, obviously suggests that the hypothesis needs to be modified somewhat, as here the referent in question is not simply *the* topic of the sentence (which in (32) clearly is Qo’). Before turning to this more intricate case, consider the following example:

- (33) *ba duru =s vanvan 'ō 'i anē*
 but 3DL =TAM7 RED:go carry LIM DEM1
- 'ō =n sōm sōm 'amigidē*
 carry =ART shell.money shell.money of.ours
- alē duru =m le sur ēn sōm anē mē diē*
 alright 3DL =TAM1 give down ART shell.money DEM1 DAT 3SG
- duru =m wōl ēn raw anē den di va'anē*
 3DL =TAM1 purchase ART pig DEM1 ABL 3SG now??

‘[“We are looking for an intra-sex pig for us.” “But I have got an intra-sex pig right down there!”] But what the two had taken with them was shell money, our traditional money. Alright, so they gave that shell money to him. ’ AS.1.014-015

In (33), the referents ‘shell money’ and ‘intra-sex pig’ have been mentioned lexically just prior to the last two clauses presented here, and they do not have competitors; hence again, no referent-retrieving function can be ascribed to *anē*. On the other hand, these two clauses seem to exhibit fairly canonical topic-comment articulation, denting the two heroes’ actions. They are realised as subject pronouns, hence typical topic expressions. However, what I believe is relevant here is the role of the shell money on the one hand and the pig on the other, namely that it is the former that changes possession from the two to the other person, and that it is the pig that changes possession from that other one to the two, in exchange for the money. This contrast is marked by the – otherwise not preferred (cf. Schnell 2012) – position of the pronominal dative PP at the end of the

clause, i.e. the typical focus position, and it has its mirror image in the predicate, namely ‘give’ vs. ‘purchase/take’. In other words, we are mainly concerned in these two clauses with how the money is relevant, and how the pig is relevant. In these two particular clauses, it seems save to consider the objects’ referents the topics of the clauses. The situation is, however, different in (32) above where Qo’ is clearly the topic of the clause in question. What (32) and (33) have in common though is that the referent of the *anē*-marked NP is of particular relevance by means of its particular properties, as described in the immediately preceding passage. It is as if we find a secondary assertion here, roughly “the *darag towla* tree has Qo’ on it”. In the subsequent context then, it is this tree that turns into a trap for Qo’ by growing in girth so that he cannot climb back down again. This seems to come close to a secondary topic function, but I will not go into more detail here. What *anē* then mainly seems to do is to mark an element of CG as particularly relevant for the current proposition, and for the resulting expansion of CG.¹⁶

This seems to be mirrored in the function of *anē* as a clause-combining or discourse-structuring device when it occurs on clause level, in final position:

- (34) a. *qe dērōl =k ōn sursur va'anē*
 finish 3TL =TAM2 lie RED:down now??
- dērōl =k ōn sursur suw anē ___ =k mi'ir va'anē*
 3TL =TAM2 lie RED:down thither DEM1 =TAM2 sleep now??

‘And then the three lay down. They lay down, and then they went to sleep’

2.PALA.033-

034

- b. *'aluwō nike kel ma*
 tomorrow 2SG:TAM2 back hither
- du =k bēr 'ān'ān ēn nak anē*
 1DL.IN =TAM2 tie.up ART canoe DEM1
- wo =m disivie nik mak rēv lu =n ko-m va'anē*
 and =TAM1 do.how 2SG TAM8 drag out =ART POSS.VES-2SG now??

¹⁶ On an entirely different line of reasoning, note that PPs can function as modifiers within NPs, hence the demonstrative in examples (32) and (33) coincidentally (?) serves as a phrase delimiter. However, this function may in fact be quite in accordance with the interpretation outlined here: the demonstrative may, on this account, serve to mark the NP as referential ‘in itself’.

{“But when will it be tied together?” [...] }Ok, tomorrow you will come back here, and we will will tie the canoe together. But then how will you drag it down to the sea?”

JJQ.081-082

Thus, (34a) clearly demonstrates the possibility for *anē* to occur in clause-final position. In (34b), its position is ambiguous between NP- and clause-level positions. The latter, however, is strongly suggested by its function in this context. Namely, where *anē* occurs in clause-final position, it seems to function as a clause-combining device by marking the proposition of the first clause as the (common) (back)ground to which a new state of affairs is adjoined to yield new information, and an expansion of the CG is ‘linked’ to this proposition. Thus, in (34a), the *anē*-marked clause refers back to the going to bed, and once this is established, the following clause refers to the characters’ going to sleep. Similarly in (34b), the *anē*-marked clause establishes the tying-up of the canoe, and once this is done, it can be pulled out of the bush to the shore. In analogy to the use of *anē* on NP level, the propositions expressed by *anē*-marked clauses are clearly known and active. Impressionistically, it seems that *anē*-marked clauses in fact always repeat a proposition expressed in the immediate pretext, and more often than not the clausal construction is in fact a head-tail linkage, i.e. a non-finite type of clause which resumes a state-of-affairs for further expansion.

In analogy to its use on NP level, the propositions expressed by *anē* marked clauses are already presupposed, thus part of CG, and as they are usually expressed repeatedly, these presupposed propositions are in fact active in the respective contexts. Thus, *anē* is presumably not used to simply signal that a proposition is presupposed, as would indeed be a case of superfluous over-marking. Rather, in analogy to the function of *anē* on NP level, it seems to pick out a particular active presupposed proposition to establish it as the anchoring point for new information, i.e. new proposition.

Hence, the function of *anē* is similar to that of “topic-marking” with pre-VC constituents in the sense that a referent is established as the topic for the following proposition. But as the examples show, its functions are somewhat wider in selecting a part of the CG as particularly relevant for its subsequent expansion. This formulation of a somewhat wider function thus includes the marking of certain (non-sentence-topical) objects or oblique

arguments, as well as its marking of presupposed propositions as the anchoring point for CG expansion. These functions of *anē* seem to diverge quite significantly from those referent-identifying or –tracking functions usually ascribed to demonstratives, as already mentioned above (cf. Diessel 1999, 2006). Such identificational properties of demonstratives are – as far as IS function are concerned – associated with focus-marking functions (cf. Diessel 1999:148f.) rather than topic-marking (or similar) ones. Investigations of the functions of demonstratives in Vera’a may reveal more systematic regularities than can be presented here.

Finally, I shall mention a few examples where the identificational functions just mentioned are still clearly observable with *anē*, and where accordingly the IS relation of the respective referent seems to be that of focus rather than topic, hence providing possible counter-evidence to the “(near) topic/ground-marking” hypothesis outlined thus far. Note that all the following examples represent direct speech:

- (35) a. *ba wunva sa =n rēne a dē =m vus ___ ē*
 but probably FOC? =ART woman REL 1PL.IN =TAM1 kill DEM3
 =n *ni'igi sa ne vōruō anē =n varaba*
 =ART child-3SG FOC? NUM.ART two DEM1 =ART twin

‘But probably this woman that we have killed, her kids are really these two here, these twins.’

ANV.024

- b. *no =k kur ma-ma' nikē birīn sa =n wova'al anēē*
 1SG =TAM2 devour RED-dead 2SG with FOC? =ART pawpaw DEM1

‘[Friend, I got enough. I will climb up there right now!] And then I will eat (and kill) you right with that pawpaw!’

GABG.098

- c. *gēdu qē' ga moros*
 1DL.IN finish TAM3 want

- ba gēdu qē' du =k leg birīn sa =n lumgav anē*
 but 1DL.IN finish 1DL.IN =TAM2 marry with FOC? =ART young.man DEM1

‘Bioth of us want (that). So both of us, we should get married with that young man there!’

1.PALA.054

First of all, such examples seem to be restricted to direct speech in the corpus of narrative texts. Passages of direct speech within a narrative evoke a communicative situation together with the imagined space where it takes place, hence the originally spatial or joint attention (of the characters) properties of demonstratives may be more prominent in the contexts; thus, in all the examples (35a) – (35c), the referents of the *anē*-marked NP is located right before the interlocutors eyes.

However, in all these examples, the respective NPs appear in a typical “focus position” and seem to be accented likewise. Moreover, they are all marked with the particle *sa*, and as will be discussed in the following section, the latter has functions that may at least produce focus-marking effects. I will discuss further examples of combined *sa* and *anē* marking of NPs there. These latter examples require further scrutiny, which I will leave to future investigations.

4.2.2 Functions of *sa*

Similar to *anē*, the particle *sa* seems to have a IS marking function that does not straightforwardly correspond to a IS category canonically assumed. Unlike *anē*, however, *sa* does not have any other functions outside the domain of IS marking. In work on neighbouring languages, forms similar to Vera’a *sa* have been glossed as ‘focus’ markers, and that seems to be what *sa* marks in Vera’a. The following examples support a focus analysis of *sa*:

- (36) a. *o bul nik sa me rem ēn wova'al ē*
 INTERJ friend 2SG FOC? TAM4 climb ART pawpaw DEM3

“Oh, (my) friend, YOU will climb that pawpaw (tree)!”

GABG.062

- b. *no =k rōn ē duru sa =s ul-ulō anē*
 1SG =TAM2 hear ? 3DL FOC? =TAM7 RED-call DEM1

“I can hear: It is the two who are calling!”

2.PALA.237

- c. *n vē vē -m sa e no*
 ART mother-2SG FOC? PERS.ART 1SG

‘I am your mum!’

1.TNU.023

In (36a), two characters argue who is going to climb the pawpaw tree to get a fruit; hence, in this passage, the addressee of the direct speech is in focus, and is construed as being in sharp contrast to the alternative, the speaker. Similarly, (36b) represents the direct speech of a father stating that those calling him and his wife from outside the house are their kids, and not an evil spirit who would have come to kill them. In (36c), a mother has gained eternal youth, but her daughter does not recognise her, so she says that that mother is really, rather than someone else.

In (36), *sa* occurs in postposition with the pre-VC subject pronouns where it, and in preposition with the post-VC predicative personal NP (headed by a pronoun). Where *sa* marks a post-VC object NP, it will also precede the phrase:

- (37) a. *di =m man viaklu sa =n nes di sa nē'ē*
 3SG =TAM1 affect through FOC? =ART song 3SG FOC? DEM1?

‘He put a spell on him by means of this song, that is it.’

ISWM.276

- b. *n 'ama man kur sa e ruwa ni-ni'iduō ē*
 ART spirit TAM5 devour FOC? PERS.ART two.people RED-child-3DL.IN DEM3

‘The spirit has devoured our two children.’

2.PALA.129

And the same is true for other post-verbal arguments or adjuncts, as in the following examples:

- (38) a. [...] *kal ba'a lē =n qoro liē di ne 'aq'a sa kēnē*
 enter into LOC =ART hole cave 3SG TAM2 hid FOC? there

‘[When the devil came to the village, that woman had already left, she had already hid into a cave.] Entered a cave, she went in hid right here.’

HHAK.036’

- b. *so wunva =n kakaka siviē =n biēg anē di =m van ma*
 ?? probably =ART story how =ART breadfruit DEM1 3SG=TAM1 go hither

di =m van sa senē
 3SG =TAM1 go FOC? like.this

‘Probably, the story how the breadfruit, it came here, it goes like this [as I have told it].’

MVB.194-196

As in these two examples, the particle *sa* often marks a locative or manner preform (‘*there*’, ‘like *this*’). A most typical case of the use of *sa* is with demonstratives in predicative function, as in the following example:

- (39) *gidē =k 'ēn ēn 'erē wōmō'ō'*
 1PL.IN = TAM2 see ART PL *fish.sp*
- wunva =n gengen go .. wōmō'ō' sa nē ()*
 probably =ART food POSS.EAT *fish.sp* FOC? DEM

[As for its food, when we gut this fish we see that]¹⁷ the *wōmō'ō'* eats probably only **those foods in the sand [or all sorts of soft things and sea grass**. When we go down to the sea,] we see all the *wōmō'ō'*, and the *wōmō'ō'*'s food is **that**.’

fish04_CH.014-018

All the examples of *sa* in (36)–(39) underscore the following point: *sa* is used with discourse-given, active (or otherwise accessible) referents here. This appears to be a general rule: I am not aware of a single example where *sa* marks an expression with inactive reference. Thus, if *sa* really is a kind of focus marker, it is restricted to ‘narrow’ focus, presumably marking ‘contrastive focus’, this type of marking signalling the existence of alternative referents that are opted out (cf. Krifka 2008).

But despite the fairly clear examples for contrastive focus in (36), this last point seems to be dubious in the light of examples like those in (37)–(39) above: in all the examples, the clauses in fact express repeated propositions. In (37a), it was already said that the song will bring the hero back to life. In (38a), it is stated in the preceding clauses that the woman goes in hide in the cave, and in (39), it was already made clear where this particular fish finds its food. In all these examples, the referent in question is already focussed in a preceding proposition (the same goes for the b examples). Obviously, it is not *sa* then that ‘produces’ the effect of contrastive focus. The particle *sa* seems to have a rather ‘enhancing’ effect on the already established focus relationship in a clause that essentially repeats the relevant proposition. It is thus expected to find *sa* co-occurring frequently with pro-forms that refer back to the “already-focussed” referent.

Now consider the following examples from a passage that represents the climax of a narrative that seems to suggest a more ‘emphatic’ function of *sa*:

- (40) a. *ei nike van ē wōl ē wak ēn mē'ēmē*
 INTERJ 2SG:TAM2 go ? DIR ? open ART door

¹⁷ cf. example x below for a full account of the beginning of this example

rēne ne kalraka ___ ne van mul anē ___ wak sa =n mē'ēmē
 woman TAM2 get.up TAM2 go DIR DEM1 open FOC? =ART door

‘[They call and call their mother and father, but nothing. The father says:] “Hey, go, open the door! [I can hear that it is the two who are calling there.”] And so the woman got up and finally opened that door! [Having opened the door, the two slip in.]’

2.PALA.236-239

Here, the action of ‘opening the door’ has also been repeatedly uttered in the preceding context, but the referent ‘door’ is very certainly not in a focus relation to the proposition ‘x opens the door’. The example is part of the continuation of what was presented in (36a) above: after a longer spat, the father decides that he door should be opened to let their children in, and so the mother does exactly that.

My preliminary analysis of *sa* in the light of these examples is the following: *sa* is some type of “emphatic” marker that entails some assertive (and concluding?) force, and by this token has some “truth-enforcing” effect. This means that a referent is of particular concern, that it is important to understand that it is involved (“It is really this thing, place, person, don’t you think otherwise!”), and that the assertion it is part of is the one that holds for now, it will not be altered anymore. Though admittedly quite vague, this description of the function of *sa* does seem to account for all the uses attested in the examples above. In examples where *sa* marks post-VC expressions, its assertive force may probably be regarded as extended over the entire predicate, for instance in (37): ‘it is really true that this song had this effect (to bring him back to life)’; ‘it is really true that our children were eaten (by a devil)’. The focus-marking effect with pre-VC subjects in (36) arises from this assertive effect as well: ‘(it is really) the two, (they call)’, ‘(it is really) you, will climb’. The ‘extended’ effect with post-VC *sa* and the clefting effect with pre-VC *sa* is probably to be attributed (at least partly) to the prosodic property of *sa* as being always pitch-accented, so that *sa*-marking has a partitioning effect on prosodic level.

Finally, consider the following examples where *sa* co-occurs with the demonstrative *anē* in pre-VC NPs or pronouns, subjects or left-dislocations:

- (41) a. *ei nīke anē sa =s dada kel ēn nak mu-k*
 INTERJ 2SG DEM1 FOC? =TAM6 RED:do back ART canoe POSS.GEN-1SG
- a =s 'ir'ir kal kel ē*
 REL =TAM6 RED:stand upwards back DEM3

‘[He came back, and he saw Spider going up and down the *nanara* tree {looking for the last piece of chipping). So it’s YOU who’s been putting up (the tree for my) my canoe again, standing upright again.’

JJQ.064

b. *ne vōwal si ne vōruō ne vōwal di =n biēg mō'*
 NUM.ART one or NUM.ART two NUM.ART one 3SG =ART breadfruit Mota

wo ne vōwal anē sa di =m 'ōg wal 'i kēnei
 and NUM.ART one DEM1 FOC? 3SG =TAM1 stay once LIM there

‘[But we have a (kind of) breadfruit here that has come here.] One or two, one is the breadfruit “Mota”, but that one is the one that belongs right here..’

MVB.007-008

c. *ne gēnē ga qag-qaga ba =n 'ew'i 'a=bēne*
 NUM.ART? here TAM3 RED-white but =ART other ASSOC.SPEC=PRO.OBL

di ga kōrkōr wo =n 'ew'i lik ga lu-luwō rekso =n
qagar ē
 3SG TAM3 black and =ART other more TAM3 RED:big like =ART fish.sp DEM3

nē'ē sa (eh) ga bigbig ()
 DEM FOC? TAM3 RED:eat

‘This one here is white, but that other kind is black/dark, and yet another one is quite big like a *qagar*. It is **this latter one** that is eaten.’

fish15_AS.042-043

d. *o bul wēwē wova'al ne vōwal sarēnē ga mine*
 INTERJ friend fruit pawpaw NUM.ART NUM-one up.there TAM3 ripe

ne'ē sa gēdu me gen ē
 DEM FOC 1DL FUT eat DEM3

‘Oh, my friend, a single pawpaw up there is ripe, and this one we will eat.’

GABG.051-

052

The *sa*-marked NPs’ referents seem to contrast in these examples with other (potential) participants: in (40a), the hero Qo’ identifies the Spider spirits as the one who erects his tree for his canoe ever again; in (40b), the particular breadfruit species contrasts with

another species; in (40c), one fish species contrasts with two other ones; and finally in (40d), the ripe pawpaw fruit seems to contrast with the other ones that are not ripe yet.

However, the focus domain in these clauses seems to be elsewhere: the clause in (40a) represents direct speech again, uttered by the hero Qo'. He already experienced that the tree is upright on its spot every morning he comes back to continue his work, but he doesn't know what's going on. It is here that he finds Spider trying to put the tree together again, hence he suddenly realizes how it happens. Thus, he finds out *how* the tree happens to stand upright again and *who* does it. Similarly in (40b), we also learn about that particular breadfruit where it can be found; (40c) informs about the edibility of that particular fish; and finally, (40d) the character being the addressee is informed that this pawpaw fruit will quench their hunger.

Considering *sa* an emphatic particle with assertive force may, however, account for these examples: the left-dislocated or subject expressions are marked as topic-like or anchors for CG expansion (as outlined in XX above) by means of *anē*, and the remainder of the clause is asserted *about* this topic. At the same time, however, their marking with *sa* has the clefting effect as described above, so that we get a second assertion about the involvement of their referents, with the remainder of the clause being “topical”. This interpretation in fact requires that all the information, including all referents, in the proposition be given and active, so that both parts of it can be understood as an assertion about the other. This also yields the critical criterion for falsification of the hypothesis: it is falsified when this type of double-marking occurs in clauses which introduce new information, including new referents.

4.2.3 Functions of left-dislocation

Left-dislocation seems to have two main functions: the first one that coincides with that of marking “canonical” sentence topics in the sense of Reinhart (1981), so that the proposition expressed in the following clause will be understood as being primarily ‘about’ this entity, so that the new information is stored as linked to this referent in the resulting altered CG. I will call this function “topicalisation” (in the narrow sense), differing from the wider function of “topic-establishment” outlined above. Various examples of this kind have already been mentioned and commented on above, hence this point does not require any general elaboration here. Just consider the following example illustrating a fairly typical context for this type of topicalisation with the referent bearing the role of a possessor in the clause:

(42) n vĕvĕruō ga garaqa anē =n movu-gi ga ... ga mōnō valakra
 ART mother-3DL TAM3 new DEM1 =ART fashion-3SG TAM3 TAM3 little incorrect

‘Their new mother had a slightly bad fashion.’

MVBW.031

Left-dislocation with a co-referent possessor has the function of expressing “HAVE possession”, in which a topical possessor is attributed the possession of an item, characteristic, etc.

The second function of left-dislocation is the already mentioned frame-setting function, discussed to some extent in 4.1.4. Finally, I mention the already discussed “double-marking” constructions, involving both marking with *anē* and with *sa*. These seem to be restricted to left-dislocated position, presumably in order to support the partitioning of the structure, and the corresponding proposition.

5 Concluding remarks

In the sections above, I have provided a still fairly rough outline of the structures involved in the expression of information structure in Vera’a. While the marking of information status through the choice of referential form follows to a large extent widely attested and explained patterns, the use of particular structures does not seem to be predictable by means of the accessibility of a referent in question. Specifically, the deployment of lexical form in subsequent mentions as well as the use of numeral phrases seems to be driven by consideration of the future salience of that referent in the *subsequent* discourse, rather than its status with regard to *prior* knowledge, including preceding discourse. Similarly, the employment of demonstratives has the function to establish a referent as the presupposed information to be elaborated upon in subsequent propositions. It is these observations that do not seem to quite fit into existing models of referential choice.

The employment of demonstratives as markers of presupposed information to be elaborated can have quasi- topic-marking functions in some contexts; and the use of the emphatic-assertive marker *sa* can likewise have the effect of focus-marking in certain contexts. However, these markers can surely not be described as “having” exactly these IS-marking functions.

Future research on the topic will have to substantiate some of the claims on the basis of more systematic corpus-investigation, but also (probably) on the basis of partly

experimental elicitations, in particular on the use of demonstratives in extra-phoric functions. A number of further structures and morphological devices are probably relevant for information packaging in Vera'a, and their analysis must be left for future research as well.

ABBREVIATIONS

A	transitive subject	NUM.ART	numeral article
ART	common article	P	transitive object
CPL	complementizer	PERS.ART	personal article
DRINK	drink (possession)	POSS	possessive classifier
EAT	eating (possession)	QUOT	quotative
FOC	focus	OBL	oblique
GEN	general (possession)	TOP	topic
HOUSE	housing (possession)	S	intransitive subject
INTERJ	interjection	SPEC	specific
LIM	delimitative (aktionsart)	VAL	valuable (possession)
NMLZ	nominalization	VES	vessel (possession)
NUM	numeral		

REFERENCES

- Ariel, Mira. 1990. Accessing noun-phrase antecedents. Londo & New York: Routledge.
- Ariel, Mira. 2000. The development of person agreement markers: from pronoun to High Accessibility Marker. In *Usage-based models of language* edited by Michael Barlow & Suzanne Kemmer, 197–260. Stanford: CSLI Publications.
- Biber, Douglas, Susan Conrad & Randi Reppen. 1998. *Corpus linguistics. Investigating language structure and use*. Cambridge: CUP.
- Biber, Douglas & Susan Conrad. 2009. *Register, genre, and style*. Cambridge: CUP.
- Chafe, Wallace. 1976. Givenness, contrastiveness, definiteness, subjects, topics, and point of view. In *Subject and topic* edited by Charles N. Li, 25-55. New York: Academic Press.
- Chafe, Wallace. 1994. *Discourse, consciousness, and time*. Chicago: The University of Chicago Press.
- Diessel, Holger. 1999. *Demonstratives. Form, function, and grammaticalization*. Amsterdam & Philadelphia: John Benjamins.
- Diessel, Holger. 2006. Demonstratives, joint attention, and the emergence of grammar. *Cognitive Linguistics* 17(4), 463-489.

- Genetti, Carol & Laura Crain. 2003. Beyond preferred argument structure: Sentences, pronouns and given referents in Nepali. In *Preferred Argument Structure: Grammar as Architecture for Function* edited by John Du Bois, Lorraine Kumpf, and William Ashby, 197-230. Amsterdam, Philadelphia: John Benjamins.
- Givón, Talmy. 1976. Topic, pronoun and grammatical agreement. In *Subject and topic* edited by Charles Li, 151-188. New York: Academic.
- Givón, Talmy. 1979. From discourse to syntax. Grammar as a processing strategy. In *Discourse and syntax* edited by Talmy Givón, 81-112. New York: Academic.
- Givón, Talmy. 1983. Topic continuity in discourse: an introduction. In *Topic continuity in discourse: a quantitative cross-language study* edited by Talmy Givon, 1-42., Amsterdam & Philadelphia: John Benjamins.
- Gundel, Jeanette & Nancy Hedberg & Ron Zacharski. 1993. Cognitive status and the form of referring expressions in discourse. *Language* 69(2), 274-307.
- Haig, Geoffrey & Stefan Schnell & Claudia Wegener. 2011. Comparing corpora from endangered languages. Explorations in language typology based on original texts. In: *Documenting endangered languages: achievements and perspectives* edited by Geoffrey Haig, Nicole Nau, Stefan Schnell and Claudia Wegener, 55-86. Berlin, Boston: Mouton de Gruyter.
- Haig, Geoffrey & Stefan Schnell. 2011. Annotations using GRAID (Grammatical Relations and Animacy in Discourse). Ms. available online at: https://www.academia.edu/1234896/Annotations_using_GRAID_Grammatical_Relations_and_Animacy_in_Discourse
- Haig, Geoffrey & Stefan Schnell. under review. The discourse basis of ergativity revisited. Ms. available at: https://www.academia.edu/3403768/The_discourse_basis_of_ergativity_revisited
- Himmelmann, Nikolaus. 1996. Demonstratives in narrative discourse: a taxonomy of universal uses. In *Studies in anaphora* edited by Barbara Fox, 205-254. Amsterdam & Philadelphia: John Benjamins.
- Kibrik, Andrej A. 2011. *Reference in discourse*. Oxford: Oxford University Press.
- Krifka, Manfred. 2008. Basic notions of information structure. *Acta Linguistica Hungarica* 55(3-4), 243-276.
- Lambrecht, Knud. 1994. *Information structure and sentence form. Topic, focus and mental representations of discourse referents*. Cambridge: CUP.
- Lichtenberk, Frank. 1997. Patterns of anaphora in To'aba'ita. In *Studies in anaphora* edited by Barbara Fox, 379-411. Amsterdam & Philadelphia: John Benjamins.

- Matic, Dejan & Daniel Wedgwood. 2013. The meanings of focus: the significance of an interpretation-based category in cross-linguistic analysis. *Journal of Linguistics* 49, 127-163.
- Ozerov, Pavel. 2014. Pragmatics and information structure apart. Differential Case Marking in Burmese. Paper presented at the workshop *Case and agreement: between grammar and information structure*. Hebrew University Jerusalem, 13 January 2014. Ms.
- Palmer, Bill. 2009. Clause order and information structure in Cheke Holo. *Oceanic Linguistics* 48(1), 213-249.
- Prince, Ellen F. 1981. Toward a taxonomy of given-new information. In: *Radical pragmatics* edited by Peter Cole, 223-256. New York: Academic Press.
- Reinhart, Tanya. 1981. Pragmatics and linguistics: an analysis of sentence topics. *Philosophica* 27(1), 53-94.
- Schneider, Cynthia. 2009. Information structure in Abma. *Oceanic Linguistics* 48.1, 1-35.
- Schnell, Stefan. 2011. A grammar of Vera'a. Kiel University.
- Schnell, Stefan. 2012. Three-participant constructions in Vera'a. *Linguistic Discovery* 10. (Special issue on Three-participant constructions xxx. In memoriam Anna Siewierska edited by Eva van Lier).
- Schnell, Stefan & Geoffrey Haig. submitted. Topicalisation and the grammaticalisation of object agreement? Available at :