Gender and person agreement in Cicipu discourse

A dissertation submitted to the School of Oriental and African Studies, University of London, in partial fulfilment of the requirements for the degree of Doctor of Philosophy

Stuart John McGill

September 2009

I, Stuart John McGill, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Abstract

The Cicipu language (Kainji, Benue-Congo) of northwest Nigeria has the kind of robust noun class system characteristic of Benue-Congo languages – GENDER agreement is found on a great many agreement targets inside and outside the noun phrase. For a number of these targets, gender agreement is in competition with a separate paradigm, that of PERSON agreement. The dissertation focuses on the distribution of this alternation with respect to subject prefixes, object enclitics, and pronouns, based on a corpus of 12,000 clauses of spoken language.

The alternation proves to be complex to describe, involving a constellation of lexical, phonological, morphosyntactic, semantic and discourse-pragmatic factors. In particular, both animacy and topicality are CONDITIONS (Corbett 2006) on agreement. While inanimate or animal participants normally trigger gender agreement, if they are topics then they may trigger person agreement. Likewise while human nouns typically trigger person agreement, this is not always the case, and gender agreement is more likely if the referent is of incidental importance to the discourse. Furthermore it is argued that this alternation is sensitive to *discourse* topic (e.g. Dooley 2007) rather than sentence topic (e.g. Lambrecht 1994).

Both gender and person subject prefixes are ambiguous agreement markers according to the typology of Bresnan and Mchombo (1987) and Siewierska (1999), since both can take part in grammatical or anaphoric agreement. Thus the Cicipu data supports Culy's (2000) contention that topicality is an independent dimension for the classification of agreement markers, rather than derivative of the grammatical vs. anaphoric agreement distinction, and leads us to re-evaluate the common assumption that dependent person markers (Siewierska 2004) cannot vary with respect to their discourse function.

Since Cicipu is otherwise undescribed, a major part of the dissertation consists of a phonological and grammatical sketch.

Acknowledgements

Huge thanks are due to Peter Austin, my supervisor. From the beginning he has been a source of good advice and encouragement, with an amazing knowledge both of theoretical linguistics and of data from throughout the world's languages. He has been patient with my lack of progress or understanding, without neglecting to administer a kick up the backside when needed. Perhaps above all I am grateful that he has been so enthusiastic about the Cicipu language – it is impossible for a PhD student to overstate the amount of encouragement that this attitude brings, when often there seems to be no-one else interested!

Thanks too to the other members of my committee, Friederike Lüpke and Lutz Marten. It has been great to have specialists on both West African and Bantu languages who were willing to read and comment on various papers or chapters. I was also able to make many improvements because of informal discussions with them.

I would also like to thank my Cicipu consultants, in particular Markus Mallam Yabani who threw himself into this work with enthusiasm. Where there is insight into the language in this thesis it is often as much due to him as to me. I am also grateful to Mohammed Musa, especially for arranging interviews with Akula elders and for his Herculean transcription efforts. My Cicipu language teachers Musa Danjuma Mai Unguwa and his brother Ibrahim generously sacrificed their farm work to help me. Others who assisted with language work include Amos Bako Timothy, Ayuba Sani, Ishiaku Ibrahim, Mohammed Mallam, Sani Ɗantanyi, Ishaya Audu, and Makawo Madugu. Many others from Galadima and the surrounding villages not only helped me to understand their language, but also treated me with kindness and offered friendship beyond the requirements of hospitality. Special thanks are due to Ɗanjuma Galadima, the Mai Unguwa of Bazama, and Baba Yabani Galadima – for allowing me to live with them in their village, and for their advocacy of the work I was doing.

While carrying out fieldwork in Nigeria I had the help of many people from

> Wan Kungwa tugwede vu ana yãawana tu a wumpa ulenji. Ana tiyãana kampa komisooni, naha ivɔ n katĩi kaavu iri yina tidamana íyyũu yi ka'albarka a ahula haYeesu. Tipata vu põ iri yina tiyãana yina yiluba daidai, naha ivɔ n katĩi kaavu ísungusanu. Ki ke lee!

> *Lord God we thank you that you've brought us to this day. Now we've prepared this story, may you yourself bless the things we've told in the name of Jesus. We ask you whatever we've done that's lacking, may you bring it to fulfilment. Amen!*

Conventions

Transcription

The transcription used for the Cicipu examples in this thesis is based on the recommendations that were made in the recently-proposed orthography statement[1], with additional tone-marking. The symbol **y** stands for the palatal approximant [j], **c** and **j** for the affricates [tʃ] and [dʒ] respectively, and the apostrophe **'** for the glottal stop [ʔ]. A double vowel **aa** indicates a long vowel [aː]. Tone and nasality are marked on the first letter only (**a̋a**) but apply to the full vowel ([a̋ː]) – contour tones are realised over both vowels together. The following accents and abbreviations are used for tone:

á	H	High
à	L	Low
â	HL	Falling
ǎ	LH	Rising
↓		Downstep occurs after this point

Downstep occurs both within and between words. In general only word-internal downstep is indicated, and this *has consequences for the interpretation of the tone on object NPs* (see §3.4.7 fn. 25).

Intonation contours are indicated in some examples, especially in chapter 8. The system used here is based on Chafe (1994:xiii):

,	a terminal contour which is not sentence-final
.	a sentence-final falling pitch
?	a yes-no question terminal contour
/	intonation unit boundary (could be any of the above three)
...	intonation unit continues beyond what was transcribed

Transcriptions enclosed in [square brackets] use standard IPA symbols. Phonetic transcriptions vary in detail depending on the distinctions in question. Unintelligible speech is indicated by xxx, hesitations by (.). The start of a paragraph is indicated by ¶.

Cicipu phonology involves a good deal of underspecification, and the symbols **C** and **V** indicate underspecified consonants and vowels in phonemic representations, in addition to their normal use in indicating syllable structure. **A** indicates an affix vowel whose quality is determined by rules of vowel harmony (§3.5). **N** stands for a nasal

---

1 http://www.cicipu.org/papers/2008-08-13_orthography_proposal_v02.pdf.

homorganic with the following consonant.

Three lines are provided for the majority of examples – the transcription, a morpheme-by-morpheme gloss, and then a free translation in English provided by the author. Occasionally morphophonemic processes obscure the boundaries between morphemes, in which case an extra morphophonemic line has been added on top:

(1)   ǹtîivì
        ǹ-tâa = vì
        1s-shoot\RLS=3s.PRO
        I shot him

[Cross-reference: easm001.005]

Each example is cross-referenced to its source, as shown above. See §1.4.3 for the conventions used. Example numbering restarts at the beginning of each chapter.

A note on the glossing of nouns is in order here. As will be seen later, nouns in Cicipu are comprised of a prefix plus root e.g. **kà-táarí** 'stone', **à-táarí** 'stones'. Throughout the thesis these prefixes will be indicated in the glosses viz. NC1-stone, NC2-stone. This has been done for the reader's convenience and is not intended as a statement concerning the psychological reality of the prefix-root distinction.

When context is supplied for an example, descriptions of the context are in normal font. Paraphrases of the sentences immediately before the example are given in *italics*.

(2)   [Context: Description of a hunting party. *I didn't see Bako in the grass. When he stood up suddenly...*]
        ǹ-tâa
        1s-shoot\RLS
        I shot

In a few English examples SMALL CAPS are used to indicate sentence accent. **Bold** text, whether in Cicipu or English, is **not** used for accent or emphasis. Instead it is intended to draw the reader's attention to the features of the example under discussion.

English paraphrases

The analysis in chapter 8 involves the consideration of a number of lengthy sections of text. Rather than clutter the chapter with pages of interlinearised text, long examples are paraphrased in English, with the relevant grammatical features marked up in bold. Interested readers are encouraged to view and listen to the original texts at http://www.cicipu.org/texts.html. The following key may be useful in the interpretation of these paraphrases:

3s.PRO	Independent personal pronoun
AG1-PRO	Independent gender pronoun (also AG2-PRO, and so on)
3s.OBJ	Person-marked object clitic
3s-came	Person subject prefix on verb
AG1-came	Gender subject prefix on verb
3s.POSS	Possessive pronoun

In-text examples

When Cicipu words are discussed in the main body of the text they are printed in bold followed by the English gloss in inverted commas e.g. **dukwa** 'go'. Words from Hausa and other languages are printed in *italics*. Hausa words are generally spelled according to the standard orthography i.e. unmarked for tone or vowel length.

Abbreviations

The following abbreviations are used (based on the Leipzig Glossing Rules, Bickel et al. 2004):

1	1st person	LW	loanword
2	2nd person	NC	noun class
3	3rd person	NEG	negative
AG	agreement	NMLZ	nominaliser
ANTIC	anti-causative	P	plural
APPL	applicative	PART	particle (unknown meaning)
ART	article	PFV	perfective
CAUS	causative	PLAC	pluractional
CNFCT	counterfactual	POSS	possessive
COP	copula	PRO	pronoun
FUT	future	PROH	prohibitive
GEN	genitive	REDUP	reduplicated
HAB	habitual	REL	relativiser
IMP	imperative	RES	resultative
IRR	irrealis	RLS	realis
K.O.	kind of	S	singular
LOC	locative	SEP	separative
LCVZR	locativiser	VENT	ventive

Cicipu does not have object agreement, so all agreement markers on verbs are subject agreement markers.

Table of Contents

14

Index of tables

Index of figures

Part 1 – Introduction

Chapter 1 – Introduction

1.1 Motivation for the study and research questions

Cicipu, a Benue-Congo language spoken in northwest Nigeria, has a robust noun class system of the kind found in many parts of sub-Saharan Africa, particularly the Bantu languages spoken in the south and east of the continent. Predicates, pronouns, and most kinds of noun modifier all agree in GENDER with their subjects, antecedents, or head nouns. The Cicipu examples below show two nouns from different genders followed by a demonstrative, relativiser, and a relative clause predicate, each of which agrees in gender with the head noun/subject.

(1) **ká**-llù   **ké**-llè   **ká**-nà   **kà**-yâa-nà
     NC1-hunger   AG1-that   AG1-REL   AG1-arrive\RLS-PFV
     that hunger which arrived

                                     [sayb001.546]

(2) **mà**-wáa   **mé**-llè   **má**-nà   **mà**-sí-hùnà
     NC4-dog   AG4-that   AG4-REL   AG4-HAB-kill
     that dog which kills

                                       [tats001.001.050]

Cicipu differs from the Bantu languages, however, in that several of these targets of gender agreement have the potential to agree in PERSON instead. Thus there is a series of competing paradigms across a number of different agreement 'targets' (Corbett 2006). This can be seen in the two examples below. The referent in both cases is the same, a camel (**kà-ràkúmí**, NC1), but in (3) the verb **dooho** 'disappear' takes a *gender* agreement prefix **ko-**, while in (4) it takes a *person* agreement prefix **u-**.

(3) **kò**-dóohò
     AG1-disappear\RLS
     it disappeared

                                       [tats001.002.033]

(4) **ù**-dóohò
     3s-disappear\RLS
     it disappeared

                                       [tats001.002.035]

Similarly the camel could be referred to by the gender-marked pronoun **k-é** or by the 3PS pronoun **éví**, and demonstratives, the interrogative quantifier **-èné** 'which', the copula,

and the article all offer this choice of agreement feature. The factors influencing the alternation are complex to unravel but of considerable theoretical interest.

This study has three main aims. The most specific is to give an account of the alternation between these two different paradigms of agreement morphology, one involving the feature GENDER and the other PERSON. This account is based on a qualitative analysis of texts from different genres, and identifies a range of relevant lexical, phonological, morphosyntactic, semantic and discourse-pragmatic factors. The second goal is wider in scope but necessarily narrower in depth, and that is to provide a comprehensive description of the Cicipu noun class system. The third goal is simply to expose the Cicipu language to the academic world by providing a sketch grammar. Naturally each of these goals build on the others, with the sketch grammar in Part II providing the base for the noun class description in Part III, which in turn informs the detailed investigation of the agreement alternation in Part IV. The rest of this section summarises the theoretical context relevant for this thesis and outlines the research questions that it attempts to address. The treatment here is brief, since the research context is set out in much more detail in chapter 2, in particular §2.2 on agreement and §2.3 on topicality.

Part IV of this thesis is largely concerned with subject markers on verbs. In the last two decades the nature of such markers has been hotly-debated in the linguistics literature. In particular a distinction has been made between GRAMMATICAL AGREEMENT, where the verb argument is expressed as a lexical NP in addition to the agreement marker e.g. **kà-ràkúmí kò-dóohò** 'the camel disappeared', and ANAPHORIC AGREEMENT, where there is no lexical NP, as in examples (3) and (4) above. This latter kind of agreement marker has been associated with topical referents, especially by Bresnan and Mchombo (1987).

However verb agreement markers can differ along other dimensions apart from the grammatical vs. anaphoric one. In chapter 2 we will see that Bresnan and Mchombo's typology of agreement markers makes no predictions about what might influence the alternation of the two sets of Cicipu agreement markers when they are both found in the same syntactic environment, as is the case for (3) and (4). This is the central analytical task with which the thesis is concerned.

To illustrate the sort of questions that will be addressed in Part IV, imagine a text

involving a camel as one of the participants. If the speaker wants to say that the camel disappeared, under what circumstances might we expect to find (3), with a subject prefix inflected for GENDER, and under what circumstances would (4) appear, inflected instead for PERSON? Is the distribution of the two kinds of markers dependent on topicality – and if so, what kind of topicality? Is the concept of 'discourse topic' of any use in describing the alternation? How important is the initial referring expression in determining future agreement markers? Is it possible to discern any kind of PROGRESSION involving the two different kinds of agreement, as the referent is tracked through the discourse?

Abstracting away from particular usage events, what sorts of referents (or what kind of nouns) are more likely to trigger gender agreement rather than person agreement, and vice versa? Is animacy relevant? Are nouns from certain genders more likely to trigger person agreement than other genders? (so if the camel was referred to using the Hausa loanword **dábbà** 'animal, NC8', would this make any difference?). And if so, is this a true gender effect or is it an underlying semantic factor which merely results in a correlation with gender?

As mentioned above, it turns out that lexical, phonological, morphosyntactic, semantic, and discourse-pragmatic factors are all relevant to the alternation between gender and person agreement, and it is therefore necessary to consider all of these areas of linguistics for a comprehensive description. This is true not only for subject agreement markers, but also for other agreement targets such as pronouns and the article.

Topic *does* prove to be important, but not the concept of 'sentence topic' to which generative theories of syntax are, by their nature, restricted. Instead, a coherent analysis of this aspect of Cicipu agreement requires a theory of topic that takes into account both the way in which topics integrate a text, and the intrinsic interest which they hold for the speaker.

1.1.1 Research questions

Part III and particularly Part IV will address the following research questions:
- What are the agreement targets in Cicipu?
- Which of these agreement targets inflect for more than one agreement feature paradigm?

o Are the syntactic environments in which the different paradigms occur mutually-exclusive?

o If not, what are the factors that influence the agreement paradigm when there is a choice?

o If topicality is one of these factors, what kind of topicality is it? Is the data better explained by theories of 'sentence topic' or of 'discourse topic'?

The rest of this chapter is structured as follows: §1.2 provides some sociolinguistic background, §1.3 reviews the anthropological and linguistic works which refer to Cicipu, and §1.4 describes the fieldsite and the data and methodology used for this research. Finally §1.5 gives an overview of the thesis.

1.2 Language situation

1.2.1 The language name

The *Ethnologue* (Gordon 2005) has an entry for Cicipu under 'Western Acipa', with the ISO 639-3 code `awc`. The name 'Western Acipa' is unfortunate for two reasons. First, it suggests the language is a close relative of the nearby language known as 'Eastern Acipa'. As we will see in §1.3, this is not the case. Secondly, 'Acipa' is not a word in any language. It is an Anglicised contraction of the Hausa ethonym *Acipawa*, or of the corresponding term for the language *Acipanci*. The autonyms are **Cípù** (one person, NC8), **À-cípù** (the people, NC2), and **Cì-cípù** (the language, NC6). Although the etymology of these terms is unknown, they have no negative connotations, and since so little has been written about 'Western Acipa' there seems little point in continuing to use a confusing and inaccurate name[1].

Incidentally, Cicipu was entered as 'Sagamuk' in the *Index of Nigerian Languages* (Crozier and Blench 1992) and its accompanying language map. The location on the map is correct but the name is not. Instead it seems to be the autonym for the *Eastern Acipa* language (Dettweiler and Dettweiler 2002:35).

1.2.2 Demographic and ethnographic notes

CAPRO (1995:190) gives an estimate of 20,000 Cicipu speakers. It is impossible to evaluate this figure with any degree of certainty, since there is no reliable census data,

1 A request to change the name to 'Cicipu' was accepted in January 2009 by the ISO 639-3 governing body, and the new name will trickle down into the *Ethnologue*, probably in 2013.

the birth rate is high, and several dialects are endangered. Gunn and Conant (1960:55) quoted a figure of 3,572 adults from a 1946 census, but again this is uncertain.

The language is spoken in a contiguous area divided between Sakaba Local Government Area in Kebbi State, and Kontagora Local Government Area in Niger State.

Figure 1: Location of Acipuland within Nigeria (map courtesy of University of Texas Libraries, http://www.lib.utexas.edu/maps/)

The area is remote and much of the region can be inaccessible by car in the rainy season – partly because of swollen rivers, and partly because there is no path through the corn wide enough. The map in Figure 2 shows some of the major routes (single-track dirt roads), villages, and approximate locations of the seven dialects and the surrounding languages. As Dettweiler and Dettweiler (2002) observed, most Acipu are farmers and do not live in towns – these are populated by Hausas (and also by people who are ethnically Acipu but no longer have Cicipu as their mother tongue). Instead the Acipu live in smaller villages and hamlets. Exogamous marriage is rare, and although Acipu villages are interspersed with those of the Avaɗi Kambari in particular, mixed-tribe villages do not seem to be common. A significant number of Acipu have migrated far into Niger State to the south-west, since farmland is plentiful there, and in the village I stayed in approximately a quarter of the young men had done this. It is not known to

what extent these migrants form a cohesive community, nor whether they continue to speak Cicipu in their new settlements.
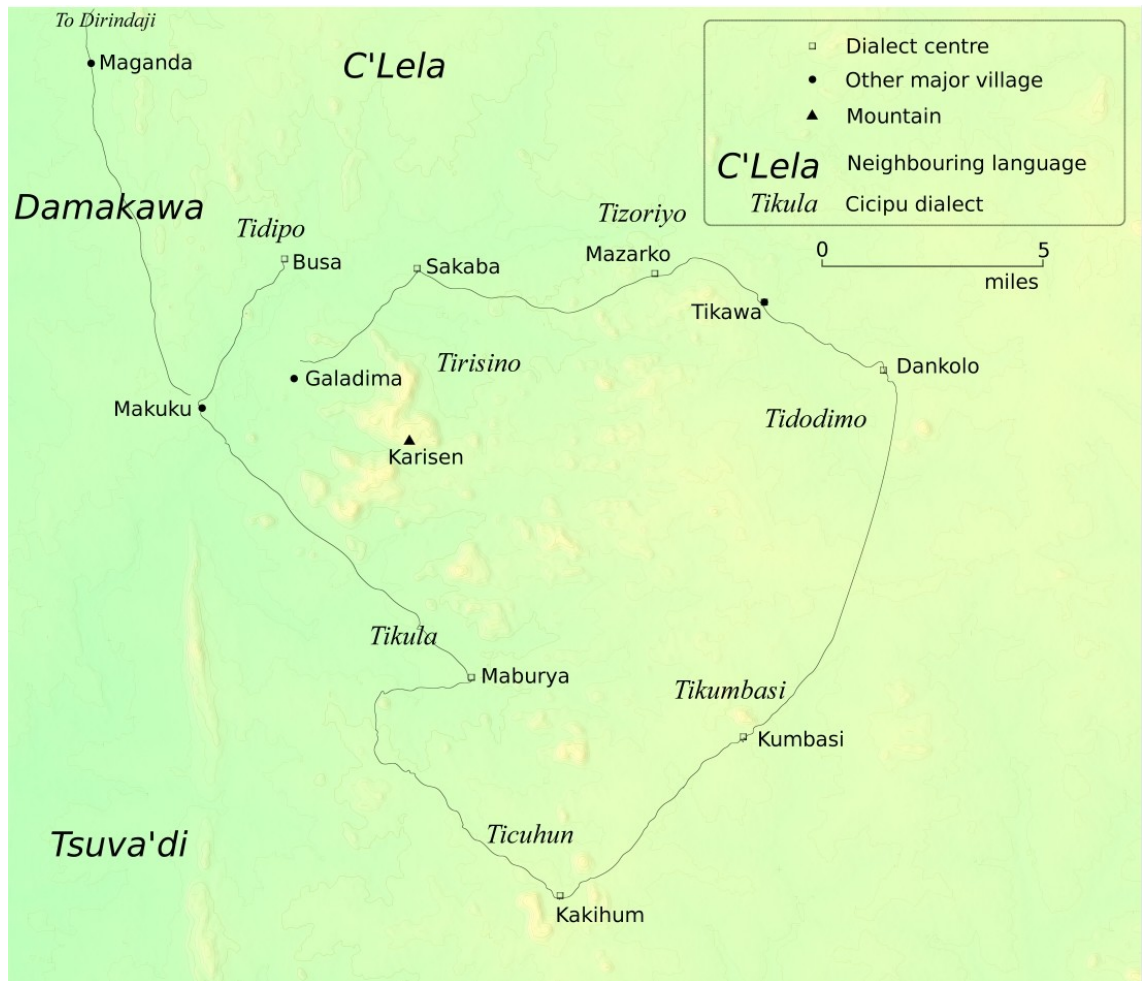
Figure 2: Location of Cicipu dialects and major towns

Kòrísìnôo hill (Hausa *Karisen*) is marked on the map in Figure 2. This is the seat of the **Wómó** 'chief' (also referred to as **D-dáa** 'king'), who lives on the highest point of the hill, accompanied only by his wife and his constant companion/advisor the **Má-llû** 'teacher' (Hausa *mallam)*. First and foremost the king is the head of the **Òrísìnôo** division of the Acipu, but he also seems to have precedence in certain matters over the leaders of the other divisions. Korisino used to be inhabited in living memory, although the old people I spoke to had moved to the valley when they were young. Eleven villages (including two **wú-uɓò** 'shrine') are still maintained on the hilltop, each corresponding to specific settlements/clans in the valley. Generally they are only used during festivals, although an old man may decide to return there to die, in which case his wife will accompany him, and family members will continue to visit and care for

28

them. The other divisions each have their own hilltop settlements, although they are not as large and well-maintained as Korisino. The only other hill I visited was Ù-kúlà, which belongs to the Àkúlà division and is located just north of Maburya. The Akula have almost all converted to Islam, and houses are no longer maintained on the hill. Only the elderly continue to carry out religious observances there.

Figure 3: Altar (rì-pízì) in Màtáarí village on Korisino

The Acipu are mostly peasant-agriculturalists. They grow guineacorn for subsistence, and sell any excess along with cash crops including beans, groundnuts, and soya beans. Farming is generally done by hand, although ploughing with oxen is becoming more common. The most common livestock are chicken and goats, and sheep are also kept. Some Acipu do own a few cows, which are generally tended by Fulani. Camels are a common sight around harvest time (December), when they are brought down from the north by Hausas in return for a share of the harvest. Horses and pigs are not kept, although the former seem to have been more common in the past. A small amount of rice is grown by women, and some dry-season gardening is done. Crops are grown in the fields next to and radiating out from the villages, and so livestock are tethered during the rainy season. There are occasional confrontations over land use with Fulani pastoralists, whose settlements are interspersed with those of the Acipu, but usually relations are peaceful. Some men enjoy hunting, which is done at night with locally-

made guns. Most of the big game has now disappeared, but monkeys are a fairly common sight in more remote areas and occasional sightings of leopards and hippos are reported. There are crocodiles on Korisino, but these are considered sacred and are not to be molested.

Traditionally huts are round, although these are now giving way to rectangular huts. Thatched roofs still seem to be the norm, but wealthier individuals are beginning to purchase zinc roofing sheets. Guest huts (Hausa *zaure*) are round, with opposing doorways to allow a cross-breeze. Compounds are fenced rather than walled, and can be entered without going through the *zaure*. Men's and women's granaries are still built using the traditional method described in Prazan (1977).

Figure 4: Traditional-style hut in Màgújí village on Korisino

The majority of the Acipu are Muslims, although above a certain age the traditional religion (**kò-rínnò**) is still dominant, particularly in the Tirisino dialect area. *Maigiro* worship (Temple 1922) is still observed (**ò-kíisó**, singular **k-kíisó**), as it is by seemingly every West Kainji people group. Each Acipu division has its own festivals held on their particular hill. At Korisino there are five major festivals, the most spectacular of which is the **Kà-zzémé** festival held towards the end of the dry season. As part of their initiation into adulthood, youths of about thirteen[2] endure beatings by a line of younger boys armed with spiked sticks. Some youths end up with heavily-lacerated backs,

---

2   And older youths and men, if they choose.

although serious injury seems to be rare.

Figure 5: Orisino children taking part in the Kàzzémé festival

Muslims also attend festivals on Korisino, although they do not take part in religious ceremonies. The influence of the Hausa culture long pre-dates actual conversion to Islam, as can be seen from the chieftancy titles listed in Mathews (1926b), and from the fact that the vast majority of people go by a Hausa name.

Christians make up a small minority, and most are relatively young. They will have nothing to do with the traditional religion, and this is a major source of conflict and persecution within families.

I was not able to elicit any kind of creation myth, but the story of the initial settlement of the Acipu hills is well-known. The outline of the story, recounted by Mathews (1926b), is that the first ruler of the Acipu, Damasa son of Damerudu son of the magician-king Kisra, fled from his original home far to the east after war with the Prophet. When they had travelled as far as the foot of Korisino, Damasa's wife was pregnant and they decided to settle on the hilltop. The 'true' Acipu (*talakawa*, Hausa for 'commoners') invited Damasa to become their king and Korisino was founded. Damasa's people are therefore the ancestors of the royal clan, the **Ò-dóndó**. The details of the story (particularly names and kinship relations) vary from dialect to dialect, both in Mathews' day and the present. Other nearby groups tell similar stories (e.g. Tsureshe, Agamalafiya 2008), and the names of Damerudu and Kisra also appear in Yoruba and

Bussawa folk histories (Mathews 1926b).

Figure 6: View from Makuku looking east, with Korisino hill on the left

In more recent times, the Acipu claim to have fought and defeated the notorious 19th century slave raider Nagwamatse of Kontagora. Nevertheless the impact of slavery on the Acipu was no doubt severe and it accounts for the fact that even within living memory they lived exclusively on the inaccessible hilltops. At the time of Mathews' (1926b) report hillside terraces were being abandoned in favour of farming in the valley below. As mentioned above the hills are now almost deserted.

For further ethnographic details see Gunn and Conant (1960) and CAPRO (1995).

1.2.3   Dialects

Dettweiler and Dettweiler (2002) identified four Acipu 'clans', although 'divisions' is a better term, since they each number hundreds or thousands of people and themselves consist of smaller clans. The names recorded by the Dettweilers were 'Arisene', 'Akumbasi', 'Akula', and 'Azakacihun' (lit. 'people of Kakihum'), corresponding to the first and the last three divisions in Figure 7. The Acipu I spoke to recognised seven sub-divisions within the ethnic group, which they said corresponded to different dialects. The names of the dialects are given in Table 1, while Figure 7 is an attempt to represent their internal structure[3].

3   Although these names do turn up in discussions about dialects, they do not seem to be as entrenched as the names for other languages such as **Tì-vádì** (Tsuvadì) or **Tì-hwǐí** (C'Lela). It is probably more

Table 1: Cicipu dialects/settlements and their Hausa names

Dialect autonym	Acipu settlement	Hausa name of settlement
Tì-rísìnôo	Kò-rísìnôo	*Karisen/Karishen/Karishin/Karissen*
Tì-dìpó	Kò-dìpó	*Kadonho*
Tì-zòoríyò	Kò-zòoríyò	*Mazarko*
Tì-dɔ́dìmɔ̂ɔ	Kɔ̀-dɔ́dìmɔ̂ɔ	*Kadedan*
Tì-kúmbásí	Ù-kúmbásí	*Kumbashi*
Tì-kúlà	Ù-kúlà	*Maburya*
Tì-cúhʉ̃u	Kwè-cúhʉ̃u	*Kakihum*
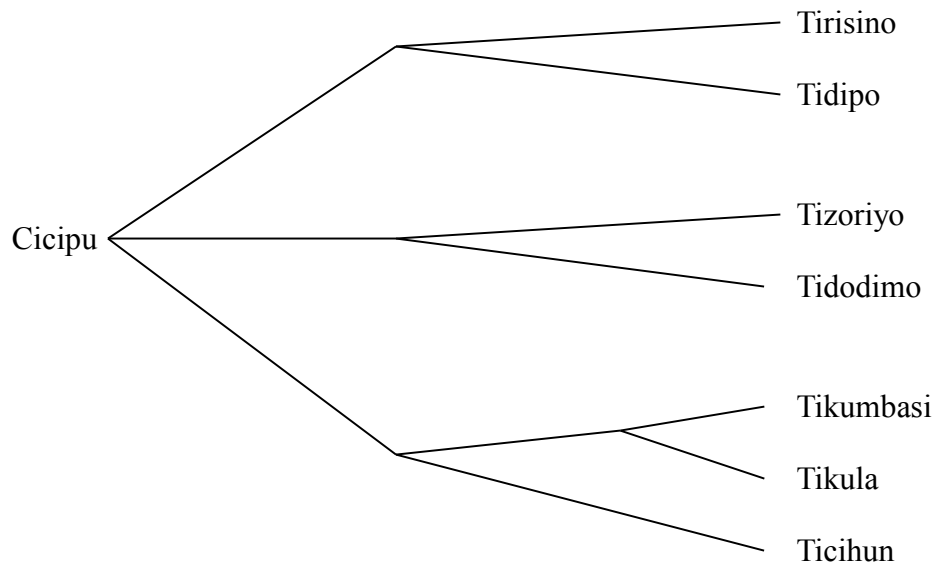
Figure 7: Reported structure of Cicipu dialects

The branching in this diagram is based on discussions with native speakers[4], but I have not yet been able to carefully check if these groupings are supported by linguistic evidence. I have first-hand knowledge of Tirisino, Tidipo, and Tikula, and certainly the first two are closer to each other than either is to Tikula. The most obvious difference between Tirisino and Tidipo is the loss of nasalisation in many words in the latter, e.g. [ʔû] vs. Tirisino [ʔʉ̃̂] 'there, far off', [pô] vs. Tirisino [pɔ̃̂] 'all'.

Support for at least the first and last groupings in Figure 7 can be found in the available wordlists. Ticuhũ is generally recognised as being the most divergent dialect,

---

common to use a phrase such as **kà-dámá kí-ivè** 'their speech', or e.g. **kà-dámá ká-aK̃úlà** 'speech of the Akula'. The word **kè-ré'è** 'tongue/language' *cannot* be used to refer to these dialects.

4 This information was first provided by Ibrahim Wasako, a Tidipo speaker, and independently corroborated by Dogo Timbidii, a Tikula speaker, and Markus Mallam Yabani and Musa Ɗanjuma Mai Unguwa, both Tirisino speakers.

and it can be seen from the data presented by Dettweiler and Dettweiler (2002) that there is a sound correspondence between **e** in Ticuhũ and **o** in the other dialects. For example, the greeting **póopò** 'hello' (*sannu* in Hausa) is pronounced [péːpè] in Ticuhũ, and **doonu** 'sit' is pronounced [deːnu]. In fact this correspondence is a stereotype of Ticuhũ speech recognised by other Acipu.

Notwithstanding these differences, the dialects are actually very close to each other, and there is no doubt that they belong to a single language. They are all mutually intelligible, and the lexicostatistical evidence points to the same conclusion – Tirisino has a lexical similarity of at least 95% when compared to Tikumbasi, Ticuhũ, Tidipo, Tizoriyo, and Tikula[5]. These dialects are clearly distinct from the closest non-Cicipu variety, Tsuvaɗi, which has <60% lexical similarity with Cicipu (McGill n.d.). Tirisino is the prestige dialect: even speakers from other dialects will admit that Tirisino is the 'best' form of Cicipu.

1.2.4 Multilingualism

Dettweiler and Dettweiler (2002:13) wrote that:

> The extent of Hausa bilingualism in the Western Acipa language community needs to be evaluated; limited bilingualism is reported, but this should not just be assumed as true.

In the nine months I spent in the Tirisino dialect area I did not find anyone who could not understand Hausa. The oldest women are at least able to understand it, even if they rarely speak it. It is my impression that the Acipu are somewhat less isolated and more 'exposed' than their Avaɗi neighbours, some of whom it appears really do not have much command of Hausa. Greetings amongst Acipu men, even within the same family, appear to be mainly in Hausa. Members of other tribes rarely speak Cicipu well, even the closely-related Avaɗi, and in mixed company Hausa is the norm. Even without the presence of outsiders, young men and boys often speak Hausa amongst themselves. Greetings amongst women seem to be in Cicipu, as do greetings between men and women.

In addition to Hausa, many Acipu speak one or more of the neighbouring West Kainji languages. From my observations in the Tirisino dialect area this is most likely to be Tsuvaɗi, followed by C'Lela. Fulfulde is not generally learnt apart from greetings and

---

5 Based on the 228-item wordlist collected by Dettweiler and Dettweiler (2002). No data exists for Tidodimo.

other very basic communication. In the course of nine months' fieldwork I encountered only two or three fluent English speakers amongst the Acipu. English is of course the medium of instruction in secondary school, which is attended by some of the Acipu from the village I stayed in. Unsurprisingly this is often to little purpose.

1.2.5 Endangerment

According to Blench (1998) there are roughly 250-400 languages in the Nigerian Middle Belt, depending on how broadly the area is conceived. Roughly half of these have less than 400 speakers. With approximately 20,000 speakers (§1.2.2), Cicipu is far in excess of the median. For Nigeria as a whole, Gordon (2005) ranks Cicipu as the 180th most populous out of the 510 living languages.

However numbers aren't everything, and even nearby languages with larger populations can be considered at risk[6], the main danger, for the larger languages like Cicipu at least, coming from shift to Hausa. The reality is that for Cicipu the level of endangerment is different for each dialect, and correlates with the degree of conversion to Islam. According to Blench (1998), there is a strong correlation in the Middle Belt between switch to Hausa and conversion to Islam, when this conversion has taken place in fairly recent times – which is the case for all of the Acipu.

The Orisino, who until recently have been almost wholly animist, have maintained their language, and children in every village I visited spoke Cicipu as their mother tongue. In contrast, the Akula have almost entirely converted to Islam, and Tikula is possibly the most endangered of the dialects, with only a few children in a couple of villages still learning Cicipu. The Tikumbasi dialect may be similarly endangered, and others likely fall between these two extremes. The influence of Christianity on language maintenance has yet to be seen, since the few converts are mostly recent. Despite the potential positive effect of Bible translation and associated literacy programmes, it should be noted that in Acipuland Hausa is the language of both the mosque *and* the church.

General factors that favour language maintenance include endogamous marriage practice and the high birth rate in the area – at least for those dialects where children are still learning the language. Speakers' attitudes to their own language are mostly positive, and initial runs of Cicipu storybooks and cassettes were received with enthusiasm. The

---

6   For example the West Kainji language Tsureshe (Blench 1998).

Acipu also seem to be proud of their culture; there is a saying amongst them that "there isn't a land in the entire world which hasn't heard of Korisino" – a belief no doubt sustained by the visits of curious Western scholars every few years.

There is no written literature in Cicipu, apart from a small number of recently-produced storybooks based on oral narratives. An orthography questionnaire/proposal has been developed based on the linguistic analysis carried out during this project, but this is still to be circulated.

1.3   Review of literature on West Kainji, Kambari, and Cicipu

Ever since the Acipu first came to the attention of scholars in the early twentieth century, they have been linked to both the Kambari people and the Kamuku people. However no linguistic evidence has ever been published to support a connection between Cicipu and any other language, and it seems that in previous language classifications cultural similarity and purported shared ethnic origin has outweighed the linguistic evidence. When the linguistic evidence is considered (McGill n.d.), it is clear that Cicipu should be regarded as part of the Kambari group.

In this section I will review the available linguistic literature on the West Kainji group as a whole (§1.3.1), on the Kambari cluster (§1.3.2), and on Cicipu (§1.3.3) – and for Cicipu I will also briefly mention the relevant anthropological sources. The geographical location of the languages concerned is given in Figure 8. To guard against confusion it is worth stressing that Cicipu (which, as mentioned in §1.2.1, has also been called 'Western Acipa') and Eastern Acipa are *not* close relatives, despite the names and the geographical proximity of their homelands.
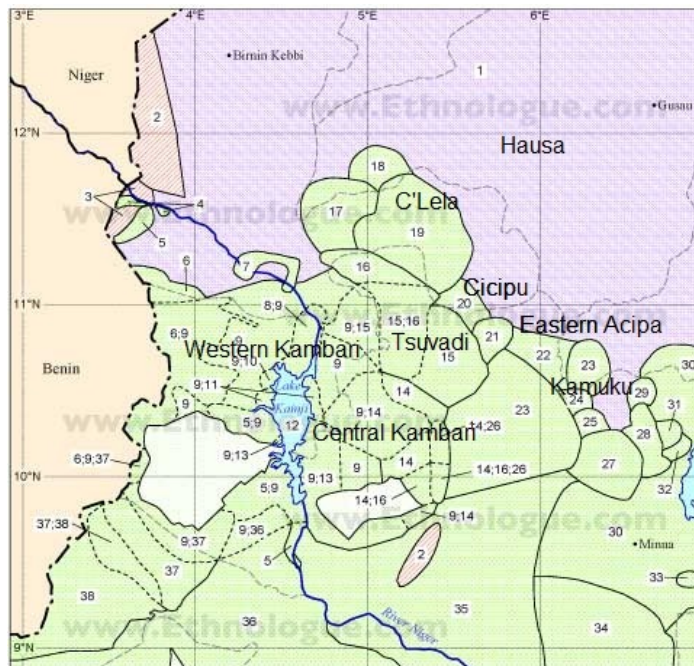
Figure 8: Cicipu and nearby languages (reproduced from Gordon 2005)

1.3.1 West Kainji

Blench (n.d. a) has observed that West Kainji languages are very diverse both in their lexicon and morphology. Consequently the genetic unity of the West Kainji languages was not recognised until Bertho (1952), who used 'Kambéri' as a cover term for the group. Greenberg's (1955) Plateau 1a grouped together what are known today as West Kainji languages. Plateau 1a and 1b (East Kainji) were then promoted to a separate Kainji group on the same level as Platoid by Gerhardt (1989).

Williamson and Blench (2000) contains the most recent published classification of Benue-Congo, based largely on isoglosses delimiting the various subgroups. It places Kainji in the Central Nigerian branch of East Benue-Congo. Blench's latest working classification is given in Figure 9:
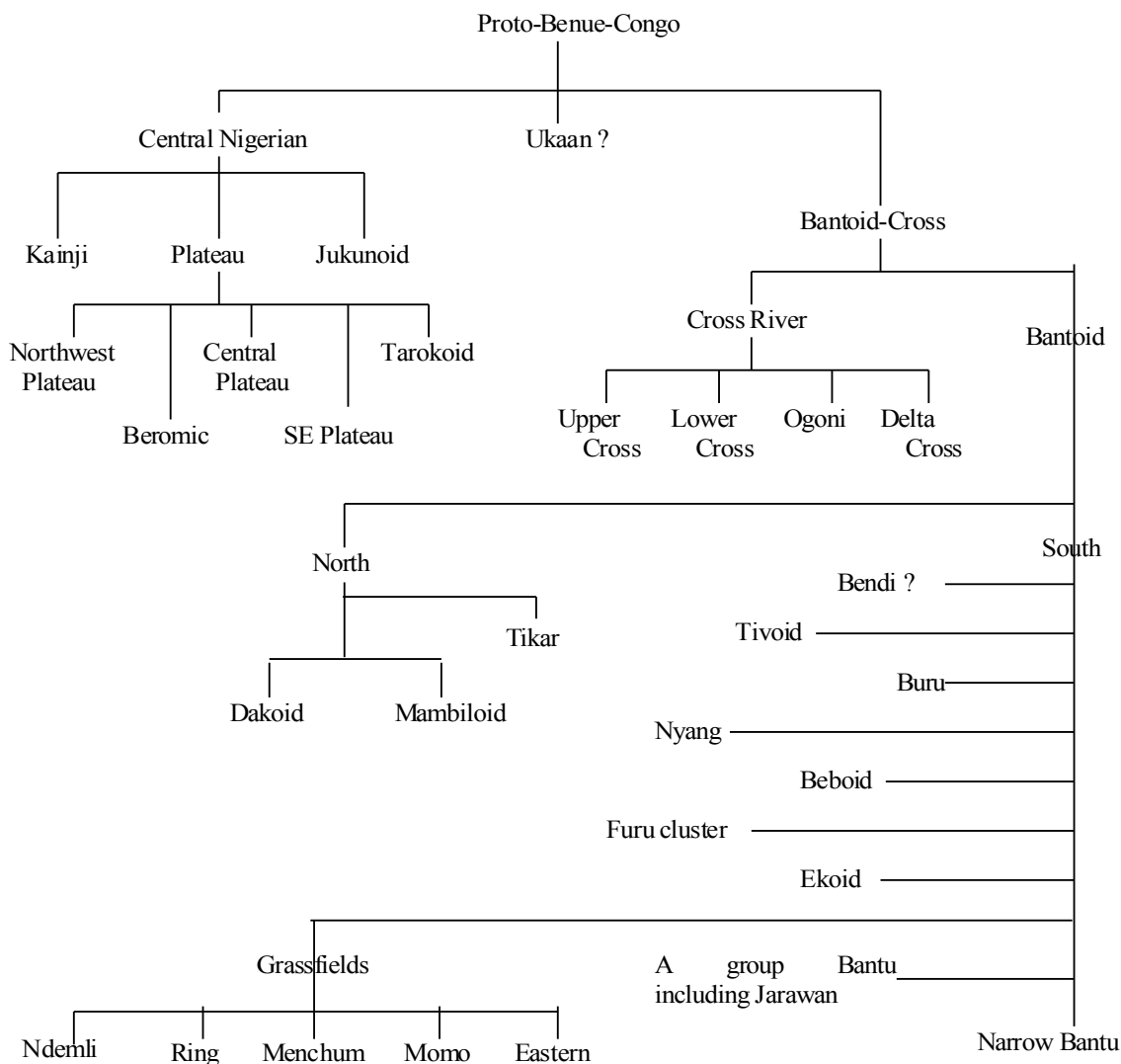
Figure 9: Classification of Benue-Congo, by kind permission of Roger Blench (May 2009)

There are twenty-five West Kainji languages according to Gordon (2005), although new languages are still coming to light (Spencer 2008, McGill 2008). The best known languages are the larger ones: Duka (or Hun-Saare), C'Lela (also known as Dakarkari), and Central Kambari, although the largest of all, Tsuvaɗi with 150,000 speakers (Gordon 2005), is yet to be the subject of any linguistic publication. Several are moribund (Gordon 2005, McGill 2008) and others still are endangered (Blench 1998).
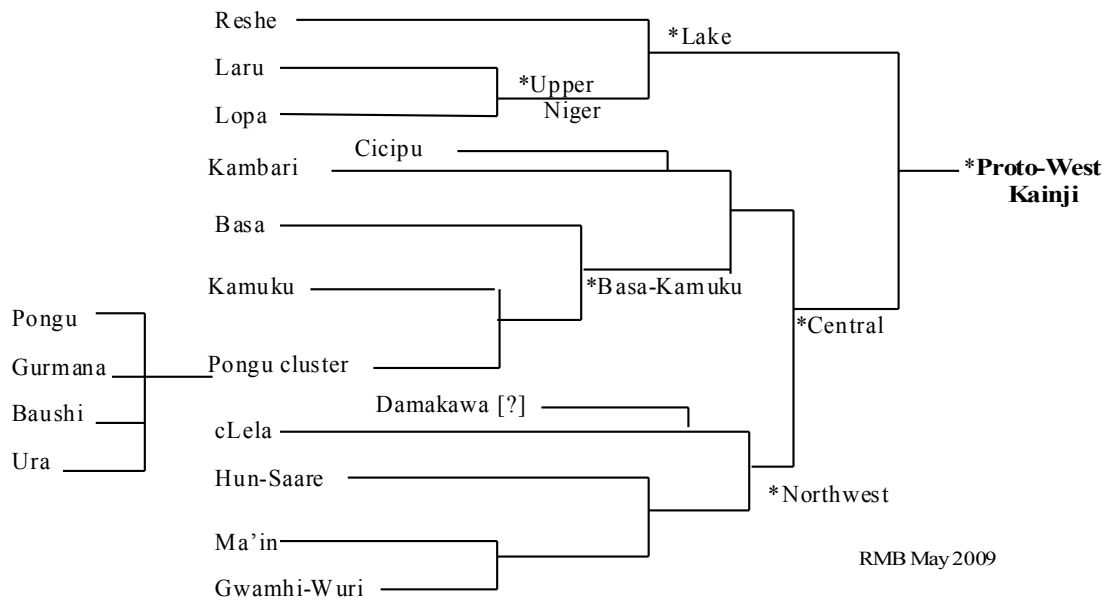
Figure 10: Subclassification of West Kainji, by kind permission of Roger Blench (May 2009)

Blench's latest working subclassification of West Kainji is reproduced in Figure 10, and is again based on lexical innovations delimiting the various subgroups (see Blench n.d. a). As mentioned above, McGill (n.d.) demonstrates that Cicipu should be regarded as belonging to the Kambari group, rather than Basa-Kamuku (as was the case in Gerhardt 1989, Williamson and Blench 2000, and Gordon 2005), and Figure 10 reflects this revised classification. The structure of the Kambari branch is presented below:
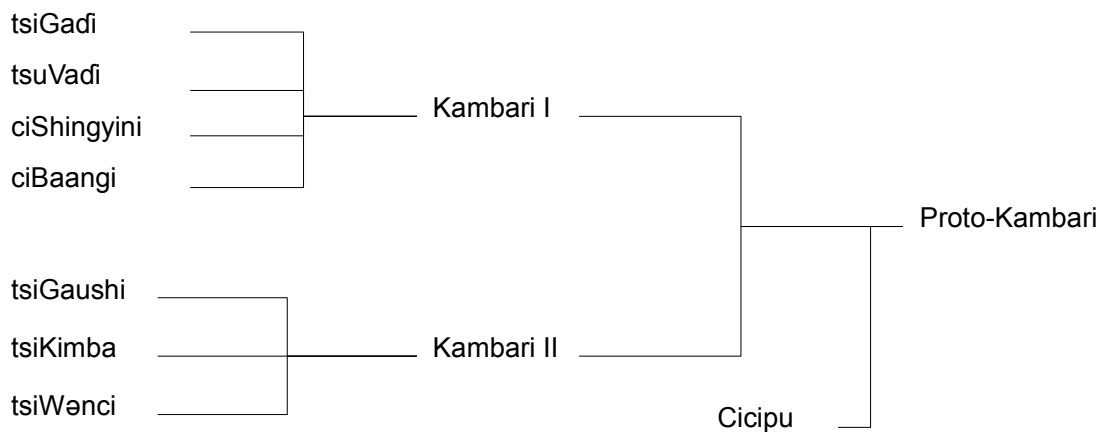
Figure 11: Revised subclassification of Kambari, Blench (n.d. a:5)

1.3.2   Kambari

Remarkably both Kambari and Kamuku feature in Koelle's (1963) *Polyglotta Africana*, first published in 1854. Kambari appears as Kámbāri (XII.E.9), and Dalby (1964)

identifies this with the modern term "Kambari". Hoffmann (1965) confirmed that the forms in Koelle's Kambari list were almost identical to those in the Central Kambari wordlist he collected.

Johnston (1919) briefly mentions 'Kambali'. He notes that it resembles the Basa-Kamuku languages Gurmana, Kamuku, and Basa, but that "apparently...[does] not make use of prefixes – at any rate for determining singular and plural" (1919:746). It is not clear what the source of Johnston's data was, and no known Kambari language lacks noun class prefixes today.

Kambari is also listed in Meek (1925 vol. 2:137) and Westermann and Bryan (1952), and Rowlands (1962) contains a 142-item wordlist and discusses its relationship with Kamuku, Duka and C'Lela. For Central Kambari Hoffmann provided a description of the noun class system (1963) and a 500-item wordlist (1965). On the basis of Hoffmann's work, De Wolf (1971) showed how the Central Kambari noun class system might be derived from his Proto-Benue-Congo reconstruction. The most important source for Kambari, however, is Crozier's (1984) unpublished PhD thesis on the discourse grammar of Central Kambari. This includes a detailed account of the language's noun phrase, verb phrase, and clause and sentence structure. More recently, trilingual Kambari-Hausa-English dictionaries have been published for three Kambari languages[7] – Tsikimba, Cishingini, and Tsishingini (Stark 2003, 2004a, 2004b)[8]. Unfortunately for the linguist, the dictionaries omit important phonological details and contain no Hausa or English finder lists.

1.3.3   Anthropological and linguistic literature on Cicipu

Anthropologists have historically paid more attention to the Acipu people than linguists. As well as general ethnographic works on the Nigerian Middle Belt, several articles have been written about the artefacts housed in the Wɔ́mɔ́'s palace at Korisino, said to have belonged to the legendary magician-king Kisra (Mathews 1960, Stevens 1975, Stewart 1980). As far as I know, the Acipu people were first mentioned in print by Duff

---

7   It is misleading to speak of Kambari as a single language, and the *Ethnologue* treats it as a cluster of six languages. In this thesis I will use 'Kambari' as a cover term when discussing the cluster as a whole (minus Cicipu), and use individual language names (e.g. Central Kambari) otherwise.

8   Confusingly, Stark (2004a) and Crozier (1984) use the term 'Cishingini' to refer to two different languages. Stark uses it to mean Western Kambari, while Crozier uses it to mean Central Kambari (which Stark has arbitrarily called 'Tsishingini' (2004b), apparently to avoid having two dictionaries with the same title). For clarity I will stick to Central Kambari and Western Kambari throughout this thesis.

(1972[1920])[9]. He comments on the *Maigirro* (ò-kíisó) religious practices (1972:66-67) and notes that the Kambari are descended from the Acipu and the Katsinawa (1972:30). The Dakarkari (Lelna) are also said to be descended from the Acipu, and Duff notes that the surrounding tribes (Zamfara, Bangawa, Dukawa, and Dakarkari) acknowledge the **Wómó** as the 'father' of the whole country, although they owe him no political allegiance. Temple (1922:30-31) includes a section on the Acipu ('Atsifawa'), apparently derived from the same source as Duff (1972). Mathews (1926b) is an unpublished 40-page "historical and anthropological report" on the Acipu ('Achifawa'), based on a two-month visit. He includes a few Cicipu words in the main text, and also in an appendix on linguistics. In two places he reports that the Acipu "deny any relationship with the Kamberri", presumably a response to elicitation based on Temple's report. Gunn and Conant (1960) contains a chapter on the Acipu ('Achipawa'), largely based on Mathews (1926b). Likewise CAPRO (1995) includes a similar section on the 'Achipawa', but also contains original research on the traditional religion and the distribution of Christians amongst the people.

As well as the Cicipu/Kambari link, early sources also mention a connection between the language referred to here as Cicipu, and a language identifiable as Eastern Acipa from the Kamuku subgroup. Temple (1922:30, 207-208) stated that 'Achipanchi' was also spoken to the east, in the Kamuku area. He writes that the two groups "know of no connection with each other", but at the same time records that the eastern group migrated there from the western area "a long time ago" – this is supported by the local traditions reported by Dettweiler and Dettweiler (2002:9), whose informants did recognise the link between the two peoples. Mathews (1926b) noted that

> There are also Achifawa in the central Makangara hills of the Niger Province [i.e. Eastern Acipa], whither they migrated from the parent body many years ago, and where Achifanchi is still said to be spoken

but this may just be Temple's remark rephrased. Gunn and Conant (1960:55) state that the Eastern 'Achipawa' (autonym 'Tochipo') "have denied any connection with the Achipawa of Kontagora [i.e. Western Acipa/Cicipu]", but again this may well be based on earlier reports.

Westermann and Bryan were the first linguists (as opposed to anthropologists) to mention the language spoken by the Acipu, commenting (1952:104) that

---

9   Duff refers to notes on the Acipu taken by C. E. Boyd, which I have been unable to locate.

It is not known what language or dialect is spoken by the *Achifawa* (*Achipawa*) among the *Dakarkari* and *Duka* [i.e. Cicipu]. They appear not to be the same as the *Achipawa* or *Achifanchi Kamuku* [i.e. Eastern Acipa].

It has already been mentioned that Bertho (1952) was the first to recognise the genetic unity of the West Kainji group. He divided the eleven West Kainji languages for which he had data into three subgroups. Based on 125-item wordlists, he placed Cicipu ('Atshefa-nshi') together with Central Kambari and Tsuvaɗi in one of three West Kainji subgroups. This grouping is significant, since although it differs from published classifications, it is supported by the evidence presented in McGill (n.d.). Bertho's other two 'Kambéri' subgroups correspond to the 'Lake' and 'Northwest' clusters of West Kainji, and he did not include Kamuku in his Kambéri group at all. Instead he assigned it to his 'Baoutchi' group, corresponding to what are now known as Plateau languages, based on evidence from noun class prefixes.

In fact there is a question mark as to the accuracy of Bertho's sources for West Kainji. Unfortunately he did not provide the actual language data on which he based his comparisons, and so it is impossible to know how accurate the material was. He made one specific comment about Atshefa-nshi: that around fifty percent of the nouns had lost their class prefixes. This is surprising, given that Cicipu has a robust noun class system. He made similar errors concerning Tsuvaɗi and Central Kambari (Hoffmann 1963:162)[10]. Nevertheless the one 'Atshefa-nshi' word he did give (*ka-teri* 'bone') is likely to be Cicipu (kè-tére in Tirisino), and there are no other obvious candidates for the language name 'Atshefa-nshi' in the Kambari cluster as it is currently understood. In any case, whatever errors there were in his data, he seems to have got the classification of Cicipu right. The next linguist to mention 'Acifanci' was Rowlands, who wrote that:

The Kamuku proper are divided into two sections…My informant called the other section [i.e. the section he didn't belong to – S.M.] by the Hausa name 'Acifawa' and their dialect 'Acifanci' (Rowlands 1962:80).

It is probable that Rowlands' informant was referring to Eastern rather than Western Acipa. Rowlands did not make this distinction, and this fact, together with the similar autonyms, may have influenced the Gerhardt/Blench classification of Western and Eastern Acipa together under Kamuku[11].

---

10 See §1.3.2 for the same error by Johnston (1919).
11 Roger Blench (p.c. 2008) has observed that cultural differences set the Kambari apart from the Acipu and Kamuku, and this may also have been a factor.

As mentioned in §1.2.1, Dettweiler and Dettweiler (2002) studied the relationship between 'Western' and 'Eastern' Acipa. They did this by collecting 228-item wordlists from three Western Acipa dialects and two Eastern Acipa dialects, and then performing a lexical comparison between them. They found that the lexical similarity between any two representatives of the two groups was always less than twenty percent, and proposed a revised classification of West Kainji with Western Acipa removed from Kamuku and allocated a branch of its own.

Finally, McGill (2007) provided an initial description of the Cicipu noun class system. The material in that paper is further developed in Part III.

1.4 Fieldwork setting and data

The analysis in this thesis is based on data collected during field trips to the Acipu area from September 2006 to March 2007, and from January to April 2008. The resulting annotated corpus consists of approximately six hours of transcribed, translated, and interlinearised audiovisual recordings, together with elicited recordings of seven hundred words from the SIL Comparative African Wordlist (Snider and Roberts 2004). The lexicon compiled from these sources amounts to around two thousand lexemes, and the corpus contains approximately twelve thousand clauses.

During these field trips I stayed in the Tirisino-speaking village of Galadima in Sakaba LGA, between Makuku and Korisino (see Figure 2). Galadima consists of ten family compounds, nine of which belong to Acipu. The remaining compound is inhabited by Reuben Acheson, a Nigerian missionary from Kaduna State, and his Hausa- and English-speaking family, and it was with them that I stayed. In total there are approximately 200 inhabitants, most of whom are children. All the Acipu in the village speak both Cicipu and Hausa, and none of them speak English.

During the first two months, elicitation sessions were conducted mainly in Hausa. By this time my Cicipu was no worse than my Hausa, and so we switched to the vernacular from then on. Outside of elicitation sessions I tried to communicate in Cicipu at all times, and living in a Cicipu village naturally allowed me to observe many things about the language and culture which would not have been possible otherwise. For example, with reference to this study, I noticed that speakers *did* use gender subject agreement in the absence of subject NPs, and that they *did* use person agreement with non-human controller referents, contrary to what I was repeatedly told in elicitation

43

sessions!

In general, I have tried to avoid using data from elicitation sessions when giving examples in the chapters that follow, preferring instead to draw on the corpus. However there are times when I have found elicited examples useful, either because they more clearly elucidate some point, or because the construction under discussion did not occur in the text corpus. My methodology is therefore "corpus-informed" rather than corpus-driven (McCarthy 2001:22), and only limited use is made of quantitative methods[12].

The examples are mainly taken from Tirisino speakers, although if a Tikula or Tidipo example makes the point better then I have not hesitated to use it. The use of such examples should not be taken to mean that the property or phenomenon being illustrated is absent from Tirisino – if that is the case, then it will be stated in the text.

The corpus contains data from several different genres as can be seen in Table 2.

Table 2: Structure of the interlinearised corpus used for this study

Text type	Genre	Duration in minutes
Observed	Conversations	5
	Sermons	24
	Songs[13]	240
Staged	Folktales	80
	Historical narratives/interviews	140
	Local event narratives	7
	Riddles	10
	Prayers	2
S**t**imulated	Topic-stimulation	82
	Fish Film	10
	Pear Film	20
Elicited		n/a

The text-types in Table 2 are based on Himmelmann's (1998) typology. He defines OBSERVED communicative events as "communicative events in which external interference is limited to the fact…that the ongoing event is being observed" (1998:27).

---

12 Frequency data is given for gender/person agreement progressions in §8.5, but no statistical analysis is attempted.

13 All the songs recorded are of the 'call-and-response' type common to sub-Saharan Africa. Often a single line is repeated for several minutes, and so the actual contribution of songs to the corpus in terms of clauses is minimal. Additionally the translation of songs is very difficult, since the meaning may be obscure even to the singers.

Naturally such events are difficult to record, and the linguist has no control over the linguistic data.

Staged communicative events, on the other hand, are "communicative events which are enacted for the purpose of recording". I have subdivided these into STAGED events proper and STIMULATED events. The difference between these two lies in the amount of control the linguist has over the proceedings. Communicative events of the former kind are less subject to the control of the linguist. Examples might be an evening of folktales organised by native speakers because they knew the linguist would want to record them, or simply someone asking the linguist to record them speak. In stimulated events, on the other hand, the speakers may be asked to perform a particular task by the linguist, for example to carry out a photo-matching experiment, to retell the story of a video they had just watched, or to simply talk about a referent of the linguist's choosing. The analyses and frequency data in Part IV are based on observed, staged, and stimulated data, but not on elicited sentences.

1.4.1 Data collection methodology

One of the major problems when investigating the effects of discourse factors on agreement is getting hold of the right kind of data for testing hypotheses. In particular, elicitation has not proved to be a very useful method of obtaining relevant data for examining the alternation between gender and person agreement, and it seems that the phenomenon in question is not amenable to analysis by native speaker introspection. Moreover, even in areas of information structure where relatively simple paradigms *can* be set up (e.g. for argument focus), elicited sentences may prove a bad guide to what actually happens in observed communication (e.g. Biber 1984, Bearth 1999). Consequently I have relied mainly on the analysis of texts rather than discoveries made in elicitation sessions.

As mentioned above, there is a middle ground between the collection of 'authentic' texts and traditional elicitation. A well-known example is the specially-made stimuli developed by the Max Planck Institute for Psycholinguistics (e.g. Danziger 1993). This technique allows the linguist to focus on a particular area of the language, for example asking questions about a set of subtly different photographs or models may illuminate how the language encodes spatial relations, ideally in a less artificial setting than a straightforward elicitation session. Most such stimuli seem to be designed for the study

of semantics rather than discourse, although the Pear Film (Chafe 1980) and Frog stories (Berman and Slobin 1994) are well-known exceptions (see also Skopeteas et al. 2006).

If the agreement alternation under discussion here was sensitive only to the information structure relevant to the form of a single sentence (in other words, if the alternation could be accounted for using Lambrecht's theory of sentence topic discussed in §2.3), then experiments could be designed to stimulate the relevant data. If, however, discourse topicality (§2.3) is involved, then the prospects are less good. The problem is that when working with an alien stimulus such as a film prepared by a foreign researcher, it is difficult to control, or even to predict with any degree of certainty which participants will be construed as more central at any one time. In fact, native speakers may well have no clear idea themselves, and this is bound to decrease the naturalness of the text.

There is also a more general problem with experiments involving stimuli that are intended to generate narrative, identified in Foley (2003). Foley used the Frog story book mentioned above with a speaker of Watam (Papua New Guinea), in order to generate an oral narrative based on the pictures and story therein. The resulting features of the text were markedly different to those of traditional Watam narratives, a distortion which Foley attributed to "the visual stimulus of the book and its illustrations [which] undoubtedly lead to a greater reflexive awareness of the participants and events of the story" (2003:95). This resulted in a normalised "literate text", despite its oral modality. Consequently much of what is interesting about the genre of narrative in Watam failed to surface in the experiment.

Despite my reservations, I did prepare some MPI-style photo-matching experiments since the data I collected in my first field trip only had a few examples of the gender/person alternation (apart from the ones I had bullied out of my language consultants). I designed an experiment based on the usual MPI photo-matching task (e.g. Danziger 1993), with two modifications. One was that the 'caller' could not see all the photos at once, but picked them one-by-one from a face-down pile. The other modification was that the caller was told beforehand that the 'matcher' had extra photos (this was not in fact true). These modifications were put in place to ensure that the callers described the scenes as carefully as possible, rather than taking shortcuts and

relying on logical deduction on the part of the matcher[14]. By using this experiment in conjunction with Chafe's (1980) Pear Film I hoped that I would stimulate examples of the alternation for referents of varying animacy (human, animal, and inanimate), and in sufficient number to discern a pattern.

The experiment used seven sets of twelve photos each. The sets fell into two basic categories – in some sets the 'figure' referent was kept constant and photographed against a variable 'ground', while in the other sets the figure was varied. Various other controls were implemented (e.g. some sets contained referents from the same noun class, others contained referents from a mix of noun classes). The caller in the experiment was instructed to carefully describe to the matcher how the objects are situated. Based on pre-fieldwork trials carried out on native English speakers, and on what I already knew about Cicipu, I expected the experiments to elicit constructions involving subject agreement prefixes – in particular verbs indicating possession, orientation, and location. It was thought that the sets with a constant referent might encourage the caller to use person ('topical') agreement, and that for the mixed sets the caller might be more likely to use gender ('non-topical') agreement – the logic being that an extended task involving only one figure would increase the likelihood of that referent being perceived as topical, no matter how inanimate it was. The instructions made it clear that it was the *same* referent (e.g. bird) in each of the twelve pictures in the constant-figure sets.

A catalogue of problems, some of which are discussed below, meant that the data was useless for what I had intended it for. To my mind, these problems argue against the use of these kind of stimuli in this field site, and perhaps others like it. Many of the problems were linked to subjects' unfamiliarity with the experimental media: physical photographs, and videos shown on a laptop. Most of the participants were under 30 and relatively 'exposed' – they would all have seen films before, and many of them had completed at least two years of primary school. However they are certainly not used to performing the tasks they were set in these experiments. People would hold the photos upside-down when describing them just as often as they held them upright, thus rendering useless oppositions that depended on orientation. No matter how many times I 'helped' by arranging the cards carefully in a 3x4 grid, after a couple more goes they

---

14 Without these modifications, the caller's attention, on viewing two similar photos, would be drawn to the *ground* (the part that differed) rather than (as desired) the *figure* (which in half the sets remained constant).

would be scattered around in a manner most unconducive to comparison. There were further problems with the identification of everyday local objects in the photographs.

Another unexpected problem was participants' inability to distinguish between mirror images. Several of the sets included two photos which were mirror-images of each other, for example a bird facing to the left, and the same photo with the bird facing to the right. When I trialled this in England, speakers had no problem differentiating the two and seeing the difference as significant with respect to the task they had been given. However Cicipu speakers either struggled to see that there was anything different about the two pictures, or even when they did realise, they struggled to communicate this to the other participant.

As well as unfamiliarity with the media, the fact that speakers are not used to taking part in experiments also causes problems. Despite having been given precise instructions, the participants did not understand that in many of the tasks each of the photos was supposed to show the same object in different situations. Instead of viewing the progression of photos as relating to a single referent increasing in discourse topicality, with each new photo they would say something like "And here's another bird". Consequently there was not even the opportunity for 'topical' person agreement to arise.

A lot of psycholinguistic experiments are repetitive and require the participant to perform the same basic tasks over and over again, something which proved more difficult than I had imagined, perhaps because sustaining this over several minutes seemed rather pointless to those taking part. In one case the participant gave an online higher-level *analysis* of the events, rather than concentrating on the task at hand!

The Pear Film experiments cannot be said to have been a complete failure with respect to my purpose in carrying them out, since the transcriptions do contain a few examples of both gender and person agreement with the same referent. However several incidents occurred which illustrate just how alien the task of retelling the events of a video is in the Cicipu culture. One of the speakers recounted that:

> When he fell, the things then fell. His friends came and helped him clear up, and they lifted him back on to the bike. Then he went and he fell again. Then they again gathered the fruit and they again put him on the bike [tapf002.003.016].

At first I found her account confusing, since there is no such repetition in the film. However when the film had been shown, the laptop crashed and part of the video had to

be replayed. That the speaker interpreted this as a continuation rather than a repetition of the action demonstrated a naivety with respect to video that I had not (although perhaps should have) anticipated. Another speaker began his account with the words "we saw a light", by which he meant the light being emitted from the laptop screen itself. The effect of these and other misunderstandings is limited in scope, but they do show that there can be problems even when the stimulus involved is one as carefully designed as the Pear Film[15].

It should be stressed that many of these issues arose when I tested the materials on English-speaking Nigerian graduates living in the Acipu area. It was fascinating to see them try to cope with the difference between mirror images, and just like the Acipu they ended up using the words *east* and *west* rather than *left* and *right*. Since they too held the photos upside-down they would have struggled whichever terms they used!

Such 'recalcitrant' behaviour could perhaps be countered by giving more training and detailed instructions before the start of the experiment, but there is a danger of imposing so many constraints that the discourse produced under these conditions becomes less and less like what the participants would actually say in everyday life, so that the experiment becomes a test of *what speakers can be made to say*. And of course if instructions are too complicated they may simply be ignored, as was the case for the photo-matching experiments.

1.4.1.1   Topic-stimulation sessions

Because of the failure of the stimuli experiments, I adopted a less sophisticated approach to investigating the gender/person alternation. Twenty-four texts were recorded, all of which are monologues, with occasional interjections by the interviewer. The speakers were asked to tell the interviewer everything that they knew about particular topics (some human, some animals, some inanimate), and the question was phrased with the topic noun in subject position e.g. "please tell me everything that tobacco does". The question was asked in Hausa to avoid making a choice between gender and person subject agreement[16], but the speaker was asked to reply in Cicipu. I was aiming for 2-3 minutes for each text – if the speaker stopped before then (which

---

15 See also §8.3.1 for differences with respect to the "light subject constraint" between these and more traditional narratives. Another difference is that the dependent imperfective aspect (§4.6.3.4) was far more frequent in the Pear Stories, perhaps because participants are giving careful attention to situations that they would normally gloss over (e.g. by reporting them with the perfective).

16 In Hausa the pre-verbal Person-Aspect-Complex is marked for both gender and person.

was rare, most texts are roughly that length, some are significantly longer), then they were prompted (this time in Cicipu) to say more.

It is normally recommended that in the case of a foreign director, any interaction with the subjects should be carried out by an appropriately-trained native speaker. Therefore I began by asking a language consultant to do the interviews, but this actually seemed to exacerbate rather than mitigate the observer paradox, and so later I resorted to conducting the interviews myself. I found people gave more comprehensive accounts to an outsider than when asked to give an account to someone who already knows everything that is going to be said. This is doubtless because they could identify a motivation for taking part in the experiment – to instruct the foreign researcher, someone already cast as a learner and well-known for his ignorance. A side benefit of trying to ensure that participants are motivated is that it helps to strengthen relationships between the researcher and the community. Speakers can have a hard enough time understanding why a fieldworker is present in a community without being given silly things to do.

The downside of this approach is obvious; the subjects are addressing themselves to a cultural outsider and a non-native speaker – the texts that arise are of a kind that would not normally occur outside of such sessions, since people are not, as far as I can tell, ever instructed in such a manner (while the Pear Film is also artificial, people do at least tell each other stories or recount something they saw happening, presumably in all cultures). I admit that this is a problem with my methodology, but I do not see how it could have been mitigated. The decision to carry out the experiments is, to my mind, vindicated by the extent to which the patterns that emerged in the experiments could also be found later, albeit more disparately, in texts which had been collected under less artificial circumstances. This will be demonstrated in chapter 8.

1.4.2   Corpus imbalance

The corpus used for this study regrettably suffers from an imbalance in two main areas. In terms of genre, conversation is under-represented. Although there are many examples of interaction between speakers, such as interviews, riddles, and folktales, and a smaller number of short 'true' conversations nested within non-conversational genres, I have no texts which are primarily conversational. Consequently the potential effect of turn transitions on gender/person agreement progressions has not been rigorously studied.

Perhaps a more serious failing is the under-representation of women's speech in the corpus. The gender/person alternation and the agreement progressions discussed in chapter 8 have been observed in female speech, but the overwhelming majority of them are in male speech. In particular it would be interesting to test whether the male tendency to refer to women with gender agreement and men with person agreement is replicated or reversed in the speech of women. The redress of these imbalances should be a high priority for future Cicipu documentation projects, regardless of any theoretical aims related to this study.

1.4.3   Cross-referencing conventions

The examples in this dissertation are cross-referenced to the source of the data. Most cross-references consist of a text identifier followed by an utterance number within that text. The text identifier contains the following components:

TEXT-TYPE − MODALITY − SPEAKER − TEXT-NUMBER.

So for example, the cross-reference `e-a-my-015.004` refers to the **fourth** (`004`) utterance in an **audio** (`a`) recording of an **elicitation** (`e`) session, the **fifteenth** (`015`) such recorded text with **Markus Yabani** (`my`) as the principal speaker. The values used in the first slot are **e** for **Elicitation**, **t** for **sTimulated** events, **s** for **Staged** events, and **o** for **Observed** events. The values used in the second slot are **a** for **audio** and **v** for **video**. If the example is in a dialect other than Tirisino then this is indicated along with the cross-reference.

The texts and accompanying metadata are in Toolbox[17] format and have been archived at the Endangered Languages Archive at the School of Oriental and African Studies. They are also available from www.cicipu.org.

Other cross-references are in the form yyyy-mm-dd.x, where the first part is the date and the *x* stands for the *x*th fieldnote made on that day. All such examples should be assumed to be the product of unrecorded elicitation sessions unless otherwise stated.

1.4.4   Contributors

The names of Acipu who contributed to the corpus used in this study are provided in Appendix B, along with their dialect, sex, and approximate age. Three people in particular were of great assistance in the painstaking task of transcribing and translating

---

17 http://www.sil.org/computIng/toolbox/

texts, as well as providing language consultancy during elicitation sessions. These were Markus Mallam Yabani and Musa Ɗanjuma Mai Unguwa, both Tirisino speakers living in Galadima village, near Korisino, and Mohammed Musa, an English-speaking Cipu from the Akula division at Maburya, who arranged, conducted, and transcribed two long interviews with members of the Akula chieftancy. Others who provided formal language consultancy were Tidipo speakers Ayuba Sani and Ishiaku Ibrahim, and Tirisino speaker Ibrahim Ɗanjuma Mai Unguwa.

Figure 12: Markus Mallam Yabani, Tirisino speaker
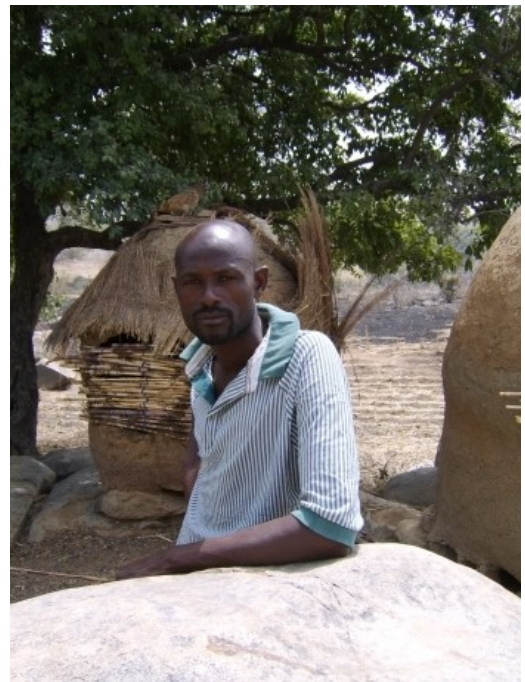
Figure 13: Musa Ɗanjuma Mai Unguwa, Tirisino speaker, and Mohammed Musa, Tikula speaker

1.5   Overview of the thesis

This thesis is made up of four parts. Part I consists of this introductory chapter together with chapter 2, which presents various theoretical preliminaries necessary for the later parts: noun class systems (§2.1), agreement (§2.2), and topicality (§2.3).

Part II (chps. 3-4) provides a sketch grammar of the Cicipu language; chapter 3 is concerned with phonology and chapter 4 with morphosyntax.

Part III (chps. 5-6) is a description of the Cicipu noun class system. Chapter 5 presents the noun classes and genders (i.e. pairings) and discusses their semantic structure and the derivational use of class prefixes. Chapter 6 is concerned with the multiple agreement targets found in the language, and includes sections on neutral agreement and the use of antecedentless agreement morphology.

Part IV (chps. 7-8) focuses on the research questions raised in §1.1 and seeks to uncover the various factors governing the alternation between gender and person agreement. Chapter 7 examines the phonological, morphological, and syntactic properties of the various pronouns and agreement markers, and identifies the agreement prerequisites which apply to each of the competing paradigms. Having secured this foundation, we are in a position in chapter 8 to investigate the distribution of gender and person agreement markers in discourse, in an attempt to determine the relevant agreement conditions for those syntactic environments which do allow a choice.

Finally, chapter 9 summarises the findings of the thesis and considers their theoretical implications.

Chapter 2 – Research context

Part III of this thesis is a description of the NOUN CLASS system of Cicipu. This description assumes certain notions and terminology from both the Africanist and typological traditions, and so these are introduced in §2.1 as a preliminary.

Parts III and IV are both concerned with the grammatical phenomenon of AGREEMENT, introduced in §2.2. Of particular concern are the domain of agreement, variation in agreement, and the typology of agreement markers. To anticipate the findings of chapter 8, we will see that the semantic and pragmatic notions usually employed in treatments of variations in agreement do not fully account for the Cicipu data, and instead we must also consider DISCOURSE TOPICALITY as opposed to SENTENCE TOPICALITY. Therefore this chapter will also introduce the notion of discourse topic (§2.3).

2.1 Noun classes and gender

It is not straightforward to define what is meant by the expression NOUN CLASS, in part due to the complexity of such systems, and in part due to the inconsistent terminology used by authors from different fields of linguistics. To get a feel for what is involved, it is perhaps best to begin with some practical examples, and so in §2.1.1 I will set out some prototypical properties of Benue-Congo noun class systems. Section 2.1.2 explains in more detail the concepts and terminology used in both typological and Africanist research.

2.1.1 Benue-Congo noun class systems

Noun class systems are found in the vast majority of Benue-Congo languages (De Wolf 1971:15), most famously in the Bantu languages of southern and eastern Africa. These systems classify nouns according to the different singular and plural affixes that they take, and usually require other constituents to agree with the noun. The exact definition of a 'noun class system' is problematic and is deferred to the next section, and so for the moment, I will build on the abstract characterisation I have just given by introducing and exemplifying some prototypical properties of Benue-Congo noun class systems.

Nouns in Benue-Congo languages usually occur with both singular and plural affixes, which are attached to the noun stem as in the following example from Cicipu:

(1)　　kà-bárá　　　　　　à-bárá
　　　　NC1-old_man　　　　NC2-old_man
　　　　old man　　　　　*old men*

The singular form of the noun for 'old man' is composed of the noun class prefix **kà-** followed by the noun root **bárá**, while the plural form of the noun is **à-bárá**, containing a different noun class prefix **à-**. This in itself simply shows that Cicipu, along with most other Benue-Congo languages, is typologically unusual in marking both singular and plural in its noun morphology. What makes it a noun class language is the fact that **kà-** and **à-** are just two out of a number of different prefixes, which can be paired in different ways. For example, the Cicipu noun root **díyá** 'hare' takes a singular prefix **mà-** and a plural prefix **ǹ-**.

(2)　　mà-díyá　　　　　　ǹ-díyá
　　　　NC4-hare　　　　　NC5-hare
　　　　hare　　　　　　*hares*

díyá is said to belong to a different noun class (in both its singular and its plural forms) from **bárá** because it takes a different pair of prefixes. Most Benue-Congo noun class affixes are prefixes, although suffixes and (rarely) infixes also occur. Conservative (i.e. unreduced) Benue-Congo systems can have as many as twenty different noun class affixes (e.g. seventeen in Chicheŵa – Mchombo 2004:6).

In addition to the system of noun class affixes, there is usually a corresponding system of agreement affixes which mark other constituents within the noun phrase. Prototypically both the head noun and the modifiers are marked. Agreement within the noun phrase occurs in all West Kainji languages studied to date, and Cicipu is no exception:

(3)　　kà-bárá　　ké-llè　　　　mà-díyá mé-llè
　　　　NC1-old_man AG1-that　　NC4-hare AG4-that
　　　　that old man　　　　　*that hare*

Here the head noun governs the form of the agreement prefix on the modifier.

Noun class agreement is not usually limited to the noun phrase. Subject (and sometimes object) noun phrases determine the agreement affix used on predicates, and anaphoric pronouns agree with their antecedent. All three types of agreement (NP-

internal, predicate, and anaphoric) can be seen in the following examples from the Bantu language Ejagham, which involve the same head noun in parallel structures: singular in (4) (with class 5 affixes) and plural in (5) (with class 8 affixes):

(4) è-yù   èj-â   j-ɔ́t   n̄j-í   n̄j-ì   ǹnàmé",
    NC5-yam   AG5-your   AG5-one   AG5-this AG5-which 1s/bought

    ń́j-ɔ́nè   é-bhìp
    AG5-it   AG5/PFV-bad

    This one yam of yours that I bought, it is spoiled.

                                              [Ejagham, Watters 2000:202]

(5) bì-yù   ìhb-â   í-bhá'é   m̄b-í   m̄b-ì   ǹnàmé",
    NC8-yam   AG8-your   AG8-two   AG8-this AG8-which 1s/bought

    ḿ́b-ɔ́nè   í-bhìp
    AG8-they   AG8/PFV-bad

    These two yams of yours that I bought, they are spoiled.

                                              [Ejagham, Watters 2000:202]

Within the front-shifted noun phrase, the head noun (*è-yù* 'yam' or *bì-yù* 'yams') triggers agreement on four other constituents: the possessive pronoun *â* 'your', the numerals *ɔ́t* 'one' or *bhá'é* 'two', the demonstrative *í* 'this', and the relative pronoun *ì* 'which'. In the main clause the anaphoric subject pronoun *ɔ́nè* '3rd person pronoun' agrees with its antecedent *yù*, while the verb is marked with the appropriate subject agreement prefix.

   In the Ejagham data just given, the noun prefixes and agreement prefixes differ significantly in their phonetic realisation, and there are a number of different sets of agreement prefixes (possessive, numeral, demonstrative, subject agreement marker). This is not always so for Benue-Congo languages, and in fact the textbook case is for all affixes to be the same or similar in form, a phenomenon known as ALLITERATIVE AGREEMENT – illustrated in this old chestnut from Kiswahili:

(6) ki-kapu   ki-kubwa   ki-moja   ki-lianguka
    NC7-basket AG7-large   AG7-one   AG7-fell
    one large basket fell

                                              [Kiswahili, Welmers 1973:171]

As well as showing that noun class agreement operates both inside and outside the noun phrase, these examples show how varied the different hosts of agreement affixes can be. Chapter 6 provides a structured account of these hosts, or agreement TARGETS, for Cicipu.

2.1.2 Concepts and terminology

2.1.2.1 Gender and noun class languages

For the overview just given a rigorous definition of terms was not necessary. However partly due to the complexity inherent in many noun class systems, and partly due to the different traditions of various groups of linguists, there is a danger of confusion if terms are not made explicit. Therefore in this section I will introduce and define a number of terms relevant for the study of noun class systems.

First of all, the obvious questions are: what qualifies as a 'noun class language'? And how is one different from a 'gender language'? Corbett's work on grammatical gender has been highly influential in recent decades, and he follows Hockett's (1958:231) definition of gender: "Genders are classes of nouns reflected in the behavior of associated words". Concerning the difference between gender and noun class Corbett (1991:146) writes that

> there is little point in maintaining a strict separation between 'gender' and 'noun class' since similar systems are described as genders in one family and as noun classes in another.

Indo-European 'gender systems' are based on what is sometimes called 'natural gender' i.e. masculine, feminine, and neuter. In contrast, Benue-Congo 'noun class systems' generally have a larger number of distinctions, and the primary semantic feature involved is animacy, rather than natural gender.

Aikhenvald's (2000) typology of noun categorisation devices contrasts noun classifiers with noun class systems. Although both systems categorise nouns, in a noun class system "some constituent outside the noun itself must agree in noun class with a noun" (2000:20). Thus by insisting on the presence of agreement she too equates 'noun class' with Corbett's 'gender', although she continues to write of Indo-European 'gender languages' and African 'noun class languages' as a concession to the respective traditions.

Most linguists working on Benue-Congo languages do not however take the terms 'gender language' and 'noun class language' to mean the same thing. De Wolf writes (1971:35):

> In order to have a true noun class system it is necessary that all nouns in a given language be MARKED for the category of gender by means of prefixes, infixes, suffixes or some of these combined, such that this gender category be a selective

one and, *regardless* of whether there exist types of class concord or class agreement in modifiers and/or substitutes such that these modifiers and/or substitutes be marked for gender with the result that gender constitutes an obligatory, inflectional category for these substitutes and modifiers. [my italics – S.M.]

While Corbett and Aikhenvald have rightly emphasised the similarities between Indo-European gender languages and African noun class languages, there is nonetheless an important difference in the way that nouns themselves are marked in these systems. Corbett (1991:62) uses the term OVERT GENDER to refer to gender which is evident from the form of the nouns themselves. It appears that all Benue-Congo languages have noun prefixes indicating their noun class, as in (1) and (2) above, and so all Benue-Congo gender languages display overt gender. If, on the other hand, the gender is not shown by the form of the noun, the language is said to have COVERT GENDER. German is such a language – there is no way of knowing from the form of the noun *Hand* 'hand' whether it is masculine, feminine, or neuter[1]. In Benue-Congo languages, 'noun class belonging' can therefore be realised both through noun affixes and through agreement affixes. These may or may not exist in a one-to-one correspondence and may or may not share the same phonetic forms, and as a result the typology of Benue-Congo noun class systems is fairly complex (see Maho 1999:127-142). Cicipu, however, has straightforward overt gender (see Part III).

2.1.2.2 Africanist terminology

The first large-scale comparative study on Benue-Congo noun classes was carried out by Wilhelm Bleek (1862, 1869). He devised a numbering system for Bantu noun classes, which was later revised by another German philologist Carl Meinhof (1932 [1899]), and became known as the Bleek-Meinhof (BM) numbering system. In this system each noun class affix (and the associated set of agreement markers) is given a number, originally ranging from 1-18. So each noun belongs to *two* noun classes (with certain exceptions such as mass and abstract nouns), one for the singular affix and one for the plural affix. To capture the overall behaviour of a noun, Africanists usually refer to the singular/plural pairings as genders[2]. So given the Cicipu example below, repeated from (1), an Africanist might speak of the KA noun class or the A noun class, but the KA-/

---

1 Zubin and Köpcke (1986) have shown that gender assignment in German is less random than was once thought. Nevertheless, for many German nouns it is not possible to predict the gender from the phonological form.

2 Confusingly, some authors refer to a noun class on its own as a gender. This will be avoided here.

A- gender.

(7) kà-bárá   à-bárá
    NC1-old_man   NC2-old_man
    old man   *old men*

Terminological confusion resulting from the two traditions may arise if a language has singular/plural noun prefixes but no agreement. Thus De Wolf writes that the Plateau language Rukuba has 'seven two-class genders' (i.e. singular/plural pairings) (1971:109), at the same time as stating that the language, as far as is known, 'is without concord' (1971:108). This usage of 'gender' is in conflict with mainstream linguistics. It would perhaps be better to use the phrase NOUN CLASS PAIRING for languages such as Rukuba, and reserve the term 'gender' for languages where an agreement relation exists. In any case it will be clear from what follows that Cicipu is unambiguously a 'gender language' (since it has agreement) as well as a 'noun class language' (since it has overt noun class prefixes). The phrase CLASS EXPONENT (De Wolf 1971) is a convenient umbrella term for any kind of class marker (noun or agreement affix) which signals gender and number. Class exponents may be null; this is the case for Cicipu class 8 (§5.5.7, §6.1.3).

A SINGULAR CLASS is a noun class whose exponents attach to, or agree with, nouns which are singular in number. KA[3] in example (7) above is a singular class of Cicipu. A PLURAL CLASS is a noun class whose exponents attach to, or agree with, nouns which are plural in number. A is therefore a plural class of Cicipu. A SINGLE CLASS is a noun class with members which do not participate in the singular/plural alternation of count nouns. Typically mass, liquid, and abstract nouns are found in single classes in Benue-Congo languages. There is obviously potential for confusion with the term 'singular class', and ideally 'unpaired class' would be a better term. However 'single class' seems to be the most widely-used term, and so I will use it here.

2.1.2.3   Controller and target genders

Corbett makes an important distinction between CONTROLLER GENDERS and TARGET GENDERS (1991:150-160). Target genders are the genders which are marked on particular agreement TARGETS (e.g. adjectives, verbs, demonstratives), while controller genders are

---

3   For non-Bantu languages it is often inappropriate to use the BM numbering system, and so noun classes of individual languages may sometimes be referred to using SMALL CAPS. When the Cicipu system is dealt with in detail a different (non-BM) numbering system will be introduced.

the genders into which nouns are divided according to their composite behaviour (similar to the Africanist use of 'gender' to mean a noun class pairing). The difference between controller and target genders can perhaps best be illustrated using Corbett's own example of Rumanian. Consider examples (8-10) from Corbett (1991:150-151):

(8) (a) barbatul e bun-**Ø**   (b) barbatii sînt bun-**i**
         man.the   is good   men.the   are   good
         the man is good   *the men are good*

                                                        [MASCULINE]

(9) (a) fata   e bun-**a**   (b) fetele   sînt bun-**e**
         girl.the is good   girls.the are   good
         the girl is good   *the girls are good*

                                                         [FEMININE]

(10) (a) scaunul e bun-**Ø**   (b) scaunele sînt bun-**e**
         chair.the is good   chairs.the are   good
         the chair is good   *the chairs are good*

                                                          [NEUTER]

Rumanian adjectives take one of two different suffixes (-**Ø** or -**a**) when agreeing with singular nouns, and one of two different suffixes (-**i** or -**e**) when agreeing with plural nouns. The problem is caused by neuter nouns like *scaun* 'chair' which have no unique agreement forms. They share the masculine form for singular nouns, and they share the feminine form for plural nouns. The situation can be summarised using an affix net[4]:

singular   plural

Ø ——————— m. ——————— i
                   n.
            f.
a ——————————————— e
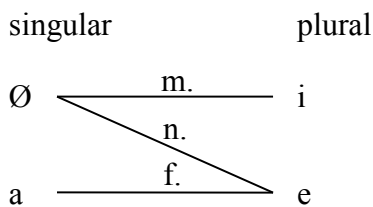
Figure 14: The gender system of Rumanian (taken from Corbett 1991:152)

Corbett's terms allow us to capture the fact that although there are three separate ways of dividing up nouns (masculine, feminine, and neuter), nevertheless the morphology is simpler than the case of Latin say, which has three independent genders. We can say that Rumanian has two target genders in the singular and two in the plural, but three controller genders.

---

4 Affix nets can be used to represent either a system of nominal affixes, or as in this case a system of agreement affixes. The lines represent the possible singular/plural pairings.

Maho (1999:142-145) provides a helpful discussion on the applicability of Corbett's methods to Bantu noun class systems. The main thrust of his argument is that neither controller gender nor target gender corresponds to an equivalent Bantuist term. At first glance it seems that individual noun classes correspond to target genders, and that Bantuist genders, or noun class pairings, correspond to controller genders. However this is only true in the case of a straightforward one-to-one correspondence between nominal affixes and sets of agreement markers. It is often the case that two or more noun classes with distinct noun affixes share a set of agreement affixes, as do the Central Kambari classes 2a and 2b (Hoffmann's 1963 numbering). Class 2a has a prefix consisting of a vowel **V-** (the exact value of **V** is determined by vowel harmony rules), with identical agreement prefixes. Class 2b shares the same agreement markers, but has a different nominal prefix **náN-** (where **N** is a nasal homorganic with the following consonant). For example **á-násá:rá** 'Europeans' (2a) has a different prefix from **nám-básárá** 'hawk' (2b), but they both trigger the same set of agreement markers. In Central Kambari these two subclasses enter into distinct noun class pairings, which is an extra reason to treat them separately. The following affix nets show that we are not just dealing with a terminological difference, since the two approaches result in conflicting number of genders.
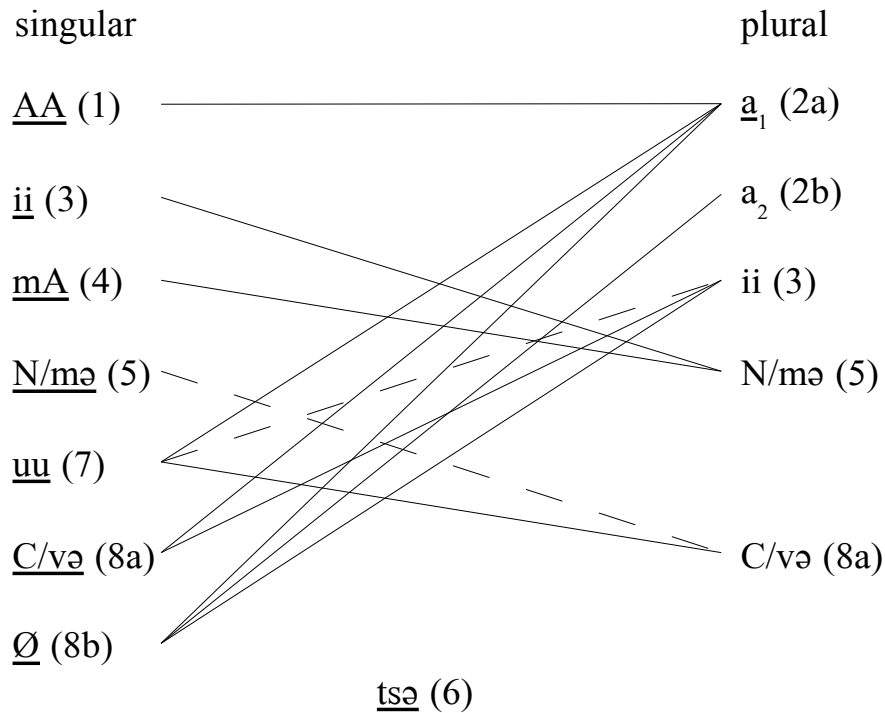
Figure 15: Kambari noun class system according to Africanist tradition (based on De Wolf 1971:60). Hoffmann's (1963) numbering is given in brackets

In Figure 15 it can be seen that there are ten noun classes and twenty-two genders (thirteen singular/plural pairings and nine single class genders[5] – the latter represented with underscores). The important thing to note is that the plural classes 2a and 2b take the same agreement markers, but their prefixes are different and so they are treated as separate noun classes. The same is true for the singular noun classes 8a and 8b. Now consider the same system diagrammed according to Corbett's approach:

---

5   Observe that classes 3, 5, and 8a can occur with either singular or plural nouns. In addition most Central Kambari noun classes also occur as single classes. Dotted lines indicate an inquorate gender (Corbett 1991:170-175) with very few members.

singular   plural

AA (1) ———————————————— a (2)

ii (3)   ii (3)

mA (4) ———————————————— N/mə (5)

N/mə (5)

uu (7)

və (8)   və (8)

tsə (6)
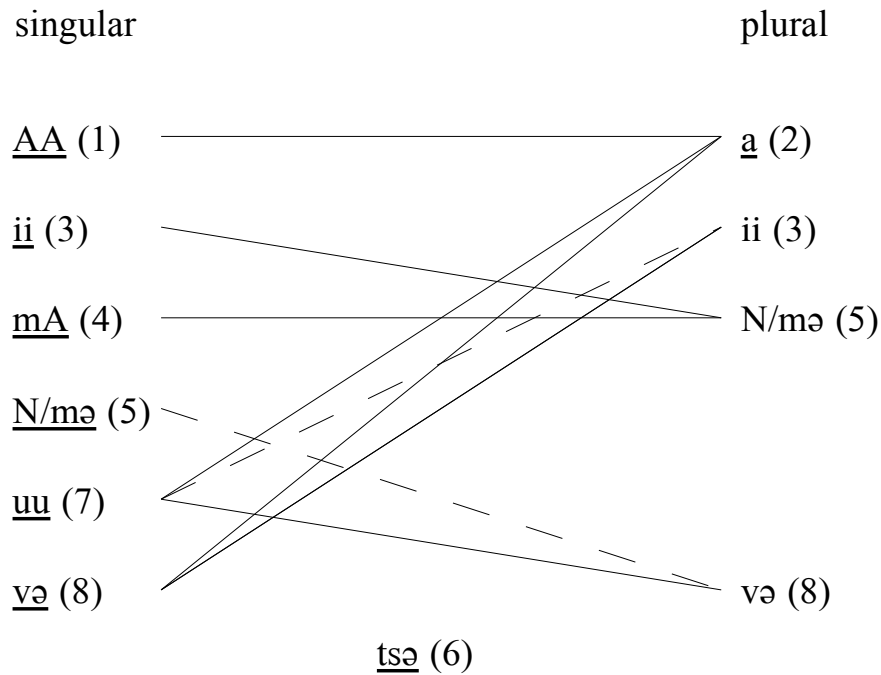
Figure 16: Kambari gender system according to Corbett's methodology (based on De Wolf 1971:60). Hoffmann's (1963) numbering is given in brackets

This diagram is similar to Figure 15, the only difference being that classes 2a and 2b have been merged into a single class 2, and classes 8a and 8b have been merged into class 8. In Corbett's terms, for the set of count nouns there are six target genders in the singular and four target genders in the plural, while for the set of mass nouns there are a further eight target genders. There are eighteen controller genders (ten singular/plural pairings and eight single class genders), four less than the twenty-two genders shown in Figure 15. Furthermore, some information about class pairings has been lost in Figure 16. For example, it is no longer apparent that class 2b is only found as the plural class for class 8b, and not for 8a.

While it is clear that Corbett's terminology does not equate with standard Bantu terminology, the difference between the two approaches proves to be minimal with respect to Cicipu, with only two classes (3a and 3b) being conflated under the former (§5.1-5.2). Nevertheless this discussion of Central Kambari has served as a useful illustration of how Corbett's analytical framework applies to Benue-Congo languages.

Before leaving this section, a note is required on the terms 'agreement' and 'concord'. They are usually taken to mean the same thing, with concord being the preferred term in the Bantu literature. Some authors do distinguish the terms but do so

in idiosyncratic ways – see Corbett (2003a:110-112) for a summary. Here I will stick to 'agreement'.

2.2   Agreement

In the previous section we saw that the presence of agreement is a definitional property of gender systems (in the terminology of mainstream linguistics) – this is what sets them apart from both classifier systems and reduced noun class systems. Corbett (2006:4) notes that agreement is hard to define satisfactorily, and several recent textbooks (Corbett 1991, 2006, Siewierska 2004) have adopted Steele's (1978:610) characterisation of agreement as "some systematic covariance between a semantic or formal property of one element and a formal property of another" as a good working definition.

I will begin by discussing the boundaries that have been proposed between agreement and other phenomena, in other words the DOMAIN of agreement (§2.2.1). The next two subsections briefly discuss the mechanism of agreement (§2.2.2) and the distinction between syntactic and semantic agreement (§2.2.3). Chapters 7 and 8 will investigate in detail a specific case of variation in agreement, namely the gender and person alternation found on several agreement targets in Cicipu. As a preliminary then, §2.2.4 covers the different ways in which agreement can vary on a given target in a given language, and what kinds of conditions trigger this variation. Much of chapter 7 is concerned with the syntactic status of the Cicipu subject agreement markers, and so the typology of agreement markers and pronominal affixes is also discussed here (§2.2.5).

2.2.1   Domain

Many linguists have attempted to limit the scope of agreement (its DOMAIN according to the terminology used in Ferguson and Barlow (1988) and Corbett (2006)). Different theorists draw the dividing-line between agreement and related phenomena in different places, and their definitions of agreement vary accordingly. In particular, there is disagreement as to whether the ANTECEDENT-ANAPHOR relation should count as true agreement. Bresnan and Mchombo (1987:743) observed that

> many current syntactic frameworks could account for differences between grammatical and anaphoric agreement (because entirely different mechanisms are postulated in these frameworks for grammatical agreement and pronominal incorporation...)

The same is true of more recent generative theories of agreement (e.g. Chung 1998). Barlow (1992:134-152)[6] comes to the conclusion that there is not a strong enough motivation for a cut-off point at the boundary between local agreement and anaphoric pronouns. Despite the differences between local (or 'grammatical') agreement and anaphoric agreement (see §2.2.5), *inter*-sentential agreement often involves the same controllers, targets, forms, properties, and conditions as *intra*-sentential agreement. Siewierska (2004:120-121) observes that:

> In most mainstream work on agreement this notion includes within its scope the determination of the form of independent person markers more commonly called anaphoric pronouns. The domain of agreement is therefore not restricted to the clause, or even sentence, but may be a larger discourse unit such as a thematic paragraph.

A further reason to admit anaphors as agreement targets is that they form a natural extension to Corbett's Agreement Hierarchy, a language universal concerning the opposition between syntactic and semantic agreement (§2.2.3). Corbett and Barlow's view will be adopted here.

While anaphoric agreement thus qualifies as agreement under Corbett's approach, it is not 'canonical' agreement according to the framework which has come to be known as Canonical Typology (e.g. Corbett 2005a, 2006). Corbett uses a number of criteria to characterise the ways in which various kinds of agreement depart from a central canon (loosely defined as the "best and clearest examples" 2006:9), and one of these criteria (C-15) is that a 'local domain' is more canonical than a 'non-local domain'. One consequence of adopting a fairly relaxed view of the domain of agreement is that in the case of non-canonical agreement involving anaphoric pronouns, controllers can be a considerable distance from their agreement targets (pointed out by Corbett (2006:41) and Ferguson and Barlow (1988:4-10)). However it does not necessarily follow that controllers can therefore operate at unlimited distances. The further the distance between the target and its antecedent, the harder it is to maintain that the form of the target is controlled solely by formal properties of the antecedent word. Instead it becomes difficult to distinguish between the influence of the controlling word, and the influence of the referent itself – in other words, there comes a point where anaphoric reference blurs into deictic reference. As Corbett (1991:243) puts it:

---

6   According to Corbett (2006:21) this is "the only extended discussion of the issue".

The fact that there is no determinable maximum distance between antecedent and pronoun means that sometimes it is not fully clear which noun phrase is the antecedent of a particular pronoun. And when a pronoun is widely separated from possible antecedents it may be being used deictically rather than anaphorically.

More typically, of course, we think of deictic reference as introducing a referent into a discourse for the first time, as in the following Cicipu sentence:

(11) (kà-táarí)   ká-mpà
      NC1-stone   AG1-this
      this *(stone)* [said while pointing to a stone]

This could be said without the noun, in which case the **ká-** prefix on the demonstrative is (to a certain extent) determined by the gender of **kà-táarí**, the basic-level term for the concept STONE (see Rosch 1978 and the discussion in Corbett 1991:244). It is not clear to me whether pronouns used in this way can be said to 'agree', since there is no textual controller for them to agree with. To be sure, pronouns tend to show the same 'agreement' features whether they are used deictically or anaphorically (Bresnan and Mchombo 1987:748-752). However they do not satisfy Steele's definition, since there is only one 'element' involved, the 'target'. Note too that there is no 'systematic covariance' here. Although the basic-level term for a concept undeniably influences the referring expression that the speaker will choose, nevertheless referents and referring expressions do not exist in one-to-one correspondence. Often there are synonyms with different genders, and failing that, every referent belongs to a superordinate category, the name for which may also belong to a different gender. For example in Cicipu a smallish stone (**kà-táarí**, NC1) may be pointed out exophorically using any one of **ká-mpà** (this, AG1), **má-mpà** (this, AG4, based on **mà-táarí** 'pebble, NC4'), **yí-mpà** (this, AG3, based on the hypernym **ì-rí** 'thing, NC3') or perhaps most likely of all **é-mpè** (this, 3PS). The actual term chosen will depend on the text-external context as well as the speaker's communicative purpose.

Corbett (2006) does not directly address deictic use, although his C-1 criterion (2006:10) states that 'controller present' is more canonical than 'controller absent', implying of course that 'controller absent' agreement still qualifies as agreement. In a recent paper on Canonical Typology, Bond (2009) argues that a strict definition ('base') is required to determine the domain of investigation for phenomena investigated in the Canonical Typology framework. It is not clear whether these two positions can be reconciled without loosening the definition of agreement too much. In §8.8 I will

discuss the possible relevance of the distinction between agreement and deixis to the Cicipu gender/person alternation.

2.2.2 Mechanism

The traditional generative understanding of agreement is that it is a redundant and asymmetric syntactic operation, where features are copied from the controller to the target. However, as highlighted by Barlow (1992:22-45), certain examples cannot be properly accounted for by a feature-copying model. For example in Russian the second-person pronoun has only one form *ty*, but triggers different agreements depending on the addressee's sex:

(12) (a) ty čital-Ø   (b) ty čital-a
          2S read.PST-MASC   2S read.PST-FEM
          You were reading [said to male]   *You were reading* [said to female]
                                                              [Corbett 1991:128]

The only way round the problem is to posit two homonymous pronouns *ty*, one marked for masculine and the other for feminine. In the light of examples such as these, theories such as Generalised Phrase Structure Grammar and Lexical-Functional Grammar have developed 'feature-merging' or 'feature-unification' models[7]. In such approaches, features may occur independently on both controllers and targets. Provided the feature values are compatible (i.e. not conflicting), then the sentence will be grammatical without requiring the proliferation of homonyms. So for examples (12a-b) the second-person pronoun would be marked for person but not gender, whereas the verb would be marked for gender and not person. At some point in the representation of the sentence the two structures are unified, giving the values for both person and gender.

2.2.3 Syntactic and semantic agreement

Steele's definition of agreement quoted above states that agreement can be with either a "semantic or formal property" of the controller. Corbett distinguishes between SYNTACTIC AGREEMENT and SEMANTIC AGREEMENT (1991:225-260, 2006:206-237) – targets showing the

---

7 Chung (1998) is an attempt to model a unification approach to agreement within the Principles and Parameters framework. Unfortunately much of the motivation for her account of agreement relies on theory-internal evidence. For example, evidence for the existence of her 'Associate relation' depends on a particular analysis of the derivation of VSO word order in Chamorro – the subject is generated within VP (the 'Internal Subject' hypothesis), raises to Spec IP in order to get case, and then lowered again to an adjunct of the verb in order to produce VSO word order. Regardless of the advantages of this approach within the Principles and Parameters framework, it is hard to evaluate it generally because it relies on many theory-specific assumptions.

former agree with a formal property of the controller, those showing the latter agree with a semantic property. Nouns which offer a choice depending on the agreement target are called HYBRID NOUNS (Corbett 1991:183-184). For example in Kiswahili a number of kinship nouns belong to morphological noun class 9 in the singular and 10 in the plural. As expected, these trigger gender 9/10 agreement within the noun phrase. However when it comes to subject and object agreement, they trigger gender 1/2 (i.e. human) agreement on the verb (Welmers 1973:175, Katamba 2003:113). In Chicheŵa 12/13 (diminutive) and 7/8 nouns denoting humans can trigger either syntactic (12/13 or 7/8) or semantic (1/2) agreement, with the latter becoming more likely the further the separation between the controller and target (Corbett 1991:248-250).

It should be stressed that, in contrast to the two types of agreement triggered by these Bantu hybrid nouns, the alternation between Cicipu gender and person agreement does not reduce to the syntactic/semantic agreement distinction. Cicipu *does* have hybrid nouns e.g. **má-gàjì** 'priest' (§6.3), but this involves two different feature values for gender rather than two completely different feature paradigms.

2.2.4 Variation in agreement

Agreement in noun class languages is usually taken to be a syntactic phenomenon, and as such it is generally viewed as obligatory – if constituents have the possibility of agreeing with the noun, then they do. This is the overall picture of gender agreement that emerges from introductions to African linguistics (e.g. Welmers 1973:159-183, Watters 2000), which understandably present a 'normalised' account to their readers, and are not able to focus on the difficulties posed by individual systems. Grinevald and Seifart (2004) argue that typologists without first-hand exposure to African noun class languages have been influenced by such simplified models. In fact, even more general definitions may require agreement to be obligatory, as in Aikhenvald (2000:20): "Noun classes are defined syntactically. They constitute a closed obligatory grammatical system". Nevertheless there are a considerable number of languages where gender agreement is not obligatory. Corbett implicitly acknowledges the existence of optional[8] agreement when he introduces the term 'enforced' gender agreement (1991:203):

> If a particular target type can mark agreement in gender then in many languages it must. A Latin adjective, which distinguishes masculine, feminine and neuter (in a given case and number combination), must do so: the ending cannot simply be

8   In the sense of 'subject to pragmatic constraints'.

omitted. This may be called 'enforced' gender agreement.

'Unenforced', or optional, gender agreement is relevant for this thesis, especially Part IV where I deal with the alternation between gender and person agreement in Cicipu. In preparation for those chapters, this section considers the different kinds of variation in agreement that are found cross-linguistically, and the conditions that trigger them.

2.2.4.1 Pre-requisites and conditions

Ferguson and Barlow (1988:4) discuss both CONFLICT and VARIATION in agreement. Two or more patterns of agreement may be in conflict with each other, such as when two conjoined pronouns trigger agreement in the predicate, or a plural pronoun is used to politely refer to a singular referent. If in a case of conflict more than one agreement pattern is grammatical, then this results in alternative possibilities, or variation in agreement. Corbett (2006) uses the term CONDITIONS to refer to the factors that determine which of the available agreement patterns occurs. Examples of Corbett's conditions include animacy, precedence (i.e. word order), topic, and focus. He is careful to distinguish between conditions and PRE-REQUISITES (2006:183):

> The essential difference is that prerequisites specify what is necessary for agreement,...while conditions affect the use of an agreement form where the prerequisites are met.

Pre-requisites for agreement are found at all levels of linguistic structure (for examples see Corbett 1991:133-135, 2006:78-85, 183-184). For example with respect to syntax, in some languages agreement may be limited to, or excluded from, certain constructions. In Kabyle (North Berber, Laoust 1928:40, cited in Aikhenvald 2000:39) demonstratives agree in gender when used as pronouns, but do not when they are used as modifiers. Conversely adjectives in German agree when used as modifiers, but do not when they occur as predicates (Corbett 1991:124). Agreement may also interact with the categories of definiteness, tense, person, number, case (Corbett 1991) and animacy (Siewierska 2004:155, Givón 1976 for Kiswahili). Also, within a particular word class, some lexical items fail to show agreement. For example, in some eastern Bantu languages there is a subset of adjectives ('invariant words') that do not take concords (Maho 1999:106-107), often borrowed from other languages. In chapter 7 we will see examples of Cicipu agreement pre-requisites involving both phonology and morphosyntax.

Conditions, on the other hand, seem to be limited to the syntactic, semantic, and pragmatic levels of linguistic structure (Corbett 2006:183-184). Examples involving pragmatics are problematic operationally, however, since in the absence of explicit formal coding of pragmatic relations it can be difficult to tell whether variation in agreement is best analysed as a prerequisite or a condition. Unlike features such as case or grammatical relations, it seems to be relatively rare to find an independent formal marker of topicality that can be counted on for evidence when it co-occurs (or does not, as the case may be) with a particular agreement pattern. Comrie (1988:271) writes "My impression is that it is extremely rare across languages to find a formal device that literally, in one-to-one correspondence, encodes some pragmatic distinction..."[9], and as Dooley (2007:100) points out, even well-known candidates such as the Japanese 'topic' marker *wa* often seem to be coding the *setting* of a discourse unit rather than anything of intrinsic interest, whether a referent (topic) or other kind of theme. So when linguists say that a particular case of agreement is linked to the topicality of the controller, what they often mean is that there is a correlation between the occurrence of agreement on the one hand, and on the other what they perceive to be the discourse salience of the controller referent from their subjective understanding of how the speaker intended the hearer to construe the text. Usually there is no independent marking of topic which can be relied upon, and so in such cases the variation with respect to agreement could equally well be viewed as a pre-requisite *or* a condition. Under the former analysis, agreement is not possible unless the controller referent is topical. Under the latter analysis, agreement is possible with non-topical controllers, but not as likely.

Corbett (2006:197-198) cites the case of long-distance agreement with topic in Tsez (Polinsky and Comrie 1999, Comrie 2003) as a clear instance of an agreement condition. The language has two kinds of gender agreement, with controllers of varying syntactic relation – either straightforward grammatical agreement with the absolutive argument, or long-distance agreement with an absolutive argument in an *embedded* clause, as in (13) where the verb **iy** 'know' agrees with the NP **magalu** 'bread' rather than the expected absolutive argument of the main clause (which in this case would be the entire subordinate clause).

---

9   See also Lambrecht (1994:119).

(13) eni-r   [už-ā **magalu**   b-āc-'ru-λi]   **b**-iy-xo
     mother-DAT boy-ERG **bread:ABS:3** CL3-eat-PSTPRT-NMLZ CL3-know-PRS
     The mother knows that the boy ate bread

[Polinsky and Comrie 1999:117]

The situation can be diagrammed schematically as in Figure 17, where C stands for controller, S for agreement slot, and P for the feature paradigm involved. For Tsez there is only one paradigm (gender) and one agreement slot on the verb, but there are two competing controllers.

Figure 17: Schematic representation of the Tsez agreement system
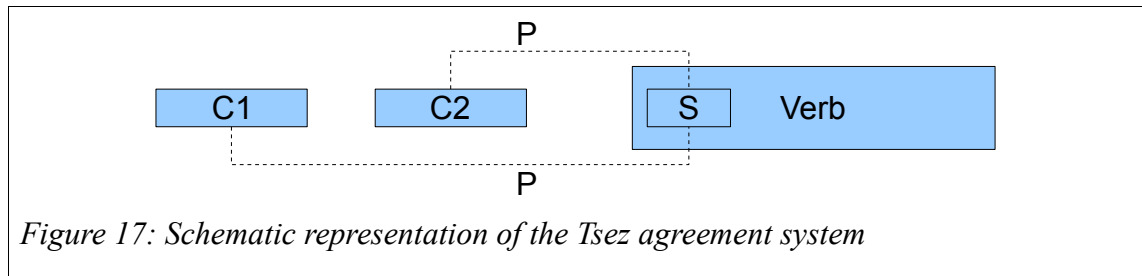
According to Polinsky and Comrie this long-distance agreement "must occur when the referent of the absolutive noun phrase is the main internal topic of the embedded clause", and "can be interpreted as a signal that the main topic of the embedded proposition is expressed by the absolutive NP" (Polinsky and Comrie 1999:122). It seems then that Polinsky and Comrie understand clause topic in Tsez as a prerequisite (using Corbett's terminology) for long-distance agreement rather than a condition. This analysis is supported by what happens when the absolutive NP is overtly-marked with either of the two topic particles found in the language. According to Polinsky and Comrie (1999:125):

> If the absolutive argument in the embedded clause is marked by either particle, LDA [long-distance agreement] is the only option...However if some constituent other than the absolutive is marked by a particle as topic, LDA becomes impossible.

Although it is not in doubt that Tsez shows variation in agreement which is dependent on discourse properties, there is an operational problem of deciding between pre-requisites and conditions. Despite this kind of difficulty, I will make use of Corbett's terminology in the description of Cicipu person agreement in Part IV, since it provides a useful point of departure for the analysis. In §8.4 I will show that in addition to morphosyntactic and semantic conditions, there is also a pragmatic dependency on subject and pronominal person agreement in Cicipu involving topicality. This

dependency will be analysed as a condition rather than a prerequisite, since there are occasions when the controller referent is clearly topical but person agreement does not occur.

2.2.4.2　Variation and topicality

Like Tsez, many other languages show variation in agreement depending on discourse properties. In every case of optional agreement involving discourse pragmatics that I am aware of, the occurrence of agreement is associated with topical, identifiable, or specific referents, or definite noun phrases, while the absence of agreement is associated with non-topical, unidentifiable, or non-specific referents, or indefinite noun phrases. This point has been made by a number of theorists and typologists. According to Aikhenvald (2000:321) "Noun class agreement often occurs only if the noun is topical, or definite." Barlow (1992:91-92) makes the same point for agreement in general, rather than just gender:

> The classification of an object in terms of its position on a definiteness scale by virtue of agreement is quite common. In general, high definiteness induces agreement, whereas low definiteness is associated with absence of agreement.

Dalrymple and Nikolaeva (2005:71) write:

> Conditions on verb agreement are often assumed to be definable in purely syntactic terms. In some languages, however, the verb shows more agreement with topical arguments than with nontopical arguments.

These theoretical statements can be supported by data from several unrelated languages, as shown below.

2.2.4.2.1　Optional agreement

The best-known case of variation in agreement in Niger-Congo is the optional cross-referencing of third-person objects on the verb in many Bantu languages (e.g. Bresnan and Moshi 1990, Morimoto 2002). Wald's (1979) study on Kiswahili object agreement identifies definiteness, animacy, and topicality as relevant factors. Bresnan and Mchombo (1987) describe subject and object gender agreement in Chicheŵa, another Bantu language, and provide an analysis framed in Lexical-Functional Grammar. Subject agreement in Chicheŵa is obligatory and is analysed as ambiguous between GRAMMATICAL AGREEMENT where the agreement marker is a redundant affix expressing the gender of a co-occurring subject noun, and ANAPHORIC AGREEMENT, in which case the

agreement marker is an incorporated pronominal with full argument status, anaphorically bound to its antecedent. Object agreement, on the other hand, is only possible with topical referents encoded by non-argument NPs, and is therefore analysed as pure anaphoric agreement. We will return to Bresnan and Mchombo's analysis in the discussion of the typology of agreement markers in §2.2.5.

Topicality is also relevant with respect to subject agreement in Bantu. Demuth and Johnson (1989) argue that the subject agreement prefixes in Setawana are anaphoric pronominals which only index topical referents, just like the Chicheŵa object markers. Morimoto (2008) proposes that in some Bantu languages verb agreement is controlled by the most topical NP in the clause rather than the grammatical subject.

Further afield, West Greenlandic provides an example involving person agreement. The verbal morphology depends on the status of the object; if it is definite, then person agreement is present with both subject and object as in (14) (the affix is a fused marker). However in (15), where the object is indefinite, the verb agrees with the subject only.

(14)　tuttu　　taku-aa
　　　 caribou　see-3.SG;3.SG;INDIC
　　　 *He saw **the** caribou*

(15)　tutu-mik　　taku-vuq
　　　 caribou-INSTR　see-3.SG;INDIC
　　　 *He saw **a** caribou*

[Fortescue 1984:86, cited in Barlow 1992:92]

Definiteness also plays a role in gender agreement in Arabic (Corbett 1991:125) and Garifuna (Northern Arawak, Munro 1997:445), and in person agreement in Kambera (Austronesian, Klamer 2008). In Garifuna the verb always agrees with the subject, but only agrees with definite objects. Similarly in Kambera indefinite objects are not cross-referenced by agreement clitics (Klamer 2008:284).

Finally, the referentiality of the agreement controller can also be relevant. In the Indonesian language Menó-Mené Sasak, only referential undergoers trigger agreement on transitive or ditransitive verbs (Austin 2004:11).

2.2.4.2.2　Choice of controller

In the examples so far (with the exception of Tsez) the variation consists of an alternation between the presence of agreement in the case of a topical/definite/referential controller, and the absence of agreement in the case of a

controller lacking one of these properties. In each case the agreement is optional, in the sense that it is possible to find grammatical sentences lacking the relevant agreement morphology. The next few examples involve a different kind of variation – there is a choice of agreement, with the controller being selected according to its topicality. The positive correlation observed in the theoretical statements given above still holds – in each case the topical referent is the controller.

The Amazonian language Paumarí (Chapman and Derbyshire 1991) is particularly interesting because of the relationship that holds between the occurrence of agreement and the DISCOURSE TOPICALITY of the controller. The language displays a system with two different paradigms marking the subject as well as a small subset of the adjectives, involving oppositions from two distinct gender systems. One of these systems is based around natural gender i.e. sex, the other around shape and other structural features of the referent. Chapman and Derbyshire (1991) use the terms 'gender' for the sex-based system and 'noun class' for the shape-based system, but this is simply a matter of convenience – both would be regarded as gender systems in Corbett's terminology[10].

Paumarí's 'noun class' and 'gender' agreement systems (using Chapman and Derbyshire's terms) are found with different sets of controllers, although there are some environments where they can both occur simultaneously: namely the S argument of intransitive verbs, the O argument of transitive verbs, possessed nouns and a small subset of adjectives. The four possibilities are shown below, first for the numeral 'one', analysed by Aikhenvald as an adjective, and secondly for an intransitive verb:

(16)   kavina   hoara-**na**
       howler_monkey.MASC.NON-KA   one-AG[MASC]
       one howler monkey

                                                         [MASC, NON-KA]

(17)   jomahi   hoara-**ni**
       jaguar.FEM.NON-KA   one-AG[FEM]
       one jaguar

                                                         [FEM, NON-KA]

---

10 Corbett (2005b) originally analysed Paumarí as a four-gender system, conflating Chapman and Derbyshire's 'noun class' and 'gender' systems – presumably resulting in the values masculine, feminine, KA- marked, and a 'neuter' gender to cover the remaining nouns. Aikhenvald (n.d.) argues that it is better to view them as two independent systems, given that they can both occur together on the same agreement target in separate morphosyntactic slots, and they never occur individually in the same morphosyntactic slot – a view later accepted in Corbett (2007:256-257).

(18) kasi'i   ka-hoara-na
      crocodile.MASC.KA   AG[KA]-one-AG[MASC]
      one crocodile

[MASC, KA]

(19) kaira   ka-hoara-ni
      guava.FEM.KA   AG[KA]-one-AG[FEM]
      one guava

[FEM, KA]

(20) voroni-'i-ha   ada   kahami
      fall-ASP-TH.AG[MASC]   DEM.AG[MASC]   palm_tree.M.NON-KA
      the palm-tree fell down

[MASC, NON-KA]

(21) ka-voroni-'i-hi   ida   ojoro
      AG[KA]-fall-ASP-TH.AG[FEM]   DEM.AG[FEM]   turtle.F.KA
      the turtle fell down

[FEM, KA, Aikhenvald n.d.:10-11]

The agreement patterns in transitive clauses are particularly interesting since the two kinds of agreement behave differently with respect to the argument functions of the verb. Noun class agreement is 'purely syntactic', whereas gender agreement is 'pragmatically determined' (Aikhenvald n.d. 13). Briefly, in transitive clauses noun class agreement is only found with object NPs. Gender agreement, on the other hand, marks the 'pivot' of the discourse ('what the story is about' – Aikhenvald n.d. 11), and can be found with either the A or the O argument.

(22) makha   'o-ka-vaniha-mani-'i-hi   ida   ba'dana
      snake.M.NON-KA   1S-AG[KA]-think_wrong-CORR-ASP-TH.AG[FEM]   DEM.AG[FEM]   lizard.FEM.KA
      I mistook the lizard for a snake

[Aikhenvald n.d. 12]

In example (22) *ba'dana* 'lizard' is the pivot, or discourse topic, and so the verb agrees with it in gender through the *-hi* feminine suffix. The two verbal prefixes are syntactic rather than pragmatic agreement, with the *ka-* prefix again agreeing with *ba'dana*. The system can be represented as in Figure 18. Again P stands for feature paradigm and the dotted line indicates a choice of agreement controller, depending on which of Cx or Cy is the pivot of the discourse.

Figure 18: Schematic representation of the Paumarí agreement system
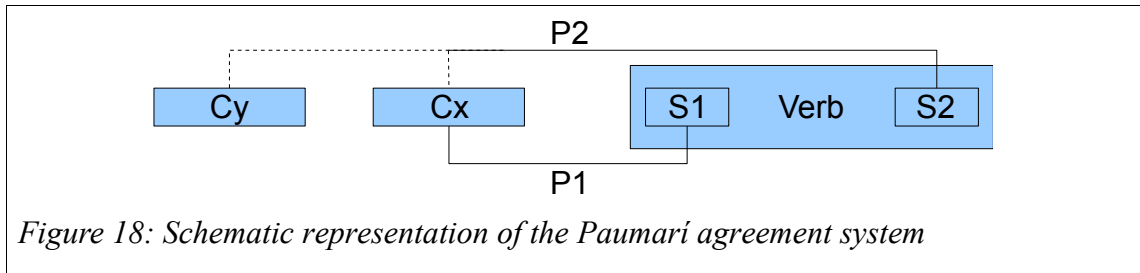
A similar situation obtains in the Papuan language Motuna (Onishi 1994, cited in Aikhenvald 2000:34), where a verb agrees obligatorily in person with its subject and object, but also in gender with the topical constituent (which may be either the subject or object).

Finally, in the Daghestanian language Dargi, when there are two third-person arguments the verb agrees with either the agent or patient, whichever one is topical (van der Berg 1999:161-165).

2.2.4.2.3   Choice of feature paradigm

We have seen topicality (and related notions such as definiteness and referential status) functioning as an agreement condition in two distinct ways in the above examples. First, we considered examples of optional agreement where agreement occurred only with topical controller referents. Secondly, we looked at examples of agreement where different controllers are selected depending on their discourse topicality. A third kind of pragmatically-conditioned agreement is found in several Kainji and Plateau languages. Instead of two potential controllers competing for a single agreement slot, this time the competition is between two different feature paradigms. The phenomenon is discussed by Hoffmann (1963) and Crozier (1984) for the West Kainji language Central Kambari[11]. Hoffmann linked the Central Kambari subject gender/person alternation to the definiteness of the subject: "A subject noun may be shown to be indefinite or definite according to whether there is class agreement in the finite verb or not" (1963:168). He gave the following examples (Hoffmann 1963:167-168):

---

11 Recall from §1.3 that Cicipu is the most distant member of the Kambari branch.

(23) (a) ma-nun u-kuwete
NC4-bird 3s-die
the bird died

(b) ma-nune **me**-kuwete
NC4-bird AG4-die
a bird died

(24) (a) ts-al u-riyete
NC6-meat_strip 3s-fall_down
the strip of meat fell down

(b) ts-ale **tse**-riyete
NC6-meat_strip AG6-fall_down
a strip of meat fell down

In (23a) and (24a) where the subjects are definite, there is no gender agreement, and instead the general third person subject prefixes (singular **u-** or plural **a-**) must be used. In contrast, in (23b) and (24b), which have indefinite subjects, the gender agreement prefix must be used instead of the person prefixes. Hoffmann's treatment of this alternation is very brief, and it is natural to wonder if there are other factors involved, particularly given the awkwardness of the English translations in the absence of any contextual information. According to Crozier (1984:215-222), although Hoffmann's analysis works most of the time, definite subjects do sometimes trigger gender agreement. The determining factor according to Crozier is the TOPICALITY of the subject referent. If the subject is topical, then it will trigger person agreement, otherwise it will trigger gender agreement. This is true to some extent for Cicipu as well, but as we will see in Part IV it is only part of the story.

One obvious question for these languages is if the two sets of agreement markers occur in the same morphosyntactic environment, why do we set up two systems at all? This question will be addressed in §7.8, but for now we just stress that person-marking is independent of gender (i.e. all nouns can potentially trigger 3PS agreement on the verb, no matter what their gender), and also that the pre-requisites for agreement differ for the two systems. The Central Kambari/Cicipu-type system is diagrammed in Figure 19:

Figure 19: Schematic representation of the Central Kambari/Cicipu agreement system
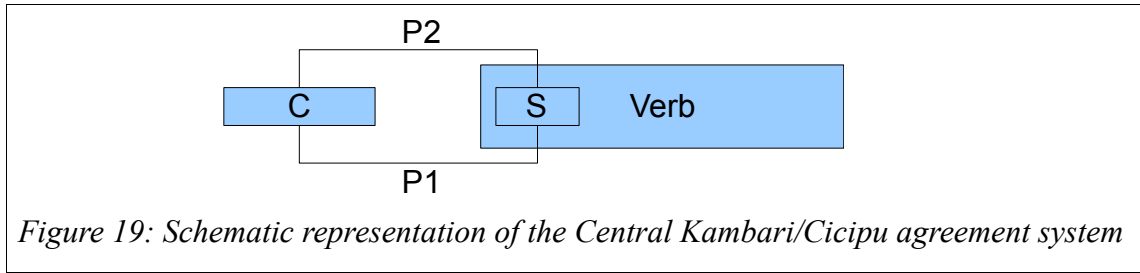
Although this alternation between gender and person agreement is best documented for Central Kambari, it is apparently present in each of the major branches of West Kainji. Agamalafiya (2007) provides an annotated folktale in the Lake language Tsureshe. Both the major participants in this story are marked, in different places, with gender and person agreement. Gender subject agreement is concentrated towards the beginning of the text, whereas person subject agreement is the norm towards the end. Both appear to be ambiguous agreement markers. Towards the end of the story a third, minor participant is introduced. Despite being 'on stage' for approximately one third of the story, it is never marked with person agreement, only gender agreement. We might hypothesise that referents are more likely to be encoded by person agreement as they become entrenched as discourse topics.

In the Northwest language u̱t-Ma'in (also known as Fakai) Smith (2007:81, 84) explicitly states that there is an alternation between gender- and person-marked subject prefixes and object pronouns for human referents. She also mentions that either marking strategy may be used to track referents through discourse, which suggests that whatever governs the alternation, it is not definiteness. Animacy can also be ruled out, since the alternation is only possible in the first place if the referent is human and third person. Topicality is an obvious alternative candidate given the patterning in the languages already mentioned.

The Basa-Kamuku language Pongu (also known as Ta̱rin, MacDonell 2007) also has separate gender and person paradigms on person markers. It is not clear from the available description when gender agreement occurs and when person agreement occurs, but one innovation with respect to the other West Kainji systems is the existence of what seem to be *portmanteau* subject agreement prefixes, marked for both gender and person. So, for example, in addition to the general 3PS subject prefix **bu-** and the gender subject prefixes **i-** and **u-**, **bjɛ-** and **bwa-** are also possible, for class **i-** and **u-**

nouns respectively (MacDonell 2007:51). So as well as the individual possibilities, Pongu seems to allow a combination of these exponents.

The alternation is not limited to West Kainji, and can be found in both the East Kainji language Amo (Anderson 1980a:157-159; see also Corbett 1991:247) and the Plateau[12] language Kaje (McKinney 1978). In Amo the alternation applies to subject 'pronouns' (probably prefixes), object pronouns, and possessive pronouns. According to Anderson, the gender-marked sets can be used with either human or non-human referents, but the person-marked sets are limited to humans. No information is given as to what governs the alternation for human referents. In Kaje the alternation is found on at least the subject prefixes. Further afield, in the Grassfields Bantu language Ngyembɔɔn (Anderson 1980b) plural nouns with human referents can trigger either gender or person subject agreement on verbs.

Finally, this competition between different feature paradigms can also be found outside Africa. The Amazonian language Miraña (Seifart 2005) has a particularly complex pronominal system whereby pronouns can be inflected for either person, "specific noun class", or "general noun class" (see Seifart 2005:259 for a textual example involving all three).

This section has illustrated the cross-linguistic correlation between the presence of agreement and topicality. It has glossed over the fact that the notion of 'topicality' is notoriously slippery and means different things to different linguists; I will take up this issue in §2.2.5.1 and §2.3. From a different perspective, topicality has also been linked to the syntactic status of agreement markers, and it is to this topic which we now turn.

2.2.5   Typology of agreement markers

The status of verbal affixes co-indexed with NPs has been the subject of much research in recent decades, with Siewierska (2004:121-127) and Corbett (2003b, 2006:99-112) providing helpful summaries. Terminologically, the question of how to refer to such affixes in a theory-independent manner is vexed. 'Incorporated pronouns', 'pronominal affixes' and even 'agreement affixes' all come with their own theoretical baggage and none of these are neutral terms. In the description of Cicipu agreement in Parts III and IV I will use the terms 'agreement markers' or 'agreement prefixes' when discussing the gender and person subject markers, but it should be borne in mind that they share

---

12 Recall from §1.3 that Plateau and Kainji are co-ordinate branches of Central Nigerian.

properties with what many people have called 'incorporated pronouns' or 'pronominal affixes'.

The central question with regard to the classification of agreement markers/incorporated pronouns is how to handle apparently 'optional' verb arguments. In a 'pro-drop' language such as Italian a sentence can be equally grammatical with or without the subject:

(25) (lui) parl-a
3s.PRO speak-3s
he speaks

In the longer sentence *lui parl-a* most linguists would analyse the pronoun *lui* as the subject and the suffix -*a* as a non-referential 3PS agreement marker. But what about the shorter sentence *parl-a*? Assuming (as many grammatical theories do e.g. Lexical-Functional Grammar, Relational Grammar, GB/Minimalism) that every clause has exactly one subject, the sentence *parl-a* must also have a grammatical subject. One way to satisfy this constraint is to posit the existence of a phonetically-empty but semantically-referential *pro*, and this is the approach taken in GB, and also in Mel'čuk's theory of agreement (1993:342-343). This empty category then bears the subject function, and the verb is said to agree with the empty category.

However many theorists are not content with the proliferation of covert arguments entailed by the above approach. Jelinek (1984) suggested a radical alternative analysis of pro-drop languages which avoided null arguments altogether. Her analysis was based on the Australian language Warlpiri, and set out to account for three properties of the language. First, Warlpiri allows null anaphora. Secondly, it is a non-configurational language with relatively free word order. Thirdly, the language tolerates a great deal of discontinuity of apparent 'constituents' such as the subject NP (e.g. the adjective can be separated from the head noun). She gave a unified explanation for all three of these properties by assuming that in sentences corresponding to the pro-drop version of (25) the affix *is* the subject. In other words, the affix is referential and bears the grammatical function of subject. The three properties mentioned above are accounted for by analysing all NPs as *adjuncts* rather than as arguments of the verb. Jelinek then extended her analysis to unrelated languages such as Spanish and Italian, and refers to such languages as 'W-type' languages.

Regarding Bantu, two of the three properties she seeks to account for through this analysis are generally absent (discontinuous expressions, and to a lesser extent free word order). Nevertheless she suggests that Bantu languages may be W-type languages, since in such languages a verb and its affixes may stand alone in a sentence, the 'subject' and 'object' NPs being optional (Jelinek 1984:70). If this is so, then it means that the apparent subject NP is never a true subject – instead it can only be an adjunct to the verb.

However the situation in Bantu is complicated by the asymmetry between subject and object (Bresnan and Mchombo 1987, Bresnan and Moshi 1990), and Jelinek's analysis does not straightforwardly apply. In particular, the object affix (under Jelinek's W-type analysis this would be the only possible candidate for the object argument) is always optional in Bantu languages and so cannot be relied upon to bear the object function[13]. As was mentioned in §2.2.4.2.1, Bresnan and Mchombo (1987) provide an alternative analysis of the Bantu language Chicheŵa in the Lexical-Functional Grammar framework. Similarly to Warlpiri, a Chicheŵa verb with its affixes can be a full sentence on its own, or alternatively nominals may occur simultaneously with agreement affixes. Rather than following Jelinek by analysing Chicheŵa as a W-type language whose verbal affixes are always the true arguments of the verb, they propose a cross-linguistic typology of agreement markers ranging from pure ANAPHORIC AGREEMENT MARKERS on the one hand (which, as in Jelinek's analysis of Warlpiri clitics, may not co-occur with an argument NP), to pure GRAMMATICAL AGREEMENT MARKERS which are non-referential and co-indexed to the true NP subject (as in the English subject agreement marker -*s*). Crucially, in between these two extremes lie AMBIGUOUS AGREEMENT MARKERS, which may take part in either anaphoric agreement (without a subject NP) or grammatical agreement (with a subject NP), and thus have a dual function[14]. The three types of agreement marker form a grammaticalisation cline (Siewierska 1999):

(26)   Anaphoric   >   Ambiguous agreement   >   Grammatical
        agreement marker   marker   agreement marker

---

13 This would be problematic in many theoretical frameworks. For example in GB, it would violate the Theta Criterion (Chomsky 1981:36), which stipulates that the theta roles of the verb must always be realised in syntactic structure. Similarly in LFG the Completeness Constraint would be violated.

14 This is made possible by the "dual structure" of LFG – grammatical functions such as SUBJ and OBJ are independently specified in the f-structure (functional structure) of a sentence rather than (as is the case in Chomskyan frameworks) being derived from the c-structure (constituent structure).

Like Jelinek, Bresnan and Mchombo extend their analysis to more familiar 'pro-drop' languages such as Spanish and Italian, and the -*a* suffix from (25) would be analysed as an ambiguous agreement marker.

Siewierska (1999:228ff) observed that the distinction betweeen anaphoric and ambiguous agreement markers is not always clear-cut, since they may behave differently depending on properties of the controller. For example, she predicts from the Nominal Hierarchy (Silverstein 1976) that "there should be instances of person agreement with an independent person marker but not a nominal NP" (Siewierska 2004:152). This kind of agreement system is apparently rare, but does exist and Siewierska gives the well-known example of Welsh subject agreement.

Siewierska also emphasises (e.g. 1999:231-234, 2004:127) the independence of a particular agreement marker's morphological and syntactic statuses. The typology in (26) is based on whether the agreement marker can co-occur with its controller in the same clause, not on its form. So for example, clitics can function as syntactic agreement markers, while affixes can function as anaphoric agreement markers (see Siewierska 1999 for specific examples).

Subsequent to the development of this typology, Austin and Bresnan (1996) reviewed Jelinek's claims about Warlpiri (and by extension, most Australian languages) and presented evidence suggesting that the verbal clitics should be regarded as ambiguous agreement markers rather than, as Jelinek argues, anaphoric agreement markers. Amongst other arguments, they observe that the ability of pronominal affixes to co-occur with indefinite or non-referential controllers results in complications for her analysis that are not suffered by the LFG account (Austin and Bresnan 1996:234-235). Austin and Bresnan's contention that the ability to co-occur with indefinite and non-referential controllers is a relevant factor in deciding between anaphoric and ambiguous markers has been debated in recent years (e.g. Baker 1996:125-129, Evans 1999/2002, Mithun 2003). Evans (2002:46), whose basic position is in agreement with Austin and Bresnan, goes so far as to contend, contrary to both Jelinek/Baker and the usual LFG treatment, that an agreement marker should be called 'grammatical' when it loses its referentiality (i.e. it can be co-indexed with non-referential NPs), independently of whether it can occur in the absence of a nominal argument.

Corbett (2003b) provides a number of tests which can be applied to agreement

markers in order to place them on this typological scale, and I will make use of these (and others) in the classification of Cicipu agreement markers in §7.6. We will see that while there is no difference morphologically between Cicipu gender and person agreement (both are affixes), and they can both be regarded as ambiguous agreement markers, they appear to be at different stages of the grammaticalisation cline, with gender agreement further towards the 'grammatical' end. Both gender and person agreement markers can co-occur with indefinite and non-referential subject NPs, which makes them grammatical agreement markers according to Evans' usage.

2.2.5.1 Anaphoric agreement and topicality

In Part IV we will see that one of the differences between the anaphoric use of gender and person agreement markers in Cicipu is that the latter are associated with topical referents. This, of course, implies that the former are less likely to mark topics. How can this be, if, as is often assumed, incorporated pronouns always link to topical antecedents? In this section I want to question this assumption and show that incorporated pronouns do not have to be topical. The Cicipu data presented in Part IV will support this claim.

Bresnan and Mchombo (1987) assume that incorporated pronominals are anaphorically linked to topics. Although their paper starts by showing that floating topics in Chicheŵa sentences can only be licensed by an incorporated pronoun, they also assume the reverse condition holds i.e. that incorporated pronouns always anaphorically link to a topic (e.g. p757 "when the SM[subject marker] is used for anaphoric binding, its antecedent within the sentence has the TOP function", p764 "the OM's, used for anaphora to a topic"). This association has often been repeated[15] but is rarely questioned.

Taking a step back for a moment, it is important to stress that there is no *a priori* logical connection between the incorporation of a pronoun and the topicality of the

---

15 Demuth and Johnson (1989:24) "B&M hypothesize that incorporated pronominals are topic-oriented; that is, they can only anaphorically link to items filling the TOP function", Uyechi (1991:438) "The topical nature of the sentence initial noun phrase of the [Navajo] *bi*-form follows directly from the pronoun incorporation analysis"), Corbett (2003:189) – "if the marker is an incorporated pronoun, it will be linked via anaphoric agreement to the 'topic' of the sentence", Morimoto (2002:294) – "[Bantu] object marking on the verb appears only when it is topic-anaphoric (like English pronouns)", Morimoto (2002:295) – "the object marker functions only as a topic-anaphoric pronoun, being in complementary distribution with a clause-internal, non-topical object NP.", and Grinevald and Seifart (2004:251) – "[Niger-Congo] object indexation is generally more discourse dependent and applies only if the argument corresponding to the syntactic function object is topicalized" amongst others.

pronoun's referents. After all, different types of independent pronouns vary with respect to their discourse functions (see Siewierska 2004:67-74 for examples). Instead, the motivation behind the putative restriction of incorporated pronouns to topics is said to be their *participation in an opposition* with independent pronouns. Bresnan and Mchombo (1987:764-765) put it like this:

> We have seen that Chicheŵa has two series of anaphoric pronouns: the OM's, used for anaphora to a topic, and the independent object pronouns, used to introduce new topics or for contrast of arguments. Kameyama 1985 has observed that all languages have two kinds of pronominals that can be used anaphorically: those used for reference recoverable from discourse, and those used for 'contrast, emphasis, or focus'.

Different languages encode the two sets of pronominals in different ways. In Chicheŵa, pronominals may be free or bound, whereas in English the difference is between stressed independent pronouns and unstressed dependent pronouns. Siewierska (2004) makes a similar point. She is actually more explicit about the association between incorporated pronouns and topicality than Bresnan and Mchombo, and her conclusion is more wide-ranging, since she is dealing with dependent pronouns as a whole (including unstressed free pronouns in English, for example) rather than just incorporated pronouns. Like Bresnan and Mchombo, she hypothesises (this time on the basis of empirical evidence) that incorporated pronouns cannot vary as to their discourse function; they always encode topical referents. The relevant passage in her book (2004:67) is as follows:

> In contrast to the typologies of person markers based on morphophonological form and syntactic function, the typology of their discourse function is quite underdeveloped and under-investigated. Moreover, in comparison to the other two typologies it is also much more restricted in scope, since only independent markers appear to vary with respect to their discourse function. *Dependent markers invariably encode referents which are highly cognitively accessible and topical within the discourse.* [my italics – S.M.]

There is an implicit assumption in these analyses that there is a neat bifurcation between pronominals that denote new or contrastive topics and those that denote non-contrastive topics. However if the notion 'topic' is to have any independent explanatory power, then it must be recognised that topics are more than just non-contrastive 'given' referents. Consider the following English example concerning a teenager's first date, especially the expressions in **bold** type:

(27) While our niece and her young suitor wandered around the store, she reached into a bin of m&ms and snagged her finger on **something**. Whatever **it** was, **it** cut her.

 Her suitor, wishing to be a gentleman, expressed concern, but also told her he didn't like blood. She tried to shield her finger from him, but when the depth of her wound finally freaked her out a little, she showed it to him.

 The boy fell.

 Hard.

 Onto his face.[16]

After being introduced with the indefinite pronoun *something*, the mystery object is twice referred to in the subsequent sentence using a definite expression, the pronoun *it*. Nevertheless it is hard to maintain that any part of the above discourse is 'about' the thing that hurt the girl. Free pronouns can be anaphoric, definite, and subjects without being topics, at least according to pragmatically-based definitions of topic (§2.3).

So if the *it* in *it cut her* is not a topic, why then is it pronominalised? As we will note in §2.3.2.5, there is a distinction between 'recent-reference mention' and topicality. The subject referent of *it cut her* is pronominalised simply because it has just been mentioned, not because there is any textual span for which it is "a matter of standing interest or concern" (Strawson 1971:97). The passages quoted above from Bresnan and Mchombo and Siewierska both conflate topicality with something else. Kameyama's (1985) observation is not the same as Bresnan and Mchombo's – not all instances of "reference recoverable in discourse" are references to topics. Similarly, Siewierska's "highly cognitively accessible" and "topical" do not mean the same thing. Referents can be "highly cognitively accessible" and "recoverable in discourse" without the interlocutors taking the slightest amount of interest in them.

If free pronouns do not have to be topics, we might wonder why it should be the case that all incorporated pronouns are topics? There do not seem to be any pre-theoretical reasons why this should be so. In fact, given that the classification of an affix as agreement marker *vs.* incorporated pronoun is rarely a black-and-white affair, we might expect typical incorporated pronouns to be *less* strongly associated with topical referents than typical free pronouns. The following example from Cicipu involves an 'indirect anaphor' (i.e. a definite referring expression without a textual antecedent – see Schwarz-Friesel 2007) and shows that incorporated pronouns can be used to mark referents which are decidedly non-topical. Example (28) is an extract from a 15-minute

_____

16 Taken from a blog entry at http://sonmislocuras.com/page/2/.

text about the Koran, which is firmly established as the main discourse topic at this point. By contrast, the Mallam is 'introduced' for the first time here, with just the 3PS agreement marker **u-** serving as the indirect anaphor.

(28) [Context: *if, say, I'm lying...*]

ǹ=ù-'úng<òs>ò   híndè   ǹ-tángù   híndè,   mú-u-yùwò
when=3s-rise<CAUS>\RLS   like_that   1s-carry_on_head\RLS   like_that   1s-FUT-fall\IRR

when he [the Mallam] *lifts* [the Koran] *up like this and I carry it on my head, I will fall*

[tats005.001.130]

In Cicipu culture it is a generally-known fact that there will be a particular Mallam associated with the mosque, and it is this that makes the example acceptable. The sentence in (28) is the limit of the Mallam's involvement in the text, and no part of the text can be said to be 'about' him[17].

If we accept that incorporated pronouns do not have to be topical, a second question becomes relevant: whether incorporated pronouns can vary with respect to their discourse functions. In other words, does it make sense to talk of a typology of incorporated pronouns with respect to their discourse functions? Culy (2000) provides an interesting account of two different third-person object agreement markers in the extinct North American language Takelma (see also Aissen 2003). One of them, which is always null (in contrast to the non-null first- and second-person forms), is analysed as an 'anti-topical' incorporated pronoun (Culy 2000:80). The other object marker, *-kʰwa*, marks paragraph-level discourse topics. Even when there is a local (sentence) topic referent such as the crows in (29), the *-kʰwa* object marker is still used rather than Ø- to denote the discourse topic Chicken-Hawk.

(29) tàːleːlákhʷ,   meːl   tʰkaː   mìˀs   texepèˀn,   ka   c'ipìn-**kʰwa**
listen_to   crows   land   one   say_so   that   address_to-**him**

He$_i$ [Chicken-Hawk] *listened to them$_j$, the Crows$_j$ covering the land said so, that speech they$_j$ addressed to him$_i$*

[Culy 2000:78]

Culy suggests that some incorporated pronouns may be sentence topics (e.g. Navajo *bi-*; Uyechi 1991[18]) while others are topics at the level of the paragraph (e.g. Takelma

---

17 The fact that incorporated pronouns may be used to introduce discourse participants should not really be seen as surprising, since in some languages new referents may even be 'coded' by zero-anaphora (Li and Thompson (1979:318-319) for Chinese, Bickel 2003 for Belhare), in which case the hearer must reconstruct the reference entirely by inference from context. See also Schwarz-Friesel (2007).

18 Although Uyechi uses the term "Discourse Topic", it is clear from the discussion that she is talking about what Lambrecht (1994) and others have called 'sentence topics' (§2.3.1)

-*kʰwa*) or even story. Comparing the function of different dependent markers *across* languages is tricky, however. When comparing the above-mentioned analyses of Navajo and Takelma, it must be remembered that Uyechi was working exclusively with what are almost certainly elicited sentences. By contrast, Culy was using texts with no possibility of elicitation. Similarly, Culy's contention (2000:79-80) that in Bresnan and Mchombo's example (30) the corn is not topical is valid, but only if we assume (unlike Bresnan and Mchombo) that discourse topic is the relevant notion, rather than sentence topic.

(30) Fîsi   anadyá **chímanga.** Á-tá-**chí**-dya   anapítá ku San Francisco
      hyena   ate   **corn(7)**   he-SER-**it(7)**-eat he.went to S.F.
      *The hyena ate the **corn**. Having eaten **it**, he went to San Francisco.*
                                    [Chicheŵa, Bresnan and Mchombo 1987:748]

Whether researchers talk about 'sentence topic' or 'discourse topic' depends as much on the grammatical framework in which they work and the data that they use as evidence, than on any principled distinction between the two notions within a single linguistic theory. The analysis applied in Part IV avoid this problem by applying the same methodology to an *intra-language* case of variation in discourse function.

   This proposal that topicality is an independent dimension relevant for the classification of incorporated pronouns conflicts with Siewierska (2004:67), who as we noted above uses the more inclusive term "dependent marker":

> There may be different dependent markers for intra- as opposed to inter-sentential antecedents, but there do not appear to be distinct dependent markers solely for different information statuses of their referents within the discourse. *This follows largely from the fact that languages tend to have only one type of dependent person marker for a given syntactic function.* [italics mine – S.M.]

Languages which *do* have more than one type of dependent person marker for a given syntactic function are of course an ideal testing ground for the claims made by Bresnan and Mchombo, Siewierska, and others. We have already seen two languages with two types of dependent marker for the same syntactic function, Takelma (object) and Central Kambari (subject, §2.2.4.2.3)[19]. Recall that Crozier (1984) analysed person subject prefixes as showing "topic" agreement, in contrast to gender subject prefixes. These languages thus seem to provide counter-evidence to Siewierska's statements, the scope

---

19 Culy (2000) suggests a further example: the Navajo 'fourth person' pronominals are (in one of their uses) "story level topic markers", in contrast to the "sentence topic marker" *bi-*.

of which is the dependent person marker (which includes grammatical and anaphoric agreement markers).

Cicipu, of course, also has more than one type of dependent person marker for a given syntactic function, and this, as in the case of Takelma and Central Kambari, makes it possible for different series of dependent markers to be associated with different pragmatic functions. In Part IV we will see that both gender and person agreement markers are capable of taking part in anaphoric agreement, with the latter far more likely to mark discourse topics. Thus the Cicipu data provides not only counter-evidence against Siewierska's characterisation of dependent markers, but also support for Culy's contention that topicality should be regarded as an independent dimension along which incorporated pronouns can be placed.

It should be stressed that the essential distinctions in Bresnan and Mchombo's typology remain unchanged even if we remove the stipulation that incorporated pronouns must denote topics. Nothing in the theoretical introductions of Bresnan and Mchombo (1987:743-752) and Bresnan (2001:144-148) require this to be the case; the latter does not even mention topicality. The fundamental distinction between grammatical and anaphoric agreement depends on whether or not the agreement marker/ incorporated pronoun can co-occur with a subject NP.

In summary, we can distinguish between two kinds of agreement – grammatical and anaphoric – and three kinds of agreement markers – grammatical, anaphoric, and ambiguous – forming a grammaticalisation cline: anaphoric > ambiguous > grammatical. The classification of an individual agreement marker is often difficult since it may behave differently in different environments. Finally, although there are syntactic constructions (e.g. sentences with floating topics) which may require the use of an incorporated pronoun in some languages, it is not the case in general that all incorporated pronouns must have topical referents. Instead, topicality is an independent dimension along which incorporated pronouns can be classified.

2.3 Topic

We saw in §2.2.4 that when there is variation in agreement, it is often the topicality of the controller referent that determines which of the agreement choices occurs. While the general pattern is clear, in some of the analyses discussed above there is not a great deal of discussion of what a 'topic' is. Much of the recent work on agreement and agreement

conditions has been done by researchers working in the LFG framework (e.g. Dalrymple and Nikolaeva 2005), who have largely adopted Lambrecht's (1994) treatment of information structure. Consequently topic is considered to be an attribute of the clause or sentence, and the ways in which topics can develop across extended discourse is often ignored. Aikhenvald (n.d.) is an exception in that she explicitly relates the occurrence of Paumarí gender agreement on the verb to the discourse 'pivot' – in her words "what the story is about". Similarly van der Berg characterises topic as "the 'theme' or 'subject', which the message is about" (1999:163). However neither author goes into any more detail.

Other researchers seem to treat topicality as the constellation of a number of factors such as animacy and definiteness, but with no independent properties. Siewierska (1984:221-222) makes the important distinction between 'inherent topicality' and 'discourse topicality', noting that "In theory any NP may function as the topic...of the clause" but that "not all NPs possess the same inherent potential for functioning as the topic of the clause" (see also Ariel 1990 and Siewierska 2004:174ff on "entity saliency", and Givón 1976 and Comrie 1981 on topic-worthiness). Inherent topicality involves definiteness, animacy, person, and the semantic role of the referent, and is independent of how the speaker might wish the hearer to construe the situation being encoded. Discourse topicality, on the other hand, involves what Dooley calls ATTENTION management – which involves (2007:10)

> expressing the speaker's current interest in, and directing the addressees' attention to, particular concepts as a strategy for construing their comprehension in a particular way.

Dooley summarises the difference between inherent and discourse topicality as follows (2007:103):

> The above factors [including egocentrism, anthropocentrism, individuation] indicate referents that are easy to perceive as topics, hence are natural candidates for discourse topics. They predispose addressees to consider certain kinds of referents as possible topics. Nevertheless, the speaker controls the construal of discourse topicality and can choose to construe as topic any of a variety of entities...

Despite her useful distinction between inherent and discourse topicality, Siewierska's (2004) treatment of variation in person agreement is an example of how, in practice, researchers can downplay the significance of discourse when considering topicality.

After stating that most of the properties of the controller relevant to variation in agreement are those related to its inherent *and* discourse saliency, she then states that "The factors determining the inherent and discourse saliency of the controllers are those comprising the familiar topicality hierarchies" (2004:149), and gives the following five dimensions – person (1 > 2 > 3), the pronominal vs. nominal distinction, animacy (animate > inanimate), referential status (definite > indefinite), and focus (not in focus > in focus)[20]. She does not, however, mention discourse topic as an independently-varying dimension relevant to variation in agreement.

Discourse topicality will be important in Part IV when we discuss gender and person agreement in Cicipu. With this in mind, this section will discuss the notion of topicality at greater length. Tomlin et al. (1997) and Dooley (2007) provide useful reviews of various treatments of topic. Tomlin et al. (1997:83) distinguish between 'clause level theme or topic' and 'higher level paragraph or discourse theme' – the former is discussed in §2.3.1, and the latter in §2.3.2.

2.3.1   Sentence topic

Lambrecht's (1994) treatment of topic, and more generally of information structure (including definiteness, identifiability, and focus), differs from many other approaches in that it is intended to be amenable to formalism (Lambrecht 1994: xv, 9-13), and to be compatible with the theoretical assumptions common to generative theories. Consequently Lambrecht's work has been endorsed by both 'formalist' and 'functionalist' theories of syntax (e.g. Lexical-Functional Grammar – Bresnan 2001, Role and Reference Grammar – Van Valin 2005). Following Strawson (1971), Lambrecht defines a sentence topic as "the thing which the proposition expressed by the sentence IS ABOUT" (1994:118), and goes on to say:

---

20 Note that although Siewierska (1984) refers to these properties as making up 'inherent' topicality, only animacy is strictly a constant property of the referent. The others do in fact depend on the discourse context. The difference between such factors and 'discourse' factors is that the speaker has much less control over the former. Given a particular speech situation, and a particular point in a discourse, the speaker has no choice but to refer to herself in the first person. Similarly if a referent is in focus pragmatically, then the speaker must formally encode this referent using a referring expression which lies within the sentence's focal domain. The speaker has a lot more 'choice' concerning grammatical techniques of attention management e.g. the decision whether or not to use a passive construction in order to keep the topical referent in the same argument position. This choice is referred to in Dooley's quote above. Payne (1997:348) divides up the factors which determine topicality using a different cut-off point, separating (i) truly inherent properties of the referent and (ii) context-imparted properties (including definiteness and person).

A referent is interpreted as the topic of a proposition if IN A GIVEN DISCOURSE the proposition is construed as being ABOUT this referent, i.e. as expressing information which is RELEVANT TO and which increases the addressee's KNOWLEDGE of this referent (1994:127).

The following passage from Strawson (1971:97) helps to explain this idea of 'aboutness':

> …stating [i.e. making statements] is not a gratuitous and random human activity. We do not, except in social desperation, direct isolated and unconnected pieces of information at each other, but on the contrary intend in general to give or add information about what is a matter of standing interest or concern. There is a great variety of possible types of answer to the question what the topic of a statement is, what a statement is "about"...and not every such answer excludes every other in a given case.

Topics are usually subjects but this is by no means always the case. For the most typical reading of (31), the subject would indeed be the topic, but this is not true for (32), where in the first clause the fronted object is the "matter of standing interest or concern".

(31) ~he went to SCHOOL

(32) FUN I LOVE, but too much fun is of all things the most loathsome

[William Blake].

Just as topics are not always subjects, neither are subjects always topics. Consider the following:

(33) (What did the children do next?) The children went to SCHOOL.

(34) (Who went to school?) The CHILDREN went to school.

[Lambrecht 1994:121]

Only in (33) does the predicate "add information about what is a matter of standing interest or concern", and so only this example can be said to have *topic-comment* articulation. Example (34) shows rather an *identificational* articulation, since the sentence identifies a missing referent in the open proposition "X went to school". Such sentences do not have topic expressions according to Lambrecht.

It is important to note that Lambrecht's definition means that the topic of a sentence cannot be determined without considering the surrounding discourse. The proposition has to be construed as being "about this referent in a given discourse". Likewise for Strawson the topic referent has to be "a matter of standing interest or

concern". Nevertheless, while acknowledging the necessity of considering the discourse context in determining sentence topics, Lambrecht deliberately restricts his treatment to sentence rather than discourse topics (1994:117):

> I will restrict my attention to SENTENCE TOPICS or CLAUSE TOPICS. I will have little to say about the notion of DISCOURSE TOPIC, which has more to do with discourse understanding and text cohesion than with the grammatical form of sentences...

For Lambrecht, discourse topics are outside of the scope of information structure, which is limited to psychological phenomena which have morphosyntactic reflexes (1994:3):

> Even though information structure is concerned with such psychological phenomena as the speaker's hypotheses about the hearer's mental states, such phenomena are relevant to the linguist only inasmuch as they are reflected in GRAMMATICAL STRUCTURE...Information structure is not concerned with psychological phenomena which do not have correlates in sentence form.

We will see in chapter 8 that the notion of discourse topic is in fact reflected in Cicipu grammatical structure, and does have a correlate in sentence form. Specifically, inanimate and lower animate referents may trigger person (rather than gender) agreement on person markers, but only if the referent is a discourse topic (i.e. not just a sentence topic). For a more comprehensive analysis of the distribution of Cicipu agreement markers it is therefore benefical to appeal to this notion.

2.3.2 Discourse topic

2.3.2.1 Givón

Givón's (1983) volume acknowledged the importance of discourse structure in determining how 'accessible' referents are at any particular time, and in it he and his collaborators sought to correlate two sets of variables (1983:13):

1. The grammatical, 'purely linguistic' devices used by the speaker to code various topics/participants in the discourse

and

2. The exact position of those topics in the discourse, in terms of thematic paragraph structure, distance from last appearance, the clustering with potential other interfering topics, persistence in subsequent discourse topics.

In one way this is similar to the task I will try to perform in Part IV − to correlate linguistic coding devices (lexical NPs, gender agreement and person agreement) with

properties of the discourse. Note however that all of the discourse properties considered by Givón are formal – they are properties of the text itself rather than of the conceptual structures that are involved in the production and comprehension of the text. Givón is of course aware of the importance of such conceptual structures (1983:12):

> In particular, we have attempted to assess the more concrete and readily measurable factors (a) ['**length of absence from the register**'] and (b) ['**potential interference from other topics**']. The fact that is is not yet possible to quantify rigorously factors (c) ['**availability of semantic information**'] and (d) ['**availability of thematic information**'] in spite of their undeniable importance creates a degree of indeterminacy in the results, so that correlations between grammatical devices and particular measurements appear to be less than categorical.

It is still "not yet possible to quantify rigorously" the availability of thematic information at a given stage of a discourse, and it is difficult to imagine how this could ever be done. Nevertheless topicality cannot simply be reduced to referential density – referential density is neither a necessary nor sufficient condition for a referent to be a topic, at least for a conceptual definition of topic involving 'aboutness'. On the one hand, discourse topics may integrate a paragraph (in the sense that the sentences in the paragraph are construed to be related to that topic) without the topic being explicitly mentioned, for example when situations are reported from the point of view of one of the discourse participants (see Duchan et al. 1995). On the other hand, a high degree of referential density does not make that referent a topic, as we will see shortly.

2.3.2.2 Dooley

Dooley (2007) attempts to bring together insights from several different theories, including Cognitive Linguistics (e.g. Langacker 2001, Fauconnier 1997), various other functional-typological approaches (Lambrecht, Chafe, and other authors including Givón and Tomlin), and cross-linguistic discourse analysis (e.g. Grimes 1975, Longacre 1996, Dooley and Levinsohn 2001). Discourse topics in Dooley's framework are defined according to their function in a DISCOURSE SPACE[21], which is the mental representation the addressee tries to construct, incrementally and through trial-and-error, in order to make sense of what he perceives the speaker's message to be. Dooley assumes that, amongst others, the following elements are always contained in a

---

21 The term 'discourse space' evokes Fauconnier's work on mental spaces (e.g. Fauconnier 1994), but Dooley's term is more inclusive.

discourse space (2007:8):

- Referential entities and propositions relating them
- Activation states (as defined by Chafe, see below)
- An overall structure that accounts for the conceptual unity of the different components of the space (called a DISCOURSE SCHEMA)
- Themes (including topics)

Dooley hypothesises that the spaces constructed by an addressee in response to a section of discourse correspond to DISCOURSE UNITS, the lowest-level of which is the PARAGRAPH (to be understood independent of modality). He defines topic as follows (2007:71):

(35)   If a discourse unit is construed in such a way that its [discourse] space is thematically integrated around a referent[22] – that is, if the relevance of each of the steps in its schema is perceived as depending on its relation to that referent and if that relation manifests as well an *intrinsic interest* in that referent on the part of the speaker – then the referent is called the TOPIC of the discourse unit [my italics].

But what exactly does it mean for a discourse space to be "integrated around a referent"? When an entity is introduced into a discourse, it brings with it (or, in Chafe's (1994) terms, makes semiactive) its DOMINION, defined by Langacker (2000:173-174) as the "conceptual region (or the set of entities) to which [a particular referential entity] affords direct access". Dooley (2007:33) notes that:

…if the entity continues to be referred to throughout a sizeable section of a discourse, the concepts in the dominion, including new ones as they are added, tend to be viewed in relation to the referential entity. In this way the growing dominion comes to be integrated by the referential entity: its component elements are viewed in relation to that entity.

Tomlin (1997:89) puts it like this: "Generally, in a stretch of connected discourse, one referent emerges as central, or the one that the propositions in the discourse are about." I have found Dooley's figurative illustration a helpful one – "the referent begins to 'attract' other concepts in the discourse space somewhat as a magnet attracts iron filings, so that they 'point in its direction'" (2007:71).

Dooley's characterisation of topic is similar to Lambrecht's in some ways, and it is useful to compare the two theories and the predictions that they make. The 'aboutness' relation is present in both approaches, but the scope is different, as shown in Figure 20.

---

22 This definition differs from Chafe's (1994:121), in that it is restricted to referential entities. Chafe's discourse topics (and Asher's e.g. 2004) correspond to what Dooley calls 'themes'.
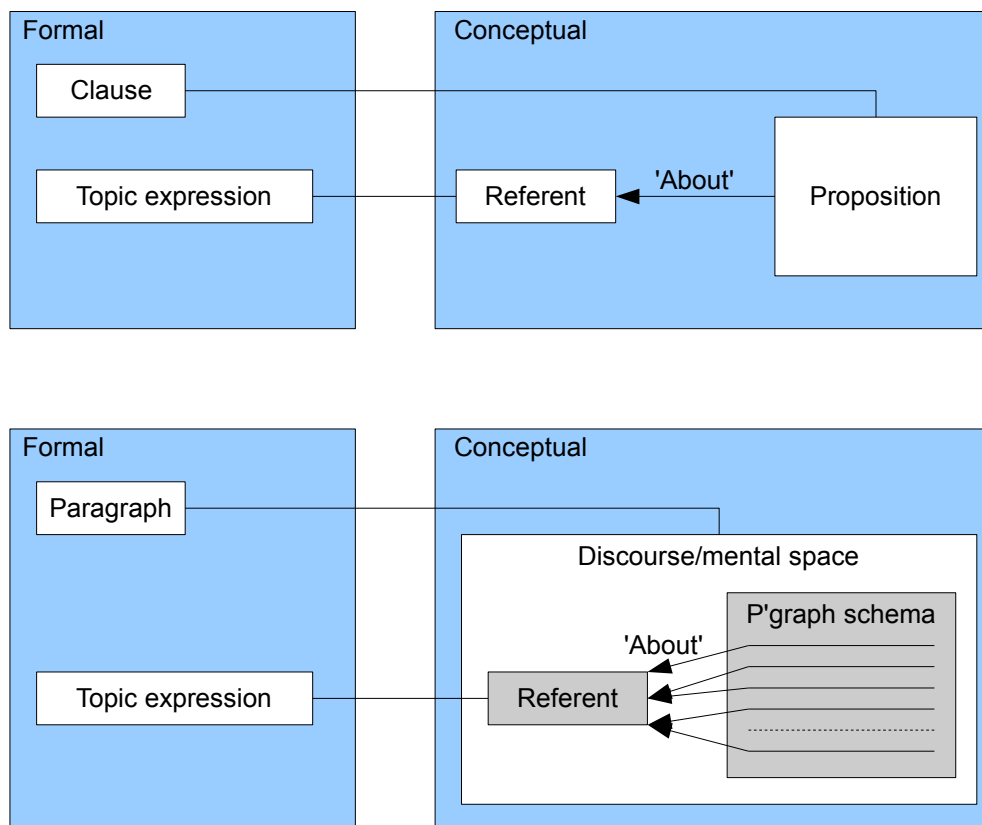
Figure 20: Informal comparison of Lambrecht's sentence topic (above) and Dooley's discourse topic (below)

According to Lambrecht, a referent is a topic if the *proposition* encoded by the *clause* is construed as being 'about' that referent. For Dooley a referent is a topic if the *conceptual schema* that we have to construct to process the *paragraph* is integrated by that referent (i.e. each step in the schema is construed as being 'about' that referent). The two approaches differ as to the status of referents with a short textual span. Consider (36) below where the grandma is only a transient participant in the discourse.

(36)　[Context: A Jewish grandfather has been talking about the fact that his grandson is difficult to please. He gives one example – outmeal]

And it's uh got good taste, its good. And the cereal – **grandma e** don't like cereal but **she** finished to the last (dish) and I enjoy – I like it too. It's tasty! And I uh...He didn't want the cereal, doesn't eat. I said "Todd, it wouldn't kill ya, taste it!"

[Lambrecht 1994:149]

The extract as a whole is clearly about cereal, and more generally the complaining grandson. Lambrecht also analyses the grandma as a topic because the two propositions concerning 'not liking cereal' and 'finishing it all up' are about her. However for Dooley

95

the grandma would not be a topic, because she fails to integrate any textual span wider than a single sentence. We will see in chapter 8 that sentence topics of this kind are much less likely to be indexed with person agreement than true discourse topics.

In Dooley's definition of discourse topic given above, the additional qualification that the speaker has an 'intrinsic interest' in the topic referent is crucial. While talking about something does increase the likelihood of that referent becoming a paragraph topic, it is not sufficient. Consider the following example of a newspaper report on a game of American football[23]:

> ...Hoffman caught **the ball**. He passed **it** to King, who punted **Ø** for fifty yards. Sedgwick caught **Ø**, and on the kick-off Stacy got **the ball**.
>
> *The New York Times*, 20th October 1895

This text consists of a sequence of five clauses, each of which contains a reference to the football. Moreover, one referential expression is an unstressed pronoun, and two of the references even involve 'sub-minimal' (i.e. null) coding. Nevertheless the narrative could not conceivably be construed as being 'about' the ball. The ball semantically integrates the text but does not thematically integrate it. According to several theories of topic (including Givón 1983, see also Dooley 2007:19-22) the ball in this example would qualify as topical. If we want to exclude such referents, then it must be recognised that discourse topicality involves more than just referential density and semantic integration. The speaker must also have an intrinsic interest in the topic referent.

2.3.2.3   Formal signals of discourse topic

Dooley notes that there is both conceptual and formal evidence for discourse structure, and that "The higher the level of discourse organization, the more the addressee can expect to depend on conceptual rather than formal signals" (2007:55). Thus Dooley is in agreement with the quotation from Lambrecht given above, but it is important to note that there are exceptions to this trend. Chafe (1994:88-89) presents evidence for the linguistic relevance of what he calls "referential importance" i.e. the "importance [of referents] to the subject matter being verbalized" (1994:88). According to Chafe, referents of "primary importance" are typically mentioned more frequently and may be

---

23 See Tomlin et al. (1997:63) and Dooley (2007:21-22) for discussion of a similar text involving an ice-hockey game. My interpretation of these kinds of passages with respect to topicality follows Dooley rather than Tomlin et al.

introduced into the text in a specific manner, especially for more structured genres such as traditional narratives. I have already mentioned Takelma and Navajo (§2.2.5.1) where special incorporated pronouns have been said to mark discourse topics. Grimes' (1978) volume contains a number of papers on 'theme oriented referential strategies', which some languages use to

> manage reference in terms of a **thematic** policy, in which one referent is distinguished from the rest when introduced, and a special set of terms refer to it no matter how many other things have been mentioned more recently (Grimes 1978: viii, emphasis in original).

Dooley and Levinsohn (2001:119-123) call this the 'VIP' (very important participant) strategy. They note that the strategy can extend across entire texts:

> ...patterns of reference sometimes make it necessary to recognize one [participant] as GLOBAL VIP. After being introduced, the global VIP is often referred to by minimum, but virtually constant, coding (2001:121).

The Bantoid language Mambila uses such a strategy in folktales (Perrin 1978). The main participant, once introduced, is referred to either by zero (when functioning as subject) or by the 3PS pronoun *bù* otherwise, with lexical re-mentions being very rare. The potential for ambiguity is minimised because, with a few exceptions, "Participants other than the main one are re-identified by a noun every time they are mentioned" (Perrin 1978:111). This 'overcoding' of non-topical participants might be though surprising, but the same strategy is used in at least two other African languages, the Grassfields Bantu language Babungo and the Omotic language Gimira (Siewierska 2004:182-183). Dooley and Levinsohn (2001:122) also analyse the thematic 'pivot' in some of the Brazilian Arawá languages as a VIP (see §2.2.4.2.2 on Paumarí). In chapter 8 I will argue that the Cicipu gender/person agreement alternation cannot be explained without an appeal to discourse topicality, and that conversely progression from gender to person marking can be viewed as a signal of discourse topicality.

2.3.2.4 Paragraphs

The definition of topic in (35) relies on the notion of 'discourse unit', of which a paragraph is the lowest-level variety. Discourse analysts are generally agreed that the production and comprehension of texts involves the construction and recognition of structural units larger than a single sentence, yet smaller than the whole text. According to Tomlin et al. (1997:66):

Discourse is neither flat nor linear in its organization; it is hierarchical, with clauses forming higher-order structures, paragraphs, which in turn combine to form larger episodes or sections of discourse.

Evidence for the existence of the paragraph can be either formal or conceptual. Formal cues of paragraph structure include high pitch at the start of the initial sentence (Lehiste 1979), longer pauses at paragraph boundaries (Chafe 1990), and the pre-boundary presence of laryngealisation (Kreiman 1982). The existence of the paragraph is also supported by at least three kinds of conceptual evidence which converge on the paragraph as a minimal unit of discourse for which the speaker can construct a coherent discourse space (see discussion in Dooley 2007:30).

Firstly, a paragraph is "a unit of speech or writing that maintains a uniform orientation" (Hinds 1979:136). Chafe (1994:128-9) identifies four orientation dimensions (see also Van Dijk 1981:177, Givón 1983:36 fn. 4):

The fact that consciousness cannot function without being oriented in space, time, society, and ongoing background events explains the characteristic provision of what is usually called a *setting* as a narrative begins. (Chafe 1994:129)

Orientation dimensions are frequently updated at the start of a new paragraph – sometimes this boundary is explicitly indicated with an initial constituent functioning as a mental 'space-builder' (Fauconnier 1997:40-41), as in "The following day,...".

Secondly, experimental evidence (e.g. Gernsbacher 1985) has shown that the ability of hearers to recall the 'surface information' (e.g. word order) of sentences is adversely affected once they have crossed a paragraph boundary. Gernsbacher (1984) demonstrated the same effect on subjects' recall of thematic information, and her explanation is that the mental representation of the previous paragraph quickly decays (or perhaps is suppressed) as a result of the "processing shift" involved in setting up the new discourse space for the new paragraph. Chafe (1994) expresses this change in terms of the ACTIVATION STATUS of entities. An ACTIVE concept is one which is presumed to be in addressee's consciousness. A SEMIACTIVE concept is one that is in the addressee's current mental representation, but whose activation level has decayed (the notion of activation status is a scalar one). Finally, an INACTIVE concept is one that is newly introduced by an utterance. A consequence of the "processing shift" identified by Gernsbacher is that previously active or semi-active referents become less active once a paragraph boundary is crossed, and therefore require extra linguistic coding (see below). Chafe (1994:138)

suggests that "semiactive consciousness is limited to the amount of information verbalized in a narrative schema, so that when more information is added it may have to be divided" i.e. into paragraphs.

Finally, there is also evidence for paragraph boundaries from experiments which test metalinguistic judgements. It has often been shown (e.g. Lehiste 1979, Gernsbacher 1985, Ji 2002) that subjects perceive paragraph boundaries in a stretch of discourse, and that their judgements concur to a certain degree with each other. These divisions are correlated with boundaries established by independent means (e.g. prosody in Lehiste 1979, availability of surface information in Gernsbacher 1985).

2.3.2.5 Coding weight

Givón (1983) and others (e.g. Ariel 1990) have observed that as the accessibility of a referent (and hence its availability as topic) decreases, so the 'coding weight' of the referring expression increases. Givón (1983:18) gives the following scale[24]:

more accessible topics [light]	zero anaphora
↕	unstressed/bound pronouns ('agreement')
	stressed/independent pronouns
less accessible topics [heavy]	full NP's

Figure 21: Givón's (1983) phonological coding weight scale

At any given point in a discourse, minimal coding (i.e. the minimum that will enable the hearer to identify the referent) is the norm. 'Undercoding' – the use of less than sufficient coding weight – is, unsurprisingly, generally infelicitous, and it does not generally occur except for stylistic effect.

Overcoding can occur for at least three reasons[25]. The first relates to what Dooley (2007:53, fn. 55) calls 'syntactic opportunism' – the syntactic slot for a referential expression provides an ideal opportunity to shoehorn in some additional information about the referent. In such cases the coding choice has a secondary discourse function which has little or nothing to do with the accessibility of the referent. Secondly, overcoding may occur when there are competing referents and therefore danger of confusion (as noted by Givón 1983:14). Thirdly, overcoding may occur at the start of a

---

24 This is just the phonological scale. Givón (1983:18ff) identifies several different scales – see also Ariel (1990) and Siewierska (2004:174ff).
25 See Bolinger (1979) for others beyond those discussed here.

new paragraph. Dooley (2007:62; see also Li and Thompson 1979, Anderson et al. 1983) notes that:

> Nominal elements, such as participants in a narrative, are also generally updated by means of a noun phrase at the beginning of a new paragraph, even if recent mention in the preceding paragraph would indicate that they are active or semi-active.

Really, however, only the first of these should be considered to be true overcoding, since in the second case while the referents concerned may both be highly-accessible, neither is sufficiently more accessible than the other to allow the use of minimal coding. The overcoding in the third case can be explained if we assume, as Gernsbacher's research suggests, that the activation statuses of referents are reset at paragraph boundaries. This can be looked at in two ways – either the fact that we have a new paragraph (signalled by the other formal and conceptual cues we have noted above) gives rise to overcoding, or as Fox (1987:168) says:

> many full NPs which occur in narratives where one could have expected pronouns are functioning to signal the hierarchical structure of the text... to demarcate new narrative units.

It is important to be clear that, within a paragraph, referring expressions do not always move straight from an NP to minimal coding. There may be a coding 'progression' within the paragraph through several stages of the coding weight hierarchy. According to Dooley (2007:99):

> Especially for high-level topics, their establishment as topic commonly goes beyond the initial introduction per se. For some topics, the introduction sentence is followed by another sentence with an overcoded reference to the topic....

He gives examples from Koine Greek and Mbyá Guarani involving a progression from lexical NP > demonstrative pronoun > minimal coding. Similarly Hinds (1984:466) indicates a similar sequence for topic establishment in Japanese:

> We see a three step progression in the identification of participants in a narrative: (1) the participant is introduced with the particle *ga*; (2) the participant is referred to with a topical noun phrase marked with *wa*; and (3) the participant is referred to by ellipsis.

Similar progressions seems to exist in Chinese (from lexical NP to pronoun to ellipsis – Li and Thompson 1979), the Tibeto-Burman language Jirel (Maibaum 1978), and the Adamawa language Longuda of northeastern Nigeria (Newman 1978). This idea of a

coding progression will be important for the analysis of Cicipu in chapter 8 when we turn to the function of gender and person agreement in discourse.

Changes in coding weight as a discourse progresses therefore provide important evidence for paragraph structure. Of course, paragraph structure does not have a *deterministic* relationship with coding progressions – as we have just noted, there are other reasons for overcoding and undercoding. In particular, the use of minimal coding for a referent within a paragraph is not sufficient evidence for it to be analysed as a discourse topic. As Dooley (2007:51) points out, minimal coding is also used for non-topical 'recent-reference centres of attention' (e.g. (27) above). Some of Lambrecht's 'secondary topics' (1994:147-150) are simply recent-reference mentions, rather than discourse topics, and in fact this notion is necessary to prevent the proliferation of topics that are not "matters of standing concern".

2.3.2.6 Access and integration functions of topics

Dooley (2007:72-73) points out that topics have both ACCESS and INTEGRATION functions, something that is not always appreciated in the literature. The first mention of a topic provides *access* to the mental space required for comprehending what is said about it; at this point in its "life history" (Langacker 2000:194) the topic has "higher informational salience, reflected in formal signals: heavier coding weight, its own intonation contour, being set apart linearly from its 'focus domain'" (Dooley 2007:77-78). Dooley calls such topics 'marked topics'[26]. Once a topic has been introduced, then its function is to *integrate* the discourse space, just as was described above. When carrying out this function topics are typically 'unmarked', and expressed by minimal coding.

2.3.3 Summary

In Part IV we will observe the usefulness of the notions discussed in this section for describing gender and person agreement in Cicipu. The linguistic coding of topics in Cicipu discourse reflects paragraph structure in two ways. First, as in many other languages (perhaps all), the resetting of the activation level at the beginning of each paragraph is reflected by increased coding weight. Secondly, and more interestingly, there is an intra-paragraph coding progression from lexical NP, to gender agreement, to

---

26 Givón (1983:9) makes a distinction between 'chain-initial' topics and 'chain-medial' topics. This is a formal correlate to Dooley's conceptual distinction between the access and integrating functions. Lambrecht's (1994:131-136) use of the terms 'marked' and 'unmarked' topic is quite different and not to be confused with Dooley's.

person agreement. We will see that the notion of discourse topic accounts for the distribution of the two kinds of agreement markers in a way that sentence topic cannot. Furthermore it will be shown that intrinsic interest is important as opposed to mere referential density – it is not the case that a referent will become topic-marked simply by mentioning it enough. Instead the speaker must have an intrinsic interest in that referent.

I will finish this chapter with a quote from Dooley (2007:56) which bears on the nature of the implication of this study for linguistic theory.

> The claim that discourse has hierarchical organization which is (partially) signalled by linguistic means does not imply that this organization is part of grammar... There may be and commonly are formal signals which help the addressee recognize discourse units, but in general discourse structure is not susceptible to judgements of 'grammatical vs. ungrammatical' in the same way that lower-level phenomena are. Judgements of discourse organisation are more often in such terms as 'clear vs. unclear or confusing', 'well-put vs. clumsily-put', etc... If judgements of 'grammatical vs. ungrammatical' reflect grammatical well-formedness, then in large part discourse organization is not a matter of grammar. So the formal signals we commonly observe in discourse must often be seen as linguistic correlates of what is fundamentally conceptual, hence extra-grammatical, structure.

To anticipate the findings of Part IV, the use of anaphoric gender agreement where anaphoric person agreement would have been more appropriate, or vice versa, is not in general "ungrammatical", but perhaps should be seen as "unclear" or "clumsily-put". The alternation between gender and person agreement cannot be fully explained by the 'grammar' of a language in the way that this term is understood in generative linguistics, and so in this case attempts to look for, say, an explanation in terms of Lexical-Functional Grammar by appealing to i-structure (King 1997) or g-structure (Falk 2006) would be misconceived. Grammatical frameworks in the generative mould such as LFG, HPSG, Minimalism, and so on, are not intended to deal with this kind of discourse phenomenon (see Newmeyer (1998:42-43) for a critique of functionalist researchers who criticise generative approaches for failing to account for this kind of data). If the alternation described in Part IV is considered to be interesting and worthy of explanation by some kind of linguistic theory, then we must look for that explanation outside of theories of autonomous syntax.

2.4   Chapter summary

In this chapter I set out the theoretical context relevant for Parts III and IV. In §2.1 I

discussed both the Africanist tradition of noun class studies and more mainstream work on grammatical gender. Section 2.2 was concerned with agreement, in particular variation in agreement and typologies of agreement markers. Finally §2.3 introduced the notion of discourse topic and various associated concepts.

Part II – Phonological and grammatical sketch

Chapter 3 – A phonological sketch of Cicipu

Cicipu has a mid-sized phonemic inventory consisting of 27 consonants and 6 vowels according to the analysis presented here, although nasalisation is contrastive for all vowels, and length is contrastive for all vowels *and* all consonants. The presentation of the phonemic charts is deferred until after the initial analysis of syllable and word structure in §3.1, which then informs the discussion on consonants (§3.2) and vowels (§3.3). The next three sections deal with three important suprasegmental topics in Cicipu: tone (§3.4), vowel harmony (§3.5), and nasalisation (§3.6). Finally (§3.7) I discuss some of the more important morphophonemic processes. For reasons of space most topics are dealt with only cursorily – as with the grammatical sketch in chapter 4 the aim is to familiarise the reader with the phonology of Cicipu, as well as to provide a point of departure for more detailed research in the future.

3.1   Syllable and root structure

I will first consider the unambivalent syllable structures found in noun and verb roots (§3.1.1), before turning to the more complex ambivalent cases (§3.1.2-3.1.6). Section 3.1.7 deals with prefixes and ideophones, which allow extra syllable types, and in §3.1.8 I will look at the structure of nominal and verbal roots.

3.1.1   Unambivalent syllables

The only unambivalent syllable types in Cicipu noun and verb roots are CV and V, although there is a strong case for admitting CVN[1] non-finally. Examples (1-2) show CV syllables in noun roots:

(1)   kà-**kúlù**   à-**kúlù**
        NC1-**hailstone**   NC2-**hailstone**
        hailstone   *hailstones*

                                                    [eamy003.1337]

(2)   s-**síró**   ì-**síró**
        NC8-**mane**   NC3-**mane**
        mane   *manes*

                                                    [eamd020.1033]

The following example shows a CV.CV verb root (also **sita** 'swell', **naha** 'leave' and

---

1   Standard abbreviations are used when referring to syllable types: C(onsonant), V(owel), and N(asal) consonant.

many others).

(3)   Ø-**kábà**
2S-**took**\RLS
you (sg.) took

[saim001.109]

V syllables may be detected root-initially because they take different sets of prefixes, and also because the root-initial vowel coalesces with the prefix vowel (a special case of a more general process − see §3.7.1). The following examples show V.CV roots, both nominal (4-5) and verbal (6-7). Note that the **vi-** prefixes in (5) and (7) could not have been predicted from the corresponding prefixes that occur before consonants in (2-3).

(4)   [kó:sì]   [ó:sì]
ko-**ósì**   o-**ósì**
NC1-**eye**   NC2-**eye**
eye   *eyes*

[eamd004.022]

(5)   [vɔ́:mɔ̀]   [jɔ́:mɔ̀]
vi-**ɔ́mɔ̀**   yi-**ɔ́mɔ̀**
NC8-**monkey**   NC3-**monkey**
monkey   *monkeys*

[eamd020.1001]

(6)   [mǎ:jà]
mì-**áyà**
AG5-**come**\RLS
they came [i.e. **ǹ-'yɔ́'yɔ́** 'fish(pl.), NC5' came]

[taff001.002]

(7)   [vǔ:wà]
vì-**úwà**
2S-**hear**\RLS
you (sg.) heard

[svtmg001.338]

V syllables are restricted to root-initial position, and even there they do not usually surface as such. If there is no prefix with which the root vowel can coalesce, a dummy approximant is added (e.g. §4.6.2.3 for the imperative).

Vowel-initial roots are rare in Cicipu (49 out of 950 nouns in the lexicon, 10 out of 471 verbs), and comparative evidence suggests they may be derived historically from consonant-initial roots, especially those beginning with an approximant.

106

CVC syllables are only possible in restricted circumstances (see §3.1.3, §3.1.5, and §3.1.7). Borrowed words with CVC syllables in the source language (invariably this is Hausa) are pronounced with a transitional schwa vowel as in [kà-húsᵊkà] 'face' from *fuska*, since CVC syllables are generally unacceptable in Cicipu.

3.1.2 Labialisation and palatalisation

A number of consonants can be labialised or palatalised:

(8)	[ù-**k**ʷárí	mɔ̀-ʔʲɔ́ʔʲɔ́	ù-ʔʷâː	ù-**h**ʲẫː]
	ù-**kw**árí	mɔ̀-**'y**ɔ́**'y**ɔ́	ù-**'w**âa	ù-**hy**ẫa
	NC7-next_year	NC4-fish	3s-pass\RLS	3s-say\RLS
	next year	*fish*	*he passed*	*he said*

There is evidence from the allomorphs of the conjunction **ǹ** 'and' (§4.4.5.4) that these should be considered single consonants. The **ǹ** allomorph occurs before short consonants, while before long consonants or consonant clusters we find **nì**. Labialised consonants pattern with short consonants, as shown by [ǹ kʷándɛ́i] 'with the dry season'. There are only six such consonants in Cicipu (/kʷ gʷ ʔʷ ʔʸ hʷ hʸ/), and so the decision to treat them as single phonemes does not greatly increase the phoneme inventory.

3.1.3 Long consonants

Long consonants occur word-initially in NC8 nouns e.g. **z-zá** 'person', **k-káa** 'woman', and word-medially in a few verbs e.g. **latta** 'sleep', **tanna** 'descend'. These are 'true' geminates (Blevins 2004:169) and should be treated as single phonetic segments on the timing tier – they never have an intervening epenthetic vowel (unlike the Hausa borrowings mentioned above), and when complex segments such as affricates are lengthened the resulting sound consists of a single long closure followed by a single frication period, rather than a repetition of the short version[2].

Nevertheless syllables ending in a geminate pattern with other heavy syllables (i.e. CVV and CVN), for example in the habitual (§4.6.3.3) tone pattern, which depends on the number of mora in the verb stem to which the habitual affix **si-** attaches. If the verb stem is CVCV (bimoraic), then the tone pattern is L L H H as in (9).

---

2   An additional test would be to check that the long consonants in verbs such as **latta** and **tanna** do not become split by the addition of the causative (§4.6.4.1) or pluractional (§4.6.5.4) infixes. In the words tested so far, geminate integrity is maintained (e.g. **kullo** 'burn', **kull<is>o** 'cause to burn').

(9) ù-sì-tá'á   CVCV   L L **H H**
     3s-HAB-want
     He wants

However for CVNCV, CVVCV, and CVGV[3] verbs the pattern is L L H(L) L:

(10) ù-sì-pɔ́ǹtɔ̀   CVNCV[4]   L L **HL L**
      3s-HAB-clap
      He claps

(11) ù-sì-wîinà   CVVCV   L L **HL L**
      3s-HAB-sell
      He sells

(12) ù-sì-hɔ́ttɔ̀   CVCCV   L L **H L**
      3s-HAB-warm_oneself_by_fire
      He warms himself by the fire

Furthermore when there is a long consonant in C_2[5] it closes off the previous syllable, so that V_1 is shorter than it would be in an open syllable. Therefore a word such as **hɔttɔ** in (12) should be analysed as CVC.CV rather than CV.CV, and thus we admit the syllable pattern CVG as well as CV.

GV is also required to cater for the word-initial geminates mentioned above. The internal structure of syllable-initial geminates has been contested in recent decades (Clements and Keyser 1983, Hayes 1989, Hume et al. 1997, Davis 1999), the main issue being whether or not they are 'long' (bipositional) or 'heavy' (monopositional and moraic). The evidence from Cicipu is puzzling and requires further investigation – word-initial geminates behave like consonant clusters in that they trigger epenthesis/reduplication (§4.4.5.4, §5.5.8), yet they also behave like single consonants with respect to phonotactic constraints (CCV with two different consonants is not allowed) and minimality constraints (§5.5.7).

3.1.4 Long vowels and diphthongs

In contrast to long consonants, long vowels in Cicipu are not subject to 'geminate integrity' (Perlmutter 1995). This can be seen from the dependent imperfective

---

3 G = geminate.
4 A few CVNCV verbs follow the 'light' pattern here e.g. **yinda** 'see', **panda** 'forget', and **kanda** 'mark'. This may be lexical idiosyncrasy, although the Cicipu verb system is otherwise highly regular. Alternatively it may reveal something about the phonological representation of the NC cluster (i.e. underlyingly /NC/ or a prenasalised /C/ – see §3.1.5 on prenasalisation).
5 C_1 = first consonant in the root, V_1 = first vowel in the root, and so on.

(§4.6.3.4) form of the verb, where the final stem vowel changes to **i**. If the verb stem is monosyllabic then the vowel is diphthongised[6] and the tone is rising:

(13)   ǹ   wàyí,   nùutǔ̃í
       ǹ   w-aya-í   ǹ-u-tũu-í
       and   3s-come-**DEP.IMPF**   and-3s-pour-**DEP.IMPF**
       he was coming, he was pouring

This is evidence that long vowels are bipositional (i.e. a sequence of vowels rather than one long vowel), since if there was only one root node linked to two timing slots (or two moras) we would not expect one half of the vowel to change independently of the other. However it does not allow us to decide between a heterosyllabic CV.V or a tautosyllabic CVV analysis.

    Historically, it seems likely that long vowels in Cicipu are derived from the coalescence of two syllables with the disappearance of the intermediate consonant, as was suggested for vowel-initial roots in §3.1.1. Many of the long vowels in Tirisino have cognates in other dialects and languages where such a consonant remains, and vice versa.

    Diphthongs differ from long vowels (and CVyV/CVwV sequences) in that their duration is not noticeably longer than short vowels, and there is no 'dip' in the waveform. When words are broken down into syllables by native speakers then the diphthong is pronounced as part of one syllable. Therefore diphthongs are considered to be a single vowel with regard to syllable structure.

3.1.5   Prenasalised stops and affricates

In addition to CV and CVG (§3.1.3), it is common to find what appear to be CVC syllables, as in [kò.dṍn.tú] 'stool'. However they always have a nasal as the coda, and this nasal only occurs *after* nasal vowels and *before* oral stops or affricates, as in (14).

(14)   [kòdṍ**n**tú   kàbṹ**ŋ**gú   kũ̀**mb**á!   kõ̀**nd**ó!]
       kò-dṍtú   kà-bṹgú   kũ̀bá!   kõ̀dó!
       NC1-stool   NC1-snake   climb\IMP   enter\IMP
       stool   *snake*   *climb!*   *enter!*

Prenasalisation affects all oral stops and affricates, with the exception of the glottal stop and its palatalised and labialised variants **'y** and **'w**. *Phonetically* it seems clear that two

---

6   In Tirisino. In Tikula long vowels change completely to [ĭː] in monosyllabic dependent imperfectives.

distinct segments are involved – the two parts can contrast in voicing (e.g. [kòdṍntú]), and the nasal component is typically much longer in duration than the preceding vowel.

However the *phonological* status of these prenasalised consonants is problematic. The clearest indication that the nasal component is absent from the underlying representation comes from the distribution of prenasalised consonants. Since Cicipu has both oral and nasal vowels, and nasal vowels do not occur directly before non-glottal stops, it is possible to regard the nasal phone in words such as [kòdṍntú] as conditioned by a combination of the preceding vowel and the following consonant.

Although most instances of prenasalisation are found root-internally, the process occasionally occurs across morpheme- and even word-boundaries:

(15) [mṹŋgʷàːnùkʷà   tʃé]
      mú-u-gwàanùkwà   cé
      1s-FUT-see\IRR   NEG
      I wouldn't know

                                                                    [sayb001.102]

(16) [kòɓ́ɔ̃ːŋ kʲè]
      kò-ɓ́ɔ̃ɔ   k-è
      NC1-axe AG1-COP
      it's an axe

                                                                    [eamd002.054]

We therefore have to assume the existence of a prenasalisation rule to account for these examples, quite independently of considerations of syllable structure in roots. The rule could then be re-used to account for root-internal prenasalisation.

There is some problematic data (e.g. the paucity of NC clusters with a voiceless C, suggesting that the C is assimilating to the N in terms of [VOICE]), but the data largely supports the claim that prenasalisation is a phonological process, and that [ṼNC] sequences are underlying /ṼC/, with the nasal segment supplied predictably[7].

Nevertheless the resultant nasal segment (which, as mentioned above, can be quite long) contributes to syllable weight in weight-sensitive processes. Two such cases are briefly discussed here. Firstly, long nasal vowels do not seem to trigger prenasalisation. There are only three examples of long nasal vowels preceding a consonant in the corpus, but they pattern consistently – none of them have a nasal intervening between

---

7   Other than in this section, the examples in the thesis are written with the nasal e.g. **kò-dóntú**. Due to the complexity of the data, particular care should be taken over any orthographic decisions relating to prenasalisation.

the vowel and the stop.

(17) kà-hhűűcí   kà-tíítíí   kɔ̀-gʷɔ̃́ɔ̃gʷɔ̃̂ɔ̃
        NC1-cloud   NC1-foreskin   NC1-crow
        cloud   *foreskin*   *crow*

There does not seem to be any obvious reason why long vowels should not trigger prenasalisation, unless the resultant nasal forms the coda of a CVVC syllable. In this case the restriction would be simply a matter of syllable weight: it is not unusual for languages to have special restrictions on 'super-heavy' CVVC syllables.

Secondly, the tone patterns on verbs are sensitive to syllable weight. The habitual tone pattern has already been mentioned in §3.1.3; recall that verbs with ṼNC sequences, as in (10), generally pattern with other 'heavy' syllable patterns such as CVVCV and CVCCV. Similarly, in the case of the realis tone pattern verbs with ṼNC sequences again pattern with other heavy syllables by taking a falling tone on the first root syllable rather than a high tone:

(18) ùbánà   ùkóò   ùkɔ̋ǹdò
        u-bana-LHL   u-koo-LHL   u-kondo-LHL
        3s-invite-RLS   3s-die-RLS   3s-enter-RLS
        he invited   *he died*   *he entered*

In summary, prenasalisation in roots is a predictable phonological process, but the resulting nasal segment is both longer in duration than might be expected and contributes to the weight of the syllable in weight-sensitive processes.

Before leaving the topic of prenasalisation, it should be noted that a small number of noun roots begin with an NC sequence, where C is a non-glottal stop:

(19) má-**nd**á   mí-**nd**á
        NC4-calabash   NC5-calabash
        calabash   *calabashes*

(20) wú-**nt**ò   ví-**nt**ò
        NC7-guest_hut   NC8-guest_hut
        guest hut   *guest huts*

The NC clusters in these words should not be considered products of the prenasalisation process, at least not synchronically.

3.1.6 Approximants

To avoid proliferation of syllable types, the ambivalent vocoids [i] and [u] are analysed

as consonants when they occur in onset position, and as vowels when they occur in nucleus position.

3.1.7   Prefixes and ideophones

The preceding discussion was concerned with only noun and verb roots, which may contain CV and V syllables (and, it was argued, CVN and CVG). Other word classes have different possibilities, in particular prefixes (nominal or verbal) and ideophones.

   Nominal prefixes and agreement prefixes are all monosyllabic, and are usually of the form V, CV, or N. The exceptions are the interesting **C-** (consonant-lengthening) allomorphs of the NC8 and 2s prefixes. The application of these prefixes results in the lengthening of the first consonant of the root, whatever this consonant happens to be. Any consonant can be lengthened in this manner; examples are given for NC8 nouns (21) and 2PS verbs (22).

(21)	**z**-zá	**k**-káa	**c**-cɔ́'ɔ̀	**s**-síró
	NC**8**-person	NC**8**-woman	NC**8**-sheep	NC**8**-mane
	person	*woman*	*sheep*	*mane*

(22)	**t**-tá'à	**'**-'ɔ́pɔ̀	**l**-láttà	**j**-jântà
	2s-want\RLS	2s-hold\RLS	2s-sleep\RLS	2s-crush\RLS
	you (sg.) want	*you (sg.) held*	*you (sg.) slept*	*you (sg.) crushed*

[eamy036.001, tats005.001.030]

We saw in §3.1.1 that the first vowel of a vowel-initial noun or verb root coalesces with the preceding prefix. If there is no prefix, as in the case of imperatives, the root is obligatorily preceded by a **w-** or **y-** 'dummy' consonant (§4.6.2.3). Vowels at the beginning of prefixes behave slightly differently. Utterance-initially they are pronounced with a (non-phonemic) preceding glottal stop, rather than a dummy consonant. Utterance-medially, prefix-initial vowels usually coalesce with the preceding vowel, although a glottal stop may be inserted, even word-internally between two prefixes (see §3.7.1 for examples).

   The NC5, AG5 and 1PS prefixes are most often pronounced as syllabic nasals homorganic with the following consonant, but before certain consonants (including /r/, /s/, /y/ and /'y/) they may surface as [ĩ] or [ũ], depending on the following vowel. They are assumed to be /m/ underlyingly, since this is the vowel that surfaces before vowel-initial stems, as in (6) above.

Cross-linguistically, ideophones (§4.4.1.3) are often phonologically 'deviant' (Childs 1994:181, Voeltz and Kilian-Hatz 2001:2). In Cicipu they are characterised by CVC syllables, which as we have seen do not generally occur in the language. So far ideophones have been found with deviant codas containing nasal consonants, the plosive **p**, and the fricative **s**. Some examples are given in (23):

(23) **vɔp**　　**pass**　　**ɗő̂oŋ**　　**pɔm**
　　splat!　　*very white*　　*very black*　　*wholly*

3.1.8　Root structure

Noun and verb roots in Cicipu are usually disyllabic, although there are a significant number of mono- and tri-syllabic roots. The monosyllabic roots almost all have long vowels, and many are demonstrably derived from former disyllabic roots (e.g. **sɔɔ** 'drink' vs. **so'o** in Western Kambari, **kò-kő̂ɔ** 'egg' vs. **kò-kő'ɔ̃** in Tizoriyo).

A few roots of words for birds and trees have four or even five syllables, but often these are reduplicated. In the following tables Hausa loanwords are omitted because they have a markedly different distribution from native Cicipu words, being more likely to have trisyllabic roots. The distribution of nouns according to number of root syllables is given in Table 3:

Table 3: Distribution of noun roots according to syllable structure

Syllables in root	No. of nouns	Examples
1	106	**kò-lúu** 'knee', **kù-tóɔ** 'hen'
2	520	**kù-cíi.nó** 'back', **kù-dá.vù** 'mortar'
3	94	**méb-bè.ríi.sè** 'swift', **cìc.cé.rè** 'star'
4	16	**kà-'ɛ́n.gà.là.mì** 'traditional bag'

Verbs follow a similar pattern to nouns, with an even higher percentage of disyllabic roots.

Table 4: Distribution of verb roots according to syllable structure

Syllables in root	No. of verbs	Examples
1	18	**caa** 'give', **yãa** 'do'
2	311	**na.ha** 'leave', **la.sa** 'greet'
3	69	**hee.pi.ye** 'ask', **mi.ri.ɗa** 'twist'
4	2	**gi.tu.wa.na** 'exhale', **ku.si.ya.nu** 'smell (i.e. with nose)'

Of the tri- and quadri-syllabic verb roots, it is likely that the majority were once bi-morphemic, and in some cases fossilised present-day derivational processes can be identified. For example **lapila** 'prepare', **bambala** 'fumble', and **tobilo** 'cool liquid by repeatedly pouring' appear to contain a 'frozen' pluractional infix <**il**>, although they cannot be used without this affix[8]. In other cases, while it is not possible to identify grammatical morphemes in present-day use, patterns still emerge. For example, **titɔmɔ** 'thresh', **ziza'a** 'shiver', and **zizaɓa** 'tickle' all combine a reduplicated prefix with an inherently repetitive motion (this combination is also found in Bantu – Schadeberg 2003:79). If we exclude such cases of 'eidemic resonance' (Bickel and Nicholls 2007:209) from the count, then there are no more than a dozen verbs with more than two syllables. The basic verb root structure in Cicipu is therefore CVCV[9].

3.1.9 Summary

While some noun and verb roots in Cicipu are analysed as beginning with a vowel, noun and verb *words* are almost always consonant-initial on the surface. Roots with prenasalised stops may be analysed as CVN.CV. A small number of noun roots begin with an NC cluster (e.g. 19-20), but with the noun prefix they syllabify as CVN.CV. If we may talk of underlying syllabification, then Cicipu allows N and V but not CVN[10] in noun and verb roots. On the surface, however, CVN occurs but not N or V. Only CV is possible word-finally.

Prefixes may be CV, N, or V, all of which can occur word-initially. Ideophones are frequently CVC.

Nouns and verb roots are typically disyllabic.

---

8   A similar phenomenon is found in Hausa and other Chadic languages (e.g. Newman 2000:518-519).
9   This contrasts with Bantu CVC – see §4.6.1 for discussion.
10 Assuming prenasalisation is predictable (§3.1.5).

3.2 Consonants

3.2.1 Phonemic inventory

The 27 consonant phonemes of Cicipu are given in Table 5.

Table 5: Consonant phonemes

	Bilab.	Lab-dent.	Dental /alv.	Post-alv.	Palatal	Velar	Labio-velar	Laryng.[11]
Stops	p b		t d			k g		' 'w 'y
						kw gw		
Impl.	ɓ		ɗ					
Affr.					c j			
Fric.		v	s z					h hw hy
Nasals	m		n					
Liquid			l					
Rhotic			r					
Appx					y		w	

Although the coronal consonants in the above table have been labelled 'dental/alveolar', at least /d/ and /t/ were found to be dental (and laminal), rather than alveolar. This was only checked for one speaker, however. It is not yet known whether /n/ and /l/ are alveolar or dental, although they might be expected to pattern with /d/ and /t/.

3.2.2 Allophones and general phonetic rules

The bilabial plosives /p/ and /b/ sometimes undergo lenition to [ɸ] and [β] when they occur inter-vocalically, especially in quick speech. For example /yapu/ 'two' may surface as [jaɸu], and /jiibo/ 'have breakfast' as [dʒiːβo].

As well as the labialised and palatalised phonemes which appear in Table 5 there are a number of non-phonemic allomorphs which have these modifications. /m/ and /v/ have labialised allomorphs [mʷ] and [vʷ] before rounded vowels, while /k/ and /g/ have palatalised allomorphs [kʲ] and [gʲ] before front vowels. Before rounded vowels /k/, /g/,/'/ and /h/ do not contrast with their labialised counterparts, and so the underlying consonant in such sequences cannot be determined. Similarly before front vowels /'/ and /h/ do not contrast with their palatalised counterparts.

The phoneme /t/ is sometimes, but not always, realised as [tʃ] before [i], for

---

11 The three glottal fricatives were missing from the corresponding chart in McGill (2007), due to an error on my part. This chart therefore supersedes the earlier one.

example with the verb **tiyo** 'get'. [tijo] seems to be considered the 'correct' form.

The fricative [ʃ] is non-phonemic but does occur as an allomorph of /s/ before [i] for some speakers.

There are only two phonemic nasals in Cicipu, /**m**/ and /**n**/. All NC clusters in Cicipu are homorganic, with [ŋ] and [ɱ] occurring before velar and interdental consonants respectively.

/**r**/ is realised as a flap/tap[12] [ɾ] utterance-medially. Utterance-initially, and when lengthened (§3.2.4), it is realised as an approximant [ɹ] (or an r-coloured vowel [ɚ]). Sometimes the flap/tap surfaces as the retroflex/post-alveolar [ɽ], especially after the vowel /**a**/, but unlike Hausa (Newman 2000:394-395) there is no phonemic distinction between the coronal and post-alveolar flaps.

The approximants /**y**/ and /**w**/ have nasalised allophones [j̃] and [w̃] which occur in the neighbourhood of nasalised vowels. Note that the former differs from the nasal [ɲ] in that it does not have a closure.

3.2.3   Distributional restrictions and examples

The most striking distributional restriction is that /**h**/ rarely occurs root-medially – only two of the 44 /**h**/'s in the lexicon are root-medial – **dooho** 'disappear' and **naha** 'leave'. The affricates /**c**/ and /**j**/ are also rare root-medially, especially in verbs. Both the examples of root-medial affricates in verbs involve trisyllabic roots of uncertain derivation: /**kucɔ'ɔ**/ 'shake off' and /**mɔnjuwɔ**/ 'glare'. Other restrictions can probably be put down to the rarity of the phonemes in question.

The next two tables show examples of all the phonemes except /**hw**/[13]. Other than this one exception, all the phonemes contrast root-initially in native Cicipu nouns, as demonstrated in Table 6.

---

12 It is not known whether this sound is a flap or tap (See Ladefoged and Maddieson 1996:230-231).

13 /**hw**/ occurs only in the (borrowed?) noun **tù-hwíí** 'C'Lela language' and in the time adverb **hwá̋yà̏**
'day before yesterday'.

Table 6: Root-initial consonant phonemes in nouns

p	ù-**pácí**	*difficulty*
b	kà-**bárá**	*old man*
ɓ	mà-**ɓásà**	*mole (on skin)*
t	kà-**tádá**	*palm (of hand)*
d	kà-**dábá**	*bush/countryside*
ɗ	ù-**ɗángà**	*tree*
k	mà-**kántú**	*knife*
kw	ǹ-**kwá'á**	*orphan*
g	ù-**gálù**	*side*
gw	mà-**gwáwá**	*deaf/mute*
'	cì-**'ádì**	*trap*
'y	mɔ̀-**'yɔ́'yɔ́**	*fish*
'w	ù-**'wîi**	*distance*
c	kà-**cá'ṹ**	*husk (of maize)*
j	kù-**jénè**	*river*
v	kà-**várá**	*goat hut*
s	kù-**sáyú**	*spear*
z	à-**zá**	*people*
h	cì-**hávì**	*scratching*
hy	à-**hyá́'à̃**	*arrows*
m	kà-**mángá**	*rope*
n	ì-**námà**	*meat*
l	kà-**lánà**	*scar*
r	kà-**rákátáu**	*heel*
y	kà-**yáyù**	*root*
w	mà-**wáa**	*dog*

Apart from /'y/, /kw/, /gw/ and /hw/ all the consonant phonemes contrast root-initially in native Cicipu verbs. /kw/ occurs root-medially in the verbs **dukwa** 'go' and **cukwa** 'praise', and /gw/ occurs root-medially in eight verbs including **langwa** 'spoil' and **hungwa** 'rest'. Neither /'y/ nor /hw/ are found in verbs at all, perhaps not surprisingly given their overall rarity[14].

---

14 /'y/ also occurs in Hausa, although only in a few high frequency words like *'ya'ya* 'children'. Although rare in Cicipu, the phoneme occurs in a variety of words e.g. **kà-sí̃'yá̃** 'breast', **hwá̃'yã̀** 'day before yesterday', **mò-rí̃'yò** 'duiker', **mɔ̀-'yɔ́'yɔ́** 'fish', **kò-'yúwó** 'grass', and **kò-cí̃'yò'ĩi** 'ram'. It is also found in other Kambari languages and it seems unlikely to have been borrowed from Hausa, where /'y/ is a recent addition to the phonemic inventory (Newman 2000:393).

Table 7: Root-initial consonant phonemes in verbs

p	pasa	*cross*
b	bana	*invite*
ɓ	ɓasa	*slap*
t	tasa	*meet*
d	dasa	*castrate*
ɗ	ɗasu	*soak*
k	kanda	*mark on wall*
g	gava	*kick*
'	'etu	*dry by hanging out*
'w	'waa	*pass*
c	ca'a	*harvest*
j	janta	*crush*
v	vasa	*hit*
s	saɓa	*embrace*
z	zaa	*find*
h	hala	*coil*
hy	hyãa	*say*
m	mata	*give birth*
n	naha	*leave*
l	lawa	*escape*
r	raa	*eat*
y	yaa	*arrive*
w	waana	*twirl*

3.2.4 Length

While many languages have long consonants, it is rare (although not unheard of) for them to occur root-initially and word-initially (Blevins 2004:181, Ladefoged and Maddieson 1996:93). As with Central Kambari (Crozier 1984:280) any consonant can be lengthened in Cicipu, including the glottal fricative and stops. The difference in length between short and long consonants can be relatively mild, with long consonants sometimes no more than half as long again as their counterparts[15], and word-initially the distinction often seems to be neutralised in normal, fast, speech. Long voiceless stops may be impossible to detect utterance-initially[16], but they can easily be heard utterance-

---

15 Cross-linguistically long consonants tend to be longer than short ones by between 1.5 to 3 times (Ladefoged and Maddieson 1996:92).

16 In some languages long voiceless plosives *can* be detected utterance-initially, perhaps from a change

medially in careful speech.

The waveforms in Figures 22-26 demonstrate length differences for /n/, /l/, and /k/. The words in Figures 22-25 all belong to NC8, which has either **Ø-** or **C-** for allomorphs[17]. The former occurs with **Ø-náatà** 'small spider' (Figure 22) and **Ø-lóokàcí** 'time' (Figure 24), resulting in a short initial consonant. The latter occurs with **n-náa** 'cow' (Figure 23) and **l-lámà** 'noise' (Figure 25), resulting in a long initial consonant.

in the amplitude of the following vowel (Ladefoged and Maddieson 1996:94). This has not been investigated for Cicipu.

17 Recall that **C** represents an underspecified consonantal weight unit, which assimilates completely to the consonant to which it is attached, resulting in a long consonant.

n	a:	t	a	

0   0.595667

Time (s)
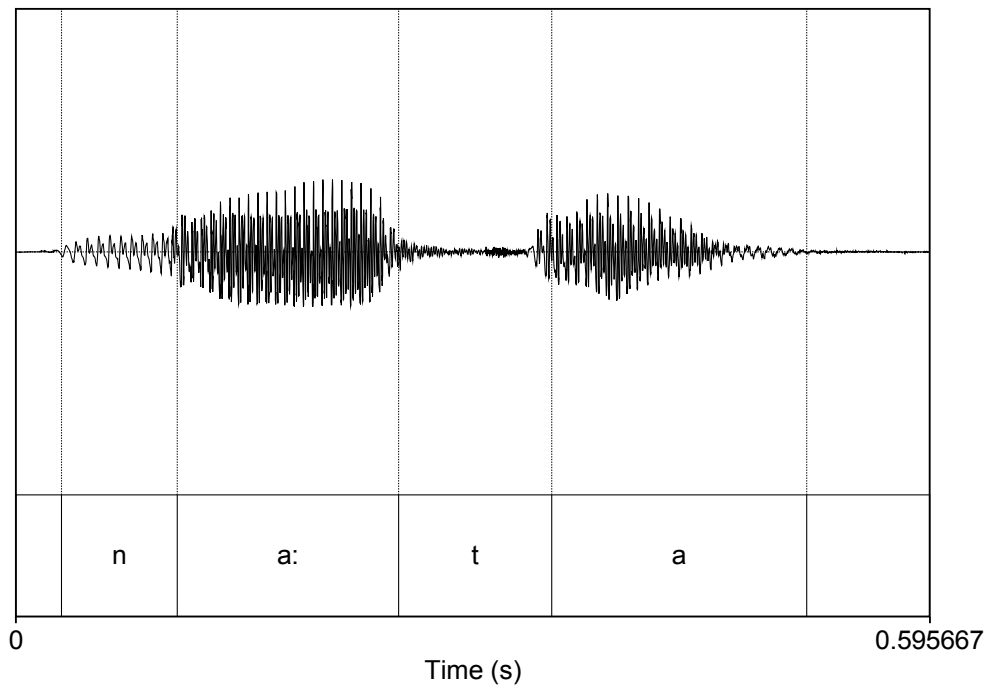
Figure 22: Waveform of **Ø-náatà** *'small spider'*

	n:	a:

0   0.597313

Time (s)
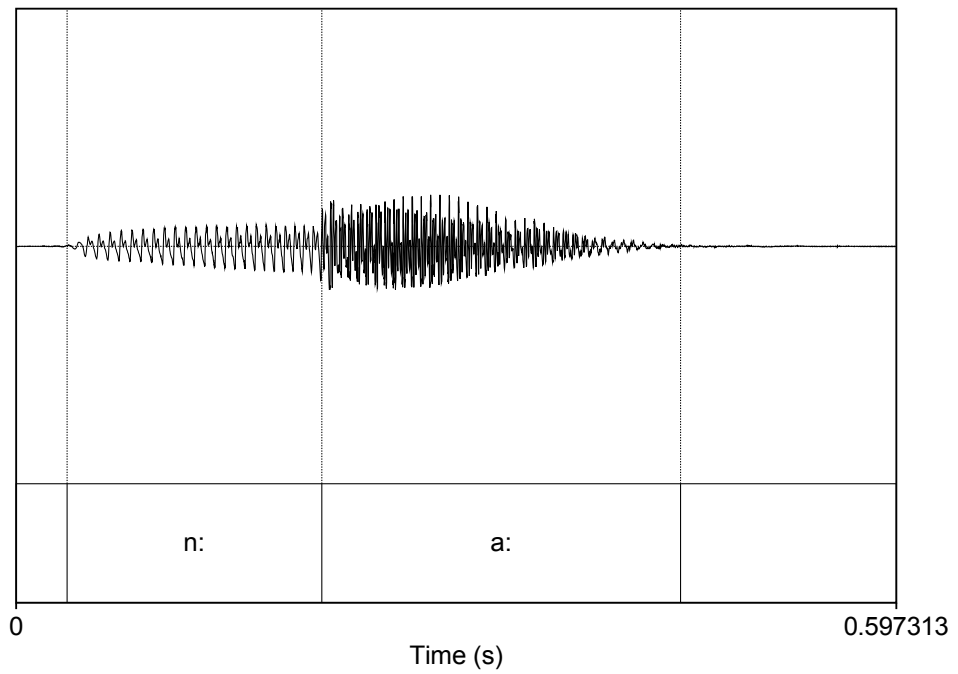
Figure 23: Waveform of **n-náa** *'cow'*
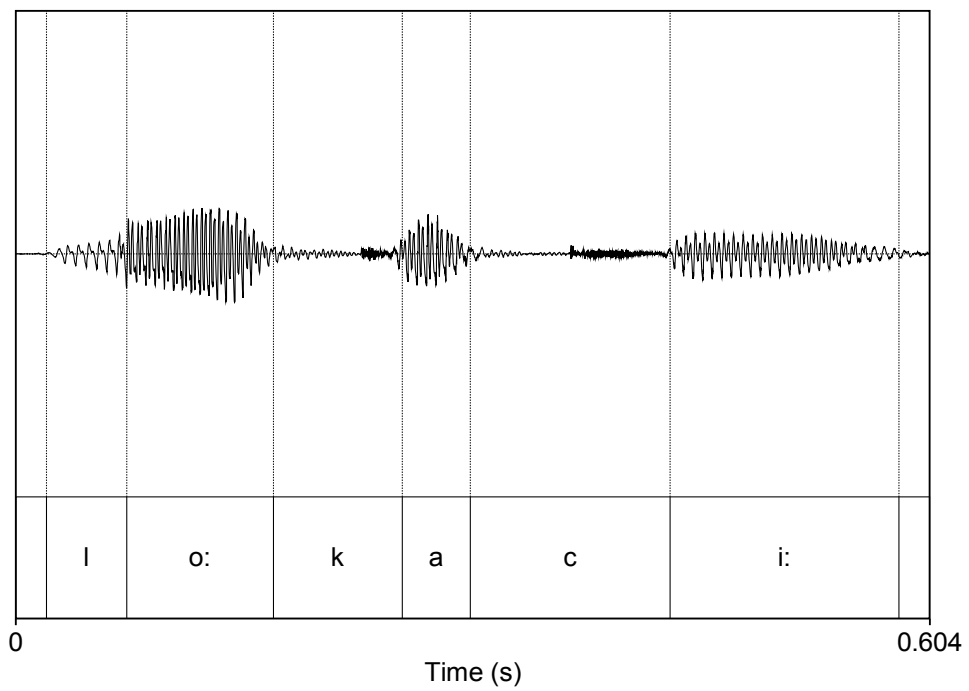
Figure 24: Waveform of **Ø-lóokàcí** *'time'*
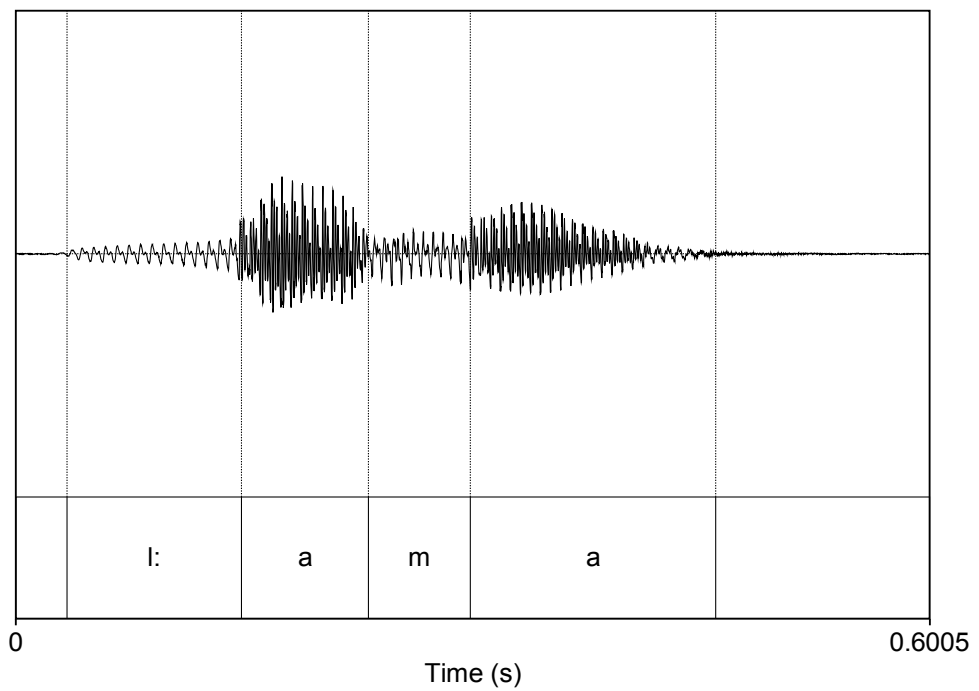
Figure 25: Waveform of **l-lámà** *'noise'*

Figure 26 demonstrates the difference between an utterance-medial short and long /**k**/:

(24) kà-dámá ká↓=k-káa
      NC1-word AG1=NC8-woman
      the word 'woman'
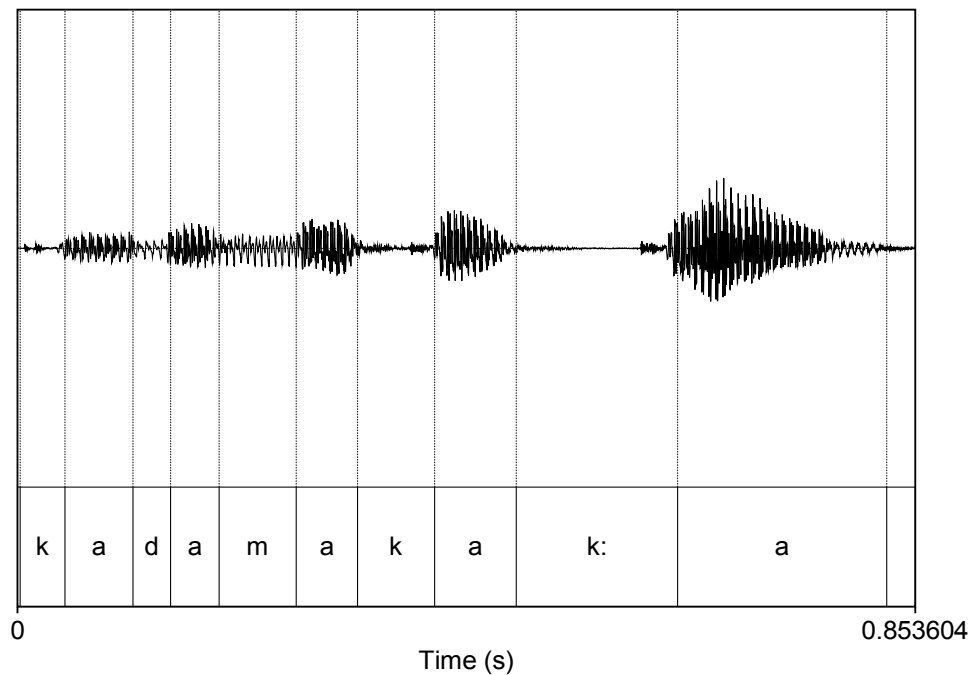
Figure 26: Waveform of **kà-dámá ká↓=k-káa** *'the word "woman"'*

Most of the long consonants are formed by an extension of some (as in /k/) or all (as in /n/ and /l/) of the short consonant. However in the case of /r/ there is a qualitative difference between the long and short variants – short /r/ is a tap/flap [ɾ] utterance-medially but long /r/ is realised as an approximant [ɹ] (or an r-coloured vowel [ɚ]), optionally followed by a flap/tap, giving [ɹɾ] or [ɚɾ] as shown in Figure 27:

(25) [ɹě̃ⁱ],   [kàdámá kéɪɾ̃ě̃ⁱ]
      r̃-rě́i,   kà-dámá ké↓=r-rě́i
      NC8-town NC1-word AG1=NC8-town
      towns, the word 'towns'
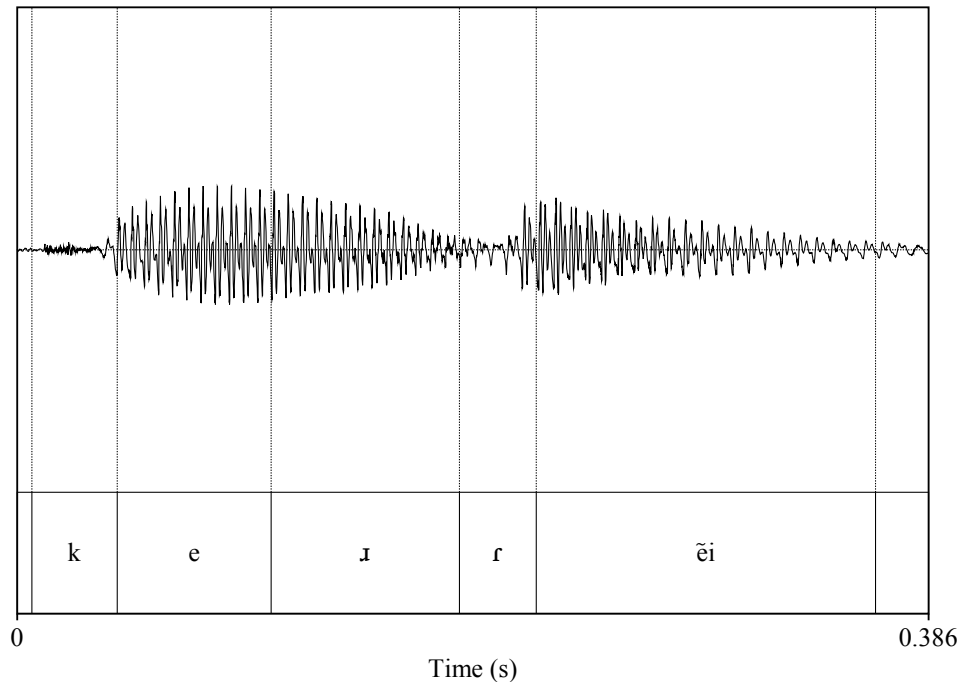
Figure 27: Waveform of **ké-r-rě̀i** *'of towns'*

The majority of long consonants in Cicipu words are ambi-morphemic, although as discussed in §3.1.3 they are true long consonants rather than simply sequences of short consonants. They arise from the application of the **C-** lengthening allomorph of the NC8 prefix to a noun root beginning with a short consonant. All the examples in the waveforms given above are of this kind.

Long consonants in noun *roots* are mostly found root-initially. Out of 626 native Cicipu noun roots, 51 start with a long consonant. It is suggested in McGill (n.d.) that these are likely to have arisen through the reinterpretation of an NC8 **C-** prefix as part of the root.

Long consonants in verbs are rare. They do not occur root-initially, and there are only a few examples root-medially – 16 out of 358 verbs. The only long consonants attested are /t/ (eight tokens), /l/ (four), /n/ (two), /w/ and /ʼw/ (one each). The form and meaning of some of the verbs suggest that the long consonants have come about as a result of the fossilisation of affixes (see §3.1.8).

123

3.3   Vowels

3.3.1   Phonemic inventory

Table 8: Cicipu vowel-inventory

	Front	Central	Back
Closed	i		u
Mid-Closed	e / ei /eu		o
Mid-Open			ɔ
Open		a / ai / au	

Although asymmetric vowel systems weighted toward the back are rare cross-linguistically (Crothers 1978:137, Schwartz et al. 1997), there is independent evidence for the analysis given here. Israel Wade (p.c.) adduced the same six vowels from his own independent research on Tirisino, and the wordlists produced by Lovelace (n.d.) for the nearby related language Tsuvaɗi also distinguish precisely these six vowels.

The Cicipu vowels are articulated more openly than the corresponding cardinal vowels, and this difference is clearly audible. The vowels of Hausa loanwords are often raised in the target form, presumably because to the Cicipu ear Hausa **a** is more like a Cicipu /e/ than a Cicipu /a/. Examples are shown below:

(26)   **dègè**   from *daga*   'from'

　　　ɓ'ɔ   from *a'a*   'no'

　　　kɔ̀-ccɔ̀kɔ́   from *jaka*   'bag'

3.3.2   Allophones and general phonetic rules

The lack of a schwa phoneme is unusual in West Kainji, where most other languages have two or even three (e.g. Central Kambari) central vowels. Vowels with the phonetic value [ə] do occur in Cicipu, but only as allophones of /a/, /o/, /ɔ/, or /e/ in the environment _CV[+high]: in other words, when the next vowel in the word is /i/ or /u/:

(27)   (a) [kòjõ̀ŋgʷɔ́lì]   (b) [dòːsə̀nû̃]
　　　　kò-yòngɔ́lì   dòosɔ̀nû!
　　　　NC1-ant   swim\IMP

　　　　ant, k.o. large   *swim!*
　　　　　　　　　[eamd022.1123]   [eamd006.049]

When such words are pronounced carefully the underlying vowel quality is never a schwa.

The distribution of the vowels /i/ and /u/ is problematic, just as in Hausa (Newman

124

2000:399-400). In some cases a neighbouring rounded vowel seems to be responsible for conditioning an /i/ to a [u]:

(28) [múː   **mú**mpà]
       m-úu   **mí**-mpà
       NC5-child   AG5-this
       these children

<div align="right">[eabg001.057]</div>

On other occasions different vowels are found in apparently identical environments, as shown by **gitu** and **gutu** in the following example:

(29) h-ăyà   è-'ésù   **gú**tù   níyù-nò   é-'èsù   **gì**tù
       3P-come\RLS 3P-again\RLS go_back\rls   send\RLS-VENT 3P-again\IRR   go_back\IRR
       they again sent for them to come back again

<div align="right">[saim001.054]</div>

Finally, for some words either /ɔ/ or /a/ is acceptable, as in **kù-lácí/kù-lɔ́cí** 'girl' and **sa'ũ̀/sɔ'ũ̀** 'touch'.

3.3.3 Nasal vowels

For each of the six vowels there is a contrasting oral and nasal pair, exemplified below. The oral vowels are considerably more common across the lexicon, with a ratio of about 5:1 in the words collected so far.

Table 9: Oral and nasal vowel contrasts

a	k-káa	*woman*
ã	ù-tã̂a	*bow (hunting)*
e	re'e	*persuade*
ẽ	sẽ́'ẽ	*carve*
i	mà-jjîi	*bird, k.o.*
ĩ	ù-jĩ́i	*value*
o	ù-kôo	*death*
õ	mò-tõ̂o	*saliva*
ɔ	ù-lɔ́ɔ	*locust bean tree*
ɔ̃	kɔ̀-kɔ̃́ɔ	*egg*
u	kɨu	*be older than*
ũ	tũ̀u	*pour*

3.3.4 Long vowels and diphthongs

Each of the oral and nasal vowels has a long counterpart as shown in Table 10. Short

125

vowels are more common, by a ratio of approximately 7:1 in the lexicon. From the rough measurements taken so far, long vowels in citation form in elicited speech are approximately half as long again as their short counterparts.

Table 10: Oral and nasal vowel contrasts

a	d**a**pa	*burn*
aa	g**aa**ta	*cut*
ã	k**ã**'ã	*scoop out*
ãa	t**ãa**	*shoot*
e	kè-r**é**'è	*tongue*
ee	cì-r**ée**nè	*fireplace*
ẽ	ùp**é**pí	*wind*
ẽe	ù-p**ễe**	*bigness*
i	c**i**ta	*sting*
ii	c**ii**ta	*squash*
ĩ	kà-y**í̃**vá	*vein*
ĩi	kà-h**í̃i**vì	*navel*
o	t**o**mo	*die out*
oo	t**oo**no	*come home*
õ	kò'ús**ỗ**	*haze*
õo	mò-t**ỗo**	*saliva*
ɔ	y**ɔ**nɔ	*pull*
ɔɔ	t**ɔɔ**mɔ	*chew*
ɔ̃	rù-k**ɔ̃́**'ɔ̃	*pool*
ɔ̃ɔ	kò-k**ɔ̃́ɔ**	*egg*
u	k**u**rɔ	*grow up*
uu	k**uu**la	*call*
ũ	kù-y**ũ̀**yũ̂	*sand*
ũu	kà-h**ṹu**cí	*cloud*

The vowels in monosyllabic roots are always long, with the exception of one verb **yo** 'be'. This verb surfaces with a long vowel when it occurs without a suffix (30), but with a short vowel before the perfective suffix **-nA** (§4.6.3.1) as in (32). It thus contrasts with **yoo** 'go' which is consistently long (31, 33):

126

(30)　ù-**yôo**　　ǹ　　Ø-móotò
　　　　3s-**be**\RLS　with　NC8-car
　　　　he has a car [lit. 'is with car']

(31)　ù-**yôo**　　Màkúukù
　　　　3s-**go**\RLS　[town]
　　　　he went to Makuku

(32)　z-zá　　　　n-nà　　**yó**-nò　　ǹ　　Ø-móotò
　　　　NC8-person　AG8-REL　**be**\RLS-PFV　with　NC8-car
　　　　the one who has a car [lit. 'is with car']

(33)　z-zá　　　　n-nà　　**yôo**-nò　　Màkúukù
　　　　NC8-person　AG8-REL　**go**\RLS-PFV　[town]
　　　　the one who went to Makuku

Other than in monosyllabic roots, long vowels are very rare root-finally in native Cicipu roots. When Hausa nouns ending in a long vowel are borrowed into Cicipu, the final vowel is usually shortened, as in **kàaká** 'grandparent'. Utterance-finally, the distinction between short and long vowels is blurred in Cicipu, just as in Hausa (Newman 2000:401).

Cicipu has four diphthongs, all of which are 'falling' (i.e. from an open to a close vowel quality): /**ai**/, /**au**/, /**ei**/, and /**eu**/.

Table 11: Diphthongs

/ai/	**mà-gɛ́i** 'sword', **má-jù'wɛ̀i** 'sexual intercourse'
/au/	**ù-yɛ́ndɛ́u** 'copaiba balsam tree', **kà-sɛ̂u** 'grave'
/ei/	**ù-sɛ́i** 'pain', **mà-dɛ́i** 'calf (body part)'
/eu/	**kè-rèzɛ̃̂u** 'cotton', **dɛ́nɛ̃̂u** 'small'

Discounting loanwords, diphthongs are only found as the last vowel of the root, and comparative evidence from Kambari suggests they may reflect historical consonant loss from root-final syllables.

3.3.5　Distribution

All of the short oral and nasal vowels are found in both V_1 and V_2 root positions. Long

vowels are largely limited to V$_1$ (especially in verbs). Vowel harmony (§3.5) strongly constrains the vowels which are able to occur together in native Cicipu roots. Furthermore there is a strong tendency (but by no means an obligation) for V$_1$ and V$_2$ to have the same quality.

3.4 Tone

3.4.1 Tone inventory

There are three contrastive lexical tones in Cicipu: H, L, and HL (falling)[18]:

(34) H v. L: kù-**sáa**   ká-**ssà**
                NC9-mountain   NC1-leaf
                mountain   *leaf*

(35) H v. HL: kà-**táa**   ù-**tâa**
                NC1-shoe   NC7-bow
                shoe   *bow (hunting)*

                t-**tíi**   **tîi**
                NC8-container   NC6:shit
                container   *shit*

(36) H H v. H L: **káayá**   **káayà**
                NC1:room   NC1:bean
                room   *bean*

                cù-**kúlú**   kà-**kúlù**
                NC6-tortoise   NC1-hailstone
                tortoise   *hailstone*

(37) L L v. L HL: ká-**kkùtì**   ká-**kkàcîi**
                NC1-shell   NC1-middle
                shell   *middle*

                mɔ́-**kɔ̀ɗɔ̀**   mà-**kùdâa**
                NC4-gossip   NC4-squirrel
                gossip   *squirrel*

---

18 Rising tones do occur phonetically, but they do not contrast with the other three tones in underlying representations.

(38) H L v. L HL: ù-**yéyù** kù-**yǔyû̂**
NC7-cold NC9-sand
cold *sand*

ì-**cíyò** kó-**ccìyûu**
NC3-beard NC1-heap
beard *heap*

Statements, commands, and yes/no questions all have their own intonational contours which are imposed on top of the lexical tones of the sentence. Statements can be analysed as having an utterance-final L tone, which means that H tones are realised as HL at the end of utterances. For commands and questions see §3.4.6 and §4.3.6.1 respectively.

As well as the statement boundary tone just mentioned, there is further evidence that HL tones should be analysed as sequences of H and L[19]. For example, the realis tone pattern on disyllabic roots (§3.4.6) is H L if the first syllable is light, but HL L if the first syllable is heavy. HL tones are mainly limited to long vowels, but they may occur on phonetically-lengthened short vowels (e.g. utterance-final imperatives §3.4.6).

In certain limited scenarios an extra-high tone seems possible. For example, in the Tikula dialect if the negator **cé** occurs after a high tone then it may surface as extra-high:

(39) ì-**ɗáná** **cé**
NC3-mark NEG
not marks

[Tikula, sagb001.026]

Other candidates for extra-high tone are the plural morpheme ** áa** 'associates of' (§4.4.2) and the topic marker **gó** (§8.3.2).

3.4.2 Downdrift, downstep, and upstep

Like many African languages, tones in Cicipu utterances undergo 'terraced' downdrift within each intonation group, whereby the pitch of each H is lower than that of the one before. Successive L tones also decline in pitch, but by less, and so the distinction between H and L is less towards the end of an intonation group than at the beginning.

Downstep occurs in several syntactic environments (see §3.4.7).

In addition to downdrift and downstep, there is also evidence for the rarer

---

19 This is generally the case for African languages (Clements 2000:153).

129

phenomenon of upstep. The relevant data is provided by the habitual tone pattern (§3.1.3). Recall that the usual tone pattern for verbs with light syllables is L L H H. Normally the negator **cɛ́** occurs with high-tone, but in (40-41) it is extra-high.

(40)　à-sì-pándá　cɛ̋
　　　3P-HAB-forget　NEG
　　　they do not forget

[samy001.069]

(41)　à-sì-tá'á　cɛ̋
　　　3P-HAB-want　NEG
　　　they didn't want

[samy001.033]

This effect is not limited to examples involving the negator, but also occurs with reduplicated verbs:

(42)　ká-nà　kà-sì-nīmā-nímā
　　　AG1-ART　AG1-HAB-bite-REDUP
　　　some [camels] bite

[tats004.001.018]

(43)　ǹ-sì-pāndā-pándà
　　　1S-HAB-forget-REDUP
　　　I forget

[tats005.002.007]

If we assume that upstep occurs after the habitual marker **si-**, then examples (40-43) can be accounted for without needing to postulate three underlying tones. The tonological representation of (43), for example, would then be L L↑ L L H L.

3.4.3　Spreading

Tone spreading is very common in Benue-Congo languages, particularly of H tones. In Cicipu H tones at the end of one word spread onto L tones at the start of the following word. Tone can spread onto both noun prefixes (44) and roots (45):

(44)　ù-kábà　**cɛ̋**　**ká**-líipì
　　　3S-took\RLS　NEG　NC1-wrong
　　　he didn't take the blame [citation form **kà**-líipì]

[saim001.044]

130

(45) t-índà   cé   Ø-káaká
      1P-see\RLS   NEG   NC8-grandparent
      we didn't see the grandparent [citation form Ø-**kàa**ká]

Gender-marked subject agreement prefixes are also affected by H-spread from preceding subjects if the verb is in the realis mood[20] – compare (46a), in which spread occurs after an H, with the unperturbed pattern in (46b):

(46) (a) mò-**ní**   **mɔ́**-kɔ́ɗ<ìl>ɔ̀-nù
          NC4-water   AG4-cut\RLS<PLAC>-RES
          the water parted

      (b) mò-ní   m-áyà   **mɔ̀**-kɔ́ɗ<ìl>ɔ̀-nù
          NC4-water   AG4-come\RLS   **AG4**-cut\RLS<PLAC>-RES
          then the water parted

By contrast, *person*-marked agreement prefixes are seemingly unaffected by preceding H tones.

It seems that H tones can spread from words of any lexical class which end in a H. Example (47) shows H-spread resulting from the imperative (§3.4.6) tone pattern:

(47) 'ìzè'ḗ   **má**-ttílú!
      wash\IMP   NC4-pot
      wash the pot! [citation form **mà**-ttílú]

The next two examples show the H from **cé** spreading further through the following word until another H is reached, at which point the spreading stops – a phenomenon referred to as the 'Plateau Principle' by Kisseberth and Odden (2003:67) (i.e. avoid a 'dip' between two Hs).

(48) t-índà   cé   **tí**-**záa**rúmà
      1P-see\RLS   NEG   NC6-flea
      we didn't see a flea [citation form **tì**-**zàa**rúmà]

(49) t-índà   cé   **ká**-**káa**nâa
      1P-see\RLS   NEG   NC1-crab
      we didn't see a crab [citation form **kà**-**kàa**nâa]

---

20 Irrealis prefixes are always H.

Certain low-tone morphemes seem to be resistant to spreading, such as gender-marked pronouns in the complement of VP (§7.2) and the person-marked object enclitics (§7.3). Compare (50a) where the H from **hán** 'where' spreads onto **mà-nnû** 'bird' with (50b) where **m-è** 'it' is unaffected.

(50) (a) hán   **má**-nnû?   (b) hán   **m-è**   ká'à?
             where NC4-bird   where AG4-PRO now
             where is the bird?   *where is it now?*

[2008-03-31.003]

Similarly, the habitual tone pattern is unaffected by H-spread, despite beginning with a L tone.

(51)   mà-wáa   **mà**-sì-cíndà
        NC4-dog   AG4-HAB-wait
        the dog guards

[tats001.001.022]

It seems there is a distinction between tones which are specified as L (and hence do not change), and tones which are underlyingly unspecified, and only surface as L if a H has not already spread on to them.

3.4.4   Polar tone

The tone on the copula is usually polar with respect to the previous syllable, just as in Hausa (Jaggar 2001:457-458).

(52) (a) ì-cíntò   y-**í**   (b) kó-ggòmbò k-**έ**
             NC3-doorway AG3-COP   NC1-bat   AG1-COP
             it's a doorway   *it's a bat*

(53) (a) p-píyá   v-**ì**   (b) kò-rúdú   k-**è**
             NC8-guineafowl AG8-COP   NC1-shelter AG1-COP
             it's a guineafowl   *it's a shelter*

There are exceptions, however, and this area is not well understood yet. In particular after a falling tone the copula is L, which is unexpected if falling tones are to be subject to the usual HL sequence analysis.

3.4.5   Lexical tone in nouns

Nouns in Cicipu have lexically-specified tone. While noun prefixes are affected by H-spread (§3.4.3), this is a phonological process and does not depend on the syntactic

configuration. There are also certain *syntactic* environments where the tone on the noun prefix changes, but these can be analysed as the result of downdrift and spreading (see §3.4.7). The following table shows the distribution of native Cicipu roots according to their citation tone pattern[21].

Table 12: Distribution of noun roots according to tone pattern

Tone pattern	No. of nouns	Examples
H	72	kò-**jóo** 'lizard', ù-**láa** 'fire'
H L	25	d-**dôo** 'horse', ù-**kwâa** 'skin'
L	4	ká-**ssà** 'leaf', wú-**ntò** 'guest hut'
H H	232	kà-**dábá** 'bush', ì-**dáná** 'line'
H L	197	kù-**cígà** 'cock', kù-**jénè** 'river'
L H	35	ì-**bèyé** 'rice', mà-**ccìjíi** 'calabash'
L HL	19	kò-**kòyôo** 'snail', kò-**ppòdôo** 'frog'
L L	18	kó-**ggòmbò** 'bat', ká-**kkùtì** 'shell'
H H H	36	ì-**cícípú** 'prayers', kà-**ppítítú** 'bubble'
H H L	3	Ø-**lówólì** 'spider', ko-**órórì** 'owl'
H L H	8	Ø-**'ísùmóo** 'stinging ant', kà-**sákkàllá** 'tree, k.o.'
H L HL	13	kò-**cí'yò'îi** 'ram', ù-**kúdìdûu** 'nettle tree'
H L L	6	kà-**cáacùwà** 'fine loincloth', ké-**kkérùkè** 'hornbill'
L H H	1	kà-**cìtúwá** 'exchange'
L H L	25	tì-**zàarúmà** 'flea', Ø-**cìccérè** 'star'
L L H	4	mò-**rìgìdó** 'navel', kà-**gàatàɓɔ́ɔ** 'praying mantis'

Every *word* has at least one H – this is a general constraint affecting all Cicipu nominal and verbal words. Note also that HL is restricted to root-final position, a common cross-linguistic pattern (Yip 2002:28).

3.4.6 Grammatical tone on verbs

In contrast to nouns, verbs are inherently toneless, and the tones with which they surface are determined entirely by grammatical properties such as mood or aspect. This is also

---

21 The tone pattern in the plural is almost always identical to that in the singular.

the case for a number of Bantu languages, and such systems are sometimes called 'predictable' (Kisseberth and Odden 2003:61). In (54-55) the verb can be found in three different moods, depending on the tone pattern.

(54) (a) ùdúkwà
u-dukwa**LHL**
3s-go**RLS**
he/she went

(b) údùkwà
u-dukwa**HL**
3s-go**IRR**
he/she should go

(55) (a) dúkwà
Ø-dukwa**HL**
2s-go**RLS**
you (sg.) went

(b) dùkwá!
dukwa**LH**
go**IMP**
go!

In each case, the segmental material remains the same, but the different tone patterns superimposed on the words give rise to different grammatical categories, and hence different meanings. In (54a) **ùdúkwà** 'he/she went' has a LHL tone pattern indicating realis mood, while in (54b) the same segmental sequence appears with a HLL tone pattern, this time indicating irrealis mood. Examples (55a) and (55b) illustrate a similar contrast, this time between realis and imperative moods, and involving the second person.

In the realis mood the basic pattern is LH(L)*, although in certain circumstances this surfaces as HL. To be precise, if the verb has a subject agreement prefix capable of bearing tone – i.e. (a) V or CV before a consonant-initial stem, or (b) a sonorant consonant before a vowel-initial stem) – then the tone is LH(L)*. If, on the other hand, the prefix consonant before a vowel-initial stem is non-sonorant (c), or if the verb has a null prefix (d), then the pattern is H(L)*. The four possibilities are illustrated below:

(56) (a) [kàdúkʷà]
ka-dukwa-**LHL**
AG1-go-RLS
it went

(b) [ẁáːjà]
w-aya-**LHL**
3s-come-RLS
it came

(c) [káːjà]
k-aya-**HL**
AG1-come-RLS
it came

(d) [dúkʷà]
Ø-dukwa-**HL**
2s-go-RLS
you (s.) went

Since there is not usually a distinct boundary between approximants and vowels, the sequence of a L-tone approximant followed by a H-tone vowel is, in practice, realised as a rising contour across both segments. On monosyllabic verbs, which always have a long root vowel (at least on the surface), L H L is realised as L HL, with a falling tone

on the verb root:

(57) ùc**âa**
      u-caa-LHL
      3s-give-RLS
      he gave

In the irrealis mood the basic tone pattern is H(L)*, with H being realised on the subject agreement prefix. For vowel-initial verbs such as **ewe** 'refuse' the sequence of H and L is realised as a falling tone HL across the first syllable. The difference between the realis (e.g. **héewè** 'they refused') and irrealis forms (e.g. **hêewè** 'they should refuse') of vowel-initial verbs can be difficult to detect, but native speakers are very clear about the distinction. The difference is clear when the pitch track of the sounds is analysed using a program such as Praat, as can be seen in Figures 28 and 29.
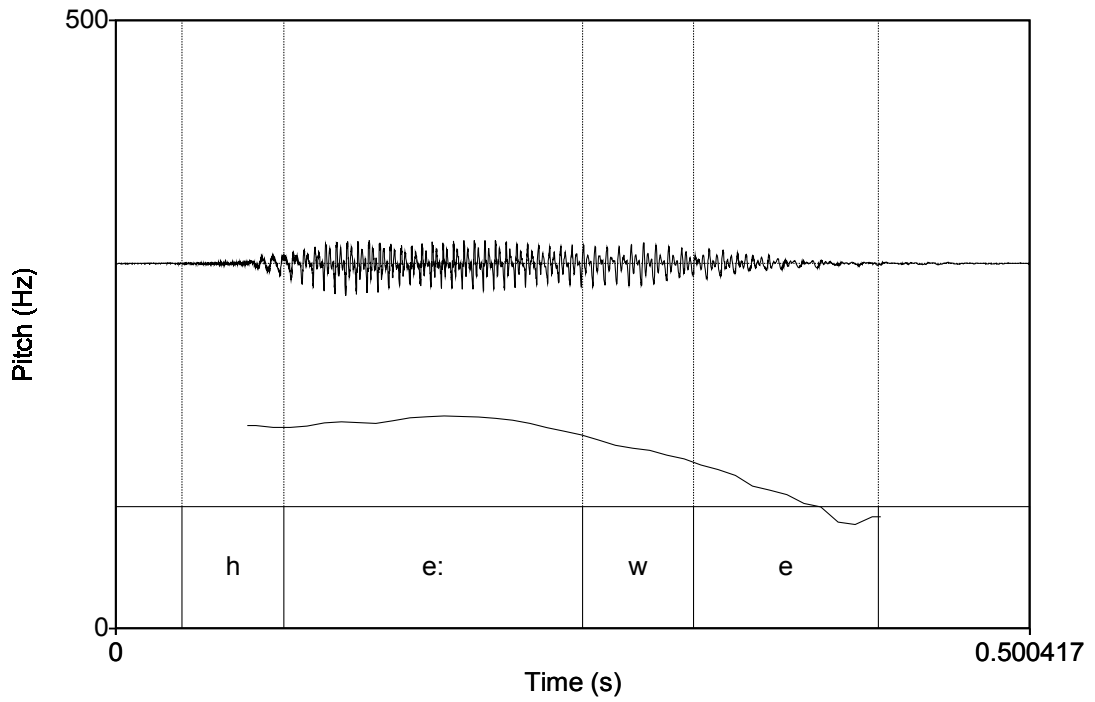
Figure 28: Pitch track for **héewὲ** *'they refused' (realis)*
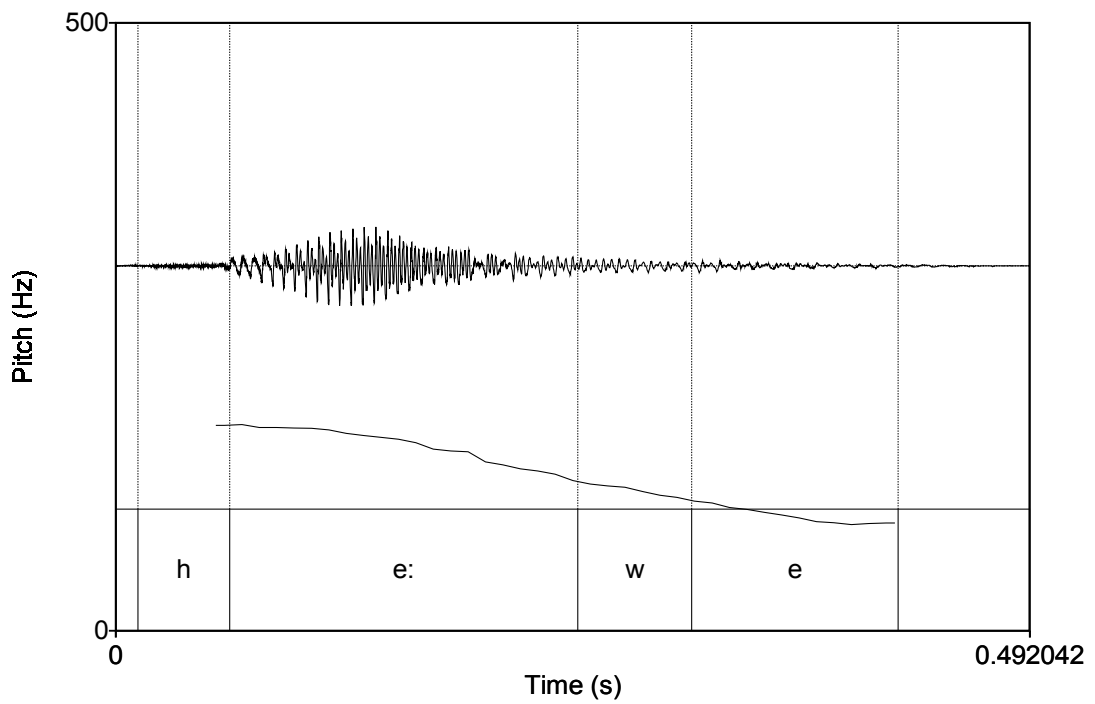
Figure 29: Pitch track for **hêewὲ** *'they should refuse' (irrealis)*

136

In the imperative mood there is no subject agreement prefix. The basic surface melody is (L)*H – in other words, the final tone is H, and any previous tones are L.

(58) dùkwá!
dukwa-LH
go\IMP
go!

[saat002.002.424]

Monosyllabic imperatives have a single H tone:

(59) sɔ́ɔ!
sɔɔ-H
drink\IMP
drink!

[saat002.002.303]

If there are suffixes after the verb root, then the H tone is realised on the last suffix, as in (60). The imperative tone pattern can therefore be thought of as 'lining up' at the right-hand edge of the verbal word, in contrast to the realis and irrealis patterns, which line up at the left[22].

(60) yàaná!
yaa-na\LH
arrive-VENT\IMP
arrive! [towards deictic centre]

[eaim003.1402]

Finally, there is an interesting interaction between vowel quality and tone in imperatives, in that if the final vowel of the verb is /u/ or /ũ/ then the tone is realised as falling rather than as H:

(61) (a) tũ̂u!   (b) cìɗònû!
    tũu\HL   ciɗonu\L L HL
    pour\IMP   bury\IMP
    pour!   *bury!*

[eamd006.004, eamd006.009]

3.4.7 Complement tone perturbation on nouns

A single complex perturbation affects nouns within several apparently unrelated syntactic constructions. The affected positions include the 'possessor' NP in the associative construction[23] (§4.4.5.1), the complement of VP, the complement position of

---

22 The L-tone plural imperative suffix **-nÀ** is an exception (§4.6.2.3)

23 See Cahill (2000:43) on Nkem, where association is marked purely by a floating L resulting in

the preposition/locative proclitic **Á** (§4.5), after the presentative **ǹdúu** (§4.3.3.4), in the NP complement of predicate locative clauses (§4.3.3.2), and optionally after the existential predicator **òkóo/àkwái** (§4.3.3.3)[24].

The following examples show the perturbation after the locative **Á**, but the same patterns are found in each of the above environments. The accents in (62-68) indicate relative *pitch* within each word, and *not* phonological tone.

(62) (a) wú-ntò   ‾ _   (b) á=wū-ntò   ‾ ‾ _
           NC7-shelter   LOC=NC7-shelter
           shelter   *in the shelter*

                                                               [eamy040.001.022]

(63) (a) kó-ggòmbò   ‾ _ _   (b) á=kō-ggòmbò   ‾ ‾ _ _
           NC1-bat   LOC=NC1-bat
           bat   *by the bat*

                                                               [eamy040.001.015]

(64) (a) d-dôɔ   ‾ _   (b) á=d-dɔ̄ɔ   ‾ ‾ _
           NC8-horse   LOC=NC8-horse
           horse   *on the horse*

                                                               [eamy040.001.011]

(65) (a) g-gwéedíbè   ‾ ‾ _   (b) é=g-gwēedībè   ‾ ‾ ‾ _
           NC8-main_road   LOC=NC8-main_road
           main road   *on the main road*

                                                               [eamy040.001.009]

(66) (a) ká-kkàcíi   ‾ _ ‾   (b) á=kā-kkàcīi   ‾ ‾ _ ‾
           NC1-middle   LOC=NC1-middle
           middle   *in the middle*

                                                               [eamy040.001.007]

In each of these examples the citation tone pattern is exactly preserved after the locative, giving rise to three pitch levels. The data can be accounted for if we assume that downstep has occurred immediately after the locative morpheme, resulting in a downward resetting of the pitch register from that point onward.

All the nouns in the above examples began with a H tone. For nouns beginning with a L tone the root tones are preserved from the first H tone onward, but initial L tones raise to the same level as subsequent Hs:

---

    subsequent downstep.

24 See Marten (2005) for a similar phenomenon in the Bantu language Herero, where complement case is found in the complement of VP and of prepositions, as well as in presentational constructions.

(67) (a) mò-ríngínó   ‾ ‾ ‾   (b) ó=mò-rìngìnò   ‾ _ _ _ _

     NC4-spine   LOC=NC4-spine

     spine   *on the spine*

                                            [eamy040.001.012]

(68) (a) k-kèeké   _ ‾   (b) é=k-kèekè   ‾ _ _

     NC8-bicycle   LOC=NC8-bicycle

     bicycle   *on the bicycle*

                                            [eamy040.001.005]

This can be accounted for by assuming, in addition to the downstep just mentioned, the H from the locative spreads rightwards – in which case the phonological representations of the (b) examples would be Á↓=mó-ríngínó and Á↓=kée-ké. The derivation is given in (69):

(69)   Underlying form: Á=mò-ríngínó   ‾ _ ‾ ‾ ‾ ‾

        H-spread: Á=mó-ríngínó   ‾ ‾ ‾ ‾ ‾ ‾

        Downstep: Á↓=mó-ríngínó   ‾ _ _ _ _ _

Since the same perturbations occur in all the other syntactic environments mentioned above, the same rules can be applied. In some of these environments the tone immediately before the perturbed noun is H, as with the locative proclitic. In other constructions, particularly the complement of VP, the preceding tone is L. If the NP in complement position begins with a H tone, then this H is the same pitch as the last L of the verb – this is therefore an example of "total downstep" (Connell and Ladd 1990:25)[25]. In this case it must be admitted that it is difficult to see what the trigger for the necessary H-spread might be. More generally, the fact that this sequence of rules has to be invoked in apparently unrelated syntactic environments suggests that the analysis presented here is missing a generalisation. It should therefore be regarded as a stepping-stone to a better solution.

3.5 Vowel harmony

There is widespread vowel harmony in Cicipu. Vowels from the set {o, ɔ, e, a} are mutually exclusive in roots, regardless of word class. So if a root contains /a/, its other vowels must come only from the set {a, i, u}, /e/ only occurs with {e, i ,u}, and so on. The vowel harmony system operates throughout the lexicon, without exception apart

---

25 As was stated in the front matter, downstep between words is not marked in the examples in this thesis. So, for example, objects occurring in the complement of VP (the unmarked position for objects) and marked as H H are actually at the same pitch as the preceding L of the verbal word.

from some compounds (§3.5.1) and loanwords (§3.5.2). Many prefixes also contain harmonising vowels (§3.5.3).

3.5.1 Distribution of vowels in CVCV noun roots

Table 13 below shows the distribution of vowels in CVCV noun roots. For simplicity the table only include roots where both vowels are short and oral, but the same restrictions hold for long and nasal vowels, for both nouns and verbs – vowel harmony is absolute in native roots. The empty cells are shaded to highlight the pattern.

Table 13: Vowel co-occurrence restrictions in CVCV noun roots where V_1 and V_2 are both short oral vowels (V_1 down the left, V_2 along the top)

	i	e	a	ɔ	o	u
i	10	1	16	2	13	14
e	5	9				4
a	9		28			12
ɔ	3			13		6
o	4				8	1
u	3	1	4		8	13

Compound words are not always subject to harmonisation – the following words both contain /e/ and /a/:

(70)   méngétàarì
       mé-ngé-t-a-árì
       NC4-child-?-NC2-man
       boy

(71)   kwákúllè
       ku-ã́'ã́-kú-llè
       NC9-day-AG9-that
       then [lit. 'that day']

Vowel harmony occasionally operates beyond the word in normal, fast speech (particularly when vowel coalescence occurs at a word boundary – see §3.7.1).

3.5.2 Loanwords

Borrowed noun and verb roots vary with respect to harmonisation, and the outcome seems to depend on how deeply entrenched the word is in the speech community.

Demonstrably modern borrowings such as **róobà** 'plastic' may occur with conflicting vowels, but in general harmonisation does take place, as illustrated in (72) and (73).

(72) **rùuká** from *roka* 'chatting'
    mè-ttégù from *taggo* 'shirt'
    kollo from *kallo* 'look'
    kwáanù from *kwano* 'metal container'

(73) **gwede** from *gode* 'thank'
    róotò from *reto* 'hanging'

Normally it is the first vowel that changes, although not exclusively as **kwáanù** in (72) shows. The examples in (73) show that true harmonisation is involved, rather than simply the vowel-raising which typically occurs when words are borrowed from Hausa to Cicipu (§3.3.1). There is no difference in vowel height between **gwede** and *gode,* or between **róotò** and *reto.* Instead the vowels have harmonised 'sideways'.

3.5.3 Affixes and clitics

Many nominal and verbal affixes harmonise according to the pattern outlined above. Nouns from three of the nine Cicipu noun classes, classes 1, 2, and 4, are formed by adding to the root the harmonising prefixes **kA-***,* **A-**, and **mA-** respectively. The quality of the **A-** vowel is determined by the root vowels as follows:

- If the root contains /e/, then the prefix vowel will be [e]
- If the root contains /o/, then the prefix vowel will be [o]
- If the root contains /ɔ/, then the prefix vowel will be [ɔ]
- Otherwise the prefix vowel will be [a]

These rules do not need to be ordered, since the strong vowel harmony operative in the lexicon ensures that the four vowels involved are mutually exclusive. It should however be pointed out that when speakers pronounce a word carefully, they often pronounce the prefix [ka] (or [a], or [ma]), regardless of the root vowels which are to follow. With that in mind, I have analysed the underlying vowel in these harmonising prefixes as /a/, and in this thesis such vowels are written **A** as in **kA-**[26].

---

26 Anderson (1980a) briefly discussed prefix-root vowel harmony in nouns for the East Kainji language Amo. He states that "Though this vowel harmony may provide a phonetic 'target', considerable variation still exists even on individual words" (1980a:157). This statement nicely captures the situation in Cicipu.

In addition to noun prefixes and the corresponding gender agreement prefixes, several other inflectional and derivational verbal affixes contain harmonising vowels:

Table 14: Harmonising affixes

Affix	Gloss	Reference	Example
A-	3PP agreement	§7.4	(74)
-wA	applicative	§4.6.4.2	(75)
-wA	anticausative	§4.6.4.3	(76)
-wA	separative	§4.6.5.3	(77)
-nA	perfective	§4.6.3.1	(78)
-nA	ventive	§4.6.5.1	(79)
-nA	plural imperative	§4.6.2.3	(80)
-kwA	suffix for borrowed verbs	§4.6.6	(81)

(74) (a) á-dùkwà
     3P-go\IRR
     they should go

     (b) ɔ́-dɔ̀nɔ̀
     3P-follow\IRR
     they should follow
     [saat001.006.011, saat001.006.044]

(75) (a) tì-yáa-**wà**
     1P-do\RLS-**APPL**
     we did to [him]

     (b) mí-dòonù-**wò**
     AG5-sit\IRR-**APPL**
     may they stay with [you]
     [saat001.006.114, saat001.002.026]

(76) (a) màsídùwà
     mà-sídù-**wà**
     AG4-heat\RLS-**ANTIC**
     it [water] spoiled [lit. got hot]

     (b) gólùwò
     Ø-gólò-**wò**
     AG8-cut-**ANTIC**
     [he/she] gets cut
     [tats005.002.083, tats004.003.005]

(77) (a) ù-hálù-**wà**
     3S-coil\RLS-**SEP**
     it uncoiled

     (b) ù-'úmbù-**wò**
     3S-close\RLS-**SEP**
     it opened
     [eaim010.114, saat001.008.097]

(78) (a) à-dúkwà-**nà**
     3P-go\RLS-**PFV**
     they had gone

     (b) kù-'íngò-**nò**
     AG9-go_home\RLS-**PFV**
     it had gone home
     [saat001.006.008, saat001.006.088]

(79) (a) sékè-**nè**
     release\RLS-**VENT**
     [he] released down

     (b) ù-yúwò-**nò**
     3S-fall\RLS-**VENT**
     he fell down
     [Tidipo, saat002.005.046, saat002.005.076]

(80) (a) ya̋a-**nà**!   (b) sèkèlè-wé-**nè**!
          do\IMP-**PL.IMP**   move\IMP-ANTIC-**PL.IMP**

          [you(pl.)] do!   *[you(pl.)] move over there!*
                                                    [saat001.002.107, saat002.002.172]

(81) (a) ù-gwáanù-**kwà**   (b) ù-tɔ́ɔrù-**kwɔ̀**
          3S-understand\RLS-**LW**   3S-push\RLS-**LW**

          he understands   *he pushed*
                                                    [eaim010.106, eabg001.124]

In every other affix the vowel is either /i/ or /u/, which are neutral with respect to vowel harmony. Harmonisation often 'passes through' these non-harmonic affixes. In (82) the 3PP subject prefix **A-** harmonises with the root **ze'ɛ̃** 'dance', despite the intervening habitual prefix **si-**. All affixes with neutral vowels in Cicipu appear to be 'transparent' rather than 'opaque' with respect to vowel harmony.

(82) è-sì-**zè'ɛ̃-zé'ɛ̃**
          3P-HAB-**dance-REDUP**

          they dance
                                                    [sahs001.002.011]

Finally, clitics show mixed behaviour with respect to vowel harmony. The locative proclitic **Á** (§4.5) and the associative agreement proclitics (§4.4.5.1) harmonise with the following word, as shown by (83) and (84) respectively. The latter example shows the same 'passing over' of a neutral vowel as (82).

(83) (a) **ɛ́↓**=k-kèeké   (b) **ó↓**=kó-ocì
          LOC=NC8-bicycle   LOC=NC9-hole

          on a bicycle   *into a hole*
                                                    [tapf001.005.011, svtmg001.177]

(84) kà-mángá **kó↓**=kú-**hóomòwò**
          NC1-rope   AG1=NC9-**tree(k.o.)**

          rope of the kuhoomowo tree
                                                    [eaim006.1454A]

On the other hand, the person-marked object enclitics **dò** (2PP) and **rè** (3PP) do not harmonise with their hosts (§7.3.2).

3.5.4 Cross-linguistic comparisons

The Cicipu vowel harmony system is essentially identical to that of Central Kambari (Hoffmann 1972:74), and very likely the other Kambari languages as well. The system

is distinct from those usually reported in the literature (inc. Dettweiler 2000 on C'Lela), in that the harmonisation rules cannot be stated straightforwardly using distinctive features. Vowels do not harmonise according to height, or roundness, or backness – instead, complete assimilation is involved. The most concise way to state the rule is as follows:

(85) *If there are two [-high] vowels in a phonological word then they must be identical*

3.6 Nasalisation

We have already seen that Cicipu has a full complement of nasal vowels, both short and long, as well as two nasal consonant phonemes /n/ and /m/. Nasalisation is not usually confined to single segments, however, and it may spread a considerable distance. This section is concerned with the phonemes affected by the process and the direction of spread.

3.6.1 Phonemes affected

In addition to the six short and long nasal vowels, the approximants /w/ and /y/ have nasal allomorphs [w̃] and [ỹ]:

(86) (a) [mãg^wá:w̃á]   (b) [ĩĵũĵû]
           ma-gwã́aw̃́á   ì-yũ̀yû̃
           NC4-bruise   NC3-fly

           bruise   *flies* [i.e. insects]

Vowels which have become nasalised due to the influence of other nasal segments do not have the same properties as underlyingly nasal vowels – for example, the former do not trigger prenasalisation (§3.1.5), as illustrated by words such as **mita** [mĩta] 'squeeze' and **kù-mócì** [kùmṍtʃì] 'old woman'. Therefore the rightward spread of nasalisation must occur after prenasalisation – in technical terms they are in a 'counter-feeding' relationship. Some roots beginning with a nasal such as /mĩto/ [mĩnto] 'shut mouth' do contain a ṼNC sequence – these are assumed to have underlyingly nasal vowels.

Root-internally no contrast has been found between oral and nasal vowels before a nasal consonant. However this contrast exists across morpheme boundaries, and since there are four verbal suffixes beginning with /n/ it is relatively common. Compare the verbs **yãa** 'do' and **yaa** 'arrive', shown before the ventive suffix **-na** in (87):

(87) (a) [ùjâːnã̀]   (b) ùjâːnã̀]
      ù-yâa-nà   ù-yâa-nà
      3s-**do**\ʀʟꜱ-ᴠᴇɴᴛ   3s-**arrive**\ʀʟꜱ-ᴠᴇɴᴛ
      he did   *he arrived*

3.6.2 Direction

The spread of nasalisation from the nasals /**n**/ and /**m**/ is predominantly to the right, although spread to the left does occur. Every vowel occurring immediately after a nasal consonant is nasalised, whereas nasalisation to the left is more variable; when it can be detected, it often seems to be weaker than nasalisation to the right. If the first vowel in a root is nasal, then the second usually is too, as in **kà-hĩ́ĩ̀** 'night' and **y-yũ̀yṹ** [j̃ːũ̀j̃ũ̀] 'fly'. This again suggests spread to the right is predominant.

Nasal spread seems to be confined to the word. Within the word, spread to the right is blocked by NC clusters as in **ù-lénjí** [ù-lẽ́ndʒí] 'sun' and **kà-búngú** [kà-bṹŋgú] 'snake'.

Although in general nasality spreads to the right, it can spread to the left from *verb* roots onto verb prefixes. This leads to contrast between nominal (a) and verbal (b) pairs:

(88) (a) kà-tã́a   (b) kã̀-tã́a
      ɴᴄ1-shoe   ᴀɢ1-shoot\ʀʟꜱ
      shoe   *he* [**kà-bárá** 'old man'] *shot*
                                                   [2008-02-05.001]

(89) (a) à-tã́a   (b) ã̀-tã́a
      ɴᴄ2-shoe   3ᴘ-shoot\ʀʟꜱ
      shoes   *they shot*
                                                   [2008-02-05.001]

3.7 *Morphophonemic processes*

3.7.1 Coalescence and elision

When vowels become juxtaposed as a result of the concatenation of two morphemes (either word-internally or across word boundaries), the vowels often coalesce, resulting in a long vowel[27], usually bearing the quality of the second.

Word-internally, we have already seen (§3.1.1) that vowel-initial noun and verb roots coalesce with prefixes to form long vowels, with the second (i.e. root) vowel

---

27 This is not the case in Central Kambari – when two short vowels come together the outcome is still a short vowel (Hoffmann 1972).

quality dominant. Another type of word-internal elision occurs in associative constructions (§4.4.5.1), in which an agreement clitic attaches to the noun prefix of the 'possessor' NP. If this prefix begins with a vowel then coalescence occurs, with the second vowel again dominant:

(90) [tʃìmɛ̃́ʔ̀ì   tá:hùlà   há:vù]
      cì-mɛ̃́ʔ̀ì   tí↓=a-húlá   há-avù
      NC6-in   AG6=NC2-name   AG2-2S.POSS
      in your (sg.) name [lit. 'the inside of your name'], citation form ʔà-húlá
      'name']

                                                              [saat001.001.011]

Occasionally coalescence fails to happen in the associative construction, and a glottal stop appears before the noun prefix as in (91), just as if it were utterance-initial. It is not known what, if anything, governs the occurrence or non-occurrence of coalescence here.

(91) [mɔ̃̀ní   mã́ʔā̀s:à]
      mɔ̀-ní   má↓=á-ssà
      NC4-water AG4=NC2-leaf
      the colour green [lit. 'water of leaves', citation form ʔá-ssà 'leaves']
                                                              [eamy001.005]

The conjunction **ǹ** 'and/with' and the locative **Á** also coalesce with vowel-initial stems, and again the second vowel in the sequence is dominant:

(92) [nà̀:ká̃́ntʃírí]
      nì=à-káncírí
      and=NC2-fingernail
      with fingernails
                                                              [eaim006.1450]

(93) [ú:rè]
      Á↓=ú-rée
      LOC=NC7-town
      in the town
                                                              [eamd003.011]

One environment where coalescence might be expected, but does not seem to occur, is pre-prefixes (§5.3.6) – when a noun prefix is attached to a stem already consisting of a prefix and noun root. In this case, the vowels never coalesce – as in (94), a glottal stop appears before the inner prefix.

(94) [mè?ìɾí]
     mè-ì-rí
     NC4-NC3-thing
     a small thing [citation form ?ì-rí 'thing']

                        [samoh001.122]

Coalescence also occurs across word-boundaries, in which case the quality of either vowel may dominate, although there is a preference for the second. Examples (95-101) show coalescence in a variety of syntactic environments, while (97-101) additionally demonstrate the elision of /w/, /y/ and /h/.

(95) [ànáhìːɗà]
     à-náhà↓   í-ɗáa
     3P-leave\RLS NC3-ground
     they left the land

                   [VERB + COMPLEMENT, samoh001.027]

(96) [súːɗāngà]
     sée   ù-ɗángà
     until   NC7-tree
     just a tree

                   [CONJUNCTION + NOUN, samoh001.068]

(97) [?ásūːnà]
     'ásù   wú-nà
     place(NC7) AG7-REL
     the place

                   [NOUN + RELATIVISER, samoh001.029]

(98) [ìnámél lê ːnà]
     ì-námà   yí-llè   yí-nà
     NC3-meat   AG3-that AG3-ART
     that meat

            [NOUN + DEMONSTRATIVE + ARTICLE, Tidipo, saat002.003.021]

(99) [ìtúmɔ́ː?úgò]
     ì-túmó   yì-'úgò
     NC3-pregnancy AG3-throw\RLS
     there was a miscarriage

            [SUBJECT + SUBJECT PREFIX, first vowel dominates, eamd014.242]

(100) [àbáráːvʷôː]
     à-bárá   hó-vôo
     NC2-old_man AG2-1S.POSS
     my old men

           [NOUN + MODIFIER, first vowel dominates, eamy005.038]

(101) [àzíːkèːké]

    à-zá   há↓=ì-kèeké

    NC2-person   AG2=NC3-bicycle

    cyclists [lit 'people of bicycles']

['POSSESSED' + 'POSSESSOR', tapf002.001.020]

In some cases (but by no means all) neither vowel predominates and the two vowels appear to have 'met in the middle'. So far, this has only been observed for the vowels /a/ and /i/ resulting in [e] as in (102-103), but it may be that the same process can occur with back vowels as well (this is the case for Central Kambari – Hoffmann 1972).

(102) [hárèːhʷáːrà]

    hárì   à-hwáarà

    until   3P-start\RLS

    then they started

[samoh001.026]

(103) [béːtáŋgì]

    báa   ì-tángì

    NEG   NC3-item

    no items

[samoh001.110]

Some of the most common types of relative clauses involve both elision and coalescence: **'ásù wúnà** [ʔásūːnà̰] 'where' (lit. 'the place that'), **ìrí yínà** [ìríːnà̰] 'what' (lit. 'the thing that'), and **àzá hánà** [àzáːnà̰] 'the ones that'. In normal speech the elided and coalesced forms are more usual, but they are readily understood by native speakers as contractions of the long forms.

3.7.2 *u*-anticipation

Cicipu has three homophonous **-wA** suffixes (applicative, anticausative, separative – see §3.5.3 for examples). On the application of any of these prefixes the final vowel of the stem to which it is attached changes to [u], regardless of its underlying quality[28].

---

28 A similar morphophonological process is found in Hausa when the feminine suffix *-aa* is attached to a stem ending in **o** e.g. *sabo* 'new (masc.)' vs. *sabuwa* 'new (fem.)' (Jaggar 2001:58).

*Table 15: **u-** anticipation before **-wA** suffixes*

Stem	Gloss	Applicative	Anticausative	Separative
dama	*tell*	damuwa		
bɔlɔ	*look for*	bɔluwɔ		
'ɔpɔ	*hold*		'ɔpuwɔ	
golo	*cut*		goluwo	
hala	*coil*			haluwa (*uncoil*)
'umbo	*close*			'umbuwo (*open*)

The same process occurs with the prefix **kwA-** found on borrowed verbs (§4.6.6):

> (104) **gwaanukwa** *understand* from Hausa *gane*
>
> **daamukwa** *worry* from Hausa *dama*
>
> **koyuko** *teach* from Hausa *koya*
>
> **dennukwe** *compress* from Hausa *danna*

This change of vowel is the result of a morphophonological rule rather than a phonological one, since it only happens on the application of a suffix beginning with a labialised or labiovelar consonant. Within lexical roots other vowels are readily found before labial consonants, as illustrated below:

> (105) **lawa** *escape*
>
> **weewe** *refuse*
>
> **siiwa** *untie*
>
> **bɔwɔ** *steal*

3.7.3 *i*-anticipation

One of the more unusual features of Cicipu grammar is the split in verb and pronominal agreement between paradigms that inflect for person and those that inflect for gender (see Part IV). Of interest here is the distinction between the AG8 noun class pronoun (§6.2.15) and the 3PS object clitic (§7.3), both **vì**. Internally, the forms are homophonous. However consider the following examples:

(106) (a) A: hán   Aúdù?
           where   Audu
           where's Audu?

     (b) A: hán   Ø-vóotò
           where NC8-goat
           where's the goat?

     B: míndìvì
        m-índà = vì
        1s-see\RLS=**3s.PRO**
        I saw him

     B: míndà   vì
        m-índà   v-ì
        1s-see\RLS **AG8-PRO**
        I saw it

                                    [2007-02-05.007]

The person-marked clitic in (106a) triggers a morphophonemic spreading process which the gender-marked pronoun in (106b) does not. The /i/ from the clitic **vì** spreads leftward, changing the last vowel of the preceding verb complex[29]. This happens without fail for every verb, even monosyllabic verbs, and the resulting contextual neutralisation can give rise to lexical ambiguity:

(107) (a) ǹtî̃ivì
         ǹ-**tẫa** = vì
         1s-**shoot**\RLS=3s.PRO
         I shot him

     (b) ǹtî̃ivì
         ǹ-**tũ̃u** = vì
         1s-**pour**\RLS=3s.PRO
         I poured (for) him

                                    [MY, 2007-03]

When combined with the ubiquitous vowel coalescence (§3.7.1) there can be considerable displacement of vowel qualities:

(108) [tʃíːvùːɾè]
     cáa = vì   ù-rée
     give\IMP=3s.PRO   NC7-town
     give him a town!

                                    [saat001.005.059]

---

29 An alternative analysis would be to assign the 3PS object clitic the underlying representation **ivi**, with the vowel at the end of the previous verb deleting. Since all verbs end in vowels and the clitics only occur immediately post-verbally either analysis is possible. However an explanation would be required for why deletion takes place rather than coalescence (compare (106a) with (95) which has a long [iː]). Also all other object clitics are CV monosyllables rather than VCV.

150

Chapter 4 – A grammatical sketch of Cicipu

4.1 Introduction

As with the phonological sketch in the last chapter, the purpose behind this chapter is twofold. Primarily it is intended to familiarise readers with the basic morphosyntactic structure of Cicipu. Having digested this chapter, they will then be better equipped to understand the more in-depth analyses in the parts that follow. The discussion does however cover more ground than is strictly necessary to meet this first goal. Since the linguistic description of Cicipu is in its infancy, a second goal is to present a grammatical sketch of the language, covering areas which, although not of direct relevance to Parts III and IV, are likely to be of interest to typologists and Africanists alike. Nevertheless, limitations of time and space have meant that much of the argumentation here does not meet the rigorous standards that (ideally) apply to the publication of full reference grammars. In particular, much of the analysis derives from inspection of the corpus rather than from metalinguistic discussion with native speakers. Consequently the sketch has more to say about what does occur than what does not.

There is a certain amount of unavoidable overlap with later chapters. Because gender agreement is so ubiquitous in Cicipu, many of the word classes introduced in this chapter will also be individually discussed in chapter 6 on agreement targets. My approach is to deal with structural matters in this chapter (e.g. modifier X takes a low-tone gender agreement prefix), and leave class-specific phenomena (e.g. the AG8 prefix for modifier X has allomorphs A and B occurring in the following circumstances...) to chapter 6.

The chapter is organised as follows. In §4.2 I give a brief typological overview of the language. I then take a more detailed look at the structure of verbal and non-verbal clauses (§4.3), nominals (including adverbs and the structure of the NP) (§4.4), prepositions (§4.5), verbs and the verb phrase (§4.6), adjectives (§4.7), numerals (§4.8), and quantifiers (§4.9).

4.2 Typological overview

Cicipu grammar has a lot in common with the well-known Bantu subgroup of Benue-Congo[1]. There is a robust noun class system with agreement on a whole host of targets,

---

1 Johnston (1919:17) wrote "Curiously enough, there are languages...in Nigeria...the syntax of

as well as a productive system of verbal extensions. In what follows, frequent reference will be made to similar or identical constructions in Bantu.

Like most Benue-Congo languages, Cicipu is head-initial at clause and phrase level, and mainly head-marking[2]. According to the traditional terminology it is highly agglutinative; in Bickel and Nichols' (2007) terms most formatives are non-fusional, non-flexive, mono-exponential, and synthetic. Cicipu is not a true polysynthetic language in the sense of Baker (1996), since there is no noun incorporation and (unlike Bantu) no object agreement.

Prefixes and suffixes occur on both nouns and verbs, with prefixes predominant on nouns. Suprasegmental modification (§4.6.2) and infixation (§4.6.1) are also found. The former is common across African languages (Creissels 2000), not just Benue-Congo. Infixation is rarer, and arguments will be given in §4.6.1 that the morphemes concerned are true infixes, rather than just word-internal affixes. Reduplication is common and found in several word classes, including nouns, verbs, adjectives, demonstratives, and ideophones. Stem modification other than by tonal change is very rare in nouns (see §5.2.2 and §5.2.7 for two examples), but certain verbal categories are consistently expressed in this way (§4.6.3.4). Morphological processes in Cicipu are in general highly regular, and suppletion is unattested.

There are formally-identifiable open classes of nouns and verbs, and a small closed (but still formally-distinct) class of adjectives. As for constituent order, Cicipu is a configurational SVO language, and is mainly consistent with respect to the correlations observed by Greenberg (1963), although NP-internal syntax is unusual (§4.4.5.6).

If one only looks at the 'core' of Cicipu grammar (e.g. verbal morphology and the noun class system) then, as mentioned above, the language is very similar to many other Benue-Congo languages. However it should be stressed how much of the grammar (in a broad sense of the word e.g. as conceived by construction grammar or cognitive linguistics – see Croft and Cruse 2004) is influenced by the lingua franca Hausa, spoken fluently by virtually all Cicipu speakers. Cicipu is a prime candidate for study for linguists interested in the influence of language contact on grammar; examples of likely contact effects can be seen in §4.3.1.2, §4.3.3.3, §4.3.5.1, §4.3.6.1, §4.3.7, §4.4.1, §4.4.5.4, §4.5, §4.7.1, and §4.8-4.9, as well as many other places in other chapters.

---

construction of which frequently recalls the Bantu idiosyncrasy; but the word-roots of the vocabulary would be found wholly dissimilar".

2   The dependent-marked associative/possessive construction (§4.4.5.1) is an exception.

Grammatical sketches of Benue-Congo languages spoken in northern Nigeria often fail to take into account possible influence from Hausa, and so for the benefit of readers unfamiliar with that language, this chapter is liberally sprinkled with references to the two main English-language Hausa grammars, Newman (2000) and Jaggar (2001).

4.3   Clause-level syntax

This section begins with a discussion of grammatical relations (§4.3.1), which then informs the subsequent account of constituent ordering (§4.3.2). The remaining sections cover non-verbal clauses (§4.3.3), negation (§4.3.4), questions (§4.3.5), and then finally clause combinations (§4.3.6).

4.3.1   Grammatical relations

There is a clear distinction in Cicipu between the subject and the other arguments of the verb, where by 'subject' I mean the single argument of an intransitive clause, or the more agentive argument of a transitive clause. Three kinds of evidence can be offered in support of this distinction[3]. Firstly, only subjects can trigger agreement on the verb. Secondly, only subjects can occur pre-verbally without an accompanying copula; for NPs with other grammatical relations to occur before the verb, they must be followed by the copula, and the verb has to have perfective aspect if the clause is in realis mood (see §4.3.2 for details). Conversely, only *non*-subjects can occur after the verb. Thirdly, the subject has to be omitted in Cicipu infinitive clauses (§5.4.1), while the object may be included[4].

NPs with a wide range of semantic roles are found as subjects in intransitive clauses including AGENT, PATIENT, RECIPIENT, THEME, and EXPERIENCER. The restriction of subject referents to topics found in some Bantu languages (Morimoto 2008) does not apply in Cicipu.

Although there is no object agreement on the verb, there are other tests for objecthood in Cicipu. If we find that the subject NP in one kind of construction consistently corresponds with the filler of a particular non-subject NP slot in a structurally-derived counterpart construction, this is evidence that the NP picked out in the counterpart construction is an object (see Croft 1991:6-8 for discussion). In Cicipu the applicative (§4.6.4.2) and the anticausative (§4.6.4.3) both provide the required

---

3   As in Benue-Congo as a whole, there is no case-marking in Cicipu.
4   See Anderson (1976) and Chung (1998:64-69) for discussion of this test.

correspondence. There is not space for detailed discussion here, but the gist of the argument with respect to the applicative is that the subject of the intransitive verb in sentences such as (1a) always coincides with the immediately post-verbal NP in corresponding applicative constructions like (1b), rather than with, say, an oblique.

(1)　(a)　**Ø-bíírò**　vì-tánnà　　　　(b)　tànnà-wá　　**Ø-bíírò!**
　　　　　NC**8-pen**　AG8-descend\RLS　　　　descend\IMP-APPL　NC**8-pen**
　　　　　the biro *descended*　　　　　　　*lower **the biro**!*
　　　　　　　　　　　　　　　　　　　　　　　　　　　[2008-02-19.001]

The cross-linguistic patterning of such constructions suggests that the correspondent in (1b) is a grammatical object. The same line of argument can be applied to the anticausative (§4.6.4.3).

　　　The binding of reflexives (§4.3.1.2) is another possible source of evidence for objecthood, but the relevant elicitation has not been carried out.

　　　Finally, two other cross-linguistically applicable tests serve to distinguish objects from other types of NPs. First, they are able to occur clause-internally without a preposition. Secondly, no other NP in the clause may bear the same semantic role[5].

4.3.1.1　Double-object constructions

As in many languages, double-object constructions exist where one of the objects bears the semantic role of RECIPIENT and the other that of THEME. The verbs **caa** and **nɔɔ**, both meaning 'give', are perhaps the clearest examples of this. In recipient-theme double-object constructions the primary object (i.e. the one closest to the verb) bears the RECIPIENT role, while the secondary object encodes the THEME.

(2)　sée　　nì　　c-câa　　[z-zá　　　n-nà　　wú-u-kàbà-wà-nà＝vù] /
　　　unless　and　2S-give\RLS　[NC8-person　AG8-REL　3S-FUT-take\IRR-APPL-VENT=2S.PRO]
　　　[ká-bbìyà]
　　　[NC1-payment]
　　　you have to give [the person who will bring it for you] / [payment]
　　　　　　　　　　　　　　　　　　　　　　　　　　　[tats001.002.022]

If the primary object is pronominal and agrees in person (rather than gender), then it takes the cliticised form (§4.4.3.3), as illustrated in (3). Pronominal secondary objects are always of the 'independent' kind (§4.4.3.1-4.4.3.2).

5　In some languages adjuncts form 'islands' with respect to relativisation, which allows them to be distinguished from the direct arguments subject and object. However in Cicipu adjunct-internal NPs are readily relativised (§4.3.4), and so this test is not applicable.

(3) kúngwá   ú-nɔ́ɔ = ɗò   v-ì
God(NC8)   3S-give\IRR=2P.PRO   AG8-PRO
May God give you(pl.) it

<div align="right">[eabg001.013]</div>

The primary object is not limited to the RECIPIENT semantic role, and may also be a GOAL (4), BENEFICIARY (5) or 'MALEFICIARY' (6).

(4) ǹ-píisànù   **á-ayà**   kwé-etú
1S-spray\RLS   **NC2-bean**   NC9-medicine
*I sprayed the **beans** [with] fertiliser*

<div align="right">[GOAL, 2007-01-10.015]</div>

(5) àgóonì**vì**   ùdángùlà
à-góonù = **vì**   ù-ɗángùlà
3P-help\RLS=**3S.PRO**   NC7-gather
*they helped **him** [with] gathering*

<div align="right">[BENEFICIARY, tapf002.002.015]</div>

(6) ú-làngwà   **z-zá**   Ø-hùuhú n-nà   Ø-yó-nò   cì-mé'ĩ̀
3S-spoil\IRR   **NC8-person** NC8-lung   AG8-REL   AG8-be\RLS-PFV NC6-inside
*it spoils [for] **a person** the lungs that are inside*

<div align="right">['MALEFICIARY', tats007.002.036]</div>

It is possible (in fact, usual) for either object to be omitted if the context is clear:

(7) ù-câa   Ø m-ɔ́ɔ   mé-evì
3S-give\RLS   NC4-child   AG4-3S.POSS
he gave his child [i.e. to someone else]

<div align="right">[PRIMARY OBJECT OMITTED, tats001.004.047]</div>

(8) ùtĩ̂ivì   Ø
ù-tũ̂u = vì
3S-pour\RLS=3S.PRO
he poured [fruit for] him

<div align="right">[SECONDARY OBJECT OMITTED, tapf001.003.027]</div>

The verbs in the examples above admit double objects without any morphological change. However a much wider range of verbs have the potential to take double objects by means of the valence-increasing applicative suffix **-wA** (§4.6.4.2).

4.3.1.2 Reflexives

There is no reflexive pronoun in Cicipu, nor are there the reflexive verbal affixes common to Bantu. Instead reflexives are expressed using the associative/possessive phrase **kà-tĩi kÁ-___**, literally 'head of____'. There is a direct analogue in Hausa (Jaggar

2001:381-385), but not too much should be made of this – after all, the 'head of' construction is also cross-linguistically fairly common (König and Siemund 2005) and so the similarity may be coincidental rather than a contact effect.

(9) ùrâasà   kàtî̂i   kéevì
     ù-râ<is>a   kà-tî̂i   kÁ-evì
     3S-eat<CAUS>\RLS NC1-head   AG1-3S.POSS
     he served himself

[samoh001.220]

Emphatic reflexives are formed by adding the preposition **ǹ** 'with' before **kà-tî̂i kÁ-__**. When used adverbially the resulting phrase has an 'exclusive' interpretation, as in (10).

(10) húupìsìlò   ǹ   kàtî̂i   kíivè
     hú-u-pìs<ìl>ò   ǹ   kà-tî̂i   kÁ-ivè
     3P-FUT-break<PLAC>\IRR   with   NC1-head   AG1-3P.POSS
     they will break by themselves

[samoh001.100]

Emphatic pronouns are formed using a construction identical to one in Hausa (Jaggar 2001:385-389): an independent pronoun, optionally followed by the conjunction 'and/with', and then the phrase 'head of___'.

(11) ká-nà   k-kúngwá   **èvì**   **ǹ**   **kà-tî̂i**   **ké-evì**   ù-câa-nà   k-è
     AG1-REL NC8-God   **3S.PRO** **with** **NC1-head**   **AG1-3S.POSS** 3S-give\RLS-PFV   AG1-PRO
     which God himself gave

[tats001.004.032]

4.3.2 Constituent order

Cicipu is a configurational language with SVO as the basic constituent order for clauses, and Head-Modifier within the noun phrase (§4.4.5), just as expected for a Benue-Congo language (Williamson and Blench 2000:30-36).

Only two clausal word orders have been observed in the corpus, SVO and OSV, with the former much more common. An example of SVO order is given below:

(12) [cì-kúlú]ₛ   [tì-kómò]ᵥ   [mó-ní   má-nà]ₒ
     NC6-tortoise   AG6-cover   NC4-water   AG4-ART
     a tortoise covered the water

[Tikula, sagb001.528]

OSV clauses are formally distinct from their SVO counterparts in two other ways. First, the object must be followed by the copula, and second, if the mood of the clause is realis then the verb must be marked for perfective aspect.

OSV word order is also *pragmatically*-marked, and is only found when the object carries a much greater 'communicative load' than the verb. All three of these stipulations also apply to the more general constituent-fronting constructions discussed below (§4.3.2.1).

Example (13) illustrates the obligatory copula, although since the verb is not realis there is no perfective marker[6].

(13)   [Context: speaker A says that when they used to go to the market they would put on different garments, not a loincloth (**kù-róonò**, NC9). Speaker B then picks up on the fact that the usual item of clothing was a loincloth...]

B: [kù-róonò]ₒ   kw-í   [z-zá]ₛ   [Ø-sí-yũ̂u]ᵥ
NC9-loincloth   AG9-COP   NC8-person   AG8-HAB-wear
it was loincloths people were wearing

[sayb001.357]

The utterance occurred during a discussion about what people wore in the past, so the idea behind the verb **yũ̂u** 'wear' was already activated (§2.3.2.4) – indeed the verb itself had been used just a few intonation units prior. The referent of **kù-róonò** 'loincloth' can be argued to be in contrastive focus here.

4.3.2.1   Constituent-fronting[7]

The basic order of elements in the clause is [S AUX V O OBL]. Speakers may depart from this basic order for pragmatic effect. More precisely, constituents may be fronted for two main reasons, because they are in focus (§4.3.2.1.1) or because they are marked topics (§4.3.2.1.2).

4.3.2.1.1   Focus

If the referent of an object or an oblique is in argument focus (§2.3.1), then it is usual to place it clause-initially:

(14)   **kà-'ázzìkí**   **kè-yìní**   k-è   vú-u-yã̀a-wà = mù?!!
NC1-prosperity   AG1-what   AG1-COP   2S-FUT-do\IRR-APPL=1S.PRO
what prosperity *will you bring me?!*

[saat002.002.392]

---

6   This was the only example in the corpus of OSV word order with two lexical NPs (OV word order without a lexical subject is more common).

7   The use of the terms 'fronting' and 'extraction' in this section is not intended to imply that a derivational process is involved.

(15) **Ø-rísìnôo**   v-ì   Ø-yó-nò   hínà
      NC**8**-[name] AG8-COP 2s-be\RLS-PFV TOP
      *it's **a Risino** you are*

[Tikula, sami001.121]

Wh-questions like (14) have focus-presupposition articulation, with the focal domain coinciding with the questioned NP, in this case a secondary object. The context for (15) was a bit of banter concerning the identity of the person being addressed in the example, and doubts were raised about his credentials as a member of the Tikula division of the Acipu. Instead, the speaker jokingly says to the addressee "[i.e. you're not a Kula but...] *it's a Risino you are*".

Example (16) involves a fronted pronoun rather than an NP, but the articulation is perhaps more obviously focus-presupposition. The speakers were discussing a time when there had been no-one willing to be the chief of the Acipu. A slave was caught, and *it was ʜɪᴍ they appointed to the chieftancy.*

(16) **ɛ́vvì**   à-náɗùkwà-nà   tí-wɔ́mɔ́
      3s.COP 3P-appoint\RLS-PFV NC6-chief
      ***it was ʜɪᴍ** they appointed to the chieftancy*

[sayb001.313]

The verbs in (15-16) are marked for perfective aspect with the suffix **-nA** (§4.6.3.1). If the clause is in realis mood, then focusing a constituent makes this suffix obligatory. It is required even when the linear order of constituents does not change, as in the focusing of the subject in (17b).

(17) (a) sáa t-èné   cì-kóotò   tì-písò
           or   AG6-which NC6-drum   AG6-break\RLS
           every drum broke

      (b) t-èné   **cì-kóotò**   t-í   tì-písò-**nò**?
           AG6-which **NC6-drum**   AG6-COP   AG6-break\RLS-**PFV**
           which drum broke?

[eamy006.168, eamy006.169]

OSV clauses and other kinds of 'focus' fronting are not strictly limited to sentences with argument focus. True argument focus (§2.3.1) is very rare in the corpus[8], and it seems to be the case that (non-topical) constituent-fronting serves a more general purpose. The following statement is an approximation of the conditions under which non-topical constituent-fronting is possible:

---

8   This seems to be the case cross-linguistically (e.g. Francis et al. 1999 on English).

(18) *for a non-topical constituent to be fronted it must carry the majority of the 'communicative load' of the clause*

This is very often the case when the object introduces a new referent:

(19) wà   'û̃   xxx cì-mé'ı̃̀ tú↓=u-táarí **kù-yŭ̃ŷ̃** kw-ì   à-zâa-nà
     REPRT there   NC6-inside AG6=NC7-stone **NC9-sand** AG9-COP 3P-see\RLS-PFV

*it is said that there inside the stony place they saw **sand***

[sayb001.230]

Strictly, this sentence does not have argument focus since while the people seeing the sand had already been introduced to the discourse, the event of 'seeing' was not yet activated. So in terms of information structure the sentence has broad (or predicate) focus, with the focal domain comprising the verb and the fronted object. However the object referent **kù-yŭ̃ŷ̃** 'sand' is being introduced for the first time. Example (20) is essentially the same.

(20) **ù-yɔ́ɔ**   w-ì   ò-cú'ò-nò   ó↓=cú-kŭ̃u   tí↓=Ø-gìgínìyà
     NC7-k.o._climber AG7-COP 3P-plant\RLS-PFV LOC=NC6-waist AG6=NC8-deleb_palm

*there was **a climbing plant** they had planted at the foot of a deleb palm*

[saat002.002.190]

Example (21) involves *two* referents introduced for the first time, but they have different statuses with respect to information structure. The utterance begins with a marked external topic **ánà àzá** 'some people', followed by the fronted object **òmódó** 'slaves', and then finally the verb. It is the object referent that is, to use Mithun's (1992) term, *newsworthy*.

(21) á-nà   à-zá   pỗ [ò-módó]ₒ h-è   [ò-'ɔ́pò̀-n-nò̀]ᵥ
     3P-ART NC2-person all **NC2-slave** AG2-COP 3P-hold\RLS-VENT-PFV

*some people they even caught **slaves***

[sayb001.192]

Given that the clause is being used to introduce the slaves into the discourse by means of a particular relationship with the referent of **ánà àzá**, the nature of that relationship ('catching') is fairly predictable. Note that only **òmódó** is part of the focal domain of the clause – **ánà àzá** is outside of the clause altogether.

4.3.2.1.2   Topic

In addition to focused constituents, topicalised constituents can be fronted (or left-dislocated). Most of the time these occur in what appears to be a clause-external position, offset from the clause by a pause:

(22) **ká-tàarà kɔ́↓=w-wɔ́mɔ́ kò-rísìnôo,** gàhí á-bìyà **k-è,**
    NC1-**fine** AG1=NC8-**chief** NC1-**[name]** before 3S-pay\IRR AG1-PRO

    the fine of the chief of Korisino, before they could pay it,

                                                                    [saim001.049]

Other times fronted topics are phonologically integrated:

(23) [Context: *chicken, he calls 'feather'...*]

    mò-yóo ù-sì-hy$\H{a}$a mò-nnó'ò
    NC4-**beer** 3S-HAB-say NC4-porridge

    beer *he calls 'porridge'*

                                                                  [tats001.003.046]

(24) [Context: *he doesn't see things correctly, through beer (**mòyóo**, NC4),*]

    sɔ́bɔ̀dà **m-í** m-è ù-d$\H{a}$a-nà n=ù-y$\H{u}$u-ní kó-dûu
    because AG4-**PRO** AG4-COP 3S-surpass\RLS-PFV and=NC7-wear-NMLZ NC1-heart

    cì-mé'ì̃ tí↓=**m-è**
    NC6-inside AG6=**AG4-PRO**

    *because **it** he sets his heart on it excessively*

                                                                  [tats002.003.063]

Even when a resumptive pronoun is present, as in (24), realis clauses with fronted topics are still differentiated from their non-topicalised counterparts by the perfective suffix **-nA** on the verb. The fact that this suffix occurs with realis verbs in fronted topic constructions as well as fronted focal constituents suggests that the occurrence of the suffix depends on the fronting of the constituent, rather than the particular pragmatic relation involved[9].

    Example (24) is also of interest because of the complexity of the syntactic structure from which the topic is extracted i.e. from the possessor of the second object of a nominalised verb, which in turn is the object of the prepositional complement of the main verb. In pseudo-English the construction is something like ***it he exceeds with putting heart inside of it***. In complex constructions such as these there is rarely a gap – instead there is a resumptive pronoun. This seems to be a common state of affairs cross-linguistically (Kroeger 2004:189), and also applies to Cicipu relative clauses (§4.3.4). The possibility of 'island constraints' on topic-fronting has yet to be investigated.

    Before leaving topicality it is worth noting in passing that as well as left-dislocated topics, Cicipu also allows truly 'external' (Kroeger 2004:137) or 'outer' (Dooley

---

9 Cicipu contrasts with Hausa in this respect – in the latter only focus-fronting results in special verb morphology, not topic-fronting (Jaggar 2001:493).

2007:82) topics, which have no anaphoric link with the following clause:

(25) **kò-rísìnôo** ká'à m-ú-gwàanùkwà cé kàm
NC1-[name] now 1s-FUT-know NEG definitely
Korisino now I really wouldn't know [about the settlement of Korisino]

[sayb001.102]

Finally, we should note that there are two explicit topicalisers **gó** and **hínà**. These are discussed in §8.3.2.

In this subsection we have seen that constituent-fronting does not correlate with any particular pragmatic relation. Instead the referents of fronted constituents may be topics, in focus, or just 'newsworthy'. This suggests that constituent-fronting in Cicipu is similar to sentence accent in English, which can indicate topic, focus, or activation (Lambrecht 1994:322-325).

4.3.3 Non-verbal clauses

This subsection discusses the following kinds of non-verbal clauses: predicate nominals (§4.3.3.1), predicate locatives (§4.3.3.2), existential clauses (§4.3.3.3), and finally presentational clauses (§4.3.3.4).

4.3.3.1 Predicate nominals

4.3.3.1.1 Nouns

Predicate nominal clauses with lexical subjects take the form NP$_1$ NP$_2$ AG$_2$-COP, with the copula agreeing in gender with the *second* NP. The copula is identical segmentally to the noun class pronouns (§4.4.3.2)[10]. Examples (26) and (27) show PROPER INCLUSION (Payne 1997:114), since the referent of the first NP is being said to belong to the class of referents indicated by the second.

(26) [kù-yúpù] [Ø-kwáarò] v-ì
NC9-crocodile NC8-creature[11] AG8-COP
the crocodile is a creature

[tats002.006.001]

---

(27) [kà-dámá   ká-avù]   [mà-gɛ̂i]   m-è
    NC1-word   AG1-2S.POSS   NC4-sword   AG4-COP

your word is a sword

<div align="right">[oamy001.024]</div>

The behaviour of the copula in agreeing with the predicating NP rather than the subject is unusual, and is an instance of 'attraction' or 'back' agreement (Corbett 2006:63-64)[12] (note that quite apart from semantic considerations, the second NP in (26-27) cannot be analysed as a subject without violating the otherwise universal restriction that Cicipu subjects come before predicates (§4.3.1)).

Examples of EQUATIVE predicate nominals involving two lexical NPs are non-existent in the corpus. Example (28) was elicited, but since the initial NP is set off from the rest of the clause by a pause, is difficult to distinguish this from the single argument identificational clauses illustrated in (29-30)[13].

(28) [à-zá-mpà],   [ɔ̀-gɔ́rɔ́mɔ̀   hò-vôo]   h-è
    NC2-person-this   NC2-elbow   AG2-1S.POSS   AG2-COP

these, they are my elbows

<div align="right">[eamy005.037]</div>

(29) z-zá   n-nà   dóonù-nà   á↓=Kángù,   w-à   **k-kɔ́gɔ́**   **v-ì**
    NC8-person   AG8-REL   sit\RLS-PFV   LOC=[name]   3S-REPRT   **NC8-Hausa**   **AG8-COP**

*and the one who settled at Kangu, they say **he was a Hausa***

<div align="right">[sayb001.185]</div>

(30) mà-hɛ̃̂u   m-è
    NC4-truth   AG4-COP

it's true [lit. 'it's truth']

<div align="right">[tats001.004.089]</div>

The position of the copula with respect to noun modifiers is variable, as shown in the following pairs of examples. The difference, if any, with respect to meaning or usage is not yet understood.

(31) kà-nnú   **k-è**   kà-pénénêu
    NC1-bird   **AG1-COP**   AG1-big:REDUP

it's a very big hawk

<div align="right">[saat001.006.086]</div>

---

12 This possibility also exists in Hausa (Newman 2000:162-163, Jaggar 2001:458-459), although agreement with the subject is more usual. The term 'attraction' is preferable to 'back' agreement, since the copula comes after both NPs, in both Cicipu and Hausa.

13 Along the same line of reasoning, an alternative analysis of (26-27) might be that, rather than being cases of attraction or 'back' agreement, they consist of a left-dislocated topic followed by a single argument identificational clause (recall from §4.3.2.1.2 that no pause is necessary after a left-dislocated topic).

(32)　m-ɔ́ɔ　　mè-ɗénêu　**m-é**
NC4-child　AG4-small　**AG4-COP**

it's a small child

[tats007.002.078]

(33)　kà-gògá　　**k-è**　　ká↓=u-tűwằ　wú↓=mó-ní
NC1-rubber_bag　**AG1-COP**　AG1=NC7-fetch　AG7=NC4-water

it's the bag for fetching water

[saat001.005.069]

(34)　tì-wɔ́mɔ́　tí↓=má-gàjì　tí↓=o-kíisó　　**t-ì**
NC6-chief　AG6=NC4-priest　AG6=NC2-maigiro　**AG6-COP**

it's the priesthood [lit. 'chieftancy of the priest'] *of the Maigiro*

[tats001.003.003]

4.3.3.1.2　Personal pronouns

When the first NP in a predicate nominal construction is a pronoun rather than a lexical NP, there at first sight seems to be a difference between the expression of (i) proper inclusion and (ii) equation. The tone of the pronoun in the former type of construction is L L, and the copula appears after the second NP:

(35)　[Context: Speaker is giving further information about some people]

èrè　　ò-módó　h-è
3P.PRO　NC2-slave　AG2-COP

they *are slaves*

[2008-03-07.003]

(36)　**ìvɔ̀**　　Ø-sáhwá　　v-ì
2S.PRO　NC8-simpleton　AG8-COP

you *are a simpleton*

[svtmg001.124]

By contrast, in the equative construction the tone on the pronoun is H L and there is no explicit copula. Additionally, the first- and third-person pronouns differ segmentally from the standard independent pronouns (§4.4.3.1):

(37)　[Context: Speaker had previously talked about some slaves. He then mentions another group of people and says...]

érè　　ò-módó　há-nà
3P.COP　NC2-slave　AG2-ART

THEY *are the slaves*

[2008-03-07.003]

163

(38) [Context: speaker is indicating the king while speaking]

évvì d-dáa
3s.cop NC8-king

HE is the king

[svgd001.002]

However this 'equation' form of the pronoun is also found in verbal clauses with focus-presupposition articulation:

(39) [Context: Summarising a discussion of the three founders of Korisino]

érè hwáarà-nà
3P.cop start\RLS-PFV

THEY started [i.e. they were the ones who started]

[sayb001.304]

Therefore it is probably better to analyse these 'equation' forms as focus forms of the independent pronouns. The paradigm is given in Table 16. These pronouns are probably diachronically derived from standard independent pronouns plus the copula (see §8.9.4).

Table 16: Focused independent personal pronouns

	SG	PL
1	ámbì	óttù
2	ìvɔ́ vì	ìdɔ́ yì
3	évvì	érè

4.3.3.2 Predicate locatives

These are formed using the focused form of the pronoun, as in (40). Example (41) shows the same structure with a gender-marked pronoun, and (42) with a lexical NP. The copula occurs immediately after the subject.

(40) évvì lêe 'ɔsɔ́nì
3s.cop there nearby

it was there nearby

[tats004.001.030]

(41) v-í v-è 'ū̀ cɔ́ɓɔ̀
AG8-PRO AG8-cop there underneath

it is there underneath

[eaim002.1388]

(42) mè-ì-rí-mpà   m-è   lêe
      NC4-NC3-thing-this AG4-COP   there
      this little thing is there

For negative locative predicative clauses see §4.3.5.1 below.

The five Cicipu demonstrative adverbs (§4.4.1.1) can also function as predicates, in which case they agree in gender with the subject. In response to the question **hán X?** 'where is the X?', one possible answer is **X AG$_x$-pâa** 'the X is here'.

(43) kà-táarí   kà-**pâa**
      NC1-stone   AG1-**here**
      the stone is here

4.3.3.3 Existential clauses

The simplest form of existential clause in Cicipu consists of the predicator **òkóo** 'there is/there are' followed by an NP. The NP complement of the predicator is optionally subject to the complement tone perturbation discussed in §3.4.7. **Òkóo** is borrowed from Hausa *akwai*, which is also used in Cicipu with the same meaning, seemingly in free variation.

(44) **òkóo**   cí-yímbì
      there_is   NC6-dark
      there was forest [lit 'darkness']

(45) ká'à **òkóo**   kà-jírígí   ká↓=g-gédù
      now   **there_is**   NC1-canoe   AG1=NC8-up
      now there are aeroplanes

Predication using **òkóo** is also a common means of attributing qualities to a referent. In such cases the referent nominal often occurs before the predicator, similar to the topicalised NPs discussed in §4.3.2.1.2 above.

(46) kè-ré'è   kí-ivè   **òkóo**   ù-pácí
      NC1-tongue   AG1-3P.POSS   **there_is**   NC7-difficulty
      their language is hard

An alternative construction uses the preposition **ǹ** 'with' to link the two NPs:

(47) vɔ́-ɔmɔ̀,   òkóo   èvì   ǹ   Ø-ɓàhánná
NC8-monkey **there_is** 3s.PRO with NC8-destruction

the monkey, it is destructive [lit. 'monkey, it is with destruction']

[tats002.001.074]

On rare occasions, the complement of ** òkóo/àkwái** is followed by the copula.

(48) **òkóo**   mò-yíló   mú↓=u-húná m-è
there_is NC4-sacred AG4=NC7-kill AG4-COP

there is a taboo against killing it [**kù-yúpù** 'crocodile, NC9']

[saim001.090]

For negative existential clauses see §4.3.5.1 below.

4.3.3.4   Presentational clauses

The presentative **ǹdúu** is used in presentational clauses, similar to its Hausa equivalent *ga* (Newman 2000:181-182, Jaggar 2001:468-469). These can either be exophoric, indicating referents present in the speech situation, or text-internal, in which case they introduce a new referent into the text-world. **Ǹdúu** may be historically-derived from the verb **inda** 'see', or perhaps even from **ǹ-dûu** 'I stretch out s.t towards'.

(49) **ǹdúu↓**   mó-ní   sɔ́ɔ!
here_is NC4-water drink\IMP

here's water, drink!

[saat002.002.303]

(50) **ǹdúu↓**   bɔ́wɔ́   pâa
here_is thief   here

here's the thief here

[saat002.001.119]

Ǹdúu can also operate at the clausal level, especially if the clause encodes a backgrounded event. It seems the event itself is being 'presented' for consideration, and the meaning can usually be conveyed by *you see* in English translation.

(51) tí-dùkwà   **ǹdúu**   t-índà-nà   'ásù ù-dḗi
1P-go\IRR **here_is** 1P-see\RLS-PFV place NC7-settlement

let's go, you see we've seen a place to stay

[sayb001.169]

4.3.4   Relative clauses

Relative clauses in Cicipu consist of three parts which occur in the following order:

(52) (HEAD NOUN) – RELATIVISER **nà** – MODIFYING CLAUSE

The relativiser agrees in gender with the head noun, but since there is often a resumptive pronoun or clitic (agreeing in gender or person) in the modifying clause, **nà** is considered to be a relativiser rather than a relative pronoun. If the modifying clause is in realis mood then the perfective aspect suffix **-nA** (§4.6.3.1) is obligatory, just as when wh-words (§4.3.6.2) or focused or topical constituents are fronted (§4.3.2.1)[14].

There are few restrictions on the function of the relativised constituent within the modifying clause, as the examples below illustrate. The relativised constituent is indicated either by a gap (shown as Ø), or a resumptive pronoun. The gap strategy can apply at any point on Keenan and Comrie's (1977) Accessibility Hierarchy (SUBJ > OBJ > OBL > POSSESSOR), as well as to time and place phrases (59-60). In the case of obliques, the preposition is gapped as well as the NP, as illustrated by (57):

(53) **ì-rí   yí-nà**   Ø   yì-lábà-nà   yì-'étě̃i
NC3-thing AG3-REL   AG3-lack\RLS-PFV AG3-fine
bad things [lit 'things which lack fineness']

[SUBJECT, GAP, oamy001.128]

(54) **mà-kǔ̃u   má-nà**   ù-kúmbà-nà   Ø
NC4-mountain AG4-REL   3S-climb\RLS-PFV
the mountain which he climbed

[OBJECT, GAP, Tikula, sami001.068]

(55) **à-zá   há-nà**   mà-gǎi   **má-nà** mà-sì-tá'á   / má-hùnà   Ø
NC2-person AG2-REL   NC4-sword AG4-ART AG4-HAB-want AG4-kill\IRR
the people the sword wanted / to kill

[EMBEDDED OBJECT, GAP, samy001.015]

(56) òkóo   **kò-móorí ká-nà**   tú-u-càa = vù   Ø
there_is NC1-task   AG1-REL   1P-FUT-give\IRR=2S.PRO
there's a task which we will give you

[SECONDARY OBJECT, GAP, saat002.002.463]

(57) **í-cìnjì   yí-nà**   wú-u-ràa   ì-rí   yí↓ = ká-ráa Ø / (**ǹ   y-í**)
NC3-money AG3-REL   3S-FUT-eat\IRR NC3-thing AG3=NC1-eat   **with AG3-PRO**
the money which he would eat food (with it)

[OBLIQUE, GAP, tats002.003.014]

---

14 The phonological identity between the relativiser/article and the perfective suffix is presumably accidental. However the perfective suffix **-nA** which occurs in relative clauses and clauses with fronted constituents is assumed to be a single morpheme (cf. Jaggar (2001:162) on a similar constellation of functions assigned to the Focus TAMs in Hausa).

(58) **mò-ní** **pồ** **má-nà** à-hyâa-nà òkóo kù-yúpù cìmé'ị̀ Ø
    NC4-water all AG4-REL 3P-say\RLS-PFV there_is NC9-crocodile NC6-inside
    all the water said to have crocodiles inside (of it)
                                        [POSSESSOR, GAP, tats002.006.014]

(59) **Ø-lóokàcí** **n-nà** ù-dôohò-nò Ø
    NC8-time AG8-REL 3S-disappear\RLS-PFV
    when it disappeared
                                        [TIME, GAP, tats004.001.036]

(60) **'ásù** **wú-nà** Yóoyò kámùkwà-nà Ø
    place(NC7) AG7-REL [name] was\RLS-PFV
    the place where Yooyo was
                                        [PLACE, GAP, sayb001.043]

The resumptive pronoun strategy is not attested within the corpus for relativised possessors or subjects. Resumptive pronouns are common for relativised obliques in matrix (61-62) clauses, but less so for relativised objects (63). Resumptive pronouns can at least be elicited for relativised possessors (64), and (in subordinate clauses) for subjects (65). Relativised time and place constituents do not have resumptive pronouns.

(61) **kw-á'à** **kú-llè** kú-nà ù-yó-nò á↓=**kw-ì**
    NC9-house AG9-that AG9-REL 3S-be\RLS-PFV LOC=**AG9-PRO**
    that house which he is at
                                        [OBLIQUE, RESUMPTIVE PRONOUN, tats002.005.011]

(62) **ì-rí** **yí-nà** Ø-lówólì Ø-yó-nò ǹ **y-í** á↓=kw-á'à
    NC3-thing AG3-REL NC8-big_spider AG8-be\RLS-PFV with **AG3-PRO** LOC=NC9-house
    kwé-evì
    AG9-3S.POSS
    what Big Spider had in his house
                                        [OBLIQUE, RESUMPTIVE PRONOUN, saat002.002.240]

(63) **ì-rí** **yí-nà** tì-zâa-nà **y-í**
    NC3-thing AG3-REL 1P-find\RLS-PFV **AG3-PRO**
    what we found
                                        [OBJECT, RESUMPTIVE PRONOUN, Tikula, sagb001.308]

(64) **ì-rí** **yí-nà** ò-nóɔ-nò=tù kà-nàbáyì [Ø/ ká=**y-ì**]
    NC3-thing AG3-REL 3P-give\RLS-PFV=1P.PRO NC1-story AG1=**AG3-PRO**
    the thing they gave us the story (of it)
                                        [POSSESSOR, GAP OR RESUMPTIVE PRONOUN, Tikula, sami001.053]

168

(65) **kà-bárá ká-nà** má-llû   mà-dámà-wà-nà   wɔ́mɔ́   w-à
     NC1-elder AG1-REL NC4-teacher AG4-speak\RLS-APPL-PFV chief   3S-REPRT

     [Ø/ **k-é**]   **kà**-húnà   kù-yúpù   ǹ   kɔ̀-ɓɔ́ɔ
        AG1-PRO AG1-kill\RLS NC9-crocodile with NC1-axe

     the elder that the teacher told the king (he) had killed the crocodile with an
     axe
                    [EMBEDDED SUBJECT, GAP OR RESUMPTIVE PRONOUN, 2008-04-01.002]

(66) **kù-yúpù   kú-nà** má-llû   mà-dámà-wà-nà   wɔ́mɔ́ w-à
     NC9-crocodile AG9-REL NC4-teacher AG4-speak\RLS-APPL-PFV chief 3S-REPRT

     kà-bárá   ká-húnà   [Ø/ **kw-ì**]   ǹ   kɔ̀-ɓɔ́ɔ
     NC1-elder AG1-kill\RLS **AG9-PRO** with NC1-axe

     the crocodile that the teacher told the king the elder had killed (it) with an
     axe
                    [EMBEDDED OBJECT, GAP OR RESUMPTIVE PRONOUN, 2008-04-01.002]

(67) **kɔ̀-ɓɔ́ɔ ká-nà** má-llû   mà-dámà-wà-nà   wɔ́mɔ́   w-à
     NC1-axe AG1-REL NC4-teacher AG4-speak\RLS-APPL-PFV chief   3S-REPRT

     kà-bárá   ká-húnà   kù-yúpù   **[Ø /ǹ   k-é]**
     NC1-elder AG1-kill\RLS NC9-crocodile **with AG1-PRO**

     the axe which the teacher told the king that the elder had killed the crocodile
     (with it)
                    [EMBEDDED OBLIQUE, GAP OR RESUMPTIVE PRONOUN, 2008-04-01.002]

In summary, as is typically the case cross-linguistically, the resumptive pronoun strategy
becomes more common with respect to gapping for relativised functions which are
lower down the NP Accessibility Hierarchy. In the same way, the deeper the level of
embedding between the head noun and the gap/resumptive pronoun, the greater the
likelihood that a resumption pronoun will occur instead of a gap.

   No potential island constraints have been tested yet.

   As in many languages (Lyons 1999:61), the Cicipu relativiser is formally identical
to the article (§4.4.5.3), but unlike the article the relativiser is obligatory and does not
depend on the identifiability/uniqueness of the head noun referent – in fact it can even
co-occur with the pre-head "specific indefinite" article[15]:

(68) òkóo **wú-nà** ká-llù   **ká-nà** à-yɛ̂a-nà   lêe
     there_is **3S-ART** NC1-hunger **AG1-REL** 3P-do\RLS-PFV there

     there was a certain hunger that happened there [lit 'they did there']
                    [sayb001.622]

Finally, it should be observed that the head noun in the schema given in (52) is optional.

---

15 In Hausa relativised NPs may also be marked by both the specific indefinite determiner *wani* and the
   definite/anaphoric linker *-n* (see Newman 2000 chp. 64 for examples).

If the referent is clear from the context then the head noun may be omitted:

(69) **ká-nà** tì-yáddà-nà
AG1-ART 1P-agree\RLS-PFV

one [**kà-dámá** 'word', NC1] *that we agree with*

[svbg001.032]

4.3.5 Negation

4.3.5.1 Clausal negation

The most common means of negating a clause is through the negation particle **cé**, which occurs post-verbally.

(70) kà-záakì kà-lápà **cé**
NC1-lion AG1-know\RLS **NEG**

the lion didn't know

[saat001.007.010]

The tone on **cé** is high and sometimes extra-high (§3.4.1). If there is a pronoun in the complement position of VP then **cé** usually follows it, but nouns always come after the negator (see §7.3.2 for details).

The clause-initial negation particle **bâa**, borrowed from Hausa, also operates at a clausal level[16]:

(71) **bâa** ú-hyȁa = vù[17]
NEG 3S-say\IRR=2S.PRO

he shouldn't say to you [i.e. 'it is not the case that he should say to you']

[tats005.001.215]

(72) **bâa** mò-cíyò à-zá há-nà
NEG AG4-get\RLS NC2-person AG2-ART

it didn't get the people

[samy001.039]

Negative locative predications (73) and existential propositions (74-75) can also be expressed using **bâa**.

---

16 **Bâa** is subject to variation in tone and vowel-length. After the conditional conjunction **ń** 'if', **bà** usually occurs with a short low-tone vowel (example (73) is an exception). **Bá** and **báa** also occur, but the factors governing the distributions of these forms are not well understood (cf. Hausa where different forms of *ba* have different functions – Jaggar 2001 chp. 6).

17 This kind of speech act is usually encoded in a different way, using the negation particle **kádà** (§4.6.2.2).

(73) ń **bâa** ìvɔ̀ lêe
if **NEG** 2S.PRO there
if you're not there

<div align="right">[tats008.004.004]</div>

(74) **bâa** z-zá n-nà wú-u-sɔ̀'ũ̀ y-ì
NEG NC8-person AG8-REL 3S-FUT-touch\IRR AG3-PRO
no-one would touch it [lit. 'there does not exist one who would touch it']

<div align="right">[sayb001.767]</div>

(75) mà-wáa **bâa** kù-cíyè
NC4-dog **NEG** NC9-hand
dogs don't have hands

<div align="right">[tats008.004.021]</div>

Negative existential **bâa** clauses can also modify an NP, in which case the interpretation is that the referent of the head noun does not posses the entity marked with **bâa**.

(76) à-yẫa ú-màtà mí-nà **bâa** à-'ízá
3P-do\RLS NC7-give_birth AG5-ART **NEG** NC2-tail
they gave birth to some [**m-úu** 'children, NC5'] *without tails*

<div align="right">[svtmg001.230]</div>

'Double negatives' may occur with both **bâa** and **cé**, but the meaning is still negative – the two negators do not cancel each other out:

(77) Niger kâi! **bâa** Nigeria **cé**
annoyance **NEG** **NEG**
Niger! not Nigeria

<div align="right">[Tikula, sami001.436]</div>

4.3.5.2 Constituent negation

Negation in Cicipu can take scope over a single constituent rather than a clause. The same negation morpheme **cé** is involved, but this time it takes an optional prefix agreeing with the gender of the constituent, as in (78). If the constituent concerned is a personal pronoun then the focus form of the pronoun (§4.3.3.1.2) is used, as in (79).

(78) cì-'ítàní ù-kôo **wù-cé**
NC6-marriage NC7-death AG7-NEG
marriage is not death

<div align="right">[ovfl001.001.001]</div>

(79) **érè**   **cé**   ò-dóon<ìs>ù-wò-nò=vì   lêe
      3P.COP **NEG** 3P-sit\RLS<CAUS>-ANTIC-PFV=3S.PRO there

it wasn't them that caused him to be settled there [i.e. he came of his own accord]

[svbg001.061]

Finally, it can be noted that Cicipu has no negative quantifiers. There are no words corresponding to *nobody* – instead the negation particle **bâa** is used to assert the non-existence of a particular kind of referent.

(80) **bâa** z-zá   n-nà   wú-u-sɔ̀'ṹ   y-ì
      NEG NC8-person AG8-REL 3S-FUT-touch\IRR AG3-PRO

there was no-one who would touch it [i.e. 'there did not exist one who would touch it']

[repeated from (74)]

4.3.6 Questions

4.3.6.1 Yes/no questions

It is likely that yes/no questions have a special intonation pattern in all languages (Payne 1997:295), regardless of whatever other formal means are available. For Cicipu the situation is similar to Hausa and many other African languages (Rialland 2008) in that the usual tonal downdrift (§3.4.2) is suspended. The extra-high final H characteristic of Hausa (Newman 2000:497-499) has not been observed.

Yes/no questions may also be marked with the optional sentence-final question particle **sṹ**.

(81) w-áa   á↓=d-dòorí   z-zá   ù-gúyà   ù-híyà   kwáanù góomà **sṹ?**
      3S-REPRT   LOC=NC8-formerly NC8-person 3S-can\RLS NC7-grind basin   ten   **Q**

it is said that in the past someone could grind ten measures?

[svtmg001.018]

Sṹ can be used across turn-transitions to question the previous speaker's statement. As the following example illustrates it can form a complete utterance by itself.

(82) A: éví   z-zá   ɗ-ɗénêu   v-vôo
          3S.PRO   NC8-person AG8-small AG8-1S.POSS

      he's my younger sibling

      B: **sṹ?**
          Q
          really?

[2008-03-07.002]

Tag questions can be formed as in Hausa (Jaggar 2001:524-525), by placing **kóo** 'or' at the end of the utterance. As in Hausa, the tone on **kóo** in this environment is actually rising viz. **kǒo**. The native Cicipu equivalent **sáa** (§4.4.5.4) followed by one of the topicalisers **gó** or **hínà** (§8.3.2) can also have this function.

Sáa can also be used to form indirect questions:

(83) nàhá   ń-jòolò-nò   **sáa**   yì-níi'wà
      let\IMP   1s-check\IRR-VENT(?)   **or**   AG3-soak\RLS
      *let me check **whether** it is soaked*

<div align="right">[saat001.008.103]</div>

No word-order changes have been observed in yes/no questions.

4.3.6.2   wh-questions

Question words are usually fronted in straightforward wh-questions, but remain *in-situ* for echo and rhetorical questions. A list of wh-words in Cicipu is shown in Table 17.

Table 17: Cicipu wh-question words

Question word	Meaning	Position in straightforward questions[18]
yìní (NC3)	*what*	fronted
yàaní / yàanú (NC8)	*who*	fronted
hán / hànú	*where*	fronted or *in-situ*
-èné	*which*	fronted or *in-situ*

Example (84) shows a direct request for information with clause-initial **yìní** 'what'. Example (85), on the other hand, was a gently-mocking rhetorical question with *in-situ* **yìní**, carrying the implication that the people being indicated didn't know anything.

(84) **yìní**   à-húlá   hé-evì
      what   NC2-name   AG2-3S.POSS
      ***what** is its name?*

<div align="right">[sayb001.340]</div>

(85) à-zá-mpà   à-lápà   **yìní**?
      NC2-person-this   3P-know\RLS   **what**
      *these ones **what** do they know?*

<div align="right">[Tikula, sagb001.076]</div>

The word for 'who' has two variants **yàaní** and **yàanú**. The former is more likely to

---

18 This kind of 'mixed' behaviour (where wh-fronting is obligatory in some cases but not in others) is rare in Benue-Congo, although it is reported for both Hausa and the West Kainji language Duka (Dryer 2005, Newman 2000:496).

occur before words whose first vowel is [+FRONT], while the latter is more likely to occur otherwise, although as (86) shows this is not an absolute rule (cf. §3.3.2).

(86)   **yàaní**   Ø-yó-nò   ú-làngwà = mù   ǹ-lávù
        who   AG8-be\RLS-PFV   NC7-spoil=1s.PRO   NC5-sleep
        who's spoiling my sleep?

                                                 [Tidipo, saat002.004.055]

There are two alternatives for the word for 'where'. **Hán** occurs in fronted position, while **hànú** is found *in-situ*.

(87)   **hán**   kú-lácí   kú-nà   Ø-yó-nò   ú-tà'à?
        where   NC9-young_girl   AG9-REL   2s-be\RLS-PFV   NC7-want
        where's the girl that you want?

                                                        [saat002.002.458]

(88)   ì-kámà   **hànú?**
        2P-be.PST\RLS   **where**
        where were you(pl.)?

                                                           [sayb001.033]

The question words so far have been 'pronominal', in that they have stood in for NPs. **-èné** 'which' is different in that it usually modifies a head noun. It takes an obligatory prefix agreeing with the head noun either in gender or in person. When agreeing in gender, the modifier can occur either before (89) or after (90) the head noun. When agreeing in person, the modifier seems to be limited to the pre-head position (91). An identical pattern with respect to agreement feature and position is found with the article and the demonstrative – see §8.9 for more details.

(89)   **kw-èné**   kú-lácí   kw-ì   kw-áyà-nà
        AG9-which   NC9-young_girl   AG9-COP   AG9-come\RLS-PFV
        which girl came?

                                                        [eamd032.185]

(90)   á↓=ká-káasùwà   **k-èné**
        LOC=NC1-market   **AG1-which**
        *to **which** market?*

                                                          [sayb001.721]

(91)   **w-èné**   Ø-'írì
        3s-which   NC8-kind
        ***which** kind?*

                                                        [sayb001.329]

Yìní 'what, NC3' can also function as a modifier in this way, allowing the questioning of particular components of a NP. In this case it too agrees with the head noun.

(92) kà-'ázzìkì   kè-**yìní**?
      NC1-prosperity AG1-**what**
      what prosperity?

<div align="right">[saat002.002.392]</div>

When **yàanú** 'who, NC8' is used in this way it takes on the meaning 'how many', as in (93). Note the contrast between the low-tone prefix and short consonant in (93) and the high-tone prefix and long consonant in (94). The latter two properties are characteristic of associative constructions with NC8 possessors (§4.4.5.1, see also §6.4.2 on neutral agreement).

(93) kò-dóntú   kà-**yàanú**   k-è?
      NC1-chair   **AG1-how_many** AG1-COP
      How many chairs are there?

<div align="right">[2008-02-21.006</div>

(94) kò-dóntú ká↓=y-**yáanú** k-è?
      NC1-chair   AG1=NC8-**who**   AG1-COP
      Whose chair is it?

<div align="right">[2008-02-21.006]</div>

To question the time of an event the composite phrase **kwá̰a kw- èné** 'which day' can be used (95). Similarly the manner in which some event was carried out can be questioned using **t-èné** 'how', which can be analysed as -**èné** 'which' together with an AG6 agreement prefix (96). When **t-èné** is fronted it is usually followed by the copula.

(95) vá-ayà   kw-ã̰a   **kw-èné**?
      2S-come\RLS NC9-day   **AG9-which**
      ***when** did you come?*

<div align="right">[Lexicon]</div>

(96) **t-èné**   t-ì   à-sì-yã̰a   kà-bíkì   ká↓='á-kkwíi
      AG6-which   AG6-COP 3P-HAB-do NC1-festival AG1=NC2-dead_person
      ***how** do they do the festival of dead people?*

<div align="right">[Tikula, sagb001.412]</div>

To question the reason for some event or state, the wh-question word **yìní** 'what' can be used followed by the verb **yũu** 'cause'.

<div align="center">175</div>

(97)   **yìní**   yì-yǔu-nà   Ø-tíyò-nò   cé?
        what   AG3-cause\RLS-PFV   2S-get\RLS-VENT   NEG

why didn't you(sg.) get [any]? [lit. 'what caused you didn't get?']

[saat001.004.022]

Finally, it should be noted that each wh-question word has a universal quantifier counterpart, formed by placing **sáa** 'whether/or' in front of it (see §4.9).

4.3.7   Clause coordination and subordination

Perhaps nowhere else in the grammar of Cicipu is the influence of Hausa seen so strongly as in the methods of clause combination. Not only are the various conjuncts and other particles often taken straight from the lingua franca, but the constructions themselves are usually identical in format. The discussion in Jaggar (2001:592-642) could to a large extent have been reproduced here verbatim, but for reasons of space (and scholarly integrity!) it is only possible to mention a few of the more common structures here.

4.3.7.1   Coordination

The usual NP-coordinator **ǹ** (§4.4.5.4) does not conjoin clauses in Cicipu. Instead it is common to find simple juxtaposition of clauses, either with intervening pauses (98) or without (99).

(98)   ù-vóndòrò   mà-díyá,   ù-húnà,   ù-tɔ́pù   á↓=kɔ́-cɔ́kɔ́   ù-dúkwà
       3S-sling\RLS   NC4-hare   3S-kill\RLS   3S-put_inside\RLS   LOC=NC1-bag   3S-go\RLS
       ù-zǔwà̰
       3S-roast\RLS

he slung [it] at the hare, he killed, he put it inside his bag he went and roasted

[Tidipo, saat002.003.131]

(99)   w-áyà   ù-júngò   rú-pɔ̃́ɔ   wù-útò-wò-nò   kú-náa   ù-hyã̂a   tò
       3S-come\RLS   3S-open\RLS   NC3-granary   3S-rise\RLS-APPL-VENT   NC9-leg   3S-say\RLS   OK

then he opened the granary he brought out the leg he said OK

[saat002.002.610]

Usually the subjects are identical in paratactic chains of this kind, but not always, as (100-101) show. There is no 'same subject' marker in Cicipu as there is in some Bantu languages (e.g. Watters 2003:254) – the fact that the two verbs in (101) have different subject referents can only be determined from the context.

(100) ká-llù   kɔ̀-'ɔ́pɔ̀=rè   à-dúkwà   á↓=kú-jénè
       NC1-hunger   AG1-hold\RLS=3P.PRO   3P-go\RLS   LOC=NC9-river
       hunger got hold of them and they went to the river

                                                                    [saat001.006.022]

(101) à$_i$-câa=rè$_j$ à$_j$-dúkwà
       3P-give\RLS=3P.PRO   3P-go\RLS
       they$_i$ gave to them$_j$ and they$_j$ went

                                                                    [saat002.002.226]

As in Hausa (Jaggar 2001:593), the particle **kúmá** 'also/and' may be used to provide a more explicit signal of coordination. The position of **kúmá** varies, but it is fairly common to find it after the subject NP (if any) and before the verb.

(102) [Context: *God said let the sun come out to give light*]

       ù-lénjí   **kúmá**   wù-útò-nò
       NC7-sun   **and**   AG7-rise\RLS-VENT
       and the sun came out

                                                                      [tats001.004.028]

If the subject referent and the verb are identical in both clauses then **kúmá** occurs post-verbally, adjacent to the constituent being contrasted.

(103) sɔ́bɔ̀   ǹ   'ínì,   ù-láa   wù-yô   ǹ=ù-'úsṹ /
       because   and   like_that   NC7-fire   AG7-be\RLS   and=NC7-power

       wù-yô   **kúmá**   ǹ   tì-lípã̀i /   wù-yô   **kúmá**   ǹ=ù-pácì /
       AG7-be\RLS   **also**   and   NC6-good   AG7-be\RLS   **also**   and=NC7-difficulty
       *because of that, fire has power / it is **also** good / it is **also** difficult*
                                                                    [tats002.002.059]

Temporally-sequenced classes are often separated by the conjunction **sée** 'until/unless/then' (from Hausa *sai*[19]). If the verb follows immediately afterwards then the tone pattern is the one indicating irrealis mood (§4.6.2.2), even if the clause is reporting an event that would (in the absence of **sée**) be expressed in the realis mood.

(104) ù-zâa   Ø-wɔ́mɔ́,   **sée**   ú-hyã̀a   ã́!
       3S-see\RLS   NC8-chief   **then**   3S-say\IRR   [surprise]
       *he saw the chief, **then** he said a!*
                                                                    [Tikula, sagb001.777]

As well as marking a temporal sequence, **sée** can also be found in the apodoses of conditional clauses (cf. English *then*).

---

19 The vowel in Hausa *sai* [ei] is phonetically very similar to the [eː] in Cicipu **sée**. See discussion in Newman (2000:402) and Jaggar (2001:11).

(105) ǹ   bâa   mò-ní   'ɔ́sɔ́nì   sée   ík-kàbà   Ø-búutà
    if   NEG   NC4-water   nearby   **then**   2S-take\IRR   NC8-water_bottle

if there is no water close by, then you take a water bottle

[tats008.003.025]

The disjunctive conjunction **sáa** 'or' may be used to conjoin clauses (see §4.4.5.4 for NP disjunction), often followed by one of the topicalisers **gó**[20] or **hínà** (§8.3.2).

(106) ù-yô   ǹ   mó-olò   **sáa**   gò   ù-yô   ǹ   Ø-gùugé?
    3S-be\RLS   with   NC4-guitar   **or**   TOP   3S-be\RLS   with   NC8-violin

Did he have a guitar or did he have a violin?

[saat001.004.039]

Sáa has a number of other uses in Cicipu that parallel those of *ko* 'or' in Hausa (Newman 2000:137, Jaggar 2001:394-395), in particular its use in contrastive-focus 'even' constructions ((107), cf. Jaggar 2001:512) and in forming indirect questions (§4.3.6.2) and universal quantifiers (§4.9)[21].

(107) **sáa**   àkwái   ìvɔ́   'ɔ́sɔ́nì
    or   there_is   2S.PRO   near

even if you are close by

[tats001.001.047]

Contrastive coordination can be expressed using the conjunction **àmáa**, again a borrowing from Hausa *amma*[22].

(108) w-wɔ́mɔ́   z-zá   t-tò   v-í,   **àmáa**   ò-wɔ́mɔ́   ò-yô   gɛ́i
    NC8-chief   NC8-person   AG8-one   AG8-COP   **but**   NC2-chief   3P-be\RLS   much

'chief' is one, but 'chiefs' are many

[eabg001.058]

4.3.7.2   Subordination

Subordinate adverbial clauses are particularly similar in Cicipu and Hausa. The following table compares the Hausa subordinators listed in Jaggar (2001:606-608) with their Cicipu equivalents.

---

20 **Gó** often has low-tone when it occurs after **sáa**. The reason why is not known.

21 **Sáa** may also be historically related to the counterfactual auxiliary **sáa** (§4.6.2.4) – this latter construction does not seem to have an analogue in Hausa.

22 In Cicipu both **àmáa** and **àmmá** are heard – the former seems to be more common.

Table 18: Hausa subordinators and their Cicipu equivalents

Gloss	Hausa	Cicipu
when	dà	ǹ
counterfactual	dàa	dàa
because	dòomin / don	dòomí / dón
up until	har	hárì
if/when	ìdan / in	ìdón / ń
before	kàafin / kàamin	kàsí / kàmí
even if	kóo	sáa
as long as	muddìn	múddìn
until/unless	sai	sée
after	bayan dà	báyán ǹ / kù-cínó kú- 'back of'
like	kàma	kàmá
even though	koo dà	sáa dà
when	lookàcîn dà / saa'àn dà	lóokàcí nnà / sá'à nnà
instead of	màimakon	màimákó ví-
because	sabòodà	sóbòdà / sóbò ǹ
until	sai dà	sée ǹ
since	tun dà	tíndà / túndà

Clauses introduced by the conjunction **ánà** 'how/like' play an important cohesive role in Cicipu discourse, in particular in tail-head linkage (Thompson and Longacre 1985:209-213), which is extremely common in Cicipu narrative. The dependent clause is introduced by **ánà** and its verb always occurs in the perfective aspect (§4.6.3.1).

(109) ù-zɯ̋wằ   kà-ká'ìlâa   ké-evì /
      3s-roast\RLS   NC1-chameleon   AG1-3s.POSS

ánà   ù-zɯ̋wằ-**nà,**   ù-râa
how   3s-roast\RLS-**PFV**   3s-eat\RLS

he roasted his chameleon / having roasted [it], he ate
    [Tidipo, saat002.003.066]

This example occurred as part of a longer chain of such back-references, diagrammed below:

Table 19: Example of intra-paragraph tail-head linkages

Sentence no.	Dependent clause	Matrix clause
1		he roasted his chameleon
2	ánà he had roasted-PFV	he ate
3	ánà he had finished-PFV eating	he rested there
4	ánà he had rested-PFV	he said...

The situation recapitulated in the dependent clause does not have to be an event; it can also be a state, as illustrated by (110). The conjunction ánà still occurs, and so cannot be properly glossed as 'when' – better would be something like 'it being the case that...'.

(110) ù-yô   nì   k-káa   yápù /
       3s-be\RLS with NC8-woman   two

       tò   **ánà** ù-yó-**nò**   nì   k-káa   yápù /
       OK   **how** 3s-be\RLS-**PFV** with NC8-woman   two

       he had two wives / OK it being the case that he had two wives /
       [saat001.004.005]

As suggested by the phrase "tail-head linkage", this construction normally serves to recapitulate information that is already known to the hearer, to serve as a point of departure for the main events to follow. A certain amount of new information may be imparted in the dependent clause, however, as long as it is fairly predictable and can therefore be treated as presupposed by the hearer.

4.4 Nominals

Nouns in most Bantu and many other Benue-Congo languages have two formal properties which set them apart from other word classes. First, nouns trigger obligatory gender agreement on other elements in and outside the noun phrase. Secondly, noun words themselves can be divided into a stem and a noun class affix, usually a prefix.

The first property, obligatory gender agreement, can be regarded as definitional for the identification of Cicipu nouns, since there are no exceptions[23]. As for the noun class prefix, there is a small number of nouns with irregular plurals (§5.2.2, §5.2.7), which means the boundary between prefix and stem is not clear-cut. The extreme example is the noun 'ásù 'place'. Despite belonging to a specific gender according to the agreement that it triggers (single class gender 7, §5.2.8.7), 'ásù has no overt morphological marking to indicate this.

---

23 i.e. there are no exceptions amongst the words which we might want to be covered by the definition because of their morphology or semantics.

The structure of the noun word is simple, consisting of a noun class prefix followed by a stem:

(111)   NC-STEM

The tone pattern on the stem does not change depending on the prefix, and the tone on the prefix largely depends on the root tones (§5.5.3). Class markers in Cicipu are obligatory and always prefixes. Thus Cicipu contrasts with the West Kainji language Pongu where noun markers are optional (MacDonell 2007:48-49), and also the northwest group of West Kainji languages which have both prefixes and suffixes (see Smith 2007 for ut-Ma'in).

This section covers adverbs (§4.4.1), proper nouns (§4.4.2), pronouns (§4.4.3), the count/mass distinction (§4.4.4), NP structure and modifiers (§4.4.5), and then finally suffixal nominalisation (§4.4.6). Discussion of deverbal nominals is deferred to §5.4 because the analysis involves individual noun class prefixes.

4.4.1   Adverbs

It does not seem possible to find purely formal criteria which distinguish adverbs from single class NC8 nouns (i.e. those without a singular/plural distinction). The difficulty is due to two related features of the Cicipu gender system, to be discussed in chapters 5 and 6 respectively. First, NC8 nouns can occur with a null prefix (§5.5.7), making them morphologically indistinguishable from adverbs, and indeed other word classes. Secondly, adverbs can trigger gender agreement in the same way as more prototypical nominals. In chapter 6 we will see that AG8 agreement can be triggered by diverse kinds of atypical controllers ('neutral agreement', §6.4), including adverbs. Furthermore adverbs can even appear with NC8 noun class prefixes (§6.4.2). Because of this noun-like morphosyntactic behaviour I have placed the discussion of adverbs under the general heading of 'nominals'. In this section I will discuss demonstrative adverbs (§4.4.1.1), time adverbs (§4.4.1.2), and ideophones (§4.4.1.3). For adverbial quantifiers see §4.9.

4.4.1.1   Demonstrative adverbs/predicates

Cicipu demonstratives show a five-way contrast in deixis. These distinctions are encoded in adverbs (which also function as predicates, §4.3.3.2) as well as modifiers (§4.4.5.2); it is the adverbs which are of interest here. Each has a H L tone melody

distributed over either one or two syllables, as shown below:

(112) pâa   *here (near speaker)*
       lêe   *there (near hearer)*
       'índè   *yonder (far from both)*
       'û   *very far away* **or** *out of sight*
       ɗôo   *here, our permanent place*

Although Cicipu has five deictic distinctions, there are only three degrees of distance involved – (i) close to some reference point (**pâa**, **lêe**, **ɗôo**), (ii) far away from one or more reference points (**'índè**), and (iii) *very* far away from one or more reference point (**'û**). **Pâa** and **lêe** are not distinguished from each other by any measure of distance, but rather by the speech-act participant functioning as the deictic anchor – the speaker for **pâa**, the hearer for **lêe**. The first three terms in (112) therefore form a "speaker and addressee-anchored" subsystem (Levinson 2004), while the meaning of the first four terms coincides with the Hausa system (Jaggar 2001:323-330, 645-647). Levinson (2004:109) writes that "Systems with more than four terms combine other semantic dimensions, like visibility or vertical distance relative to the speaker, or shape of the referent". This is the case for Cicipu, as the gloss for **'û** above suggests. If the location in question is out of sight and thus cannot be indicated by the speaker, then **'û** is used, regardless of distance.

The meaning of the fifth term **ɗôo** is difficult to pin down, but it involves an interesting combination of SOCIAL and PLACE deixis. Its range of meaning overlaps with that of the near deictic **pâa**, but there are times when only **ɗôo** is appropriate. For example if two Acipu are discussing their village at home they could use either **pâa** or **ɗôo** to refer to it. If they travelled to a nearby Kambari town, then not surprisingly **pâa** 'here' can no longer be used to refer to their home village. However they can continue to use **ɗôo** in this way, regardless of their geographical location. For inhabitants of different villages, the range of locations that can be referred to using this deictic is different, and so the precise meaning of **ɗôo** in any given usage event can only be determined by deictic anchorage. Although it is clearly deictic in meaning, **ɗôo** differs in character from the first four terms listed above, and so it is debatable whether Cicipu should be classified as having a five-term system of spatial deixis.

4.4.1.2 Time adverbs

Time adverbs can occur clause-initially (113) or clause-finally (114), but have not been observed clause-internally.

(113) **lóngò**   à-hwâarà ú-yìndà Ø-kèeké lêe
       recently 3P-start\RLS NC7-see NC8-bicycle there
       recently *they started getting bicycles there*

[sayb001.477]

(114) cì-cípù kàm, tì-yónnò ù-dóowò **ká'à**
       NC6-Cicipu definitely AG6-be_near\RLS NC7-disappear **now**
       *Cicipu, it's nearly disappearing **now***

[Tikula, sagb001.707]

The set of temporal adverbs includes the following members:

(115) ká'à   *now*   [eamy031.011]

       ká'àká'à   *just now/immediately*   [eaim003.1402]

       lóngò   *recently/in the near future*   [eamy031.063]

       lóngòlóngò   *a while ago/ quite far into the future* [eamy031.063]

       dòorí   *formerly*   [tats005.002.230]

       dègèlêe[24]   *then*   [tats001.004.071]

4.4.1.3 Ideophones

Forty-six apparently native ideophones have been identified, but the use of Hausa ideophones is also common, and several of these are significantly different from the source words phonetically, suggesting true borrowing rather than code-switching. They usually express the manner of an event as in (116), but they may also describe a referent (117)[25], or function as intensifiers in stative clauses (118).

(116) vû-uwà   ù-tâa   **kíip!**
       2S-hear\IRR 3S-shoot\RLS **sound_of_penetration**
       you would hear him shoot!

[tats005.002.088]

(117) wǎ-ayà   w-ǐndà   í-rí   pâa **gáaa!**
       3S-come\RLS 3S-see\RLS NC3-thing here **huge**
       then he saw something huge!

[tats005.002.175]

---

24 This word is very common in narrative. Diachronically, at least, it is composed of **dègè** 'from' + **lêe** 'there' (cf. Hausa *daga nan* 'then/from there').

25 Unless this is describing the manner in which something huge is encountered.

(118) yì-'úyò-nò **páss**
AG3-light-? **snow_white**
it's snow white

<div align="right">[cf. Hausa <i>fat,</i> eamy001.010]</div>

Many Cicipu ideophones are phonologically unusual in some way (e.g. syllable structure §3.1.7).

4.4.2 Proper nouns

Proper nouns are assigned gender in different ways according to the animacy of their referents. When humans are referred to by names these nouns always trigger AG8 agreement, which is found with other kinds of atypical agreement controller (§6.4). The names of towns or other locations, on the other hand, are usually assigned gender according to their morphology (or perceived morphology – see §5.6 for 'borrowed' towns).

(119) ǹ-râa   **m-ѐ**
1S-eat\RLS   **AG4-PRO**
I win it [**Màkúukù**, a local town, NC4]

<div align="right">[saat001.002.099]</div>

(120) hárì kò-rísìnóo **kà**-zámùkwà kò-rísìnóo
until NC1-Korisino **AG1**-become\RLS NC1-Korisino
until Korisino became Korisino

<div align="right">[sayb001.106]</div>

The names of people can take part in associative plural constructions (Daniel and Moravcsik 2005), which consist of the high tone morpheme **ǎa** followed by the name. The relation between the named person and their associates is usually one of kinship, but it is also commonly used to refer to a person's agemates or close friends. Occasionally the associative relation is broader:

(121) [Context: discussion of people who would travel long distances to gather *melukeci* to make ropes]
ǎa mé-lúkécí?
PL NC4-plant_k.o.
those of the melukeci plant?

<div align="right">[sayb001.540]</div>

4.4.3 Pronouns

The pronominal paradigms of Cicipu are dealt with in detail in §7.1-7.3. In this section I just make some general remarks and then list the paradigms of free pronouns and

pronominal clitics[26].

There is no distinction between inclusive and exclusive 'we' in Cicipu, unlike the Northwest group of West Kainji (e.g. Smith 2007). There is a distinction in number between singular and plural, but there is no dual.

For third-person pronouns there is of course an alternation between paradigms marked for gender and those that are not. This alternation is the main topic of Part IV.

4.4.3.1   Independent personal pronouns

Table 20: Independent personal pronouns

Person	Singular	Plural
1	ámú	ótú
2	ìvɔ́	ìɗó
3	éví	éré

The 3PS pronoun has an alternate form **ɛ́** – see §8.9.4 for discussion and possible diachronic relationship to **éví**.

4.4.3.2   Independent noun class pronouns

Table 21: Independent noun class pronouns in Cicipu

Class	1	2	3	4	5	6	7	8	9
Pronoun	ke	he	yi	me	mi	ti	wi	vi	kwi

4.4.3.3   Object clitics

These attach to the right-hand side of the verbal word.

Table 22: Object clitics

Person	Singular	Plural
1	mù	tù
2	vù	ɗò
3	vì	rè

4.4.3.4   Demonstrative personal pronouns

The singular demonstrative pronouns are composed of the contracted form **ɛ́** of the independent 3PS pronoun (§4.4.3.1) plus one of the deictic locatives (§4.4.1.1). They should not be confused with the person-marked demonstrative modifiers (**wú-mpà**,

---

26 For the pronominal subject affixes/agreement markers see §7.4-7.5.

wú-llè etc... §4.4.5.2).

Table 23: Demonstrative pronouns

Demonstrative pronoun		Corresponding locative	
é-mpè	*this one*	**pâa**	*near speaker*
é-llè	*that one*	**lêe**	*near hearer*
é-"índè	*'yonder' one*	**'índè**	*far from both*
		'û	*very far from both/invisible*[27]
ó-nɗò	*this one*	**ɗôo**	*here, our permanent place*

There is a puzzle about the direction of vowel harmony in these words, since in **é-mpè** the vowel quality [e] appears to have spread rightward from the **e-** to the following deictic **pâa**, whereas in **ó-nɗò** the situation is reversed: the [o] seems to have spread leftward from the **ɗôo** deictic to the **é-** 3PS pronoun. The prenasalised consonants in **é-mpè** and **ó-nɗò** are part of a more general problem which also affects the demonstrative modifiers (§4.4.5.2).

To express plurality, an alternative set of demonstrative pronouns is used, derived from the word for person/people **z-zá/à-zá**. The empty cells are not attested in the corpus and no attempt has been made yet to elicit these forms.

Table 24: Demonstrative pronouns derived from z-zá 'person'

Singular	Plural	Corresponding locative	
z-zá-mpà	à-zá-mpà	pâa	*near speaker*
z-zá-llè	à-zá-llè	lêe	*near hearer*
z-zá-"índè	àzá-"índè	'índè	*far from both*
z-zá-"û	?	'û	*very far from both/invisible*
?	à-zá-nɗò	ɗôo	*here, our permanent place*

4.4.4   Count and mass nouns

As with other Benue-Congo languages, count and mass nouns can be distinguished according to the noun classes to which they belong. For example the word for 'stone' **kà-táarí** takes the NC1 prefix **kà-** in the singular and the NC2 prefix **à-** in the plural. For other noun roots there is only one choice, and these tend to be either (i) abstract, (ii) nouns that denote some kind of homogeneous mass (either liquid or solid), or (iii) nouns like mosquitos or termites where it is seldom necessary to single out one entity

27 There are no occurrences of the corresponding demonstrative pronoun in the corpus.

(Welmers 1973:160). Examples of these three kinds of nouns in Cicipu are **sùpôo** 'falsehood, NC6', **mò-ní** 'water, NC4'[28], and **tú-ukò** 'cockroach, NC6'.

See §5.3.3 for the derivational relationship that exists between certain count and mass nouns (e.g. *lump of charcoal* vs. *charcoal*).

4.4.5 The noun phrase

The noun phrase in Cicipu is comprised of an obligatory head followed by optional determiners, possessors, or modifiers. The associative construction, demonstratives, and article will be discussed in this subsection. Other modifiers are covered in different parts of the chapter: for relative clauses see §4.3.4, and for adjectives and numerals see §4.7 and §4.8 respectively. Virtually all noun modifiers can also occur as the heads of NPs i.e. pronominally, and this property of Cicipu is discussed in §4.4.5.5. The subsection ends with a discussion of the typologically-unusual ordering of NP constituents (§4.4.5.6).

4.4.5.1 The associative ('genitive/possessive') construction

In many Benue-Congo languages, one particular type of noun modification takes on a heavy functional load: the associative construction (Welmers 1973:275-278), sometimes called the 'genitive' (and less often the 'possessive'). In Cicipu the associative construction is important for the understanding of possession, the attribution of nominal properties, possessive pronouns (§4.4.5.1.1), demonstratives (§4.4.3.4, §4.4.5.2), prepositions (§4.5), ordinal numerals (§4.8.3), and various deverbal nominalisations (§5.4), amongst other areas of the grammar.

The Cicipu associative construction takes the form NP_1 AG_{NP1}=NP_2, where NP_1 is the 'possessed' noun and NP_2 the 'possessor' noun:

(122) [àhúlá   hák꞉à꞉]
       à-húlá   há↓=k-káa
       NC2-name   AG2=NC8-woman
       the name of the woman

[saat002.002.030]

NP_1 occurs in its citation tone pattern, the agreement proclitic AG_{NP1} is high tone, while NP_2 undergoes a complex tonal perturbation (§3.4.7). The most salient feature of this perturbation for the present discussion is that the pitch of the NP_2 vowels is consistently

---

28 A noun with the usual plural prefix corresponding to NC4 can in fact be elicited – **ǹ-ní** 'waters, NC5' – but the same shift of meaning occurs as in English (cf. *I'll have two waters please*).

lower than that of the agreement clitic vowel, so that for monosyllabic roots, say, the tone pattern on the second phonological word surfaces phonetically as [H L].

Unlike most Bantu languages (Welmers 1973:276), there is no segmental 'associative' morpheme to which the agreement morpheme adjoins – instead it attaches straight onto the left edge of the second NP. The downstep which occurs after the associative agreement prefix may be the vestige of a previous segmental form[29].

The agreement morpheme harmonises with the noun root to which it is attached, as shown in (123) below, and this is evidence that the morpheme belongs to the same phonological word as the following noun. Note the harmonisation 'passes through' the noun class prefix **ku-** (recall the discussion on transparent affixes in §3.5.3).

(123) kà-mángá   kó↓=kú-**hóomòwò**
       NC1-rope   AG1=NC9-**tree_k.o.**
       rope of the kuhoomowo tree

                                                    [eaim006.1454A]

I have analysed the associative agreement morpheme as a proclitic rather than a prefix, in part because it does not always attach to lexical noun heads. If there is a pre-head demonstrative or article, then the associative agreement morpheme will bind to it rather than the head noun, as in (124)[30]:

(124) m-óɔ /   mé-evì /   sáa kúmá **mó²↓**=wú-nà   z-zá
       NC4-child AG4-3S.POSS or   more   AG4=3S-ART   NC8-person
       his child, or [that] of another person

                                                    [tats002.004.039]

Furthermore, the agreement morpheme may have semantic scope over an entire *clause* rather than just an NP. In the example below the reason is not the 'reason of the chieftancy', but the 'reason of the chieftancy being there at Kadaada'. The clause is not nominalised, but is a non-verbal locative clause (§4.3.3.2) with **tì-lêe** 'there' as the predicate.

---

29 See Cahill (2000) for languages with purely tonal associative morphemes, especially Nkem (2000:43), where the tonal associative morpheme also results in downstep. See also Watters (1980:111) for the same phenomenon in Ejagam.

30 The associative agreement marker **mɔ** is cut off by a glottal stop here, perhaps due to the grammatical word boundary.

(125) mà-hwáarí **má↓**=tí-wɔ́mɔ́ tì-lêe tì-'ísànùu-nà Kà-dáadà
NC4-start AG4=NC6-chief AG6-there AG6-stand\RLS-PFV? NC1-[name]
the reason why the chieftancy is there at Kadaada lit. 'the reason of [the chieftancy is there it stands at Kadaada]'

[Tikula, sagb001.636]

One final argument in favour of the clitic analysis is that the agreement morpheme attaches only once to complex NPs, unlike say Kikongo (Mel'čuk 1993:339), where multiple association markers occur.

As will be clear already, the relationship between the referents of the two NPs involved in the associative construction is by no means always possession. Welmers (1973:276) and Crozier (1984:99-101) give examples of the broad range of semantic relations covered by the associative, for Kiswahili and Central Kambari respectively. Cicipu is no different in this matter, as the following examples show.

(126) kw-á'à **kú↓**=n-náatà
NC9-house AG9=NC8-small_spider
the house of Small Spider

[POSSESSION, saat002.002.014]

(127) k-káa **ví↓**=n-náatà
NC8-woman AG8=NC8-small_spider
the wife of Small Spider

[KINSHIP, saat002.002.321]

(128) kò-zúvù **ká↓**=kú-cíyè
NC1-digit AG1=NC9-hand
finger [lit. 'digit of hand']

[PARTITIVE, eamy019.001]

(129) kà-ddírí **ká↓**=a-kúngú
NC1-stack AG1=NC2-corn_stalk
stack of cornstalks

[COLLECTIVE, tats002.002.056]

(130) ì-rí **yú↓**=ú-ráa
NC3-thing AG3=NC7-eat
food [lit 'things of eating']

[FUNCTION, tats001.003.026]

(131) ì-rí **yú↓**=ú-ráa **yí↓**=ká-hǐ̀ì
NC3-thing AG3=NC7-eat AG3=NC1-night
evening meal [lit 'things of eating of night']

[TIME OF USE, eamd018.580]

189

(132) ká-ɓàngálí **ká**↓ = ká-záakì
     NC1-slap   **AG1**=NC1-lion

the blow of the lion

[RELATIONSHIP BETWEEN EVENT AND AGENT, saat001.007.088]

(133) ù-dɔ́vɔ́   **wí**↓ = i-dá'ũ̀
     NC7-pound   **AG7**=NC3-guineacorn

the pounding of the corn

[RELATIONSHIP BETWEEN EVENT AND PATIENT, 2008-03-30.001]

There is no distinction in Cicipu between between obligatorily and optionally possessed nouns, nor between alienable and inalienable possession.

As with almost all NP modifiers in Cicipu, the right-hand side of the associative construction can occur without a noun head, as in (134) below. Here **mú**↓=**u-gálù** might be translated 'that of the side'.

(134) [Context: one hunter was hunting along the river bank, doing **ǹ-kácí** (NC5) **mí**↓=**t-túdù** lit. 'hunting of bank']

é-mpè sí-yãã kúmá **mú**↓=**u-gálù** kù-jénè
3s-this HAB-do more **AG5**=NC7-**side** NC9-river

and that one was doing [hunting] *of the side of the river*

[Tidipo, saat002.004.043]

4.4.5.1.1 Associative ('possessive') pronouns

There are six personal associative (or 'possessive') pronouns, which are similar in form to the independent personal pronouns (§4.4.3.1). These pronouns take regular agreement prefixes and occur after the head noun.

(135) (a) **kà-bárá**   **kó-vôo**
       NC1-elder   AG1-1s.POSS

my elder

[eamy005.035]

    (b) **k-káa**   **vá-avù**
       NC8-woman AG8-2s.POSS

your (sg.) wife

[saat002.002.611]

The full paradigm is shown below:

Table 25: Possessive pronoun paradigm with gender agreement prefixes

	1s	2s	3s	1P	2P	3P
1	ko-vôo	ká-avù	ké-evì	kó-ttù	kó-ɗɗò	kí-ivè
2	ho-vôo	há-avù	hé-evì	hó-ttù	hó-ɗɗò	hí-ivè
3	yi-vôo	yá-avù	yé-evì	yí-ttù	yí-ɗɗò	yí-ivè
4	mo-vôo	má-avù	mé-evì	mó-ttù	mó-ɗɗò	mí-ivè
5	mi-vôo or n-vôo	má-avù	mé-evì	mí-ttù	mí-ɗɗò	mí-ivè
6	ti-vôo	tá-avù	té-evì	tí-ttù	tí-ɗɗò	tí-ivè
7	wu-vôo	wá-avù	wé-evì	wú-ttù	wú-ɗɗò	wí-ivè
8	vi-vôo or v-vôo	vá-avù	vé-evì	ví-ttù	ví-ɗɗò	ví-ivè
9	ku-vôo	kwá-avù	kwé-evì	kú-ttù	kú-ɗɗò	kwí-ivè

The 2s, 3s, and 3P forms all have long vowels together with C- prefixes, which suggests the pronouns are vowel-initial. As with the combination of noun prefixes and nouns (§3.1.1), these long vowels are interpreted as the product of coalescence between the prefix vowel and the first root vowel. Again, the root vowel quality is dominant.

The CV- prefixes (i.e. 1s, 1P, and 2P) triggered by class 1, 2, and 4 nouns all have the prefix vowel [o], which is best thought of an underspecified **A** vowel harmonising with the possessive pronoun:

(136) (a) **kà-bárá**   **kó-vôo**
        NC1-elder   AG1-1s.POSS
        my elder

                                                   [eamy005.035]

      (b) **ò-ɗûu**   **hó-ttù**
        NC2-heart   AG2-1P.POSS
        our hearts

                                                   [oamy001.150]

      (c) **mà-díyá**   **mó-ɗɗò**
        NC4-hare   AG4-2P.POSS
        your (pl.) hares

                                                   [eamd032.062]

The other class prefixes retain their usual vowels:

(137) (a) **cì-kóotò**   **tì-vôo**
NC6-drum   AG6-1S.POSS

my drum

(b) **c-có'ò**   **ví-ttù**
NC8-sheep   AG8-1P.POSS

our sheep (sg.)

(c) **kù-lácí**   **kú-ɗɗò**
NC9-girl   AG9-2P.POSS

your (pl.) girl

Finally, the tone on the agreement prefix on the 1PS possessive pronoun is unspecified, assimilating to the tone immediately to the left. So in examples (136a) and (137a) we have high-tone **kó-** after **kà-bárá**, but low-tone **tì-** after **cì-kóotò**.

As a general point, it may be observed that the usual H L tone pattern on the associative pronoun matches that of the lexical associative construction examples given in the previous subsection, once the effect of downstep has been taken into account. This provides us with formal evidence for considering these pronouns to be a special case of the associative, just as Welmers (1973:276) does for Bantu – this argument will be taken up again in the discussion of neutral agreement in §6.4.2.

The discussion above is concerned with the person paradigm. The gender-marked paradigm of pronouns (§4.4.3.2) may also occur in the second part of the associative construction (i.e. the 'possessor' position). These pronouns do not change their form when they occur in this position, and so no additional paradigm has been set up here[31].

4.4.5.2 Demonstrative modifiers

The five deictic adverbs discussed in §4.4.1.1 above also have modifier counterparts encoding the same distinctions. These occur after the noun and take a high-tone CV-agreement prefix. The modifier roots are shown in **bold** below:

---

31 There is no under-differentiation of gender in this environment, and pronouns of all noun classes may be found (cf. Kiswahili where only class 1 and 2 pronouns occur in this position – Welmers 1973:276).

Table 26: Demonstrative modifiers

Demonstrative modifier			Corresponding locative	
ká-ayá	ká-**mpà**	*this hut*	pâa	*near speaker*
ká-ayá	ké-**llè**	*that hut*	lêe	*near hearer*
ká-ayá	ké-**"índè**	*that hut*	'índè	*far from both*
kò-dóndó	kó-**"û**	*that garden*	'û[32]	*very far from both/invisible*
kwá-a'à	kú-**ɗɗô**	*this house*[33]	ɗôo	*here, our permanent place*

There are also person-marked demonstratives **wú-mpà** and **wú-llè**[34] which may be used to modify any noun, regardless of gender.

(138)   **wú-mpà** z-zá   tíyò-nò   **wú-mpà** ì-rí   hànú?!
      3s-this NC8-person get\RLS-VENT **3s-this** NC3-thing where
      this person got this thing where?!

                                             [saat002.002.285]

(139)   **wú-llè** ì-rí   yí↓=ká-ráa
      3s-that NC3-thing AG3=NC1-eat
      that food

                                             [tats007.002.093]

Demonstrative modifiers may occur either before or after the head noun, although the resultant difference in meaning is hard to pin down. In some languages there is a distinction between post-head and pre-head adjectives with respect to the occurrence of agreement (Rijkhoff 2002:21), but this is not the case for Cicipu: agreement must occur, regardless of position. There is however a distinction with respect to agreement *features* – post-head demonstratives must agree in gender, whereas pre-head demonstratives may be inflected either for gender or for person alone – just as we found for the question word -**ènέ** ('which') in §4.3.6.2, and just as we are about to see for the article in the next subsection.

    Lyons (1999:115) notes that there are also pre- and post-nominal demonstratives in Kiswahili (see Welmers 1973:288 for details), with the former construction indicating that the NP referent is topical. We will consider this possibility for Cicipu in §8.9.2 – in anticipation of that discussion, note that person agreement in Cicipu is only possible on

---

32 There are no occurrences of demonstrative pronouns encoding the 'invisible' **'û** deictic in the corpus.

33 cf. **kwá-a'à kú-ɗɗô** 'your (pl.) house' from the previous section. Note the tonal difference on the modifier, cf. Anderson (1980b:43) for a similar correspondence between deictic and 2PP possessive pronoun in Ngyembɔɔn-Bamileke (Grassfields Bantu).

34 The other three deictic possibilities are not attested in the corpus, but they have not been ruled out by elicitation either.

the demonstrative if it occurs in the position corresponding to topicality in Kiswahili.

The fact that long consonants occur root-initially in four of the five modifiers is striking, as is the high-tone on the agreement prefix. In §6.4.2 on 'neutral' agreement it will be suggested that the demonstrative modifiers in Cicipu are derived diachronically from instances of the associative construction (§4.4.5.1). This sets Cicipu apart from the Bantu languages, for which Welmers states "The associative...is not used in constructions of noun plus modifying numeral or noun plus demonstrative" (1973:277).

As with almost all NP modifiers in Cicipu, the gender-marked demonstratives can occur as noun heads themselves, as in the example of cataphoric reference below. The AG1 agreement morphology seen below is a form of 'antecedentless' agreement (§6.3) which occurs when the agreement controller is a proposition.

(140) ké-llè   mà-hũu m-è:-
     AG1-that   NC4-truth AG4-COP
     this is the truth:-

<div align="right">[ovkz002.020.001]</div>

The person-marked demonstratives, however, do not occur as noun heads, and constructed examples such as (141) are rejected. Instead the demonstrative pronoun é-mpè (§4.4.3.4) would be required in this example[35].

(141) *wú-mpà, kà-bárá   kó-vôo   k-é
     3S-this   NC1-old_man AG1-1S.POSS AG1-COP
     [Intended meaning: *this, it's my old man*]

<div align="right">[eamy005.032]</div>

4.4.5.3 Article

Just like the demonstratives, the article -nà can occur either before or after the noun head. This time, however, the difference in meaning is clear. Post-head articles indicate anaphoric reference to an NP, while pre-head articles mark an indefinite but specific NP. In both cases the article takes a high-tone agreement prefix. As we have already seen for a number of other agreement targets, the prefix on the article may agree in either person/ number, or gender – see §8.9.1 for further details.

The difference in meaning is particularly clear in the conventionalised expressions kú-nà kw-ã́a'ã́ 'one day' and kw-ã́a'ã́ kú-nà 'the next day'. The more complex examples below illustrate that the pre-head article cannot occur in cases of non-specific reference; that is, when the referent is not identifiable to either the speaker or the hearer. Example

---

35 Note that in Cicipu, as in Hausa, it is not impolite to refer to humans simply as 'this' or 'that'.

(142) shows this with person agreement, (143) with gender agreement.

(142)   dòorí   ǹ   (*wú-nà)   z-zá   yôo   á↓=kà-káasùwà,
       formerly   when   3s-ART   nc8-person   go\RLS   LOC=nc1-market

       ù-sì-cíyó   cé   Ø-kèeké
       3s-HAB-get   NEG   nc8-bicycle

       in the past if someone went to the market, he wouldn't get a bicycle
                                                         [2008-02-01.001]

(143)   dòorí   ǹ   (*má-nà)   m-ɔ́ɔ   mò-yôo   á↓=kà-káasùwà,
       formerly   when   AG4-ART   nc4-child   AG4-go\RLS   LOC=nc1-market

       ù-sì-cíyó   cé   Ø-kèeké
       3s-HAB-get   NEG   nc8-bicycle

       in the past, if a child went to the market, he wouldn't get a bicycle
                                                         [2008-02-01.001]

The following constructed dialogues further illustrate the difference. In (144) speaker B is looking for a particular stone, perhaps one that he lost earlier. The referent of **ká-nà ká-táarí** 'a certain stone' is identifiable to the speaker, but not to the hearer. When introducing such referents to a discourse, the pre-head article is very common[36].

(144)   A: Ø-yúu   yḕa   yìní?
            2s-CONT   do   what

            what are you doing?

       B: ǹ-yúu   bɔ̀lɔ̀   **ká-nà**   ká-táarí
            1s-CONT   look_for   AG1-ART   nc1-stone

            I'm looking for a [particular] *stone*

            [Some time later...]
       B: tò, m-índà   ká-táarí   *(ká-nà)
            OK   1s-see\RLS   nc1-stone   AG1-ART

            OK, I've found the stone

                                                        [2008-02-01.001]

In (145) on the other hand, speaker B does not care what sort of stone he finds. The referent of **ká-táarí** 'a stone' is identifiable neither to the speaker nor to the hearer. In such cases the pre-head article cannot occur.

---

36 It is not however obligatory. If the general class of referents is already a topic of discussion, then the introduction of an individual referent of that class will usually be accompanied by the pre-head article. Conversely, if the referent has little intrinsic importance (despite being specific) then the pre-head article will usually not occur. This is similar to Jaggar's (1988) finding that the Hausa specific indefinite determiner *wani* is more likely to mark the introduction of referents which persist for longer in the discourse (see §8.2).

(145) A: Ø-yúu yằa yìní?
    2s-CONT do what

    what are you doing?

    B: ǹ-yúu bɔ̀lɔ̀ (**ká-nà**) ká-táarí
    1s-CONT look_for AG1-ART NC1-stone

    I'm looking for a stone [any old one]

    [Some time later...]
    B: tò, m-índà ká-táarí (**ká-nà**)
    OK 1s-see\RLS NC1-stone AG1-ART

    OK, I've found the stone

[2008-02-01.001]

Turning to speaker B's second statement in these dialogues, I was told that the post-head article was obligatory in (144) but optional in (145). In fact it is hard to square this with the evidence from the corpus, where NPs are frequently used to refer anaphorically without the post-head article, although the delay while searching for the stone may be a factor; once the activation status (§2.3.2.4) of the stone has decayed enough the article may be required as a 'reminder' to the hearer not to introduce a new referent. Perhaps the most we can say for sure is that NP-anaphors are less likely to be marked with the post-head article if the original reference was non-specific[37].

Crozier (1984:80-85) argues that the Central Kambari enclitic **'V** is not simply a definite article but an "anaphoric" article (Lyons 1999:53), a more restrictive category which is used when referents have previously been mentioned in the text. He contrasts its overwhelming anaphoric use in Central Kambari with Halliday and Hasan's (1976:73) contention that the English definite article is primarily exophoric or cataphoric. Lyons (1999:52-53) makes a similar point regarding the Hausa definite article (see also Newman 2000:143, and the discussion in Jaggar (1985:149-153)). Although the Cicipu article is not diachronically related to either of these, it seems to be equivalent in function when it occurs after the noun.

Entities which are by their nature identifiable/unique such as the moon or sun never occur with the article.

Lyons (1999:26-30), following Postal (1970), argues that personal pronouns are the pronominal counterpart of definite articles. This analysis is not appropriate for Cicipu, since the article may itself function pronominally (see 148-149 below).

---

37 In this respect the Cicipu 'definite' (i.e. post-head) article is different to English *the*, since *I found the stone* could not, I think, be used in the setting given for (145).

Moreover, personal pronouns (at least in the third person) may themselves be followed by the article[38]:

(146) [Context: someone walks into a compound and asks 'Which ones are the slaves?']

à-zá-mpà,   **érè**   **n-nà**   ò-módó   há-nà
NC2-person-this **3P.COP**   **AG8-ART**   NC2-slave   AG2-ART
these ones, they are the slaves

[2008-03-07.003]

Unsurpisingly given their contrasting pragmatic functions, the pre-head and post-head forms of the article have not been observed modifying a single noun. However the relativiser (which is formally identical to the article) readily co-occurs with the pre-head article:

(147) sáa Ø-ɓàa **ká-nà**   kà-nàbáyì   **ká-nà**   vú-u-gùyà
or   2s-have **AG1-ART**   NC1-news   **AG1-REL**   2s-FUT-can\IRR
ù-dán-dámá-wà＝tù?
NC7-little-speak-APPL＝1P.PRO
do you have any particular story you can tell us?

[Tikula, sagb001.089]

This distributional fact supports the analytic distinction I have made between article and relativiser, as does their different internal positions within the NP (even when both are post-head – see §4.4.5.6).

Both pre-head and post-head articles can occur pronominally i.e. without a lexical noun head. It may seem strange to retain the distinction between "pre-head" and "post-head" when the article itself *is* the head, but there is a clear difference in meaning between examples such as (148), where a new (specific) referent is being introduced into the discourse, and those such as (149) which involve anaphoric reference to previously-introduced referents.

(148) [Context: speaker was talking about a specific camel (his own) which doesn't bite]

àmáa **ká-nà**   kà-sì↑-nùmà-númà
but   **AG1-ART**   AG1-HAB-bite-REDUP
but others [**kà-ràkúmí** 'camel, NC1'] *bite*

[PRE-HEAD ARTICLE, tats004.001.018]

---

38 This is also true for Hausa (Newman 2000:145).

(149) **má-nà** mà-'wâa m-áyà mò-yôo mò-yúwò
AG4-ART AG4-pass\RLS AG4-come\RLS AG4-go\RLS AG4-fall\RLS

the [**m-ɔ́ɔ** 'child', NC4] *passed and then went and fell*

[POST-HEAD ARTICLE, tapf001.004.008]

Just like the person-marked forms of the demonstratives, **wú-nà** does not occur pronominally – at least, of the 113 instances in the corpus, all of them occurred with a head noun immediately afterwards.

Finally, example (148) above illustrated one further property of the pre-head article: although it occurs when introducing a new referent, the context often involves contrast with a previous referent. Another example is given below:

(150) [Context: speaker was talking how dogs recognise members of the household.]

àmáa ìdón **wú-nà** z-zá v-ì
but if **3s-ART** NC8-person AG8-COP

*but if it's **another** person*

[tats001.001.085]

In this respect the article is similar to the modifier **-mbɔ̃** 'another':

(151) z-zá n-nà dɔ́nɔ̀-nɔ̀ / Ø-síyà **ví-mbɔ̃,** b-bɔ́wɔ́ v-ì
NC8-person AG8-REL follow\RLS-PFV NC8-side **AG8-another** NC8-thief AG8-COP

*the one who follows the **other** side / is a thief*

[tats001.004.104]

However **-mbɔ̃** also has an 'incremental' use as in (152)[39], a shade of meaning which seems to be absent from the article.

(152) [Context: after recovering from a hangover]
wú-u-yòo ú-bɔ̀lɔ̃ **mɔ́-mbɔ̃**
3s-FUT-go\IRR NC7-look_for **AG4-another**

*he'll go looking for **more** [**mò-yɔ́o** 'beer, NC4']*

[tats002.003.052]

As illustrated in (152), **-mbɔ̃** takes a high-tone gender agreement prefix, occurs as the head of an NP as well as a modifier, and can modify mass nouns as well as count nouns.

4.4.5.4 Nominal coordination

Nominals are conjoined using a single proclitic meaning 'and'. This takes the form **ǹ** before short consonants and **nì** before long consonants or a consonant cluster. If the stem is vowel-initial, then that vowel lengthens so that **ǹ=à-zá** 'and people' becomes

---

39 cf. Quirk et al. (1985:389) on *another.*

[nàːzá]. Example (153) shows **nì** before first a consonant cluster and then a vowel-initial stem, and then **ǹ** before a short consonant. Example (154) shows **nì** before several words beginning with a long consonant.

(153) à-ràkúmí **nì**=n-jàkíi,   **nì**=i-náa   **ǹ**=kà-kíngí kí-ivè
NC2-camel **and**=NC5-donkey **and**=NC3-cow **and**=NC1-rest AG1-3P.POSS

camels, and donkeys, and cows, and the rest of them

[samoh001.194]

(154) **ǹ**=kà-ràháazì **nì**=k-kɔ́gɔ́   **nì**=v-vádí **nì**=h-hwí'í
and=NC1-Fulani **and**=NC8-Hausa **and**=NC8-Vadi **and**=NC8-Lelna

the Fulani and the Hausa and the Vadi and the Lelna

[saim001.029]

Singular personal pronouns may not occur as the first member of a conjoined NP; instead the corresponding plural pronoun is used – forming what is sometimes called an 'inclusory pronominal construction' (Lichtenberk 2000)[40]. So to express the meaning "he and his wife" in (155) the phrase **éré nì=kkáa véevì** 'them and his wife' (perhaps better 'them *including* his wife') is used. Similarly to express "me and my wife" one would say **ótú nì=kkáa vvôo** 'us and my wife', likewise for second person. Using a singular personal pronoun as the first member of such a construction is considered incorrect, and will be corrected by native speakers[41].

(155) cù-kúndú **éré**   **nì**=k-káa   **vé-evì**   á↓=ú-dúkwá
NC6-hyena **3P.PRO** **and**=NC8-woman **AG8-3S.POSS** LOC=NC7-go

*hyena **him and his wife** were going*

[saat001.008.110]

There is no such restriction on the second member of a conjoined NP, which suggests that the two NPs are in a comitative relation rather than a co-ordinate one.

While on the subject of 'anticipatory plurals' it may be noted that, as in Hausa (Newman 2000:136-137, Jaggar 2001:393), verbs may take a plural agreement prefix if the referents encoded in the clause are performing a reciprocal action. The question in the following example was addressed to a single person, yet plural agreement occurs on the verb.

---

40 Newman (2000:136) and Jaggar (2001:393) use the term 'asymmetric coordination' for the same phenomenon in Hausa, while Watters (2003:245) uses the term 'incorporative bond' for Grassfields Bantu.

41 All the examples in the corpus are of personal pronouns. It would be interesting to investigate if a similar restriction applied to the independent *gender* marked pronouns.

(156) ì-síriyà   ǹ   kw-í
  2P-prepare**RLS** with **AG9**-**PRO**

have you (sg.) settled with her?

As well as conjoining two NPs, **ǹ** can also be used to express the **INSTRUMENT** semantic role, as in (157). In the case of subject NPs a prepositional phrase expressing **ACCOMPANIMENT** with the subject may occur after the verb complex as in (158).

(157) wú-u-gàvà kúmá, ǹ kù-náa kú-nɗò
  3S-**FUT**-kick**IRR** more with **NC9**-leg **AG9**-this

he would kick again, with this leg

(158) sée ú-'wàa ǹ kà-ká'ìlâa ké-evì
  then **3S**-pass**IRR** with **NC1**-chameleon **AG1**-**3S**.**POSS**

then he went with his chameleon

This conjunction **ǹ** cannot be used for conjoining clauses or verb phrases. However contrary to Welmers' observation (1973:305) that nominal conjunctions rarely have any other use in African languages, **ǹ** also introduces 'when' dependent clauses (§4.6.3.4) and thus covers the same range of meanings as Hausa *da* 'and/with/when'.

Disjunction is expressed using the disjunction **sáa** 'or'. Unlike **ǹ**, **sáa** may conjoin NPs (159), PPs, or clauses (see §4.3.7.1).

(159) ì-ɓócì **sáa** yí↓=kú-náa, **sáa** yí↓=kó-túmó, **sáa** yí↓=kú-cíyè,
  NC3-illness **or** **AG3**=**NC9**-leg **or** **AG3**=**NC1**-belly **or** **AG3**=**NC9**-hand

illness either of the leg, or of the stomach, or of the hand,

Finally, both **ǹ** and **sáa** may optionally occur at the beginning of a series of conjoined NPs, just like Hausa *da* and *ko* (Newman 2000:132, Jaggar 2001:391, 394). This was illustrated above in (154) for **ǹ**, and in (159) for **sáa**.

4.4.5.5 Pronominal function of modifiers

Almost any noun modifier in Cicipu can also function as the head of an NP. This has been observed for relative clauses, the 'back end' of the associative construction, demonstratives, and the article, and will also turn out to be the case for adjectives (§4.7) and numerals (§4.8). This ability to refer independently of a head noun may of course be linked to the obligatory agreement prefix on these modifiers, which increases their

potential to refer unambiguously.

Given this is the case, it is tempting to try to preserve the left-headedness of the Cicipu NP by analysing the various pre-head modifiers we have encountered as NPs in their own right, so that a more literal translation of, say, **ká-mpà kà-táarí** 'this stone' might be *this one, a stone*. Unfortunately this does not seem to be possible. Recall that each of the pre-head modifiers (**-èné** 'which' (§4.3.6.2), the demonstratives (§4.4.5.2), and the article (§4.4.5.3)) can be marked for person instead of gender. In none of these cases can the person-marked forms function as independently-referring constituents; rather they can only occur immediately before a noun. If we have to admit that **wú-mpà kà-táarí** with person agreement on the demonstrative has the Modifier-Head order, then there is no reason not to assume the same for the gender-agreeing **ká-mpà kà-táarí**.

4.4.5.6 NP-internal syntax

Rijkhoff (2002:24) made the cross-linguistic observation that, outside of elicitation sessions, it is rare to find NPs with more than two modifying constituents. This is true for Cicipu, although NPs with three modifying constituents do occur in the corpus (e.g. ex. 168 below). When multiple modifying constituents do occur, they often seem to involve phrasal modifiers (see Rijkhoff 2002:24 for discussion), and nested or multiple associative constructions are not all that rare.

When multiple constituents do occur in an NP, they follow a regular (although typologically-unusual) order:

(160) (ART/DEM) **NOUN HEAD** (ART) (DEM) (POSS) (NUM/ADJ) (RELATIVE CLAUSE)

Exceptions to this pattern can be analysed as compounds, where the modifying associative construction/adjective is an intrinsic part of the referring expression, rather than merely adding further detail. So in (161), for example **z-zá ɗ-ɗénêu** does not mean 'small person', but 'younger sibling'. Similarly there is a contrast between the straightforward modification 'children of the future' in (162), and the idiomatic compound 'water babies' (a malicious kind of spirit) in (163).

(161) z-zá　　　ɗ-ɗénêu　　vé-evì
　　　　NC8-person　AG8-small　AG8-3s.POSS
　　　　his younger sibling

[Unexpected ADJ before POSS, sayb001.669]

(162) m-úu   mí-mpà mí↓=záamàní
      NC5-child AG5-this AG5=future(NC8)
      these children of the future

<div align="right">[<small>DEM</small> before <small>POSS</small>, svtmg001.290]</div>

(163) m-úu   m=mó-ní   mí-llè
      NC5-child AG5=NC4-water AG5-that
      these water babies [lit. 'children of water']

<div align="right">[Unexpected <small>POSS</small> before <small>DEM</small>, tats005.002.144]</div>

At the moment I am not aware of any *morphological* evidence that the expressions in (161) and (163) are compounds, however cross-linguistically compounding often has this kind of effect on ordering rules, and the analysis suggested here is supported by the non-compositional meanings. If such compounds are assumed to fit into the noun head slot above, then almost all the exceptions disappear[42]. It should be observed that nominal combinations with non-compositional meanings do not *necessitate* this kind of treatment when they occur with modifiers – they can also be consistent with the normal ordering within the NP. The idiomatic NP **ìhìtílà yɔ́ɔkɔ̃́ɔ** 'kerosene lanterns' in (164) below is interrupted by the demonstrative **yí-mpà**, in violation of certain theories about idioms (see Croft and Cruse 2004:225ff) but in accordance with the order set out in (160) above.

(164) ì-hìtílà   yí-mpà   yɔ́↓=ɔ-kɔ̃́ɔ
      NC3-lamp   AG3-this   AG3=NC2-egg
      these kerosene lanterns [lit. 'these lamps of eggs']

<div align="right">[<small>DEM</small> before <small>POSS</small>, sayb001.315]</div>

The position of articles and demonstratives in-between the noun head and the numeral/adjective/relative clauses is contrary to what Rijkhoff (2002) predicts for configurational languages, and is said to be found in 'few/very few languages' by Cinque (2005:319-320)[43]. Interestingly, it also differs significantly from the typologically more frequent order given by Crozier (1984:74) for Central Kambari:

(165) **NOUN HEAD** (NUM) (ADJ) (RELATIVE CLAUSE) (POSS) (DEM) (ART)

The ordering given above for Cicipu is partly based on elicitation, and partly based on inspection of the hundreds of relevant examples in the corpus. More detailed elicitation

---

42 There is one exception in the corpus which does not seem to involve a compound: in **ù-hwáarà [wú↓-tú-wɔ́mɔ́] [wú-nà]** 'the start of chieftaincy' the order is [POSS] [ART]. It may be possible to explain this as a case of "trigger-happy" agreement (§6.5) on the article i.e. 'start of [the chieftaincy]'.

43 Cinque assumes, following Kayne (1994), that the order derives by movement of the noun from the underlying order Dem Num Adj N.

is needed to determine the relative ordering of numerals and adjectives, and to test the extent to which deviations from (160) are judged acceptable. Here I will just list sufficient examples to permit the deduction of the post-head ordering given in (160), repeated below for convenience:

(166) NOUN HEAD (ART) (DEM) (POSS) (NUM/ADJ) (RELATIVE CLAUSE)

(167) ká-ayá   [ká-nà]   [ké-"índè]
NC1-hut   AG1-ART   AG1-that
that [aforementioned] *hut*

[ART before DEM, saat002.002.450]

(168) mà-k-kù-náa   [mé-llè]   [má↓=k-káa   vé-evì]   [má-nà
NC4-?-NC9-leg   AG4-that   AG4=NC8-woman   AG8-3S.POSS   AG4-REL

ù-câa-nà   Ø-Lɔ́wɔ́lì]
3S-give\RLS-PFV   NC8-big_spider
that little leg of his wife that he gave Big Spider

[DEM before POSS, POSS before REL, saat002.002.596]

(169) kà-bárá   [ká-mpà]   [kà-yápù]
NC1-old_man   AG1-this   AG1-two
these two old men

[DEM before NUM, eamy005.007]

(170) m-ɔ́ɔ   [mé-evì]   [mò-tò]
NC4-child   AG4-3S.POSS   AG4-one
his one child

[POSS before NUM, tats001.004.079]

(171) à-vá'ã̀   [hí-ivè]   [hè-pénèpénêu]
NC2-bundle   AG2-3P.POSS   AG2-big:REDUP
their large bundles

[POSS before ADJ, saat002.002.237]

(172) ù-nú   [wù-tò] [wú-nà   wù-kúsù-nò]
NC7-work   AG7-one   AG7-REL   AG7-remain\RLS-PFV
the one work which remains

[NUM before REL, saat002.002.186]

(173) cì-kóotò [tì-dénêu] [tí-nà   à-hyã̂a-nà=vì   ú-vàsà]
NC6-drum   AG6-small   AG6-REL   3P-say\RLS-PFV=3S.PRO   3S-hit\IRR
the small drum which they had told him to beat

[ADJ before REL, saat001.008.071]

4.4.6   Nominalising suffixes

Most deverbal nominalisations in Cicipu are derived by adding one of a number of

different noun class prefixes to a verb stem. Accordingly these are deferred to §5.4, once the noun class system has been properly introduced. There are however two nominalising *suffixes* which are independent of the noun class system, and these are mentioned below.

4.4.6.1 Stative -ni

Suffixation of **-ni** to a verb stem in conjunction with the addition of the NC7 noun prefix **u-** (§5.3.4) gives rise to a nominal, usually denoting the state arising from the activity encoded by the verb as in **ù-dóonúu-ní** 'seated' from **doonu** 'sit' and **ù-mánnánúu-ní** 'stuck' from **mannanu** 'stick'. The final vowel of the verb is lengthened (if short) before adding the suffix.

The resultant words have certain nominal properties. For example they trigger gender agreement on modifiers (174), and can also function as subjects.

(174) ùkóoní   wúYēesù   nùu'úngóoní   wéevì
      ù-kóo-ní   wú↓=Yéesù   n=ù-úngó-ní   wé-evì
      NC7-die-NMLZ AG7=[name]   and=NC7-rise-NMLZ AG7-3S.POSS
      the death of Jesus and his resurrection

[tats001.004.108]

However they can also function as 'adjectival' predicates in a way that straightforward nouns cannot. Consider the following two examples. The predications are similar semantically, but the constructions differ in that the deverbal nominal **ùgónjɔɔní** 'curved' can occur as a direct complement of the verb **yo** 'be' in (175), whereas the noun **ù-'wîi** 'distance, depth' in (176) must be preceded by the preposition **ǹ**.

(175) ùyô   ùgónjɔɔní
      ù-yô   ù-gónjɔ́-ní
      3S-be\RLS   NC7-be_curved-NMLZ
      it is curved [**gɔnjɔ** = 'be curved']

[eati001.1533]

(176) mò-ní   má-mpà   mò-yô   (n)*=ù-'wîi
      NC4-water AG4-this AG4-be\RLS **with=NC7-distance**
      this water is deep [lit 'is with depth']

[eati001.1524]

Finally, **-ni** can also occur after all five of the demonstrative adverbs (§4.4.1.1), in which case they have H tone rather than HL, and the suffix has a L tone rather than its normal H e.g. **pâa~páanì** 'here', **lêe~léenì** 'there'. Whether the bare adverbs and their

'nominalised' counterparts differ in their syntax or semantics is not yet known.

4.4.6.2 Locative -tu

Deverbal nominals formed by adding the suffix **-tu** denote the place where the action encoded by the verb characteristically takes place, as in **kà-pítáatú** 'sole, footprint, track' from **pita** 'step', i.e. 'place involved in stepping', and **kà-ɓángáláatú** 'upper part of the cheek' from **ɓangala** 'slap', both basic-level terms for body parts.

Verbal affixes can be included in the nominalisation, as shown by (177) for the anticausative (§4.6.4.3) and (178) for the pluractional (§4.6.5.4).

(177) mɔ̀'ɔ́púwɔ́ɔtú
      mɔ̀-'ɔ́pɔ́-wá-**tú**
      NC4-hold-ANTIC-LCVZR
      handle [lit 'place which gets held']

      [eamd018.631]

(178) màzáaní   màtísà   kò'úmbólóotú
      mà-záaní   mà-tísà   kò-'úmb<il>ó-**tú**
      NC4-pen   AG4-join\RLS   NC1-close<PLAC>-LCVZR
      the pen joins the lid [lit 'place which repeatedly closes']

      [eaim007.1463]

The tone on the resulting nominal is all H except a L on the prefix, and as with **-ni** the final vowel of the verb is lengthened before adding the suffix[44]. The noun is most often assigned to the 1/2 (KA/A) gender (§5.4.5), but 4/5 and 8/3 have also been observed.

4.5 Prepositions

Cicipu has prepositions but no postpositions, in line with the observations made by Greenberg (1963) for VO languages. This is a small category in Cicipu since the general locative clitic **Á** covers a wide range of semantic relations.

The most common 'preposition' is actually a proclitic: the locative **Á**. This has been analysed as a clitic rather than a free form because it harmonises with the following word[45], and is therefore within the domain of vowel harmony. The tone is always high, and results in downstep[46] for the rest of the intonation unit (§3.4.7). The following examples demonstrate the different phonetic forms of **Á**, as well as a sample

---

44 **kò-dóntú** 'stool, chair' from **doonu** 'sit' is an irregular exception.

45 Recall that **A** stands for a vowel which harmonises with the word to which it is attached.

46 The Bantu language Akɔ́ɔ̄sē has a very similar locative prefix **á↓-**, also causing downstep (Hedinger 1980:10).

of the wide variety of locative meanings it can express.

(179)   **á↓**=cí-kû̃u   tú↓=u-ɗángà
     LOC=NC6-waist   AG6=NC7-tree
     ***at** the base of the tree*

<div align="right">[LOCATIVE, tats005.002.168]</div>

(180)   **á↓**=á-zá   Kwángwà
     LOC=NC2-person   [town]
     *[living] **amongst** the people of Kwangwa*

<div align="right">[LOCATIVE, tats005.002.168]</div>

(181)   **é↓**=k-kèeké
     LOC=NC8-bicycle
     ***on** a bicycle*

<div align="right">[ON SURFACE, tapf001.005.011]</div>

(182)   **ó↓**=ko-ócì
     LOC=NC9-hole
     ***into** a hole*

<div align="right">[INSIDE CONTAINER, svtmg001.177]</div>

(183)   ù-yôo   **á↓**=ká-dábá
     3S-go\RLS   LOC=NC1-countryside
     *he went **to** the bush*

<div align="right">[TOWARDS GOAL, saat001.002.007]</div>

(184)   sée   ǹ=ɔ̀-kɔ́dɔ̀-nù   ká-tî̃i   ké-llè   **ó**=rú-'ú   ya-ávù
     unless   and=3P-cut\RLS-RES   NC1-head   AG1-that   LOC=NC3-body   AG3-2S.POSS
     *you have to cut off that head **from** your body*

<div align="right">[AWAY FROM SOURCE, tats004.002.045]</div>

(185)   àtî̃ivì /   ɛ́ríhyɑ́'ɑ̀
     à-tɑ̃̂a=vì /   **é↓**=rí-hyɑ́'ɑ̀
     3P-shoot\RLS=3S.PRO   LOC=NC3-arrow
     *they shot him / **with** arrows /*

<div align="right">[INSTRUMENTAL, svmy001.022]</div>

(186)   **á↓**=á-náa
     LOC=NC2-leg
     ***by** foot [i.e. means of transport]*

<div align="right">[INSTRUMENTAL?, sayb001.374]</div>

While (180) has a noun denoting human referents as the complement of the preposition, this only seems possible with plural nouns. To express a locational relation involving a singular human such as e.g. *coming towards him*, the noun **'ásù** 'place' is used instead,

with the human as possessor e.g. úuwàyà 'ásù wéevì (lit. 'in coming place of him').

Given that Á seems to be bound to the following word, the question arises as to whether it should be analysed as a case marker rather than a preposition. There are two reasons why it has been called a preposition here, despite its morphological status. First, there are no unambiguous case markers in Cicipu such as an accusative. Secondly, the preposition is not 'governed' by the verb – it seems that no verb requires one of its arguments to be marked with Á[47].

The other common preposition is ǹ 'and/with' (§4.4.5.4).

Tí 'like/similar to' and cúnṹ 'without' are only attested a few times in the corpus, but they are also candidates for prepositions. Alternatively, tí may just be the AG6 associative agreement marker in a headless associative construction (§4.4.5.1, see §6.3 for the connection between gender 6 and 'manner'). In that case structurally-faithful translations of (187) and (188) would be something like *they will bring it out [in the manner] of a mirror*, and *back then it wasn't like [the manner] of this thing*.

(187) h-ú-utò-wò   tí   Ø-mádúubí
       3P-FUT-go_out-APPL   like NC8-mirror
       they will bring [it] *out **like** a mirror*

                                                                [sayb001.408]

(188) Ø-lóokàcí ví-llè   ánà   tí↓=i-rí-mpàa-nì
       NC8-time   AG8-that how   like=NC3-thing-this-NMLZ
       back then it wasn't like this

                                                                [sayb001.430]

(189) ǹ=à-kábà   ù-dá̃ngà   à-sì-yẫa   cúnṹ   Ø-bíndígà
       and=3P-take\RLS NC7-tree   3P-HAB-do   without   NC8-gun
       *when they took a piece of wood they would do **without** guns*

                                                                [samoh001.067]

Other prepositions include dègè 'from', sée 'until', hárì 'until', borrowed from Hausa *daga*, *sai* and *har* respectively. The first two of these are illustrated in the following idiomatic expression (which is structurally equivalent in Hausa).

(190) dègè àmú   sée   kà-mẫyã̀   kò-vôo
       from 1S.PRO   until NC1-elder_sibling   AG1-1S.POSS
       just me and my elder brother

                                                                [sayb001.003]

---

47 **Duwa** 'put' does occur with a locative PP after the object but this is not obligatory as in the case of English *put*. The PP, when it does occur, could be analysed as as an adjunct rather than a complement.

4.6 Verbs and the verb phrase

This section describes the Cicipu verb phrase. After a preliminary section on the structure of the verbal word (§4.6.1), the core of the section is organised according to the major grammatical categories distinguished in the verb – mood (§4.6.2), aspect and tense (§4.6.3), and valence (§4.6.4). There follows a discussion of several other miscellaneous verbal suffixes (§4.6.5). The next sections discuss borrowed verbs (§4.6.6) and semi-grammaticalised auxiliaries (§4.6.7). The final two sections discuss the relative ordering (§4.6.8) and co-occurrence possibilities (§4.6.9) of the various affixes.

The overall organisation of this section is functional rather than formal. So, for example, the subsection on aspect and tense covers both morphological and syntactic expressions of aspect. Within each subsection, however, morphological exponents are treated before phrasal ones.

4.6.1 Verbal word template

Although Cicipu is spoken in the north-west of Nigeria, it is typologically very similar to the Bantu languages of southern and eastern Africa. This similarity manifests itself in two very obvious ways. First, in the robustness and regularity of its noun class system, as set out in Part III, and secondly in the structure of the verbal word. Cicipu is highly agglutinative; not only is there a large number of verbal affixes, many of them can occur simultaneously, resulting in verbal words consisting of up to ten concurrent morphemes (including the verb root and the object enclitic):

(191) zzá   nnà   ùtóbìlìsìsùwòwònòmù   sháyì
      z-zá   n-nà   ù-tób\<ìl>\<ìs>\<ìs>o-wò-wò-nò=mù   Ø-sháyì

NC8-person AG8-REL 3S-cool\RLS\<PLAC>\<CAUS>\<CAUS>-ANTIC-APPL-PFV=1S.PRO NC8-tea

the person who has caused tea to become cooled down in a forceful and iterative fashion for me

[2008-02-11.002]

While such monsters are vanishingly rare in everyday speech, it is common to find three or more segmental affixes on a verb, in addition to the ubiquitous tone pattern which expresses the grammatical mood.

The structure of the verbal word in Cicipu is similar to, but not identical to, the pattern found in Bantu. The Cicipu verbal word follows the following template:

(192)   AGR-   (HAB/FUT-)   **root**   (<extension(s)>)*   (-suffix)*   (=OBJ.PRO)
        (C)V-   si-/u-   **CVCV**   <VC>   -CV   =CV

AGR is a single slot for a subject agreement prefix, which may come from either the gender-marked paradigm or the person-marked paradigm. These prefixes are discussed at length in chapter 7 and so no more will be said about them now. Other than these agreement prefixes, only two affixes can come before the root, the future **u-** (§4.6.3.2) and the habitual **si-** (§4.6.3.3). The extension slot can be filled by either the pluractional infix <**il**> (§4.6.5.4), the causative infix <**is**> (§4.6.4.1), or both. The 'suffix' slot can be filled by various diverse verbal affixes, including aspectual and valence-changing devices. In addition to the segmental affixes, the tone pattern on the verbal word expresses the mood of the clause. More is said about ordering and the possible combinations of affixes in §4.6.8-4.6.9. Following any suffixes, a pronominal object clitic (§7.3) may attach to the end of the verbal word.

One important difference between the Cicipu and Bantu verbal word derives from the structure of verb roots. Bantu verb roots have a CVC shape, with a dummy 'final vowel' being supplied according to rule. Verbal 'extensions' are VC-shaped morphemes which come between the root and the final vowel. Although such extensions are word-internal, given that the final vowel is predictable both it and the extension can be analysed as suffixes rather than infixes (e.g. Schadeberg 2003).

For Cicipu, however, extensions seem to be true infixes[48]. In approximately half (128 out of 276) of the disyllabic verb roots, V_1 and V_2 differ in vowel quality. Various rules of increasing complexity and diminishing returns could be set up to deal with some of the remainder, but in many cases V_2 must be lexically-specified. Examples (193-194) show the verbs **pina** 'shave', **pino** 'boil', **dooho** 'disappear', and **goonu** 'help' together with the causative <**is**> and pluractional <**il**> morphemes. The final vowel in these examples cannot be determined by any phonological rules, and so must be considered as part of the root. Thus the causative and pluractional morphemes interrupt two portions of the root and must be analysed as true infixes.

---

48 And other West Kainji languages (e.g. C'Lela – Steve Dettweiler p.c.).

(193) (a) ù-pín<ìl>à
3s-shave<PLAC>

he shaved [many times]

(b) ù-pín<ìl>ò
3s-boil<PLAC>

it boiled [many times]

[2008-02-12.008]

(194) (a) ù-dôoh<ìs>ò
3s-disappear<CAUS>

he caused s.t. to disappear

(b) ù-gôon<ìs>ù
3s-help<CAUS>

he had s.o. to help

[2008-02-12.008]

4.6.2 Mood

The fundamental distinction in Cicipu verbs is mood. There are a number of aspectual morphemes but none of them are obligatory, and very often the verbal word consists simply of a subject agreement prefix together with the verb root. Mood, however, must always be specified, and this of course is tied into the way in which this category is encoded – through the tonal melody of the verbal word. No speaker can produce a pitchless utterance, and so almost all clauses have a specific value for mood: realis, irrealis, imperative[49].

4.6.2.1 Realis

In the realis mood the basic tone pattern is LHL(L)* as in (195), although in certain circumstances it may be HL(L)* – see §3.4.6 for details.

(195) ù-dúkwà
u-dukwa\LHL
3s-go\RLS
he went

[saat002.002.113]

In the absence of a past/present tense distinction in Cicipu, clauses headed by stative verbs marked with realis mood are ambiguous between past and present interpretation e.g. **ù-cúwò** can mean 'he/she was full' or 'he/she is full'.

---

49 Clauses with habitual aspect (§4.6.3.3) are an exception, since they have another tone pattern altogether. Dependent imperfective clauses (§4.6.3.4) also have a distinct tone pattern.

In certain restricted contexts the realis can be used to encode events which have not yet been realised, in particular those expected to happen in the imminent future[50]. For example the leave-taking expression in (196) is used when the speaker knows they will see the addressees again shortly.

(196) sée   n=ì-tôonò
       until   and=2P-come_home\RLS
       until you (pl.) come home

4.6.2.2 Irrealis

In the irrealis mood the basic tone pattern is H(L)*, with H being realised on the subject agreement prefix. For further details on the tone, see §3.4.6. Irrealis mood can mark verbs in independent clauses, as in (197), in which case – as in Hausa (Jaggar 2001:185) – it expresses the speaker's desire that something should or should not happen. When used in the second person this is politer than using the imperative (§4.6.2.3).

(197) úcàa
       u-caa\HL
       3S-give\IRR
       let him give

[oamy001.228]

Example (198) shows the use of the irrealis with the prohibitive **kádà** (from Hausa *kada*)[51].

(198) kádà   ɔ́dɔ̀nɔ̀
       kádà   ɔ-dɔnɔ\HL
       PROH   3P-follow\IRR
       they shouldn't follow

[saat001.006.044]

Alternatively, irrealis mood can occur in the subordinate clause complements of certain verbs (e.g. **ta'a** 'want', **kamaata** 'be fitting').

Clauses marked with irrealis mood normally encode events that are yet to be realised, although two classes of exceptions can be mentioned. The first involves habitual events, which can be encoded using irrealis mood. This is especially true for past habitual events (200), for which the habitual aspect prefix **si-** (§4.6.3.3) is rarely used.

---

50 cf. the Hausa perfective/completive (Jaggar 2001:158).

51 Negative prohibitions can also be expressed by using **bâa** to negate a positive irrealis clause, but this is much rarer in the corpus (see §4.3.5.1 for an example).

(199) n=à-kúɗùwà,   ɔ́-hɔɔmɔ̀, é-'ètù
when=3P-peel_off\RLS 3P-rub\IRR 3P-dry\IRR

when they've trimmed [the leaves], they rub [them], they dry [them]

[svtmg001.007]

(200) **tí-yũ̀u**   í-yáyíbá / **tí-yũ̀u** á-dábálá **tí-yã̀a** ké-kù-zè'ũ̀ /
1P-wear\IRR NC3-girdle_k.o. 1P-wear\IRR NC2-bead_k.o. 1P-do\IRR NC1-NC9-dance

we would wear 'iyayiba' girdles / we would wear decorative beads we would dance

[svtmg001.297]

A comparable constellation of functions involving habituals (past or present) and subjunctive mood is found in Hausa (Jaggar 2001:191-193). Jaggar, following Wolff (1993)[52], suggests there are two separate homophonous paradigms. If so, the parallel use of the irrealis for the same two functions in Cicipu is remarkable. Either this coming together is less anomalous than has been supposed by Hausa scholars[53], and is an independent property of the Cicipu irrealis, or it is the result of contact with Hausa, suggesting that in Cicipu, at least, the irrealis tone pattern should be assigned a unitary function.

The second exception is the use of the irrealis after the sentence connective **sée** 'then' (from Hausa *sai*) in narratives to encode non-habitual past events:

(201) sée   úcàa   Ø-Lɔ́wɔ́lì
then 3S-give\IRR NC8-large_spider

then he gave Big Spider [water]

[saat002.002.040]

The conflict between the irrealis marking and the semantics in (201) suggests that the H L pattern on the verb in such examples might be a morphotonological process caused by **sée**, rather than an expression of the category of irrealis mood.

4.6.2.3 Imperative

In the imperative mood there is no subject agreement prefix. The basic melody is (L)*H – in other words, the final tone of the verbal is H, and any previous tones are L. Again, for further details, see §3.4.6.

---

52 See also the discussion in Newman (2000:593-596).
53 In fact this combination is also found in the Austronesian language Kubokota (Chambers 2009) – see Palmer (2001) for further Austronesian and Papuan examples.

(202) dùkwá!
dukwa**LH**
go**IMP**
go!

If the verb root is vowel-initial, then a semi-vowel is added to the start of the word, as in (203). This semi-vowel surfaces as /**w**/ unless the root-initial vowel is /**i**/, in which case /**y**/ occurs[54].

(203) (a) **w**èewé   (b) **y**ìndá
ewe\LH   inda\LH
refuse\IMP   see\IMP
refuse!   *see!*

If an imperative is addressed to more than one person, then the plural imperative suffix **-nÀ** is obligatory. The imperative tone pattern is also present, but this time the H tone falls on the penultimate syllable, and the plural imperative suffix is always L tone.

(204) dùkwánà
dukwa\LH-**nà**
go\IMP-**IMP.PL**
[you (pl.)] go*!*

Prohibitions are not expressed using the imperative mood, but rather with second person irrealis mood and the prohibitive **kádà** (§4.6.2.2).

4.6.2.4 Counterfactual

Counterfactual mood can be expressed analytically using the auxiliary **sáa**, as in (205-206). The main verb occurs after the auxiliary and has no subject prefix. The tone on the main verb is often H H (especially for CVCV roots – cf. §3.1.3 on the habitual tone pattern), but is not yet fully understood.

(205) sáa kò-tôo tì-**sáa**   tíyó
or   AG1-one 1P-**might** get
we might even have got every one [if you hadn't messed up]

---

54 **àyá** 'come!' without any initial semi-vowel is an exception.

(206) ǹ-**sáa** húná ì-rí yí-vôo
     1s-**might** kill NC3-thing AG3-1s.POSS
     I ought to have killed my animal

<div align="right">[saat002.002.514]</div>

4.6.3 Aspect and tense

4.6.3.1 Perfective

The suffix -**nA** occurs in backgrounded clauses which encode situations viewed as completed. It is more-or-less obligatory in dependent clauses beginning with the conjunction **ánà** 'how/like' (207) and in relative clauses encoding situations viewed as complete (208).

(207) ù-dúkwà ń-kácí / ánà ù-dúkwà-**nà** ń-kácí...
     3s-go\RLS NC5-hunting how 3s-go\RLS-**PFV** NC5-hunting
     he went hunting / when he had gone hunting...

<div align="right">[Tidipo, saat002.003.031]</div>

(208) bá z-zá n-nà gúyà-**nà** Ø-yáarì ví-ivè
     NEG NC8-person AG8-REL can\RLS-**PFV** NC8-language AG8-3P.POSS
     there's no-one who knows their language

<div align="right">[svtmg001.261]</div>

What appears to be the same suffix -**nA** also occurs whenever a constituent is fronted through focus or marked topicalisation (§4.3.1-4.3.2). Again the event encoded has to be viewed as complete.

(209) k-èné kà-bárá k-è ká-ayà-**nà**?
     AG1-which NC1-old_man AG1-COP AG1-come\RLS-**PFV**
     which old man came?

<div align="right">[eamy005.027]</div>

(210) kà-bárá ká-mpà k-è ká-ayà-**nà**
     NC1-old_man AG1-this AG1-COP AG1-come\RLS-**PFV**
     this old man came

<div align="right">[eamy005.028]</div>

As was mentioned in §4.3.7.2, the perfective is very often found in tail-head linkage constructions.

4.6.3.2 Future

The future tense in Cicipu is encoded by the **u-** prefix, which occurs between the subject prefix and the verb root. It always occurs with the irrealis (H L) tone pattern.

(211) mú-**u**-yòo   cí-níyáa
    1S-**FUT**-go\IRR   NC6-farm
    *I **will** go to the farm*

The exact glossing of the future morpheme in Cicipu is problematic. Speakers usually translate it with the Hausa future*,* and it mostly (but not always) refers to events in the future according to some reference time. An exception is the use of the future to indicate habitual events in the past, as in (212) (cf. *would* in English). In these cases it is not clear how the meaning differs from the basic irrealis form (cf. §4.6.2.2).

(212) à-zá   há↓=d-dòorí   hú-**u**-gùyà   ù-kábà   ì-tángì
    NC2-person   AG2=NC8-formerly 3P-**FUT**-take\IRR   NC7-take   NC3-item
    the people of before would be able to carry things

4.6.3.3 Habitual

Habitual meaning can be expressed using the prefix **si-**, which occurs in the same slot as the future prefix. Additionally there is a specifically habitual tone pattern overlaid on the verbal word, distinct from each of the mood tone patterns (§4.6.2). As already discussed in §3.1.3, the exact tone pattern depends on the weight of the first root syllable[55]. The pattern also depends on syntax – in particular whether a constituent has been fronted for focus or relativised. The different possibilities are summarised in Table 27 and exemplified in (213-219).

Table 27: Tone on the habitual prefix plus root, according to root structure

Root structure	Default tone pattern	Fronted tone pattern
CVCV	L H H (ex. 213)	H L L (ex. 217) or H H L
CV:CV/CVNCV	L HL L (ex. 214)	H HL L (ex. 218)
CVC:V	L H L (ex. 215)	H H L?
CV:	L H *or* L HL (ex. 216)	H HL (ex. 219)

(213) ù-**sì**-**tá'á**
    3S-**HAB**-**want**
    he wants

---

55 For mono- or disyllabic stems. Trisyllabic stems show the 'heavy' pattern e.g. **ù-sí-kàbà-nà** 3PS-HAB-take-VENT 'he/she brings'.

(214) ù-sì-wîinà
　　　3s-**HAB**-**sell**
　　　he sells

<div align="right">[2008-04-01.004]</div>

(215) ù-sì-hɔ́ttɔ̀
　　　3s-**HAB**-**warm_by_fire**
　　　he warms himself by the fire

<div align="right">[2008-04-01.004]</div>

(216) (a) ù-sì-hyɑ́a　　　　　　(b) ù-sì-câa[56]
　　　　　3s-**HAB**-**say**　　　　　　　3s-**HAB**-**give**
　　　　　he says　　　　　　　　　　　*he gives*

<div align="right">[2008-04-01.004]</div>

(217) lêe　　v-ì　　　à-sí-kùdɔ̀
　　　there　AG8-COP　3P-**HAB**-**meet**
　　　it's there they meet

<div align="right">[Tikula, svsdt001.005]</div>

(218) lêe　　v-ì　　　à-sí-wâayà
　　　there　AG8-COP　3P-**HAB**-**come**
　　　it's there they come

<div align="right">[Tikula, svsdt001.003]</div>

(219) yìní　ì-sí-yǚu?
　　　what　2P-**HAB**-**wear**
　　　what did you (pl.) wear?

<div align="right">[sayb001.318]</div>

Most of the time **si-** occurs in utterances with a habitual or generic interpretation:

(220) tá-abà　　　tì-**sì**-húná　á-ngá
　　　NC6-tobacco　AG6-**HAB**-kill　NC2-tooth
　　　tobacco kills the teeth

<div align="right">[tats007.002.058]</div>

However this is not necessarily the case, and the prefix may also occur in clauses denoting a single event. In these cases the prefix seems to indicate continuous aspect, with a meaning similar to that of the analytic structure described in §4.6.3.6.

4.6.3.4 Dependent imperfective

Dependent 'when' clauses may begin with one of two conjunctions – **ánà** 'how' or **ǹ** 'and/with'. As noted in §4.6.3.1, in clauses beginning with **ánà** the verb is in realis

---

56 Some monosyllabic verbs pattern like **hyãa**, others like **caa**.

mood and almost always carries the perfective aspect suffix -**nA**. In clauses beginning with **ǹ**, on the other hand, the perfective affix cannot occur. If the verb carries the realis tone pattern with the usual final vowel, then the event is interpreted as perfective. If, on the other hand, the verb has a (L)*H tone pattern and the lexically-specified final vowel is replaced by either **i** or **u**, then the interpretation is imperfective, as in (221):

(221) naàmìní   sée   àzâanà   dádã   ìrímpàanì
      ǹ = a-mina-í   sée   à-zâa-nà   dádã   ì -rí -mpà-nì
      and=3P-skin-**DEP.IMPF** then   3P-see**RLS-PFV** moreover **NC**3-thing-this-**NMLZ**
      while they were skinning, they saw something

                                                    [Tidipo, saat002.004.064]

If the last vowel of the verb stem is **u** as in (222-223a) then it does not change. As with the final vowel in imperatives (§3.4.6), a H tone on an **u** vowel is realised as HL. Example (223) demonstrates that it is the last vowel of the verb *stem* that changes, rather than the last vowel of the root.

(222) nùujòolû
      ǹ-u-joolu-í
      and-3S-check-**DEP.IMPF**
      while he was checking

                                                    [2008-04-13.002]

(223) (a) nùufèllènû
          ǹ-u-felle-nu-í
          and-3S-cut-**RES-DEP.IMPF**
          while he was skimming [i.e. stones, lit. 'cutting']

      (b) nùufèllèní
          ǹ-u-felle-na-í
          and-3S-cut-**VENT-DEP.IMPF**
          while he was cutting down

                                                    [2008-04-13.002]

As mentioned in §3.1.4, if the verb stem is monosyllabic then the vowel is diphthongised and the tone is rising.

There is no constraint involving the subject of the dependent clause and the grammatical functions in the main clause – in other words, dependent imperfectives are not 'control' structures.

4.6.3.5 Progressive

The progressive suffix -**su** can attach to a few verb roots, resulting in a verb appropriate

for describing actions done repeatedly (224) or to excess (225).

(224) sée kúmá á-gìtù   á-ɗàsù-**sù**   kù-yáamà
       until more 3P-go_back\IRR 3P-soak\IRR-**PROG** NC9-name_of_month
       then they further soak the Kuyaama beer

[Tikula, sami001.280]

(225) Ø-'wâa-**sù**   kà-tíi   tì-gɛ́i
       2S-pass\RLS-**PROG** NC1-head NC6-length
       your head is too long [lit. 'you exceed head length']

[saat001.006.187]

4.6.3.6 Continuous

Continuous aspect is expressed using the auxiliary verb **yo** 'be' followed by the locative **Á** (§4.5) and then the infinitive/gerundial form of the verb (§5.4.1).

(226) **ùyô**   úubɔ̀lɔ̀   mòhíˆî
       ù-**yô**   á↓=u-bɔ́lɔ́   mò-híˆî
       3S-be\RLS   LOC=NC7-search NC4-blood
       it was looking for blood [lit. 'it was at looking for blood']

[2008-02-11.001]

This combination is also found in some Bantu languages (Welmers 1973:167), and indeed similar constructions occurs frequently in the world's languages (Bybee, Perkins, and Pagliuca 1994:130-132).

Most of the time the vowel of the auxiliary coalesces with the initial vowel of the infinitive to form what may be glossed as a continuous aspect morpheme **yuu**, as in (227). The **u** vowel is usually long but may occur shortened. Native speaker consultants perceive the full and reduced forms as identical, and when an adverbial phrase such as **ǹ màhɛ́u** 'really' is added as in (228), the auxiliary is clearly identifiable.

(227) ù-**yúu**   bɔ̀lɔ̀   mò-híˆî
       3S-CONT   search NC4-blood
       it was looking for blood

[samy002.018]

(228) **ùyô**   ǹ   màhɛ́u   úubɔ̀lɔ̀   mòhíˆî
       ù-**yô**   ǹ   mà-hɛ́u   á↓=u-bɔ́lɔ́   mò-híˆî
       3S-be\RLS with NC4-truth LOC=NC7-search NC4-blood
       it was looking for blood [lit. 'it was at looking for blood']

[2008-02-11.001]

It is possible, for at least some speakers, to separate the coalesced continuous morpheme **yuu** from the main verb, as in (229), which is in fact my main motivation for glossing it

as a separate aspect morpheme:

(229) **yúu**   lêe   **úu**dùkwà   sű?
Ø-**yúu**   lêe   **á↓=u**-dúkwá sű
2s-CONT   there   LOC=NC7-go   Q
Are you about to go?

This last example demonstrates the use of the continuous aspect in conjunction with **lêe** 'there' to express an imminent event; literally *are you there in-going?*

Finally, it should be noted that the **i/u** alternation (§3.3.2) applies to this morpheme. The basic vowel seems to be **u**, which is not surprising given **yuu** is a contraction of **yo** + **u**. However in the neighbourhood of an **i** vowel the continuous morpheme often surfaces as **yíi** instead.

4.6.3.7   Perfect

Past events with present relevance are often coded with the deictic **'û** 'there far off' or 'there (invisible)'. Despite the use of the most distal deictic, the event encoded is almost always recent, and the category marked is best described as a resultative perfect rather than an experiential one (Dahl and Velupillai 2005). Consequently this construction is found more often in conversation than in other genres. In narratives of past events it is usually felicitously translated by the English pluperfect, as in (233).

(230)   Ø-'úngò   **'û**?
2s-rise\RLS   **there**
you've risen? [standard opener in morning greetings]

(231)   ù-dúkwà   **'û**
3s-go\RLS   **there**
he's gone [said in response to a query about someone's whereabouts]

(232)   ǹ-kɔ́tɔ̀   **'û**
1s-finish\RLS   **there**
I have finished [said at the end of a speech]

(233)   ánà   ù-dúkwà-nà,   èsée,   Ø-náatà   Ø-níi'wà   **'û**
when   3s-go\RLS-PFV   actually   NC8-spider   AG8-melt\RLS   **there**
when he went, actually, Spider had already re-hydrated

'û always comes after the object enclitic (§7.3), if there is one. If there is an object NP in the complement of VP then the deictic can come either before (234) or after (235) it.

(234) ánnà   ǹ-húnà   **'û**   **k-káa**   **v-vôo**
today 1s-kill\RLS **there** NC**8-wife** AG**8-1s.POSS**
today I have killed my wife

[saat002.002.346]

(235) ù-tábbàtà   Ø-lɔ́wɔ́lì   Ø-râa   **ì-námà**   **yí-llè**   **'û**
3s-touch\RLS NC8-large_spider AG8-eat\RLS NC**3-meat** AG**3-that** **there**
he felt sure Big Spider had eaten that meat

[saat002.002.624]

4.6.4   Valence

Underived Cicipu verbs can be classified as intransitive, (mono)transitive, or ditransitive. Subject agreement is obligatory, although the lexical subject may be omitted. This state of affairs can be observed across Bantu (Bearth 2003:122), and the details with respect to Cicipu are discussed in §7.6. There are no object agreement markers.

Objects are frequently omitted[57], as in (236), even when they are specific and bear an important role in the semantic representation of the verb. The motivation seems to be economy of expression. As in Hausa (Jaggar 2001:602) object NPs with human referents are not normally omitted.

(236) ánà   ù-bɔ́dɔ̀-nɔ̀=**vì**   mò-tôo   w-áyà   ù-kúɗùwà Ø   dáɗà
when 3s-apply\RLS-PFV=**3s.PRO** NC4-saliva 3s-come\RLS 3s-peel\RLS   moreover

ù-kábà   Ø ù-dúkwà-wà   Ø   á=kù-jénè   ù-yô   ù-ɗásù   Ø
3s-take\RLS   3s-go\RLS-APPL   LOC=NC9-river 3s-be\RLS NC7-soak

*when he had covered **it** [Spider's body] with saliva he peeled [**it**] off, he took [**it**] and went with [**it**] to the river and was soaking [**it**]*

[saat001.008.098]

Object omission is also common in Bantu languages, but according to Bearth (2003:123) it serves a different purpose, that of "limiting what is said to the main point". In other words, speakers omit the object not because the hearer will be able to recover the reference from the context, but to avoid violating Grice's (1975) maxim of quantity – they simply don't want to specify referents if they are not relevant (cf. *I've eaten Ø already*). Cicipu is closer to Hausa in this respect – in (236) the identity of the unexpressed arguments is perfectly clear (in this case it is the body of Spider).

---

57 See §4.3.1.1 for ditransitive verbs.

Many Benue-Congo languages also have several derivational affixes which affect the valence of the verb, and Cicipu is no exception: the causative infix <is> and the applicative suffix -wA increase the valence of the verb by one, while the anticausative suffix -wA decreases the valence by one.

4.6.4.1 Causative

The causative <is> was first mentioned in §4.6.1, along with the reason for analysing it as an infix. This morpheme is doubtless derived from the same source historically as the causative verbal extension -*is* found across Bantu (Schadeberg 2003:73), although in Cicipu it may only be added to intransitive stems. It can co-occur with transitive *roots*, but only if the anticausative suffix -wA is also present so that the valence remains at 2. This combination differs subtly in its semantics from the bare root, as discussed in §4.6.4.3 below.

(237) **hina**   *ripen*   **hin**isa   *cause to ripen*

   dooho   *disappear*   **dooh**iso   *cause to disappear*

   sukulu   *move*   **sukul**usu   *cause to move*

   yuwo   *fall*   **yuw**oso   *cause to fall*

   'ungo   *rise*   **'ung**oso   *raise*

   he'we   *dry*   **he'w**ise   *cause to dry*

   ruɓa   *sink*   **ruɓ**isa   *cause to sink*

   kullo   *burn*   **kull**iso   *cause to burn*

For monosyllabic verbs the [i] seems to be absorbed in the root vowel, as in the pairs **sõo~sõosõ** 'cry~make s.o. cry', and **raa~raasa** 'eat~serve s.o. food'.

 As in certain Bantu languages (Schadeberg 2003:74), what appears[58] to be the same infix has an additional function: rather than affecting the valence of the verb, applying the <is> infix can have an intensifying effect instead:

(238) (a) ɔ'ɔ́pɔ̀   (b) ɔ'ɔ́pànù   (c) ɔ'ɔ́pìsànù

    ɔ-'ɔ́pɔ̀   ɔ-'ɔ́pɔ̀-nù   ɔ-'ɔ́p<ìs>ɔ-nù

    3P-hold\RLS   3P-hold\RLS-RES   3P-hold\RLS<CAUS>-RES

    they hold   *they grab*   *they grab with great force*
                       [tats005.002.089]

---

58 As Schadeberg notes, it is hard to see the motivation for the shift in meaning.

(239) kònd\<**ìs**\>ì-nó   á↓=ká-ayá!
      enter\<**CAUS**\>\\IMP-VENT   LOC=NC1-room

come into the room! [said with urgency]

The causative morpheme can be repeated as in example (191) at the beginning of this section. The effect of this reduplication is not well understood, although it may have the effect of intensification in addition to causation.

There are alternative means of expressing causation in Cicipu apart from the \<**is**\> causative[59]. One way is using a periphrastic construction involving the verb **yũu** 'wear, appoint, cause'. This construction has a direct analogue in Hausa using the verb *sa* 'put, wear, appoint, cause' (Jaggar 2001:552), and given the similarity in lexical meaning it seems likely Cicipu has borrowed the construction.

A causative meaning can also be obtained by replacing the final vowel of the root with /**u**/: e.g. **kondo/kondu** 'enter/cause s.t. to enter', **koyo/koyu** 'learn/teach'. It is not known how productive this process is, nor how or if these causatives differ in meaning from the \<**is**\> ones.

4.6.4.2 Applicative

The most obvious effect of the applicative suffix -**wA** is to increase the valence of the verb by one. If the basic verb is intransitive, then the derived verb has a single object. If the basic verb is transitive, then the object of the basic verb (240a) becomes the secondary object, the primary object slot being filled by the 'dative' object as in (240b). The subject is unaffected in all cases. I was not able to elicit applicative forms of any basically-ditransitive verbs such as **caa** 'give' (e.g. *I gave food to the child for its mother*).

(240) (a) ù-dámà   [ì-rí   yì-'étéi]
            3S-speak\\RLS   NC3-thing AG3-fine

            he spoke good things

       (b) ùdámùwà   [àzà]   [ìrí   yì'étéi]
            ù-dámà-**wà**   **á-zá**   ì-rí   yì-'étéi
            3S-speak\\RLS-**APPL**   NC**2-person** NC3-thing AG3-fine

            he told the people good things

The semantic role of the primary object is variable, as was the case for the morphologically-unmarked double-object constructions (§4.3.1.1). Often it is the

---

59 See also §4.6.4.2.

RECIPIENT or BENEFICIARY of the encoded event, as in (241). Equally often the primary object encodes a participant adversely affected by the action – a 'MALEFICIARY', as in (242). Other semantic roles include LOCATIVE (243), COMITATIVE (244), and THEME/PATIENT (245) – but *not* INSTRUMENTAL. The primary object is often omitted, as in (244).

(241) mú-u-yẽ̀a-**wà**=[vù]   [kà-'ázzìkí]
1S-FUT-do\IRR-**APPL**=2S.PRO   NC1-prosperity
I will make you (sg.) prosper [lit. do you prosperity]

[BENEFACTIVE, saat002.002.393]

(242) z-zá   n-nà   sì-yẽ̀a-**wà**=[vù]   [ù-bɔ́wɔ̀]
NC8-person   AG8-REL   -HAB-do-**APPL**=2S.PRO   NC7-steal
the one who's been stealing from you [lit. doing you stealing]

[MALEFACTIVE, saat002.001.104]

(243) àdúkwùwàvì   ǹ   wúllè   ìrí
à-dúkwà-**wà**=[vì]   ǹ   wú-llè   ì-rí
3S-go\RLS-**APPL**=3S.PRO   with   3S-that   NC3-thing
they went to it [the spirit] *with something*

[LOCATIVE, saat001.008.099]

(244) ùdúkwùwà   ákūjēnè
ù-dúkwà-**wà** [Ø]   á↓=kú-jénè
3S-go\RLS-**APPL**   LOC=NC9-river
he went with [the corpse] *to the river*

[COMITATIVE, saat001.008.099]

(245) tànnùwá   bíirò
tànnà-**wá**   [Ø-bíirò]
descend\IMP-**APPL**   NC8-biro
lower the biro!

[THEME, eamy012.1428]

When the applicative is applied to an intransitive verb with a theme or patient argument as in (245), the effect of adding -**wA** is very much like a causative. Some intransitive verbs take <**is**> and others -**wA**, but I have not found any which allow a choice. There is no obvious semantic difference between verbs which form causatives with -**wA** and those which form causatives with <**is**> – compare (245) with **'ungo/'ungoso** 'rise/raise'.

4.6.4.3 Anticausative

The anticausative suffix -**wA** is formally identical to the applicative suffix just discussed. In terms of valence, however, they have the opposite effect, and they can co-occur in the same verbal word (as in (191) above). Therefore they are considered to be

homonyms. The function of the anticausative is to downplay the role of the agent/causer in the event denoted by the verb, so much so that it cannot be expressed at all, even in a prepositional phrase. This inability to refer to the agent is why the suffix has been glossed as anticausative rather than passive (Dixon and Aikhenvald 2000:7).

Not all verbs can take the anticausative suffix – the ones that do seem to encode events in which the patient is changed in some way:

(246) (a) ù-sídù   mó-ní   (b) mò-ní   mà-sídù-**wà**
          3s-heat\RLS NC4-water   NC4-water AG4-heat\RLS-**ANTIC**
          he heated water   *the water got heated*
                                                              [2008-02-11.011]

The anticausative may combine with the causative <is>. The combination of a valence-decreasing device with a valence-increasing one does not result in any net change in valence, but the doubly-derived verb differs slightly in its meaning. ?**ù-lénjí wù-sídù mó-ní** 'the sun heated the water' sounds strange because it implies agency on the part of the sun. However **ù-lénjí wù-síd<is>ù-wà mó-ní** 'the sun caused the water to get heated' is fine, because it has no such implication. This fits in with the established cross-linguistic pattern that analytical causatives tend to imply less agentivity/volition on the part of the causer than lexical causatives (Payne 1997:182).

4.6.5   Other verbal suffixes

4.6.5.1   Ventive

The ventive suffix **-nA** indicates that the action encoded takes place in the direction of the speaker or some other deictic centre, for example the current centre of attention in a narrative. It is best to view this as a homonym of the perfective aspect suffix discussed in §4.6.3.1 (and indeed of the plural imperative suffix, §4.6.2.3).

The effect of the suffix is similar to the Hausa grade 6 ventive ending *-o* (Jaggar 2001:256), and they correspond in translations given by native speakers. The suffix is very productive: common examples include **wuuto-no** 'come out' vs. **wuuto** 'go out', **kaba-na** 'bring' vs. **kaba** 'take', and **kondo-no** 'come in' vs. **kondo** 'go in', but one can also just 'do' (**ỹãa-na**) something in a particular direction.

When talking about the movement of physical objects in the speech situation (rather than, say, a narrative with a displaced 'storyworld'), the deictic centre is usually the position of one of the speech participants. However for certain actions a different

deictic centre can become conventionalised (e.g. the ground for falling objects – when dropping a pen from below head height, **-nA** is consistently used (**bíírò vì-yúwò-nò** 'the pen fell'), even though the movement is away from the interlocutors).

The action encoded by verbs with the ventive suffix does not actually have to involve a direction. Instead the action may benefit the speaker (or person at the deictic centre) in some way:

(247) nì   Ø-yôo   ń-kácí,   sée   ú-hùnà-**nà**   ì-námà
      and 2s-go\RLS NC5-hunting until 3s-kill\IRR-VENT NC3-meat
      when you (s.) go hunting, then it kills meat [for you]
      [tats008.004.024]

The perfective and ventive **-nA** suffixes may occur together, in which case they are usually realised by **-nnA**, as in (248). The verb **yaa** 'arrive' is exceptional in that **yaa-na-na** is found rather than **yaa-nna**.

(248) ùdánùwànàmù   kàtákàddá /
      ù-dánà-wà-nà = mù   kà-tákàddá
      3s-write\RLS-APPL-VENT=1s.PRO   NC1-book

      ánà   ùdánùwà**nnà**mù   kàtákàddá...
      ánà   ù-dánà-wà-**nà-nà** = mù   kà-tákàddá
      when   3s-write\RLS-APPL-**PFV-VENT**=1s.PRO   NC1-book

      she wrote me a letter / when she had written me a letter...
      [2008-02-12.003]

For the verb **dɔnɔ** 'follow' the ventive form is **dɔnnɔ** rather than the expected **dɔnɔnɔ*. The perfective form is **dɔnɔ-nɔ**. When both the ventive and perfective suffixes are applied the result is **dɔnɔ-n-nɔ**, which suggests (perhaps surprisingly) that the perfective affix occurs closer to the root than the ventive.

4.6.5.2 Resultative

If the encoded event reaches some sort of natural endpoint then the resultative derivational suffix **-nu** may be used. Consider the following pairs:

(249) **kɔdɔ**   *cut*   **kɔdɔnu**   *cut off*

      golo   *cut*   **golonu**   *cut up*

      sɔɔ   *drink* **sɔɔnu**   *drink up*

      vayu   *drop*   **vayanu**   *drop in* [i.e. into a container]

      yuwo   *fall*   **yuwonu**   *fall in*

The use of the verbs **kɔdɔ**, **golo**, and **sɔɔ** does not imply that the action being encoded

was completed, or that the object referent was wholly cut or consumed. The derived verbs **kɔɗɔnu**, **golonu**, and **sɔɔnu** do however lead to this interpretation. The verbs **vayanu** and **yuwonu** are appropriate when the theme ends up enclosed in some way, for example in a pit or a pool.

-**nu** occurs obligatorily with many verbs such as **ciɗonu** 'bury' **hintonu** 'tie', **jungonu** 'shut'[60], **sɔttɔnu** 'urge', **yaddanu** 'abandon' (from Hausa *yad da* 'throw away'). Most of these verbs are 'accomplishments' (Vendler 1957) with an inherent endpoint; one cannot partially tie a knot, or abandon a child.

The affix often serves to indicate an action carried out with more force than normal, as in **'ɔpɔ~'ɔpɔnu** 'hold~grab' and **koɗo~koɗonu** 'tap~tap forcefully'. Related to this, the suffix may also be used to indicate a certain urgency – when playing the board game *dara* players have been observed to call **dùwànú!** when they are losing patience, rather than the usual **dùwá!** 'place!'.

Sometimes the effect of adding the affix is even more unpredictable as in **naha** 'leave' vs. **nahanu** 'leave s.o. an inheritance', **dama** 'speak' vs. **damanu** 'whisper', and **gama** 'join' vs. **gamanu** 'mix'. The argument structure possibilities of the verb may also change – in (250a) with **konto** 'bump into', the wall can only be expressed as the object of the verb, whereas in (250b) with **kontonu** 'collide' it is also possible to use the locative **Á**.

(250) (a) Ø-kóntò   kò-pûu
         2s-bump_into\RLS   NC1-wall
         you bumped into the wall

     (b) Ø-kóntò-**nù**   (ɓ↓=) kó-pûu
         2s-bump_into\RLS-RES (LOC=) NC1-wall
         you bumped into the wall

[eaim006.1448]

4.6.5.3 Separative

A few verbs take the separative suffix -**wA** (again, presumed to be homonymous with the valence-changing suffixes discussed in §4.6.4). The only examples attested are **hala~haluwa** 'coil~uncoil', **'umbonu~'umbuwo** 'close~open', **ciɗonu~ciɗuwo** 'bury~uncover earth', and **kuntonu~kuntuwo** 'clench~unclench fist'. In each case the verb with the -**wA** suffix involves separation as part of its meaning. As argued by

---

60 Interestingly **jungo** means 'open', but adding -**nu** is not a productive means of opposite formation.

Schadeberg (2003:79) for Bantu, 'separative' is a better gloss than reversive, since it predicts which member of the pair will occur with the suffix.

4.6.5.4 Pluractional

In Cicipu verbs with the highly productive pluractional[61] infix <**il**> can encode several different kinds of situations, depending on the source of the 'plural' component of meaning. This can be supplied by multiple actors, multiple patients, or multiple events. The various possible sources of pluractionality and different interpretations are summarised in Table 28:

Table 28: Sources of pluractional marking

Multiple agents	Multiple patients	Interpretation	Example
✘	✘	Iterative	(251)
✘	✘	'Disintegrative'	(252-253)
✘	✔	Distributive	(254)
✔	✘	Collective	(255)
✔	✔	Distributive	(256)
✔	✔	Reciprocal	(257)

If there is only one agent and (at most) one patient, then pluractional marking can give rise to two interpretations. The more straightforward is the 'iterative' reading shown in (251), where the event is carried out repeatedly. The second possibility is what I have called the 'disintegrative' reading, in which the theme becomes divided into several parts as a result of the event. This is illustrated by (252-253).

The pluractional is also used when a single agent acts on multiple patients, as in (254). Conversely, it is also used when multiple agents act on a single patient (as in the collective reading shown in 255). Finally, the pluractional can encode multiple agents acting on multiple patients, as in the distributive reading in (256). If the agents/patients are acting on *each other* (e.g. the reciprocal event encoded by 257) then pluractional marking can be used even though there are only two referents in total.

(251) wú-u-dèmp<ìl>è
       3s-FUT-punch<PLAC>\IRR
       he would pummel [him]

[saat002.002.544]

---

61 Pluractional verbs are found in Hausa as well as Benue-Congo. The term was suggested by Newman (e.g. Newman 1990).

(252) mò-ní   m-áyà   mɔ̀-kɔ́ɗ<ìl>ɔ̀-nù
NC4-water AG4-come\RLS AG4-cut<PLAC>\RLS-RES
then the water parted [lit 'cut']

<div align="right">[sahs001.003.015]</div>

(253) kà-jírígí k-áyà   kà-ɓát<ìl>à
NC1-vessel AG1-come\RLS AG1-tear<PLAC>\RLS
then the aeroplane disintegrated

<div align="right">[Tidipo, saat002.005.025]</div>

(254) ǹ-sɔ́<ɔ̀l>ɔ̀   mó-ní ǹ y-í pɔ̃̂
1S-drink<PLAC>\RLS NC4-water and AG3-PRO all
I drank water from each of them [e.g. cups]

<div align="right">[2008-02-21.004]</div>

(255) ì-zíyá̌ yí-cít<ìl>a   cì-lúu
NC3-bee AG3-sting<PLAC>\RLS NC6-leopard
the bees stung the leopard

<div align="right">[saat001.006.104]</div>

(256) à-zá   há-nà à-mát<ìl>à-nà=ɗò
NC2-person AG2-REL 3P-father<PLAC>\RLS-PFV=2P.PRO
the people who fathered you (pl.)

<div align="right">[sayb001.797]</div>

(257) sée nì=i-tás<ìl>à-nù
until and=2P-meet<PLAC>\RLS-RES
unless you (pl.) meet [i.e. meet each other]

<div align="right">[tats004.001.035]</div>

There is at least one verb whose pluractional form takes on an unpredictable meaning: **sa'a~sa'ila** 'wash hands or face~wash legs'. In some words, the infix seems to have fossilised, and they can no longer occur without it e.g. **havila** 'scratch', **piɗolo** 'roll around in the dirt', **titila** 'clear away' and **ɗangula** 'gather'. The iterative component of meaning can clearly be seen.

For monosyllabic roots the vowel of the pluractional infix is absorbed by the first vowel of the root (e.g. (254) above, also **tãa~tãalã** 'shoot' and **dãa~dãalã** 'stretch'). If the last consonant in the root is **l** then the infix vowel may disappear leaving a double **ll** (cf. **dɔnɔ~dɔnnɔ** in §4.6.5.1). Examples are **lapila~lapilla** 'prepare' and **dala~dalla** 'dip food in soup'. In these cases the shortened double-**l** form seems to be more common than the full form.

4.6.6   Borrowed verbs

Verbs borrowed from Hausa are often conveniently indicated by the suffix **-kwA**.

(258)   **gwaanukwa**   *understand*   from *gane*

daamukwa   *worry*   from *dama*

haɗukwa   *join*   from *haɗa*

laatukwa   *perish*   from *lalace*

macukwa   *close in on*   from *matsa*

taɓukwa   *do s.t. once*   from *taɓa*

taarukwa   *meet*   from *tare*

yaahukwa   *forgive*   from *yafe*

koyuko   *teach*   from *koya*

dennukwe   *compress*   from *danna*

Some borrowed verbs are never heard with this suffix. Often this is the case for what seem to be well-established borrowings (e.g. **gwede** 'thank' from *gode,* **biya** 'pay' from *biya*). In other cases the lack of **-kwA** may be a result of code-switching rather than borrowing; speakers have been observed correcting the bare Hausa form to 'proper' Cicipu by adding the suffix.

The historical derivation of this suffix is mysterious. It has no other synchronic function, and there is no obvious source candidate. Moreover the trisyllabic verbs with **-kwA** actually look *less* like native Cicipu verbs than the bare stems.

4.6.7   Auxiliary and aspectual auxiliary verbs

A number of phrasal constructions involve semi-grammaticalised auxiliary or quasi-auxiliary verbs. Some can be used as the only verb in a clause, while others are limited to the constructions discussed here. These auxiliaries inflect for subject and sometimes other verbal categories but their semantic contribution is adverbial, in that they do not express the main conceptual event or state of the clause.

Three kinds can be identified, according to the way in which the (semantically) main verb is marked. First, following some auxiliaries the verb has no prefix at all, as with **'esu** 'add' or 'do s.t. again' and **ɗan** 'do s.t. a little'[62], illustrated below. **Vesi** 'do s.t. too' is also of this kind.

---

62 *Ɗan* is a Hausa borrowing; in fact, the whole construction is borrowed cf. Jaggar (2001:152).

(259) mò-ní   mè-**ɛ́sù**   gítà-nà
NC4-water   AG4-**again**\RLS   go_back-VENT
the water came back again

[samoh001.128]

(260) ù-**ɗán**   tòsù
3S-**little**\RLS   push
he pushed a little

[eaim006.1457C]

Other auxiliaries are followed by the NC7 infinitive (§5.4.1). Examples are given below for **'ɔpɔ** 'hold' which can also mean 'do s.t. repeatedly/excessively', **hwaara** 'start' with the "aspectual auxiliary" (Binnick 1991:174) meaning 'start doing s.t.' or 'do s.t. first' (Hausa *fara*), and **naha** 'leave' or 'stop/refrain from doing s.t.'. Others of the same kind include **guya** 'be able to', **jungo** 'open' or 'start doing s.t.', and the Hausa loans **saaba** 'be used to doing s.t.' (*saba*) and **danganu** 'keep doing s.t.' *(dinga)*.

(261) ù-**'ɔ́pɔ̀**   ù-pídǎ
3S-**hold**\RLS   NC7-lick
*he **repeatedly** licked*

[saat002.002.298]

(262) ù-**hwáarà**   ù-wútò-nò
3S-**start**\RLS   NC7-go_out-VENT
*it **started** coming out*

[tats005.002.085]

(263) à-**náhà**   ù-húná wí↓=i-námà
3P-**leave**\RLS   NC7-kill AG7=NC3-meat
*they **stopped** the killing of meat*

[samoh001.140]

The third kind of auxiliary is followed by a main verb with a subject agreement prefix. The construction is nevertheless distinct from a straightforward clause juxtaposition, since only one event is encoded. Examples include **bã'ã** 'do s.t. already' (264) and **aya** 'come' or 'then' (265). The latter auxiliary is very common on the main event line in narratives, especially when the situation encoded by the clause involves a temporal progression with respect to that of the previous clause.

(264) tì-**bã́'ã̀**   cé   tí-sírìyà
1P-**already**\RLS   NEG   1P-prepare\RLS
*we hadn't **already** prepared*

[saat001.005.058]

(265) kà-táarí ká-mpà k-**áyà**   kú-utò
NC1-stone AG1-this AG1-**come**\RLS AG1-go_out\RLS

then *this stone went out*

4.6.8 Order of affixes

The order of verbal affixes is strict, and no exceptions to the sequence set out below has been found in the whole corpus, nor could any be elicited.

(266)   AGR-{FUT/HAB}-**root**<PLAC><CAUS >-ANTIC-APPL-{PFV/PL.IMP}-VENT=OBJ.PRO

The future and habitual prefixes do not seem to co-occur, and so they have been placed in the same slot. No examples have been found containing both the perfective suffix and the plural imperative suffix, and it is probable that the two categories are logically incompatible. The object clitics (§7.3) are bound to the right-hand edge of the verbal word.

4.6.9 Co-occurrence

The various affixes and suprasegmental morphemes discussed in this section can be divided into three groups, according to their behaviour with respect to co-occurrence.

The first group is comprised of the tonal patterns associated with realis, irrealis, and imperative moods, together with the habitual aspect marker **si-**. Because each of these categories requires a particular tone pattern for their expression, it is logically impossible for any two of them to be expressed at the same time.

The second category consists of the perfective aspect marker **-nA** and the future tense prefix **u-**. **-nA** can co-occur with the realis mood tone pattern, but not seemingly with the others. **u-**, on the other hand, only occurs with the irrealis tone pattern. These two morphemes therefore cannot co-occur.

The third category consists of the various derivational affixes – the resultative, separative, applicative, anticausative, and ventive suffixes, together with the pluractional and causative infixes. Each of these can occur with one from the first group together with one from the second. Moreover it seems that in theory they could all be found in a single verbal word – recall example (191) above.

In addition to their co-occurrence possibilities, the third category forms a natural class with respect to deverbal nominalisations (§5.4), which can only contain affixes from the third category. There is also a correlation between the above three classes and

their semantics, which in turn correlates with the productivity of the affix. Morphemes from the first and second categories express inflectional TAM categories and are maximally productive, whereas those from the third category express derivational categories and are less productive.

Little information is available on the co-occurrence possibilities of the dependent imperfective (§4.6.3.4) and the progressive -su (§4.6.3.5), although given the former requires a specific tone pattern we can assign it to the first group, and the restricted productivity of the latter makes it a candidate for the third group.

4.7   Adjectives

There are very few adjectival roots in Cicipu, and many 'adjectival' meanings are instead expressed using nouns (e.g. tì-lípầi 'goodness, NC6') or verbs (e.g. gɔnjɔ 'be curved'). As might be expected, true adjectives in Cicipu share properties of both nouns and verbs. Like nouns they carry lexical rather than grammatical tone, and when they function as predicates they can be followed by the copula (e.g. (269) below, cf. §4.3.3.1). On the other hand, like verbs they do not have lexical gender, and instead agree in gender with their head noun. Some examples of adjectives are given below:

Table 29: Examples of adjectives

'étêi	*fine*
pénêu	*big*
ɗénêu	*small*
sáavínà	*new*
kúnó	*old*
'úyónò	*light*
rúmònò	*dark*
sílánà	*red*

An adjective can occur either as a modifier (267) or as the head of an NP in the absence of a noun (268), as with most other noun modifiers (§4.4.5.5).

(267)   mɔ̀-'yɔ́'yɔ́   **mà-sílánà**
        NC4-fish   **AG4-red**
        ***red** fish*

[Tikula, taff002.016]

232

(268) à-hyɑ̂a=mù   ń-vàsà   **tì-ɗénénêu**?!
      3P-say\RLS=1S.PRO   1S-beat\IRR   **AG6-small;REDUP**

*they told me to beat the **small** one* [**cì-kóotò** 'drum, NC6']*?!*

[saat001.008.073]

When functioning as the heads of NPs, adjectives can take part in predicate nominal constructions, as in (269). There is no formal distinction between attributive and predicative adjectives in Cicipu. Note that in addition to the lexical tone pattern on adjectives, the (optional) presence of the copula after the complement NP distinguishes these constructions from verbal clauses.

(269) kà-tákàddá ká-mpà   kò-**'úyònò**   k-é
      NC1-book   AG1-this   AG1-**light**   AG1-COP

this book is light-coloured [lit. 'this book is a light-coloured one']

[eamy001.012]

4.7.1   Reduplication

The reduplication of adjectives has different effects, depending on the quality involved. The meanings of some adjectives are intensified under reduplication, as in **pénêu/pénénêu** 'big/very big', **ɗénêu/ɗénénêu** 'small/very small' and **'étéi-'étéi** 'very fine', whereas with colours the meaning is attenuated, as in **sílánà-sílánà** 'reddish, pink'. Cicipu thus contrasts with Hausa, where adjective reduplication *always* seems to give rise to attenuation (cf. *babba-babba* 'biggish' – Newman 2000:27-29, Jaggar 2001:144-146).

4.8 Numerals

4.8.1 Cardinal numbers

Table 30: Cardinal numbers in Cicipu

1	tôo	21	ùkúppá wùyápù ǹ vìtò
2	yápù	22	ùkúppá wùyápù ǹ vìyápù
3	tâatù	30	ùkúppá wùtâatù
4	nósì	40	ùkúppá wùnósì
5	tẫu	50	ùkúppá wùtẫu
6	tóríhǐ̃	60	ùkúppá wùtóríhǐ̃
7	tíndàyà	70	ùkúppá wùtíndàyà
8	kùríllò	80	ùkúppá wùkùríllò
9	kùtíttí	90	ùkúppá wùkùtíttí
10	ùkúppá	100	ɗèrí (H.) ~ ùkúppá sùkúppá
11	ùkúppá ǹ vìtò	200	ɗèrí yápù
12	ùkúppá ǹ vìyápù	1000	dúbù (H.)
20	ùkúppá wùyápù		

Cicipu has a straightforward decimal numeral system with very few irregularities – it is possible to construct any number analytically using only the words for 1-10. It is interesting to note that Mathews (1926b) gives *hi* as the Cicipu for 1, with *zattu* (**z-zá t-tò** 'one person') in parentheses. While I never encountered this word in my fieldwork, clear cognates of *hi* exist in other West Kainji languages, particularly in the Basa-Kamuku branch (Blench n.d. a). The form of **tórí-hǐ̃** 'six' suggests that **tôo** is a later replacement for *hi*[63], since the words for 6 and 7 are historically derived from the phrases "5 and 1" and "5 and 2" in other West Kainji languages (e.g. Smith (2007:69) for u̱t-Ma'in), and more generally in West Africa (Welmers 1973:294).

Kùríllò 'eight' and kùtíttí 'nine' begin with what looks like the AG9 prefix **ku-**. I have observed speakers omitting the **ku-** (e.g. **m-ɔ́ɔ mò-ríllò** 'eight pieces'), but on querying I was told this was not correct Cicipu[64].

---

63 Cross-linguistically the word for 'one' is hardly ever borrowed, and even when it is it co-exists with the original term (Souag 2007). **Tôo** may be a language-internal innovation derived from **z-zá t-tò** 'alone' lit. 'one person' – cf. the equivalent attributive numeral construction in Central Kambari, **z-zá t-tò**, where the cardinal number is **íyyá̌** (Hoffmann 1963:23).

64 In the East Kainji language Amo the numerals 6 and above take an invariant 'concord' prefix **ku-** (Anderson 1980a:163), so that 8 and 9, for example, are **ku-lír** and **ku-tírí** respectively (in their attributive forms). It is not hard to see how such an invariant prefix could be reanalysed as part of the

Ù-kúppá 'ten' can occur with or without the noun class prefix.

In general, the Cicipu numbers are being replaced by the Hausa numbers, and younger speakers often have difficulty with the numbers above 5. Even adults may not know these numbers, and they often count in Hausa instead, even when the rest of the conversation is being conducted in Cicipu.

4.8.2 Attributive numerals

Attributive (enumerative) numerals have a lot in common with adjectives (§4.7). They can occur as noun modifiers or as pronouns, they are lexically-specified for tone, and they take the same low-tone gender agreement prefixes. It may also be the case that they occur in the same syntactic slot in the NP (see §4.4.5.6). Examples are given below:

(270) kà-mángá **kò-tò**
NC1-rope **AG1-one**
one rope

[saat001.002.059]

(271) wɔ́-ɔtɔ̀ **wù-yápù**
NC7-month **AG7-two**
two months

[2008-03-21.002]

See §6.2.2 for more details on gender agreement on numerals.

4.8.3 Ordinal numbers

If a numeral occurs in the 'possessor' slot in an associative construction (§4.4.5.1) then it is understood to be ordinal, just as in Bantu (Welmers 1973:277) and many other Niger-Congo languages (e.g. Igbo, Welmers 1973:283). There is also a clear parallel with Hausa, where the possessive linker occurs between the head noun and the numeral in ordinal constructions (Newman 2000:386, Jaggar 2001:363). Compare (272) with (271) above:

(272) wɔ́-ɔtɔ̀ **wú↓=y-yápù**
NC7-month **AG7=NC8-two**
the second month

[2008-03-21.002]

The ordinal forms differ from the enumerative forms in two ways – the tone on the prefix is high, and the neutral NC8 lengthening allomorph **C-** is inserted between the

---

number itself, and this may be what has happened in the Cicipu forms for 8 and 9.

agreement prefix and the numeric root (see §6.4.2).

4.8.4 'Adverbial' numerals

The number of times an event has occurred can be expressed using a numeral with an AG9 prefix but no head noun (see §6.3 for further uses of AG9 agreement morphology).

(273) ǹ-dóonù   **kù-tò**
     1s-sit\RLS   **AG9-one**
     I sat down once [i.e. a single time]

                    [eamy003.1358]

(274) sáa   ù-wáyáa-ní   **kù-yápù**   pâa
     even NC7-come-NMLZ **AG9-two**   here
     even coming here twice

                    [sayb001.579]

Ku- may also occur on the wh-word **yàanú** 'how many' (§4.3.6.2) in the absence of an agreement controller, in which case the interpretation is usually 'how many times':

(275) **kù**-yàanú kw-ì   sée   ì-sí-ya̋a-wà   kó-kkwíi   kà-bíkì
     AG9-who   AG9-COP unless 2P-HAB-do-APPL NC1-corpse NC1-festival
     how many times do you have to do a festival for a corpse?
                  [Tikula, sagb001.415]

4.8.5 Distributive numerals

Finally, as is common in African languages, the reduplication of a numeral expresses distributive meaning:

(276) à-zá   yápù yápù
     NC2-person two   two
     together in twos

                  [samoh001.213]

4.9 Quantifiers

Each of the Cicipu wh-question words (§4.3.6.2) has a universal pro-form/quantifier counterpart, formed by placing **sáa** 'whether/or' in front of it.

Table 31: Universal quantifiers

Question word	Gloss	Universal quantifier	Gloss
yìní (NC3)	*what*	sáa yìní	*everything*
yàaní/yàanú (NC8)	*who*	sáa yàanú	*everyone*
hán/hànú	*where*	sáa hànú	*everywhere*
-èné	*which*	sáa AGx-èné 'X'	*every 'X'*

The first three entries in the table are pro-forms and are therefore used in place of a lexical NP. Sáa AG-èné, on the other hand, only occurs before a noun. As with plain -èné, the agreement prefix in the sáa AGx-èné phrase may inflect for either gender or person. The following example illustrates gender agreement:

(277) **sáa**   k-**èné**   ká-'íngùwà
       or   AG1-**which** NC1-village
       every village

[Tikula, sami001.145]

This construction appears to be a calque from the Hausa *ko-wa, ko-me*, etc... (Newman 2000:622-626, Jaggar 2001:370-375).

Other quantifiers include **pɔ̂** 'all', corresponding to *duk* in Hausa, the paucal quantifier **címmài** 'little, small', and the multal quantifier **gɛ́i** (sometimes **g-gɛ́i**). The word **kù-pílù** is also often used to express quantity (cf. Hausa *yawa*) but since it can occur either as the subject of a clause or as the head of a complex NP triggering AG9 agreements (as in 279), it is better analysed as an NC9 noun.

(278) sée   mà-yɑ̂̃a   **kù-pílù**   ó↓=kó-túmó
       unless   AG4-do\RLS   **NC9-many** LOC=NC1-belly
       until it [**mò-yóo** 'beer, NC4'] *fills the stomach* [lit. 'does a lot in the stomach']

[tats002.003.029]

(279) kù-pílù   **kú**-mpà   kú-nà   tì-yɑ̂̃a-nà   pɑ̂a
       NC9-many **AG9**-this AG9-REL   1P-do\RLS-PFV   here
       the many that we are here [i.e. the large numbers of people we are]

[sayb001.634]

Part III – The noun class system of Cicipu

Chapter 5 – The Cicipu noun class system

With respect to the main research questions of this thesis, the chapters so far have been preliminary, even if they have been detailed in places. Chapters 1 and 2 (Part I) introduced the Cicipu language and set out the research context to which this thesis relates. The next two chapters (Part II) consisted of a sketch phonology and grammar of Cicipu, thus laying the foundations for the more focused study which begins here.

This chapter and the next, which together form Part III, describe the Cicipu noun class system. As was discussed in §2.1, noun class or gender systems can be approached from two angles. On the one hand, the way in which the nouns of the language are divided up ('noun classification') is of interest both to Africanists and typologists. This topic forms the subject matter for the current chapter. On the other hand, noun class systems, especially in Benue-Congo, are also interesting because of the ubiquitous agreement involved. As we saw in §2.2 this has been the subject of considerable theoretical interest in recent decades. Chapter 6 therefore describes the gender agreement system of Cicipu, covering the expression of gender on manifold agreement targets as well as various other topics of current theoretical relevance.

5.1 Overview

As demonstrated in McGill (n.d.), the Cicipu noun class system is formally very similar to those found in the Kambari languages, and so here I have followed the (arbitrary) numbering system used by Hoffmann (1963) and Crozier (1984) for Central Kambari. It should be noted that the class numbers bear no relation to the Bleek-Meinhof system used for Bantu languages (e.g. Welmers 1973:165). The following table lists the Cicipu noun classes and corresponding prefixes, together with illustrative examples of both C- and V-initial noun stems (the various pairings i.e. genders are given in §5.2). Recall that **C** and **A** represent consonant and vowel weight units respectively, the phonetic values of which are determined by the stem to which they are attached. **N** stands for a nasal homorganic with the following consonant.

Table 32: Cicipu noun classes

Class	Noun prefix	Agreement prefix	Example	
1	kA-	kA-	kà-bárá	*elder*
			kò-kɔ́ɔ	*egg*
			kò-jóo	*lizard*
			kè-téré	*bone*
			kɔ́-ɔɓí	*he-goat*
2	A-	A-	à-bárá	*elders*
			ɔ̀-kɔ́ɔ	*eggs*
			ò-jóo	*lizards*
			è-téré	*bones*
			ɔ́-ɔɓí	*he-goats*
3a	i-/y-	i-/y-	ì-námà	*meat*
			ì-ɗǎa	*ground*
			ì-dɔ̂ɔ	*horses*
			yɔ́-ɔmɔ̀	*monkeys*
3b	ri-		rì-pízì	*altar*
			rì-hyá'ɑ̀	*arrow*
			rú-usì	*rainy season*
4	mA-	mA-	mà-díyá	*hare*
			mɔ̀-tɔ́ɔ	*chick*
			mò-kóotó	*kitchen hut*
			mè-pésé	*twin*
			mà-nnú	*bird*
5	N-, mi-	N-, mi-	ǹ-díyá	*hares*
			ǹ-tɔ́ɔ	*chicks*
			ǹ-kóotó	*kitchen huts*
			m-pésé	*twins*
			mì-nnú	*birds*
6	ti-, tu-, ci-, cu-	ti-, tu-	tì-síì̀	*hair*
			tù-mócì	*friendship*
			cì-lúu	*leopard*
			cù-kúlú	*tortoise*
			cí-iyá	*flour*
7	u-/w-	u-/w-	ù-pépí	*wind*
			ù-kɔ́mɔ́	*salt*
			ù-lénjí	*sun*
			wɔ́-ɔvɔ́ɔ	*fear*
8	Ø-, C-, v-	Ø-, C-, v-	Ø-cìccérè	*star*
			Ø-híi'ò	*insect, k.o.*
			c-cɔ́'ɔ	*sheep*
			d-dɔ̂ɔ	*horse*
			z-zá	*person*
			vɔ́-ɔmɔ̀	*monkey*
9	ku-/kw-	ku-/kw-	kù-cígà	*cockerel*
			kù-cínó	*back*
			kù-míná	*cloth*
			kwé-etú	*medicine*

It can be seen from the table that the number of noun classes depends on the criteria we use to distinguish them. If the prefix borne by the noun is the deciding factor, then there

are ten (morphological) noun classes. However two of these classes (3a and 3b) share the same agreement markers, and so it can also be argued that there are nine (syntactic) noun classes. In Benue-Congo studies syntactic noun classes which share the same prefix are generally assigned the same number, but distinguished by a following letter (e.g. 3a, 3b). This is the approach that has been followed with class 3 here.

It will be clear to anyone familiar with the Benue-Congo or Bantu literature that, superficially at least, the Cicipu system is very different to both the suggested Proto-Benue-Congo (PBC) reconstructions (e.g. De Wolf 1971) and the present-day Bantu systems. There are fewer classes, and the forms of the original PBC prefixes have in some cases changed beyond recognition (see McGill n.d. for a comparison). Nevertheless, there are also striking similarities, in particular the robust and ubiquitous alliterative agreement to be discussed in chapter 6, which was no doubt the motivation for Johnston (1919) to include Cicipu's West Kainji relatives (Gurmana, Kamuku, and Basa) in his Semi-Bantu. Much the same could be said about the other Kainji languages for which we have data[1] – the prefixes and class pairings are much changed from PBC, but the mechanics of the agreement system have been retained.

The rest of the chapter is organised as follows: §5.2 lists the noun class pairings (genders) which occur in Cicipu and characterises their semantic structure (as far as is possible). The next two sections look at the derivational function of noun class prefixes – when applied to both noun stems (§5.3) and verb stems (§5.4). Section 5.5 describes the prefix allomorphs for the various classes in more detail, and finally §5.6 investigates how loanwords fit into the system.

The basics of this chapter were presented in a somewhat different form in McGill (2007). None of the statistics have been changed, and the illustrative examples are often the same, but it has been possible to significantly expand and improve the analysis.

5.2 Genders

The basic structure of the Cicipu gender system is most easily demonstrated through an affix net, as shown in Figures 30 and 31 below. As described in §2.1 the lines represent double class genders (i.e. those containing paired singular and plural nouns), with solid lines being used for well-established genders and broken lines for inquorate genders

---

1 For West Kainji, see Hoffmann (1963) on Central Kambari and Hoffmann (1967) on C'Lela, as well as the unpublished papers listed in McGill (n.d.). For East Kainji, see Anderson (1980a) on Amo.

with only a few members (Corbett 1991:170-175). The underlined prefixes represent single class genders. Figure 30 shows the *noun class* prefixes while Figure 31 shows the corresponding *agreement* prefixes – the lines are identical in both diagrams, with the only difference being the treatment of class 3 – recall from §5.1 that nouns from classes 3a and 3b trigger the same agreement prefix **yi-**, although they take part in different pairings, with 3a occurring with plural nouns, and 3b with singular nouns.

The 'alliterative' nature of the agreement is clear from the diagrams, with the two sets of prefixes either identical or very similar. Regarding the pairings of noun class (i.e. genders) it can be seen that Cicipu has a 'crossed' system (Heine 1982:197):

- Class 1, 3b, 4, 6, 7 and 9 prefixes occur with singular nouns
- Class 2, 3a and 5 prefixes occur with plural nouns
- Class 8 prefixes can occur with either singular or plural nouns

singular   **plural**

kA (1) ——————————————————————— A (2)

ri (3b) - - - - - - - - - - - - - - i (3a)

mA (4) ——————————————————————— N/mi (5)

ti/ci (6)

u (7)

C/Ø/v (8)   C/Ø/v (8)

ku (9)
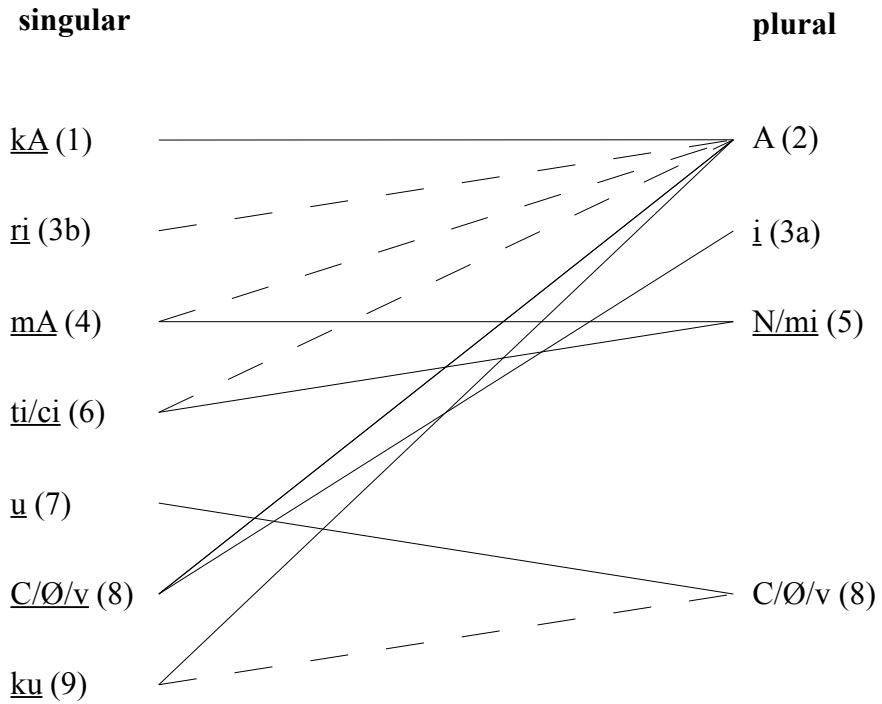
Figure 30: Cicipu genders according to noun prefixes

singular   **plural**

kA (1) ——————————————————————— hA (2)

yi (3) - - - - - - - - - - - - - - yi (3)

mA (4) ——————————————————————— mi/N (5)

ti (6)

wu (7)

C/Ø/vi (8)   C/Ø/vi (8)

ku (9)
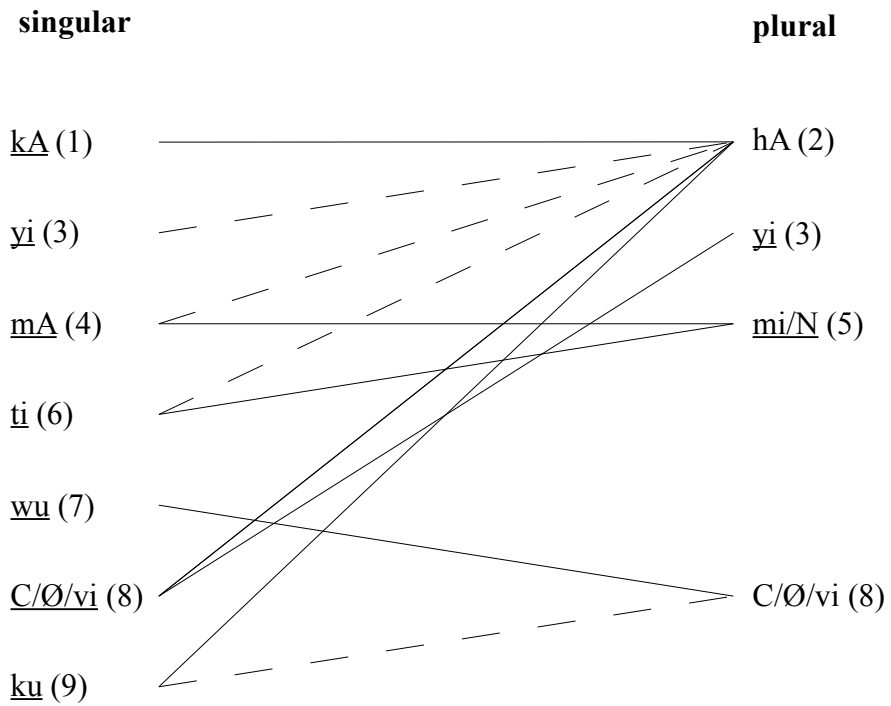
Figure 31: Cicipu genders according to agreement prefixes

In total there are eleven double class genders, seven of which are fairly well established (seventeen or more items). The other four are inquorate. As well as the double class genders, classes 1, 3a, 3b, and 4-9 make up nine single class genders[2]. One noun **à-húlá** 'name' takes a class 2 prefix and triggers class 2 agreement, regardless of whether it is being used with singular or plural meaning. Therefore it should be considered as belonging to an inquorate single class gender.

Table 33 below provides gives a rough idea of the distribution of Cicipu nouns within the lexicon according to their gender. The figures should be treated with caution, for two main reasons. The database of nouns used in this study (see §1.4) was not compiled using a statistical sampling technique, and so some genders may be under- or over-represented with respect to the lexicon as a whole[3]; in particular a high proportion of nouns in gender 7/8 have only been observed during elicitation sessions on tree names. Also the derivational function of noun class prefixes (§5.3-5.4) means that it is often difficult to judge whether the properties of a particular prefix-stem combination are sufficient idiosyncratic to warrant treating it as a separate lexical entry. Nevertheless, the table is included as some may find even these rough indicators useful. Proper nouns and identifiable loanwords have not been included, even though they make up a significant proportion of some genders (especially single class genders 1 and 8). The seven paired genders which were represented with solid lines above are in **bold**.

Table 33: Distribution of Cicipu genders (singulars down, plurals across)

Class	2 A-	3a i-	5 N-	8 C-	Single	
1 kA-	**227**				8	
2 A-					1	
3a i-					19	
3b ri-	4				7	
4 mA-	3		**103**		17	
5 N-					5	
6 ci-	2		**17**		49	
7 u-				**50**	18	
8 C-	**17**	**75**			11	
9 ku-	**64**		1		15	(Total: 713)

---

2 According to standard Bantuist terminology. There are only *eight* single class 'target genders' according to Corbett's approach, since Bantuist genders 3a and 3b would be conflated into the single target gender 3 (see §2.1.2.3). The number of double class genders is the same under either analysis.

3 Of course, the 'lexicon as a whole' is itself a very slippery notion.

The remainder of this section provides examples from each of the genders, along with some discussion of their semantic structure. In fact, the extent of the overall semantic cohesion of African noun class systems has been a matter of long-standing disagreement[4]. Hoffmann set out the two extreme views with respect to semantic cohesion in his paper on the Central Kambari noun class system (1963:169):

(a)   Noun classes are only morphological categories and are void of any meaning whatsoever.

(b)   Noun classes have an inherent meaning and nouns are put into the appropriate classes according to their meaning.

With regard to Proto-Bantu, Richardson (1967:378) was firmly in favour of the first position:

> …it is impossible to prove conclusively by any reputable methodology that nominal classification in Proto-Bantu was indeed widely based on conceptual implication. In the absence of any such proof one might equally well assume that the assignment of nominals to classes was for the most part an arbitrary grammatical device.

Similarly, concerning the much earlier Proto-Niger-Congo Williamson (1989) concluded that since the correlation between formal classes and meanings is never complete in conservative Niger-Congo systems,

> while the noun classification system arose on a semantic basis in pre-Niger-Congo, it had already become a grammaticalised, essentially formal system in proto-Niger-Congo (1989:32).

Advocates of the second position include Denny and Creider (1986), who attempted to find a single abstract meaning behind each Proto-Bantu noun class. Subsequently, others have applied techniques from Cognitive Linguistics[5] to a number of Niger-Congo languages (e.g. Hendrikse 1997, Selvik 1997, Sagna 2008). This approach seems to offer a more robust methodology than Denny and Creider's earlier work, but it is still open to the charge that the analyses are not falsifiable – a different linguist may well come up with a different analysis using the same techniques.

Contini-Morava (2002) applies a similar "semantic network" approach to Kiswahili, but her analysis also takes into account the use of class markers and concords in reference-tracking, the functional motivation most often proposed for noun

---

4   See Dingemanse (2006) for a recent review.
5   Following Lakoff's (1987) seminal analysis of the noun class system of the Australian language Dyirbal.

classification (e.g. Heath 1983, Corbett 1991:320-323). The categorisation of nouns into genders permits a parallel classification of relational elements that make reference to nouns (adjectives, demonstratives, numerals, verbs etc…), and thus restricts the set of possible referents when relational elements appear in discourse. Contini-Morava notes that this pragmatic function could be fulfilled *whether or not* there was any semantic basis to the classification[6], and she suggests that speakers construct noun class semantic networks as a language-learning strategy to ease the burden of memorisation. So "it is not necessary to believe that all…speakers construct identical networks" (2002:37), in part because "different speakers are exposed to different vocabulary items in different contexts of use" (2002:9). The distribution of various nouns across different genders, on the other hand, is shared by all speakers.

There seems little doubt that these recent approaches can be fruitful when the investigator has an in-depth knowledge of the language concerned (e.g. Sagna 2008)[7]. However I approached this study as a cultural outsider with no previous knowledge of Cicipu, and based on only nine months of fieldwork I can claim no special insight in this matter. I could draw out some Cognitive Grammar 'semantic networks', but given my present limited knowledge of Cicipu semantics, such a formal treatment would merely be 'cookbook' in nature, adding little or nothing of theoretical interest to existing studies.

With that in mind, Table 34 provides a summary of some of the semantic regularities to be found. For convenience the derivational effects of the various genders have also been listed (see §5.3-5.4 for details). Genders 1/2 and 9/2 are miscellaneous in nature, and most of the single class genders are too small for any perceived pattern to be worth mentioning. The gender with the most obvious semantic coherence is 8/2, which contains only nouns denoting humans or spiritual beings[8].

---

6 In fact Heath (1975) argues that it is actually advantageous for the system *not* to have a semantic basis, since competing referents very often share semantic properties (e.g. human, animate).
7 See also the discussion in Katamba (2003).
8 The converse is not true – nouns for humans are found in several of the other genders. Gender 8/2 is more flexible in Central Kambari, where it is not exclusively human (Crozier 1984:70).

Table 34: Semantic content of the larger Cicipu genders

Gender	Semantic content	Derivational effect
1/2	*miscellaneous*	augmentative (§5.3.2), town / settlement (§5.3.5), individuative (§5.4.5)
4/5	small things, possibly round things and hand-held things	diminutive (§5.3.1)
6/5	unpaired body parts, possibly unusual animals and things to do with farming	
7/8	trees, possibly places and long things	
8/2	humans and spirits only	agent nominalisation (§5.4.4)
8/3	birds and animals	
9/2	*miscellaneous*	
4	liquid nouns	continuous action (§5.4.3)
6	non-liquid mass nouns, collective nouns	institution / manner of (§5.3.3), language (§5.3.5)
7	abstracts	locative and de-adjectival abstracts (§5.3.4), infinitive (§5.4.1)
9	*miscellaneous*	repeated action (§5.4.2)

In the list that follows, double class genders are treated first, then single class genders.

5.2.1   Gender 1/2

(1)   Gender 1/2: **kA-/A-**
kà-búngú   *snake*   à-búngú   *snakes*
kè-téré   *bone*   è-téré   *bones*
kò-cílò   *abdomen*   ò-cílò   *abdomens*
kɔ̀-ɓɔ́ɔ   *axe*   ɔ̀-ɓɔ́ɔ   *axes*

Gender 1/2 has the largest membership by far, as is the case in the Kambari cluster generally. Historically there has been a lot of movement into this gender from the others (De Wolf 1971:62), and any semantic coherence it may have once had has disappeared. Nevertheless, 1/2 does have a derivational function as an augmentative (§5.3.2) and an individuative (§5.4.5).

5.2.2   Gender 4/5

(2)   Gender 4/5: **mA-/N-**
mà-kwá'á   *orphan*   ǹ-kwá'á   *orphans*
mà-lálá   *puppy*   ǹ-lálá   *puppies*
mɔ̀-'yɔ́'yɔ́   *fish*   ǹ-'yɔ́'yɔ́   *fish (pl.)*
mà-sí'yá̰   *nipple*   ǹ-sí'yá̰   *nipples*

The gender has a clear semantic coherence. Most obviously, it contains a large number of nouns relating to small objects. In particular words for small animals and birds, and words for young animals and humans occur in this gender. Other examples are **mé-bbèríisè** 'swift [i.e. the bird]', **mè-gémí** 'knuckle', and **mà-díyá** 'hare'.

Hand-held tools and round things may also cluster in this class. The word **mà-gɛ́i** 'sword' denotes a relatively large object, suggesting that gender 4/5 extends to hand-held implements regardless of size. Compare also **kà-gísì** 'stick' (1/2) with **mè-gísì** 'walking-stick' (4/5). Similarly **mà-kápá** 'large calabash' may be classified here because its denotatum is round rather than small. Some words for disabled humans (**mò-gútù** 'leper/cripple', **mà-gwáwá** 'mute') also belong to this class.

The borrowed words **mà-kárántáa** 'school', **má-llû** 'master', **mà-búdî** 'key', **má-gàjì** 'priest', **mà-záaní** 'pen', **mɔ̀-zɔ́ɔrí** 'cat', **mò-gúugí** 'brush' and **mà-'áná** 'meaning' all begin with *ma-* in Hausa, and these were presumably assigned to this gender because of their phonological resemblance to existing 4/5 words. Other than loanwords, it is hard to find words that do not fit semantically. **Mò-kóotó** 'kitchen' is a candidate, although a kitchen hut is usually smaller than a normal hut. Similarly **mè-ttégù** 'shirt' (from Hausa *taggo*) may have originally referred to an item of clothing shorter than the Hausa *riga*. I do not have an explanation for **mà-sítá** 'large wooden support'.

A small group of irregular nouns belong to gender 4/5 according to their agreement properties, although they are not straightforwardly composed of prefix plus root:

(3) **méngétìkká**   *young girl*   **míntìkká**   *young girls*
     méngétàarì   *young boy*   **míntàarì**   *young boys*
     mɔ́ɔ   *child*   **múu**   *children*

The word for 'child' and other kinship terms often seem to be irregular in Benue-Congo (e.g. Anderson 1980b:40).

This gender also has a derivational function as a diminutive (§5.3.1).

5.2.3 Gender 6/5

(4) Gender 6/5: **ci-/N-**
     cì-yɔ́ɔ   *yam*   **ǹ-yɔ́ɔ**   *yams*
     cì-cîi   *forehead*   **ǹ-cîi**   *foreheads*
     cì-réenè   *fireplace* **ǹ-réenè**   *fireplaces*
     cì-kɔ́'ừ   *chest*   **ǹ-kɔ́'ừ**   *chests*

This gender is not found in Kambari, and is a relatively small one in Cicipu. It contains

unpaired body parts (**cì-cî͠i** 'forehead', **cì-kɔ́'ṵ́** 'chest', **cù-kṵ̂u** 'waist', **cí-llú** 'neck'), certain miscellaneous (and perhaps unusual) animals (**cù-kúlú** 'tortoise', **cù-kúndú** 'hyena', **tì-zàarúmà** 'flea', **cì-yúu** 'porcupine', **cì-lúu** 'leopard', **tì-jímà** 'male warthog'), a few words to do with farming (**cì-níyáa** 'farm', **cì-yɔ́ɔ** 'yam', **cì-yú'ú** 'big hoe') plus three others: **cì-yímbì** 'dark', **cì'-ádì** 'trap' and **cì-réenè** 'fireplace'.

5.2.4 Gender 7/8

(5) Gender 7/8: **u-/{C-, Ø-, vi-}**

ù-lénjí	*sun/day*	**l-lénjí**	*days*
ù-pépí	*wind*	**p-pépí**	*winds*
ù-yáa	*road*	**y-yáa**	*roads*
ù-ɗángà	*tree*	**ɗ-ɗángà**	*trees*

Gender 7/8 contains several nouns denoting places (**ù-rếi** 'town', **ù-yáa** 'road', **ù-rédù** 'nest' and the word for 'place' itself **'ásù**[9]), some abstracts (**ù-jî͠i** 'expense', **ù-kôo** 'death', **ù-nếî͠** 'weight'), and some 'elemental' nouns (**wɔ́-ɔtɔ̀** 'moon', **ù-lénjí** 'sun', **ù-pépí** 'wind', **ù-láa** 'fire'). It also contains the word for 'tree' **ù-ɗángà**, along with many words for different species of trees and several other objects that are long in shape (e.g. **ù-sî͠i** 'pestle', **ù-hóiyú** 'stream', **ù-tâa** 'bow'). There are a few exceptions whose class membership does not appear to have any semantic motivation (**ù-kwấa** 'skin', **wú-ntò** 'guest hut' and **ù-kɔ́mɔ́** 'salt').

In addition to the basic place words, the class 7 prefix **u-** also has a locative derivational function (§5.3.4).

5.2.5 Gender 8/2

(6) Gender 8/2: **{C-, Ø-, vi-}/A-**

k-káa	*woman*	**à-káa**	*women*
z-zá	*person*	**à-zá**	*people*
w-wɔ́mɔ́	*chief*	**ɔ̀-wɔ́mɔ́**	*chiefs*
vá-arì	*man*	**á-arì**	*men*

Gender 8/2 has the most clearly-defined semantic structure of any of the genders, since all members denote humans or spiritual beings. It is the default gender for Hausa loanwords which denote humans (see §5.6), and is also the gender used to form agent deverbal nominalisations (§5.4.4).

---

9 The word **'ásù** is also interesting because it does not have a noun prefix. Therefore it does not display its gender overtly, but only through the agreements it triggers. This is very rare in Cicipu, and the only other nouns which could be said to have covert gender are a few semi-regular NC6 nouns (see §5.5.6).

5.2.6 Gender 8/3

(7) Gender 8/3: {C-, Ø-, vi-}/i-

d-dɔ̂ɔ	*horse*	ì-dɔ̂	*horses*
s-síró	*mane*	ì-síró	*manes*
g-gā́a	*plait*	ì-gā́a	*plaits*
vɔ́-ɔmɔ̀	*monkey*	yɔ́-ɔmɔ̀	*monkeys*

A high proportion of the nouns in this gender denote animals or birds (32 out of the 75 non-borrowed words). This is also the default gender for loanwords referring to inanimate objects e.g. **bíndígà** 'gun' from Hausa *bindiga* (see also §5.6).

5.2.7 Gender 9/2

(8) Gender 9/2: **ku-/A-**

kù-tívì	*ear*	à-tívì	*ears*
kù-béyí	*hoe (weeding)*	è-béyí	*hoes (weeding)*
kù-cínó	*back*	ò-cínó	*backs*
kù-tɔ́ɔ	*hen*	ɔ̀-tɔ́ɔ	*hens*

Similar to 1/2, there is no obvious semantic unity to this gender. **Kwā́-a'ā́** 'day/morning' [kʷā́ːʔá] has a variable irregular plural, either **ā́a'wā́** or **'wā́a'wā́** 'days/mornings' (the expected form would be *ā́a'ā́); note the labialisation *within* the stem in the plural[10]. It triggers regular 9/2 agreement.

5.2.8 Single class genders

5.2.8.1 Gender 1

(9) Gender 1: **kA-**

kà-rìmâi	*pleasure*
kò-rísìnôo	*Korisino hill*
kà-'àlbárkà	*prosperity* (from Hausa *albarka*)
kà-làahíyà	*well-being* (from Hausa *lafiya*)

This gender consists almost entirely of Hausa loans, especially ones denoting abstract concepts. It also contains a number of proper nouns for settlements (see §5.3.5).

5.2.8.2 Gender 3a

(10) Gender 3a: **i-**

ì-cícípú	*prayers*
ì-sípá	*song*
ì-sáyàtú	*comb*
ì-ɓócì	*illness*

---

10 See also the irregular 9/8 noun **kwá'à** in §5.2.9.

5.2.8.3   Gender 3b

(11) Gender 3b: **ri-**
      rù-wé'è   *valley*
      rù-'ú   *body*
      rù-húngwá   *resting*
      rú-usì   *rainy season*

5.2.8.4   Gender 4

(12) Gender 4: **mA-**
      mò-hí̋í́   *blood*
      mò-yóo   *beer*
      mà-hṵ̂u   *truth*
      mò-zóngó   *rejoicing*

Several gender 4 nouns denote liquids such as **mò-tôo** 'saliva' and **mò-ní** 'water', and this **mA-** prefix can be traced back to Proto-Niger-Congo (Williamson 1989). While some of these liquid nouns do pluralise (taking class 5 prefixes), the referent changes its nature considerably, as is the case in other languages (cf. *I like beer* vs. *I want two beers*)[11]. I have counted all such liquid nouns in this gender rather than 4/5, regardless of whether a countable plural could be elicited or not.

5.2.8.5   Gender 5

(13) Gender 5: **N-**
      ǹ-háalú   *journey*
      ǹ-kácí   *hunting*
      ǹ-lávù   *sleepiness*
      mí̋-isĩ̀   *ululation*
      ǹ-sándá   *between the legs*

This is a small gender with only five members and could be classified as inquorate.

5.2.8.6   Gender 6

(14) Gender 6: **ci-**
      cì-cínà   *charcoal*
      tì-hí'ó   *straw*
      tì-yéyí   *dirt*
      cù-kùyâa   *earwax*

Gender 6 contains non-liquid mass nouns. It also contains abstract nouns derived from verbs or from other nouns (§5.3.3, §5.3.5). At least five nouns are irregular, or possibly vowel-initial[12] roots:

---

11 Corbett (2000:81) refers to this process as 'recategorisation'.

12 Without a corresponding plural, it is not clear whether **tî̋i** and **tâa** are vowel-initial or not. For example we could analyse **tâa** as **ti + â**, in which case it would no longer be irregular, but simply the result of the normal rules for noun prefixes before vowel-initial roots. The same holds for **tî̋i**. This would be a

(15) tîi   *faeces*
 tâa   *food*
 síngí̧   *hair*
 sùpõ̂o   *falsehood*
 súmbó   *drink made from red sorrel sepals (Hausa* zobo*)*

5.2.8.7 Gender 7

(16) Gender 7: **u-**
 ù-yéyù   *cold*
 ù-kíisó   *the hereafter*
 ù-kwárí   *next year*
 wɔ́-ɔvɔ́ɔ   *fear*

Gender 7 contains mostly abstract nouns, which may have motivated the derivational use of this gender for forming de-adjectival abstract nouns (§5.3.4).

5.2.8.8 Gender 8

(17) Gender 8: {**C-, Ø-, vi-**}
 l-lámá   *sound*
 c-cá'á   *harvest*
 Ø-kúngwá   *God*
 Ø-rúhũ̀   *last year*

Most gender 8 nouns are Hausa loans (e.g. **Ø-bɔ́kɔ̀i** 'week' from *bakwai*, **Ø-mènkéní** 'cocoyam' from *mankani*). The remaining nouns do not seem to have any semantic coherence.

5.2.8.9 Gender 9

(18) Gender 9: **ku-**
 kù-hɨ́ivì   *breath*
 kù-yũ̀yũ̂   *sand*
 kù-sṹu   *smell*
 kù-pílù   *many*

As with the double class gender 9/2, there is no apparent semantic coherence here.

5.2.9 Inquorate genders

The following examples illustrate the genders with less than five members:

(19) Gender 3/2: **ri-/A-**
 rì-hyã́'ã̀   *arrow*   à-hyã́'ã̀   *arrows*
 rì-pízì   *altar*   è-pízì   *altars*
 rù-pṍɔ   *granary*   ɔ̀-pṍɔ   *granaries*

---

more appealing analysis if there were candidates for vowel-initial monosyllabic roots in other noun classes as well.

(20)　Gender 4/2: **mA-/A-**
　　　mɔ̀-rìgìdɔ́　　*navel*　　　　　ɔ̀-rìgìdɔ́　　*navels*
　　　mɔ̀-ríngínɔ́　　*spine*　　　　　ɔ̀-ríngínɔ́　　*spines*

(21)　Gender 6/2: **ci-/A-**
　　　cì-kɔ́otɔ̀　　*drum*　　　　　ɔ̀-kɔ́otɔ̀　　*drums*

(22)　Gender 9/8: **ku-/{C-, Ø-, vi-}**
　　　kw-á'à　　*house*　　　　á"wâa *or* 'wá"wâa　*houses*

(23)　Gender 2: **A-**
　　　à-húlá　　*name/names*

5.3　*Derivational use of noun class markers*

As pointed out by Mufwene (1980), the role of noun class prefixes in Bantu is more than simply indicating grammatical gender and number. They also have a 'derivational' use, similar to suffixes such as *-ment* and *-ness* in English. The same is true for Cicipu, as will be clear from this section and the next.

In this section I will consider the application of various class prefixes to *noun* stems. The resulting semantic effects have all been observed in other African languages; Heine (1982:199) commented that "Augmentative, Diminutive, Human, and Place are among the most productive derivative genders in 'nature-based' languages". Heine also observed that "even within a given language genders tend to differ considerably with regard to derivative productivity" (1982:198). This is also true for Cicipu.

Section 5.4 is concerned with DEVERBAL nominalisations, which are formed by the application of a number of different class prefixes to a *verb* stem.

5.3.1　Diminutive (gender 4/5)

Nouns belonging to gender 4/5 usually denote small things (§5.2.2). In addition, nouns typically occurring in some other gender may be given 4/5 prefixes if the speaker wants to suggest a smaller object than the hearer would otherwise expect. For example **y-yénè** 'catfish' (8/3) is a generic term covering catfish of all sizes, but **mè-yénè** (4/5) denotes a small catfish, and it would be inappropriate to use this word to refer to any other kind. This process is very productive, allowing speakers to readily create novel words. Often the diminutive prefix occurs as a pre-prefix (§5.3.6).

The semantic relation involved in this kind of derivation is sometimes non-compositional and unpredictable, as the following examples show:

253

(24)

mɔ̀-ccɔ̀kɔ́	*pocket*	**kɔ̀-ccɔ̀kɔ́** (1/2)	*bag*
mà-sɨ́'yá	*nipple*	**kà-sɨ́'yá** (1/2)	*breast*
mɔ̀-tɔ́ɔ	*chick*	**kù-tɔ́ɔ** (9/2)	*hen*
mɔ̀-zúvù	*pinkie*	**kɔ̀-zúvù** (1/2)	*finger*

For example **mɔ̀-tɔ́ɔ** 'chick, 4/5' means more than just an undersized **kù-tɔ́ɔ** 'hen, 9/2'. However the 4/5 noun almost always has connotations of 'smallness' in some way, perhaps only with respect to its counterpart.

It should be stressed that there many noun stems which inherently 'belong' to gender 4/5 and do not occur with other prefixes, and so gender 4/5 is not a "minor target gender" in the sense of Corbett (1991:160). Indeed Cicipu has no minor target genders at all – the various derivational functions discussed in this section and the next are all accomplished by pressing into service the lexical genders already introduced in §5.2.

5.3.2   Augmentative (gender 1/2)

Despite its overall lack of semantic coherence, gender 1/2 has a derivational function as an augmentative. The resulting nouns show a similar variety with respect to semantic compositionality as the diminutive just discussed.

(25)

kè-yénè	*large catfish*	**y-yénè** (8/3)	*catfish*
kà-ttílú	*burial pot (large)*	**mà-ttílú** (4/5)	*pot*
kà-ɗángà	*large piece of wood*	**ù-ɗángà** (7/8)	*tree* **or** *piece of wood*
kà-nnú	*hawk*	**mà-nnú** (4/5)	*bird*

This does not seem to be as productive as the diminutive, and I have only observed a few cases, usually involving a stem normally occurring in gender 4/5. As with the diminutives, augmentatives may also be formed with pre-prefixes (§5.3.6); in fact for augmentatives this may be more usual than simply replacing the prefix.

5.3.3   Abstract quality (gender 6)

The abstract nouns found in this gender tend to be derived either from verbs, or from other nouns denoting humans. This is in contrast to abstract nouns from gender 7 (§5.3.4), which do not usually have counterparts in other genders.

(26)

cì-'ítàní	*marriage*	**'ita**	*marry*
cì-pátí	*request*	**pata**	*beg/plead*
tì-zɔ́'ɔ	*laughter*	**zɔɔsɔ**	*laugh*
tì-wɔ́mɔ́	*chieftaincy*	**w-wɔ́mɔ́**	*chief* (8/2)
tù-mɔ́cì	*friendship*	**m-mɔ́cì**	*friend* (8/2)
tì-yírì	*witchcraft*	**y-yírì**	*witch* (8/2)
tí-gàjì	*priesthood*	**má-gàjì**	*priest* (4/5)

Often the derived gender 6 noun can be given the gloss 'institution of___' or 'manner/behaviour typical of___' (cf. Mufwene 1980:250-251). This derivation is relatively productive and several words I tried to coin were readily accepted e.g. **tì-bɔ́wɔ́** 'thievery' from **b-bɔ́wɔ́** 'thief' (8/2). Abstract Hausa loanwords may also enter this class e.g. **tì-búkáncí** 'the practice of native medicine' from *bokanci*, **tì-túréncí** 'English' from *turanci*[13].

All language names belong to gender 6 (see §5.3.5).

The association with 'manner' also comes out in some specialised uses of gender 6 agreement morphology (see §6.3).

Finally, there is also a derivational relationship between several of the 'collective'[14] mass nouns in gender 6 and their individuated counterparts in gender 1/2. For example **cì-cínà** (6) is 'charcoal', but **kà-cínà/à-cínà** (1/2) means 'lump/s of charcoal'. Similarly **tì-hí'ó** (6) is 'straw' but **kò-hí'ó/ò-hí'ó** (1/2) means 'stem/s of grass'.

5.3.4   Locative and de-adjectival abstract nouns (gender 7)

In addition to the basic place words, the class 7 prefix **u-** may sometimes be prefixed to a noun root to indicate a location, as seen in the paradigms below. This pattern does not appear to be very productive (although see §5.3.5 on ethnic groups and their locations).

(27)   **kà-táarí**   *stone* (1)
         à-táarí   *stones* (2)
         mà-táarí   *pebble* (4)
         ǹ-táarí   *pebbles* (5)
         t-táarí   *bead* (8)
         ì-táarí   *beads* (3)
  ⇒ **ù-táarí**   *stony place* (7)

(28)   **k-kíisò**   *spirit* (8)
         ò-kíisò   *spirits* (2)
  ⇒ **ù-kíisò**   *the hereafter* (7)

The NC7 prefix can also be attached to a motion verb to form a nominal with a locative meaning (§5.4.1).

Mufwene (1980:248-249) mentions the possibility in Bantu of deriving nouns by adding a noun class prefix to an adjective. Although there are not many adjectives in Cicipu (§4.7), a derivational relationship exists between at least three of them and corresponding nouns in class 7. These are **ù-pêe** 'bigness' ~ **pénêu** 'big', **ù-rùmó**

---

13 The *-ci* suffix in Hausa has a similar semantic effect as the gender 6 nominalisation being discussed.
14 See Corbett (2000:117-119) on collective nouns.

'darkness' ~ **rùmónò** 'dark', and **ù-sílá** 'redness' ~ **sílánà** 'red'.

5.3.5 Ethnic groups

For ethnic groups, languages, and settlements there is a paradigm involving classes 1, 2, 6, 7, and 8, shown below. The gaps in the table are unattested.

Table 35: Ethnic group paradigm

Class	Acipu	Karishen	Kadonho	Maburya	Hausa	Gloss
8	c-cípù	Ø-rísìnô	d-dípó	k-kúlà	k-kɔ́gɔ́	*person (sg.)*
2	à-cípù	ò-rísìnô	ò-dípó	à-kúlà	ɔ̀-kɔ́gɔ́	*person (pl.)*
6	cì-cípù	tì-rísìnô	tì-dípó	tì-kúlà	tì-kɔ́gɔ́	*language/dialect*
1		kò-rísìnô	kò-dípó			*name of settlement*
7				ù-kúlà		*name of settlement*

Karishen, Kadonho, and Maburya are the Hausa names of three of the Cicipu dialect centres (§1.2.3); the same paradigm applies to the other dialects as well[15].

Note that **tì-kɔ́gɔ́** can mean either the Hausa language or religion (i.e. Islam), or perhaps just 'in the Hausa manner' – similar to the suffix *-ish* in English. Welmers (1973:452) makes a similar point for the equivalent class in Kiswahili, *ki-*.

5.3.6 Pre-prefixes

Mufwene (1980:253) distinguishes between primary derivations, such as the ones that have been illustrated so far involving a single prefix, and secondary derivations, which involve the addition of a 'pre-prefix' to an already delimited noun. In Cicipu pre-prefixes seem to be restricted to augmentatives (29) and diminutives (30).

(29) ká-k-kà-híilí   kó-cì-kóotò   ké-kù-zè'ũ̀
      NC1-?-NC1-field   NC1-NC6-drum   NC1-NC9-dance
      big field   *big drum*   *a lot of dancing*
         [samoh001.144]   [saat001.008.073]   [svtmg001.322]

(30) má-k-kà-nàbáyì   má-k-kà-síngí   mé-'-ì-rí-mpà-nì
      NC4-?-NC1-story   NC4-?-NC1-hair   NC4-?-NC3-thing-this-NMLZ
      small story   *little hair*   *this little thing*
         [svtmg001.129]   [saat002.002.160]   [saat002.002.315]

We have now seen that diminutives and augmentatives can be formed either by replacing the normal prefix (§5.3.1-5.3.2) or by adding a pre-prefix[16]. The non-

---

15 Five of the seven dialect centres are formed with the NC1 prefix. The exceptions are **ù-kúlà** and **ù-kúmbàsi**.

16 Or both, in fact, for example **ké-k-kè-yénéyénè** denotes a catfish large enough to swallow a human. Note the inner prefix with the standard low tone, and then the pre-prefix with characteristic high tone

compositionality found in cases of prefix replacement has *not* been observed with pre-prefixes – for example adding the ɴᴄ1 prefix to **ù-ɗángà** 'tree' leads to a straightforward augmentative **kà-ˀù-ɗángà** 'large tree', but replacing the existing prefix to form **kà-ɗángà** results in the related, but unpredictable meaning 'large piece of wood'.

When it comes to agreement, the behaviour of nouns with pre-prefixes is perhaps a little surprising – as discussed in §6.5, agreement may be with either the inner or outer prefix.

5.4   Deverbal nominalisation

In the previous section we looked at examples of the derivational effect of noun class prefixes when applied to noun stems. In this section we consider their effect when applied to verbs. There are five productive processes involving different prefixes (§5.4.1-5.4.5), but also a few less productive uses of prefixes (§5.4.6). Judging by Crozier's (1984:112-117) discussion of Central Kambari, in Cicipu there seems to be greater meaning associated with the individual class prefixes.

5.4.1   Infinitives and locatives (gender 7)

The most productive deverbal nominalisation is the infinitive/gerund/participle formed by prefixing the class 7 prefix **u-** to a verbal stem. It seems that any verb can be so modified, and the process is perhaps more inflectional than derivational. The tone on the resulting noun is either LHL or less commonly LHH (i.e. with L on the prefix), and is apparently lexically-determined but otherwise arbitrary.

(31) **ù-dámà**   *speaking*   **dama**   *speak*
      ù-húllò   *blowing*   **hullo**   *blow*
      ù-pídǎ   *licking*   **piɗa**   *lick*

These words have properties of both verbs and nouns, which of course justifies the label 'participle'. In (32) the participle **ù-dɔ́vɔ́** 'pounding' functions as the head noun of an associative construction, and also triggers agreement on the relativiser **wú-nà**.

(32) **ù-dɔ́vɔ́**   **wú↓**=í-dá'ũ̀   **wú**-nà kù-mócì   kù-yó-nò   á↓=ú-yǎa
     ɴᴄ7-**pound**   ᴀɢ7=ɴᴄ3-g_corn   ᴀɢ7-ʀᴇʟ ɴᴄ9-old_woman   ᴀɢ9-be\ʀʟs-ᴘғᴠ   ʟᴏᴄ=ɴᴄ7-do
     the pounding of guineacorn that the old woman is doing
                                                          [2008-03-30.001]

On the other hand, in (33) the same participle occurs with a pluractional verbal

---

and consonant-lengthening.

extension <il> and a bare object NP **ì-dá'ṳ̀** 'guineacorn' filling the PATIENT semantic role of the nominalised verb **dɔvɔ** 'grind'. As well as the pluractional, the applicative, resultative, and ventive suffixes have also all been observed occurring in participles.

(33) ò-yô   á↓ = u-dɔ́v<íl>ɔ́   ì-dá'ṳ̀
    3P-be\RLS   LOC=NC7-pound<PLAC>   NC3-guineacorn
    they are repeatedly pounding guineacorn

                                                              [2008-03-30.001]

Further evidence of the participle's verbal nature is shown by (34), where it is modified by an adverb.

(34) tá-abà   tì-sì-yṳ̰́u   **ú-kùngwà**   **mázámázá**
    NC6-tobacco   AG6-HAB-put   **NC7-age[verb]**   **rapidly**
    *tobacco causes **rapid ageing***

                                                              [tats007.002.081]

These participles in Cicipu have a restricted distribution compared to prototypical verb forms – they cannot head main clauses, nor do they occur as dependent 'sentence adverbial' clauses. Instead they can only occur as the complements of main verbs. Particular subordinating verbs (§4.6.7) seem to consistently select either this form of the verb, or a finite verb form i.e. there is no choice. The subject NP must be omitted, and the (logical) subject of the participle must be the same as that of the main clause. For these reasons it seems appropriate to refer to this NC7 participle as an infinitive (cf. the Bantu class 15 *ku-* infinitive – e.g. Visser 1989).

    Since the word forms under consideration show properties of both verbs and nouns, it is worth asking whether they form an independent lexical category. Concerning the gerund in English, Kroeger (2004:45-50) argues against a 'mixed' category analysis. Instead the gerund is ambiguous, with individual uses being classifiable as either clearly nominal (i.e. a deverbal nominalisation) or clearly verbal (i.e. a true gerund). This does not seem to be the case for the Cicipu participle. Examples (35-36) were readily accepted in elicitation sessions. In (35) the participle takes an object like a verb, but also functions as the head of a relative clause like a noun.

(35) **ù-dɔ́vɔ́**   **ì-dá'ṳ̀**   **wú-nà** kù-mócì   kù-yó-nò   á = ú-yãa
    NC7-pound **NC3-guineacorn** **AG7-REL** NC9-old_woman   AG9-be\RLS-PFV   LOC=NC7-do
    ***the pounding guineacorn that** the old woman is doing*

                                                              [2008-03-30.001]

Similarly in (36) the participle again takes an object, but itself functions as the fronted

object of the verb **ỹãa** 'do', as well as triggering AG7 gender agreement on the accompanying copula.

(36) **ù-dɔ́vɔ́**   **ì-dá'ʉ̌**   **w-ì**   ù-yó-nò   á=ú-yə̃̀a
    NC7-pound   NC3-guineacorn   AG7-COP   3s-be\RLS-PFV   LOC=NC7-do
    it's pounding guineacorn *she is doing*

                          [2008-03-30.001]

These elicited examples therefore suggest that the participle should indeed be analysed as a single mixed category, rather than ambiguous between noun and verb – certainly it appears to offer better evidence for this than the English gerund. Unfortunately the relevant constructions are absent from the corpus[17], so it offers no evidence either way.

Even if the participle in the associative construction of (32) is of the same category as the one in the participle + object construction in (33), there is still a semantic difference between the two constructions. The former seems to be more general in meaning, and there is a wide variety of ways in which the activity encoded by the participle can be modified. In the case of a transitive verb the 'possessor NP' in the associative construction can bear the PATIENT (as in (32)), AGENT (37), or INSTRUMENT (38) role. The participle + object construction, on the other hand, is naturally limited to the same semantic roles as straightforward objects.

(37) ù-yíndá   wú-ttù
    NC7-see   AG7-1P.POSS
    our seeing

                          [tats002.007.024]

(38) ù-yáahùkwà-nù   wú↓=u-nóo
    NC7-forgive-RES   AG7=NC7-mouth
    verbal forgiveness [i.e. not of the heart]

                          [oamy001.115]

One further difference applies in cases where the following NP corresponds to the object of the finite construction. If the NP is definite then it is far more likely to appear in the participle + object construction than the associative construction – in fact the latter has only been observed with indefinite NPs. The contrast can be seen in the following pair of examples:

---

17 i.e. a participle based on a transitive verb, with an expressed object and a nominal modifier or other evidence of AG7 agreement morphology. It is not really surprising that constructions of this complexity are missing from the corpus.

(39) ń ǹ=ù-hwáarà ù-sɔ́ɔ wú↓=tá-abà
    if when=3s-start\RLS NC7-drink AG7=NC6-tobacco
    if he starts smoking tobacco

[tats007.002.013]

(40) ń ǹ=ù-hwáarà ù-sɔ́ɔ tá-abà tí-llè
    if when=3s-start\RLS NC7-drink NC6-tobacco AG6-that
    if he starts smoking that tobacco

[tats007.002.079]

Some NC7 deverbal nominalisations have quite different meanings, as shown below – for example **ù-káa** 'message' denotes the THEME of the verb **kaanu** 'send', and **ù-tə̂a** 'bow' is the INSTRUMENT of the verb **tə̃a** 'shoot'. These should be regarded as homophonous with the corresponding infinitives rather than vague.

(41) **ù-kôo**   *death*   **koo**   *die*
    ù-káa   *message*   **kaanu** *send*
    ù-kúrè *error*   **kure**   *err (from Hausa* kuskure)
    ù-tə̂a   *bow*   **tə̃a**   *shoot*

NC7 deverbal nominalisations are not always infinitives – they can also denote locations. Recall from §5.3.4 that locative nouns can be formed using the NC7 prefix. This prefix can also be attached to a motion verb to form a nominal denoting the GOAL of the movement. Examples include **ù-kúmbù** 'up' from **kumba** 'climb' and **ù-pásù** 'across' from **pasa** 'cross' (note the change in the final vowel).

(42) ù-pásù   wú-mpà   [**pasa** = 'cross']
    NC7-across AG7-this
    this side [of the river]

[Tikula, sami001.435]

Finally, it should be noted that NC7 deverbal nominalisations may also take the stative **-ni** nominalisation suffix (§4.4.6.1).

5.4.2   Repeated action (gender 9)

The NC9 prefix **ku-** can be applied to a large number of verbs, perhaps the majority, and the resulting nominal denotes a repetitive action. For some verbs the final vowel may change to [u]. In some cases this change seems to be obligatory, as with **ze'ẽ** 'dance' in (43), in other cases such as **kollo** 'look at' it is optional. If the original vowel was nasal then the change is to a nasal **ũ**.

(43) ò-kóndò **kù-zé'ṹ**   [**ze'ẽ** = dance]
3P-enter\RLS NC**9-dance**
they started dancing

<div align="right">[sahs001.002.009]</div>

In the corpus, NC9 nominalisations very often involve verbs conducive to iteration, as in (44). Stative verbs (e.g. **cuwo** 'be full'), on the other hand, do not seem to allow the nominalisation at all.

(44) **kù-bɔ́lù**   *searching*   **bɔlɔ**   *search*
      kù-'ízè'ṹ   *washing* (i.e. the action)   **'ize'ẽ**   *wash*
      kù-yímà   *forging*   **yima**   *forge*
      kù-zízá'ṹ   *shivering*   **ziza'a**   *shiver*

Like the NC7 infinitive discussed above, these NC9 nominalisations have properties of both nouns and verbs. They trigger AG9 agreement (45) and can function as the head of an associative construction (46), but can also occur with verbal extensions (47).

(45) kù-cíɗù kw-ì   [**ciɗa** = pluck]
NC9-pluck AG9-COP
a lot of plucking [e.g. of fruit]

<div align="right">[2008-03-14.002]</div>

(46) ù-yô   kù-cíndà kú-ttù   [**cinda** = wait]
3S-be\RLS NC9-wait AG9-1P.POSS
he's looking after us [lit. 'he is our waiting']

<div align="right">[2008-03-14.002]</div>

(47) kù-ɗís<íl>ú   [**ɗisa** = spill]
NC9-spill<PLAC>
drizzle

<div align="right">[eamy003.1336]</div>

Perhaps the most common way of expressing repetitive action is through a construction consisting of the verb **yãa** 'do', followed by **kúu-** and then the verb denoting the repeated action, as illustrated below.

(48) à-yẫa   kú↓=u-bámbálá
3P-do\RLS AG9=NC7-fumble
they fumbled

<div align="right">[saat001.003.031]</div>

(49) ùdángànù   **ùyã̀a** **kúubɔ́dì** vì   ǹ   mòtôo
      ù-dángànù   **ú-yã́a** **kú↓=u-bɔ́dɔ̀**=vì ǹ   mò-tôo
      3S-keep_doing\RLS NC**7-do** AG9=NC**7-moisten**=3S.PRO with NC4-saliva
he kept on moistening him with saliva

<div align="right">[saat001.008.096]</div>

(50)   à-yâa   kú↓=u-sírìbà   cì-kóotò
3P-do\RLS   AG9=NC7-play_frenziedly   NC6-drum
they played the drum frenziedly [lit. 'they did playing']

The long vowel and the fact that the tone pattern on [**kúu-** + verb] is identical to that of the 'back end' of the associative construction (§4.4.5.1) suggests the breakdown provided in the glosses to (48-50), that of an AG9 associative agreement marker followed by the NC7 infinitive, familiar from §5.4.1. This would explain the presence of the bare object NP in (49-50) – it is an infinitival object just as in (40) above. However it also raises the question as to the identity of the missing 'front end' of the associative construction. One speculative candidate is **kù-pílù** 'much/many'.

5.4.3   Continuous action (gender 4)

NC4 deverbal nominalisations are similar to the NC9 ones just discussed, in that they both require there to be multiple occurrences of the activity concerned. Nonetheless they are quite distinct, and while NC9 nominalisations *highlight* (or to use a term from Langacker 1987, 'profile') the iterative nature of the event, NC4 nominalisations seem to 'smooth over' the individual multiple occurrences, focusing instead on the homogeneous[18] state of affairs which arises – hence the term 'continuous action nominalisations'. For instance, the NC4 nominalisation in (51) below could not be used to refer to a single instance of kicking, but it could denote someone's taking part in a game of football. To refer to a repeated kicking motion, the NC9 nominalisation would be more appropriate. If one did want to use this construction to refer to a simplex event, for example while watching a slow-motion replay of a player kicking a ball, then the NC7 prefix **u-** (§5.4.1) would have to be used instead.

(51)   ù-yô   á=**má**-gávà
3S-be\RLS   LOC=**NC4**-kick
he is playing football [lit. 'he is at kicking']

It should be noted that not all NC4 deverbal nominalisations are of this kind. If the initial consonant of the verb is lengthened the resulting nominal denotes a diminutive physical object. These are the diminutives of the 'individuative' nominalisations (§5.4.5).

---

18 It is noteworthy that NC4 prefixes also occur with liquid mass nouns (§5.2.8.4), which have this same property of homogeneity.

(52) mɔ̀-k-kɔ́'ɔ̀ *a ladle*   kɔ'ɔ   *scoop*
      mɔ̀-g-gɔ́dɔ́ *small lump, pimple*   gɔdɔ   *make into ball*

5.4.4 Agent (gender 8/2)

Agent nominalisations are formed by adding prefixes from the human gender 8/2 to a verb stem, a pattern found in other Benue-Congo languages e.g. Akɔɔ́sē (Hedinger 1980:5) and Esimbi (Stallcup 1980:151). In Cicipu the tones on the resulting word are all H and the final vowel is nasalised, although the exact form is subject to variation – the final vowel of all verbs can be either (i) diphthongised, (ii) lengthened, or (iii) followed by an **m**. So for example the verb **pata** 'request/beg' can be nominalised as either **pátáu**, **pátáa**, or **pátám**, all meaning 'beggar'. Further examples of the nominalisation are given below. These nominalisations are apparently fully productive, although they are actually rather rare in the corpus.

(53) **sé'ém**   *carver*   **se'ẽ**   *carve*
      bɔ́wɔ́m   *thief*   **bɔwɔ**   *steal*
      pád̃ám   *butcher*   **pada**   *slaughter*
      pínám   *barber*   **pina**   *shave*
      mátám   *parent*   **mata**   *give birth*
      lápám   *wise person*   **lapa**   *know*
      kúmbám   *rider*   **kumba**   *ride*
      dãnám   *writer*   **dana**   *write*
      húnám   *killer*   **huna**   *kill*
      gúyám   *one with ability*   **guya**   *can*
      wáyám   *someone coming*   **waya**   *come*
      dúkwám   *someone going*   **dukwa**   *go*
      ráwám[19] *eater*   **raa**   *eat*

The plurals are usually regular NC2 plurals, e.g. **à-mátáu** 'parents', but there is at least one irregular plural: **túmám/à-támám** 'farmer/farmers'. It is not known whether any verbal extensions can be included in the nominalisation, nor whether this nominalisation (or an 8/3 equivalent) occurs with inanimate actors.

   In some cases there is a contrast with another morphologically-related word denoting an agent. For example, as well as the regular form **bɔ́wɔ́m** 'thief' there is also a word **bɔ́wɔ́**. The semantic difference seems to be that the former is someone whose life is characterised by thieving, whereas the latter can be applied to somebody who has only committed a single act of theft.

   A couple of other agent nominalisations have been observed, but it is not known how productive they are: they are **sɔ́ɔmú** 'drunkard' from **sɔɔ** 'drink', and **gúyáyé** from

---

19 **Ráam** was not accepted.

guya 'can', apparently meaning 'human being'.

5.4.5 Individuative (gender 1/2)

Gender 1/2 nominalisations usually denote a physical object, often a product arising from the action. Sometimes the first consonant of the verb stem is lengthened, just like the pre-prefixes discussed in §5.3.6, and the final vowel may change to **i**. It does not seem possible to predict the tone on the nominal.

(54) **ká-g-gìtà**   *a snapped off piece*   **gita**   *snap*
　　　kɔ́-k-kɔ̀'ì   *a scoop* (i.e. a measure) **kɔ'ɔ**   *scoop*
　　　kɔ̀-k-kɔ́ɗì   *a cut piece*   **kɔɗɔ**   *cut*
　　　ká-ɗìsíi   *a spill*   **ɗisa**   *spill*
　　　kɔ́-kùdɔ́   *a meeting*   **kudɔ**   *meet* (i.e. gather together)
　　　kó-kùntònû   *a fist*   **kuntonu** *make a fist*

Other semantic roles are possible including AGENT (**kó-g-gòoní** 'helper' from **goonu** 'help') and LOCATIVE (**kɔ̀-t-tɔ́pú** 'hole/pit' from **tɔpu** 'put inside'). See also §4.4.6.2 for the locative nominaliser **-tu**, which very often results in a noun with 1/2 gender.

5.4.6 Miscellaneous

Deverbal nominals belonging to classes 1, 3a, 3b, 5, and 6 do occur, but they are less productive and their semantic properties are not yet understood. Various stem changes may also occur, in particular a change in the final vowel to **i** or **u**.

Table 36: Miscellaneous deverbal nominalisations

Class	Nominalisation		Verb	
NC1	kà-ráa	*eating*	raa	*eat*
	ké-ttépì	*talking*	tepe	*talk*
NC3a	ì-túmà	*farming*	tuma	*cultivate*
	ì-sípá	*song/songs*	sipa	*sing*
	ì-háttú	*yawning*	hatta	*yawn*
	ì-póntú	*clapping*	pɔntɔ	*clap*
	ì-'óɔrí	*tiredness*	'ɔɔ	*tire*
	ì-pándárí	*forgetfulness*	panda	*forget*
NC3b	rù-húngwá	*resting*	hungwa	*rest*
NC5	ǹ-háalú	*journey*	haala	*walk*
	ǹ-lávù	*sleep (n.)*	latta	*sleep (v.)*
NC6	cì-ɓáyá	*discipline (n.)*	ɓaya	*discipline (v.)*
	tì-sôɔ	*drink (n.)*	sɔɔ	*drink (v.)*
	cì-kàllá	*clearing (i.e. location)*	kalla	*clear*
	cì-dámí	*whisper*	dama	*speak/speak against*
	cì-párì	*stalking*	para	*stalk*
	cì-tái	*shooting*	tãa	*shoot*
	cì-kání	*cutting*	kana	*cut*
	cì-dópì	*seam,* or *weaving*	dopo	*sew*

5.5 Prefixes and allomorphs

Cicipu nouns follow the prototypical Benue-Congo pattern of noun class prefix plus
stem (§4.4). In this section I will discuss the phonological form of the noun prefixes.
General properties are dealt with first, followed by a discussion of the class-specific
allomorphs.

5.5.1 Vowel-initial roots

Vowel-initial roots (§3.1.1) can be found in each class:

Table 37: Noun class prefixes occurring with vowel-initial roots

1	[ká:dándá]	*thorn* (kA- + ádándá)
2	[á:dándá]	*thorns* (A- + ádándá)
3a	[yó:mɔ̀]	*monkeys* (i- + ómɔ̀)
3b	[rú:má]	*war* (ri- + úmá)
4	[má:bà]	*anger* (mA- + ábà)
5	[mĩ́:sĩ̀]	*crying* (mi- + ísĩ̀)
6	[cé:dũ̀]	*top of the head* (ci- + édũ̀)
7	[wó:vó:]	*fear* (u- + óvɔ́ɔ
8	[vó:mɔ̀]	*monkey* (vi- + ómɔ̀)
9	[kʷé:tú]	*medicine* (ku- + étú)

265

The prefix vowel quality is usually subsumed by that of the root vowel, apart from class 9 where the labialised [kʷ] retains a trace of the /u/ prefix vowel, and class 3b, where some words have been found with optional palatalisation – 'rainy season' can be either [rúːsì] or [rʲúːsì].

5.5.2   Vowel harmony

Unlike the north-western West Kainji languages (cf. Hoffmann 1967 for C'Lela), where prefixes often just consist of a consonant followed by a short transitional schwa, Cicipu prefixes follow the Kambari/Kamuku pattern whereby the prefixes always have phonetically full (but still short) vowels. Nouns from classes 1, 2, and 4 are formed by adding to the root the harmonising prefixes **kA-**, **A-**, and **mA-** respectively. The quality of the vowel in these classes is determined by the rules which were set out in the discussion on vowel harmony in §3.5:

- If the root contains /**e**/, then the prefix vowel will be [e]
- If the root contains /**o**/, then the prefix vowel will be [o]
- If the root contains /**ɔ**/, then the prefix vowel will be [ɔ]
- Otherwise the prefix vowel will be [a]

While these rules predict the vowel quality of the prefix in the majority of cases, there is an additional factor involved when it comes to roots which only have close vowels. Roots with only [i] or only [u] sometimes occur with prefixes containing the mid vowels [e] and [o] respectively, rather than the expected [a]. This can be viewed as an assimilatory process, with the underlying prefix vowel /**a**/ raising in the environment of a high root vowel, but there does not seem to be any way to predict whether or not this process will occur.

(55)   **kè-bìmbîi**   *buzzing insect, k.o.*
       mè-gísì   *walking stick*
       kè-bíkí   *celebration (from Hausa* biki*)*

       kà-yíví   *cold meal (tuwo)*
       ká-dìsíi   *spot*
       kà-gísì   *stick*

(56) kó-cìyû   *heap*
      kò-ɗûu   *heart*
      kò-lúu   *knee*

      kà-búngú   *snake*
      kà-gúutù   *buttock*
      kà-núu   *head of corn*

There remain some nouns whose prefixes cannot be accounted for either by harmonisation or by assimilation. The words in (57) should have the prefixes **ma-/ka-** according to the former, and **me-/ke-** according to the latter. The roots in all such exceptions have contained only /**i**/ vowels, and the prefix vowel has always been /**o**/.

(57) mò-hírì   *blood*
      mò-síi   *shame*
      mò-ní   *water*
      kò-zíní   *ghost*

Finally, the word **mà-déi** 'calf (of leg)' may be an exception to the otherwise universal application of vowel harmony, since it seems to have a consistent [a] vowel in the prefix. It may be that the usual rules do not apply to diphthongs – this is the only attested example of a monosyllabic root involving /**ei**/ or /**eu**/. Words containing the diphthongs /**ai**/ and /**au**/ also take prefixes with [a] vowels, but as discussed in §3.5 this could just be the default value of the prefix.

5.5.3 Tone

Concerning the tone on noun class and agreement prefixes, Welmers (1973:179) wrote:

> For the most part, in a given [Bantu] language, the noun and concord prefixes all have the same tone, far more commonly low than high. Subject concords may have high tone in some verbal constructions, but this is a function of the verbal construction, not of the concords themselves. In some languages, however, certain concords for classes 1 and 9 have low tone, while all other concords have high tone.

The situation in Cicipu is similar to Welmers' first group of Bantu languages, in that the tone on the noun class prefix is generally low and does not vary according to class. Instead the tone on the noun class prefix is usually predictable from the structure of the noun root. It can be derived as follows:

- If the noun root is vowel-initial, then the prefix vowel merges with the first root vowel to form a long vowel, as illustrated in §5.5.1. The prefix tone is almost always high (**kàa'ɛ́a** 'rat' is the only known exception).

(58) [ká:dándá]   *thorn* (**kA-** + **ádándá**)
     [kɔ́:ɓí]   *he-goat* (**kA-** + **ɔ́ɓí**)
     [kó:cì]   *hole* (**ku-** + **ócì**)
     [kà:'ɛ̃́:]   *rat* (**kA-** + **à'ɛ́a**)

- If the noun root contains only low tones then the prefix is high.

(59) **kó-ggòmbò**   *bat*
     mɔ́-kkɔ̀ɗɔ̀   *rumour-mongering*
     ká-dàngì   *testicle*
     wú-ntò/ví-ntò *guest hut/guest huts*

- Otherwise the prefix is low.

(60) **kà-bárá**   *elder*
     kò-dóndó *garden*
     ù-kúwò   *baobab tree*

Nouns with a L H root tone pattern are an exception to the above rules, and these may have either high (61) or low (62) prefixes. There does not seem to be any way to predict which.

(61) H - L H   **ká-kkàcɛ̃́i**   *middle*
                ká-ɗìsíi   *spot*
                kɔ́-kùdɔ́ɔ   *meeting*

(62) L - L H   **kɔ̀-ɓìyɔ́ɔ**   *pumpkin*
                kà-kàanáa   *crab*
                mà-cìijíi   *calabash*

5.5.4 Classes 3a and 3b

Class 3a and class 3b nouns both trigger **yi-** agreement prefixes. However they have been divided into two morphological noun classes, according to whether the noun takes an **i-** (3a) or a **ri-** (3b) prefix. There is no difference between the two in terms of agreement, but there is with respect to gender pairings. Class 3a nouns are either members of the single class gender (3a), or plurals of class 8 nouns (8/3a). Class 3b nouns likewise have their own single class gender (3b), but they also pair up with class 2 plurals (3b/2). The distribution is summarised below:

Table 38: Distribution of class 3 prefixes according to gender

Prefix	Singular	Plural	Single class gender
3a (**i-**)	✘	✓ (8/3a)	✓
3b (**ri-**)	✓ (3b/2)	✘	✓

The history of these two classes can be illuminated by considering comparative data from Central Kambari, which has an interesting counterpart to class 3b. There are at least two Central Kambari nouns which, although they trigger AG8 agreement, have an unexpected *li-* prefix (De Wolf 1971). One of these nouns (**lì-yâ** 'arrow') has a cognate **rì-hyá'ằ** in Cicipu class 3b. Furthermore, the *li-* nouns pair with class 2 plurals, just like Cicipu class 3b. De Wolf views these two Central Kambari roots as a remnant of the Proto-Benue-Congo gender **li-/*a-* (corresponding to Proto-Bantu 5/6), which includes words for 'arrow', 'body', and 'war', glosses which can all be found in Cicipu 3b/2 (although the roots themselves are not obviously cognate). Class 3a, on the other hand, is most likely a reflex of Proto-Benue-Congo **i-* (corresponding to Proto-Bantu 10). Thus despite their syntactic equivalence, classes 3a and 3b probably have different ancestor classes.

5.5.5 Class 5

The NC5 prefix has an interesting set of allomorphs, just as in Central Kambari (Hoffmann 1963):

- **N**[20]- before short consonants: **ǹ-kántú** *knives*
- **mi-** before long consonants and vowels: **mì-nnú** *birds*
  mí-ĩsĩ̀ [mí:sĩ̀] *crying*

There are in fact two ways of forming NC5 nouns from roots with long initial consonants. One method involves adding an epenthetic vowel to the **N-** allomorph to form **mi-**:

(63) **mà-nnú** *bird*   **mì-nnú** *birds*
    má-llû *master*   **mí-llû** *masters*

The other method, which is apparently not found in Central Kambari, is to reduplicate the first syllable of the root before adding the **N-** prefix. We will come back to these two methods in §5.5.8.

---

20 **N-** represents a nasal homorganic with the following consonant. Sometimes this surfaces as a nasalised vowel [ĩ] or [ũ] – see §3.1.7 for details.

5.5.6   Class 6

The NC6 prefix takes a variety of forms: **ci-**, **cu-**, **ti-**, or **tu-**. The variation between **i** and **u** is in this case largely predictable (see McGill 2007:68-70 for discussion). The **c/t** alternation is more complex. Of the 19 paired nouns (from genders 6/5 and 6/2), 17 begin with [c], and only two (**tì-zàarúmà** 'flea' and **tì-jímà** 'male warthog') with [t]. This can be treated as an idiosyncratic property to be handled in the lexicon. The unpaired gender 6 nouns are divided roughly equally between [c] and [t], with no discernible phonological pattern to the distribution. The reason for this variation is unknown, and if it were not for the phonetic similarity between the two phones[21], I would have no hesitation in setting up morphological subclasses 6a and 6b, just as has been done for 3a and 3b.

5.5.7   Class 8

The NC8 prefix has a most interesting set of variants, the distribution of which is as follows:

- If the root is consonant-initial, then C_1 may simply be lengthened[22], for example:

(64) **z-zá**   *person*   **à-zá**   *persons*   8/2
  k-káa   *woman*   **à-káa**   *women*   8/2
  d-dɔ̂ɔ   *horse*   **ì-dɔ̂ɔ**   *horse*   8/3
  c-cɔ́'ɔ̀   *sheep (sg.)*   **ì-cɔ́'ɔ̀**   *sheep (pl.)*   8/3

- In other cases, C_1 does not lengthen and so there is no prefix:

(65) **Ø-cìccérè**   *star*   **ì-cìcérè**   *stars*   8/3
  Ø-móotò   *car*   **ì-móotò**   *cars*   8/3
  Ø-híi'ɔ̀o   *ant, k.o.*   **ì-híi'ɔ̀o**   *ants, k.o.*   8/3

- If the root is vowel-initial, then **vi-** is added[23]. As with all vowel-initial roots, the prefix vowel coalesces with V_1 giving a long, high-tone vowel with the same quality as V_1:

---

21 Certain words even show 'free' variation between the two phones, such as **ciyo/tiyo** 'get', although this is not the case generally. Concerning the related language Central Kambari, David Crozier (p.c.) has remarked that the NC6 prefix consonant may be either [c] or [ts]. This time, however, the alternation is principled, with [c] occurring before roots with a front V_1, and [ts] otherwise. Interestingly this alternation does not apply more generally in the language, only to the NC6 prefix, and /c/ and /ts/ are found in contrast just as /c/ and /t/ are in Cicipu.

22 I have not found a long C_1 in any class 8 *roots*.

23 Since the prefix vowel is always replaced in the coalescence, it is not possible to be sure that the underlying value is /i/. However the corresponding AG8 prefix is **vi-**, and in all other classes the prefix vowels are the same in their respective noun and agreement prefixes.

(66) vi- + ɔ́mɔ̀ → vɔ́ɔmɔ̀   *monkey*   8/3
      i- + ɔ́mɔ̀ → yɔ́ɔmɔ̀   *monkeys*

The same possibilities exist for the NC8 plurals of nouns in gender 7/8 such as
ù-lénjí~l-lénjí 'day~days' and wɔ́-ɔtɔ̀~vɔ́-ɔtɔ̀ 'moon~moons'.

The distinction between the lengthening prefix and no prefix at all is not always
clear-cut, especially at the beginning of the phonological word. Certain nouns always
seem to start with a long consonant (e.g. k-káa 'woman'), but others are more variable,
even within the speech of an individual. There seems to be a correlation between short
words and the occurrence of the C- prefix, and a tentative hypothesis might be that (a)
there is a trimoraic minimality constraint for noun words, and (b) syllable-initial
geminates are moraic. Thus if an NC8 noun root is short (i.e. bimoraic), then a
lengthened prefix would be necessary to satisfy the minimality constraint[24].

There are three constructions in which the consonant is almost always lengthened,
even for words such as those in (65) which otherwise have a null prefix. These are after
the associative construction proclitic (§4.4.5.1), after the co-ordinating proclitic nì
(§4.4.5.4), and after the locative proclitic A (§4.5). It seems then that the distinction
between the C- and Ø- sets of NC8 nouns illustrated in (64) and (65) becomes blurred in
a (phonological) word-medial environment. In these positions every consonant has been
observed lengthened.

Illustrative examples of nouns with the C- allomorph are shown in Table 39. Long
voiceless plosives are not easily detectable utterance-initially[25]; in these cases the
examples given are utterance-medial.

---

24 The high frequency semi-grammaticalised z-zá 'person' would be an exception.

25 Possibly there is a phonetic difference in the following vowel which would betray its presence, such
  as an increase in amplitude (Ladefoged and Maddieson 1996:94). This has not yet been investigated
  for Cicipu.

Table 39: Examples of NC8 consonant-lengthening prefix

b	b-bíndù	*tadpole*	eamy043.004
ɓ[26]			
c	c-cɨ́i	*python*	eamy040.002.021
d	d-díyò	*grasscutter*	saat002.004.050
ɗ	ɗ-ɗángà	*trees*	tats001.002.010
g	g-gálù	*sides*	eaim003.1391
gw	g-gwándà	*pawpaw*	tapf001.005.038
h	h-hóiyú	*streams*	eamd024.1292
hy			
hw	nì = h-hwí'ì	*and a Lelna person*	saim001.029
j	j-jévé	*an antelope*	eamy032.012
k	k-kúwò	*baobabs*	sayb001.040
kw	nì = k-kwɛ́a	*and skins*	eamy032.035
l	l-lénjí	*days*	eamd002.002
m	m-mácì	*friend*	eamd015.380
n	n-náa	*cow*	saim001.058
p	p-píyá	*guineafowl*	eamd002.135
r	r-rée	*towns*	eamd002.095
s	s-sé'ɨ́	*co-wife*	eamd015.364
t	t-tɨ́i	*container*	tats001.001.007
v	v-vɔɔcíyɔ̀	*slimes*	eamy003.1304
w	w-wɔ́mɔ́	*chief*	saim001.108
y	y-yúmù	*iron*	eamy040.001.008
z	z-zá	*person*	eamd002.063
'	'-'írì	*kind* (from Hausa *iri*)	taff002.053
'w			
'y			

For Central Kambari Hoffmann (1963) set up two subclasses, 8a for nouns with the **C-**/**vi-** prefix, and 8b for those that have no prefix. Given that this alternation is often not predictable in Cicipu, as well as being neutralised word-medially, I have decided not to split the class in this description.

---

26 There are no attested NC8 nouns whose stems begin with **ɓ**, **hy**, **'w**, or **'y**. These are all fairly rare phonemes so this should not be surprising.

5.5.8 Epenthesis and reduplication

A number of Cicipu words form their plurals by duplicating the first consonant of the root together with the normal method of changing the prefix. These words have so far only been found in genders 1/2 and 4/5, and their roots always begin with a long consonant. This is significant because roots beginning with a long consonants are relatively rare in Cicipu. Some examples are given below:

Table 40: Reduplicated plurals with root-initial long consonants

Singular	Plural	Gloss	Gender
kɔ̀-ddɔ́ɔ	ɔ̀-dí-ddɔ́ɔ	*edge*	1/2
kà-ɗɗáyí	à-ɗí-ɗɗáyí	*tear [i.e. a rip]*	1/2
kà-hhṹucí	à-hṹ-hhṹucì	*cloud*	1/2
ká-kkàcĩ́i	á-ká-kkàcĩ́i[27]	*middle*	1/2
kà-llávù	à-lì-llávù	*dream*	1/2
mɔ̀-ccɔ̀kɔ́	ǹ-cí-ccɔ́kɔ́	*bag, small*	4/5
mɔ̀-ggɔ́dɔ́	ǹ-gú-ggɔ́dɔ́	*bump*	4/5
mɔ̀-kkɔ́'ɔ̀	ǹ-kú-kkɔ́'ɔ̀	*ladle*	4/5
mɔ̀-ppɔ̃́ɔ	m̀-pí-ppɔ̃́ɔ	*granary, small*	4/5
mè-ttégù	ǹ-tí-ttégù	*shirt*	4/5
mɔ́-kkwètè	ń-kù-kwètè	*gecko*	4/5

An epenthetic short (usually high-tone) [i] or [u] is inserted between the first (short) consonant and the original (long) consonant. For 1/2 nouns with initial long consonants the reduplication is usually optional, and there does not seem to be any semantic difference between the two forms. Nouns from gender 4/5, on the other hand, only have the reduplicated possibility in the plural[28].

The majority of 4/5 nouns form their plurals by reduplication, while just a few use the prefix **mi-**: recall from §5.5.5 that this is an allomorph of the NC5 prefix **N-** applying to roots beginning with a long consonant. These two processes of reduplication and epenthesis can be seen as means to the same end – the avoidance of consonant clusters with three timing units[29]. The effect of reduplication is to split the **N-** prefix from the long consonant and place it to the left of the shortened reduplicant instead, resulting in a consonant cluster of only two timing units. Epenthesis in NC5 plurals avoids the consonant cluster altogether, so that we have, for example, **mì-llú** 'necks' instead of ***ǹ-llú**.

There is an interesting parallel to the **N-**/**mi-** alternation from outside the noun

---

27 Unusually the reduplicated vowel in **á-ká-kkàcĩ́i** is [a], rather than [i] or [u].

28 I have come across one exception: **mé-bbèríisè** 'swift', **ń-bèríisè** 'swifts'.

29 For a similar phenomenon in the Austronesian language Leti see Hume et al. (1997:375).

class system. The conjunction **Ǹ-** 'and/with' (§4.4.5.4) has a variant **nì-,** and the two allomorphs are distributed in just the same way as the NC5 prefixes: **nì-** occurs before long consonants and vowels, and **Ǹ-** before short consonants[30].

For completeness' sake two quite different kinds of reduplication should be mentioned. The first involves the plurals of nouns from gender 9/2:

(67)   **kù-ɓáa**   *marsh*   **á-ɓàaɓà**   *marshes*
      kù-sáa   *mountain*   **á-sàasà**   *mountains*
      kù-sɔ́ɔ   *lake*   **ɔ́-sɔ̀sɔ̀**   *lakes*
      kù-tâa   *debt*   **á-tàatà**   *debt*

In this construction the prefix is high tone instead of low and there is no association with long consonants. Another difference is that with 9/2 reduplication the root syllable is duplicated, not just the first consonant. It does not seem possible to predict which 9/2 nouns will undergo reduplication, although all the examples so far have been monosyllabic. For those that do reduplicate in the plural, the process is obligatory.

The second kind involves the plurals of nouns from gender 7/8, although this time the reduplication is optional[31].

(68)   **ù-rée**   *town*   **r-réerèe**   *or* **r-rée**   *towns*
      ù-yáa   *road*   **y-yáayàa**   *or* **y-yáa**   *roads*
      ù-hóiyú   *stream*   **Ø-hóihóiyú**   *or* **h-hóiyú**   *streams*

5.6   Loanwords

Loanwords can be found in most genders, although they cluster in genders 8, 8/3, and 1/2. Precise distributional percentages are impossible to provide, mainly due to the difficulty of distinguishing between borrowing and code-switching. As mentioned in chapter 1, virtually all Cicipu speakers are fluent in Hausa. The figures below give a rough picture of the distribution, but like the numbers in §5.2, they should be treated with caution.

---

30 Cicipu does not generally allow syllables with codas, and so it has not been possible to find any further parallel situations that might give rise to a consonant cluster.

31 This reduplication process is a potential source of the NC8 lengthening prefix. However it should be observed that NC8 prefixes are usually exponents of *singular* nouns. See McGill (n.d.) for further discussion.

Table 41: Proportion of loanwords in genders

Gender	Prefix	Total	Loans	Percentage borrowed
1/2	kA/A	282	56	20
4/5	mA/mi	115	12	10
8/3a	vi/i	118	43	36
9/2	ku/A	64	0	0
7/8	u/vi	51	1	2
8/2	vi/A	21	4	19
6/5	ti/mi	17	0	0
1	kA	15	7	47
3a	i	20	1	5
3b	ri	7	0	0
4	mA	21	3	14
5	mi	5	0	0
6	ti	55	6	11
7	u	18	0	0
8	vi	44	33	75
9	ku	15	0	0
2	A	1	0	0
3b/2	ri/A	4	0	0
4/2	mA/A	2	0	0
6/2	ti/A	1	0	0
Total		876	166	19

Leaving aside genders 1 and 1/2 for the moment, genders 8, 8/2 and 8/3a are the most over-subscribed. It is probably not a coincidence that the noun prefix for class 8 is either null or a **C-** allomorph; Hausa words can 'fit in' to class 8 without looking out of place[32]. Some examples are:

(69) **róobà** (8)   *plastic* (roba)
      kàràatúu (8)   *reading* (karatu)
      z-zénì (8)   *cloth* (zane)

      káatáakó (8/3)   *plank* (katako)
      mìsáalì (8/3)   *example* (misali)
      d-dágá (8/3)   *belt* (daga)

Although it is presumably the **Ø** allomorph that provides the strongest attraction to class 8, loanwords also occur with the **C-** allomorph, as can be seen in (69). Other examples include **h-hùuhú** 'lungs' (*huhu*), **c-cáawá** 'noise' (*tsawa*) and even calqued idioms such as **k-kái ví-↓ń-háalú** 'on a journey' (*kan tafiya*, lit. 'head of a journey').

    Hausa words may also be assigned to other Cicipu noun classes if they start with a syllable resembling a noun class prefix, as shown below. The three examples referring to towns illustrate that the assignment is purely formal and not semantic[33].

---

32 See Corbett (1991:72-73) for discussion with regard to loans in Kiswahili.
33 This reinterpretation seems to be productive, although I have not yet investigated gender assignment

(70) **mà-kárántáa** (4/5)   *school* (makaranta)
      má-llû (4/5)   *master* (mallami)
      mà-'ánáa (4/5)   *meaning* (ma'ana)
      mà-búɗî (4/5)   *key* (mabuɗi)

      tá-abà (6)   *tobacco* (taba)

      Màkúukù mé (4)   *it's Makuku town*
      Kónòo ké (1)   *it's Kano town*
      Dirindaji vì (8)   *it's Dirindaji town*

It seems that in Cicipu borrowed words are not assigned gender on the basis of their meaning. This contrasts with the situation found in ṵt-Ma'in from the Northwest branch of West Kainji, where loanwords are frequently assigned gender because of their semantic properties (Smith 2007:37-66). One reason for this difference may be that the null-prefix class in ṵt-Ma'in is restricted to humans, and is therefore unavailable as a general destination for formal assignment.

There remains a large number of loanwords in genders 1 and 1/2, the presence of which cannot be explained by either formal or semantic assignment rules (recall that this gender has no coherent semantic structure).

(71) **kà-làahíyà** (1)   *peace* (lafiya)
      kà-'àlbárkà (1)   *prosperity* (albarka)
      kà-hákùrì (1)   *patience* (hakuri)
      ká-húusì (1)   *anger* (haushi)

      kà-ɓáuná (1/2)   *buffalo* (ɓauna)
      kò-wàndó (1/2)   *pair of trousers* (wando)
      kà-káasùwà (1/2) *market* (kasuwa)
      kà-húskà (1/2)   *face* (fuska)

There are at least two ways to account for this. First, it may be that as the largest gender in Cicipu, 1/2 (and by extension 1) functions as a default, if the formal assignment rules evident in (69) and (70) fail, for whatever reason, to come into effect[34].

An alternative and perhaps preferable explanation is that the existence of borrowed words in 1/2 is due to a general gravitation of nouns towards that gender. This trend was observed for Central Kambari by De Wolf (1971), and the same pattern is found in Cicipu. It may be that these loanwords started their lives in gender 8/3 as we would expect, but over time they have shifted to gender 1/2 along with other Cicipu words

---

for towns unknown to the Acipu.

34 The idea of a default gender has been criticised by Corbett (1991:77-80), since linguists may propose one even when native assignment rules already account for the data. According to Corbett's (1991:43-49) analysis of Bantu-type systems i.e. gender is assigned according to the noun class pairing of the noun, we should say that there is a *default noun class pairing* 1/2 for loanwords, which in turn means they are assigned 1/2 gender in the normal way.

(perhaps by means of some of the derivational processes discussed in §5.3 and §5.4). The nouns **kɔ̀-ccɔ̀kɔ́** 'bag' (from *jaka*), **kà-ttílú** 'pot' (*tulu*) and **kɔ̀-ttúdù** 'ridge' (*tudu*) may be evidence of this shift – they all have root-initial long consonants which suggests an earlier stage of their existence in 8/3[35].

If the loanwords in 1/2 have moved there from 8/3, then we would expect to find more established items in 1/2 and newer borrowings in 8/3, and this is indeed the case. Words for modern objects such as **móotò** 'car', **róobà** 'rubber'[36], and **kácâa** 'bicycle chain' all belong to 8/3, while words denoting universal or more traditional concepts are more often found in 1/2.

---

35 See McGill (n.d.) for the suggestion that gender movement out of 8/2 or 8/3 is responsible for most root-initial long consonants in nouns, whether borrowed or not.

36 All English loanwords must have come through Hausa – as mentioned in chapter 1, hardly any Acipu speak English.

Chapter 6 – Agreement

In the previous chapter I presented the various noun classes and genders (pairings) of Cicipu, together with the derivational functions of the noun class prefixes. In this chapter we will be concerned with the gender agreement system.

Although complex enough from the perspective of many other language families, gender agreement in Cicipu is less 'messy' than in many Bantu or Grassfields Bantu languages. There is no distinction between primary and secondary concord (Welmers 1973:172-175), and the system is very close to true 'alliterative' agreement, where the prefixes found on the various agreement targets are identical both to each other and to the corresponding noun prefix. Thus there is no need here for the daunting tables of gender agreement typical of Bantu descriptions.

The chapter is arranged as follows: §6.1 is concerned with the form of the agreement prefixes, especially the AG5 and AG8 allomorphs. Section 6.2 provides a comprehensive listing of all gender agreement targets, both inside and outside the NP. Section 6.3 deals with two related phenomena: antecedentless agreement and semantic agreement; in both cases the prefix on the agreement target is determined by something other than syntactic properties of the controller. Section 6.4 investigates neutral agreement, in other words agreement triggered by atypical controllers which do not have their own gender. Section 6.5 deals with complex NPs which offer a choice of agreement feature values on the target, and finally §6.6 briefly mentions gender resolution.

As with chapter 5, this chapter is based on material previously published in McGill (2007). Once more the basic data is largely unchanged, but the analysis is expanded and improved.

6.1   Agreement prefix allomorphs

6.1.1   Vowel harmony

Recall the vowel harmony shown by the NC1, NC2, and NC4 prefixes (§5.5.2). The AG1, AG2, and AG4 prefixes behave in an identical way, with the value of the prefix vowel being determined by the vowels of the agreement target to which it is prefixed. This is demonstrated below for the AG4 prefix on three different demonstratives.

(1) (a) mà-díyá **má**-mpà
NC4-hare AG4-this

this hare [near speaker]

(b) mà-díyá **mé**-llè
NC4-hare AG4-that

that hare [near hearer]

(c) mà-díyá **mó**-"û̂
NC4-hare AG4-yonder

that hare [very far off]

[eamd032.058]

6.1.2 AG5 allomorphs

The AG5 agreement prefix has two allomorphs:

- **mi-** before stems beginning with a vowel[1] or a long consonant, and sometimes before stems beginning with a short consonant
- **N-** sometimes before stems beginning with a short consonant

The same allomorphs are found for the NC5 prefix (§5.5.5), but the distribution is slightly different. In the case of the NC5 prefix, **mi-** and **N-** are in complementary distribution, but here there is a choice before short consonants, as can be seen from the extract in (2) below. The speaker uses the **ǹ-** allomorph as the subject prefix on **'ugo** 'fall' in the first intonation unit, but switches to **mi-** for the second occurrence of the verb.

(2) **m-úu**   mí-nà mí↓=ɗ-ɗángà m-áyà   ǹ-'úgò /
NC5-child AG5-ART AG5=NC8-tree   AG5-come\RLS AG5-fall\RLS

ánà **mì**-'úgò-nò...
when AG5-fall\RLS-PFV

then the fruit fell / when it fell...

[tapf002.004.024]

Similarly, the next two examples show the two allomorphs before the adjective **ɗénêu** 'small'.

(3) m-úu   **mì**-ɗéné **mì**-ɗénêu
NC5-child AG5-small   AG5-small

young animals [lit. 'small children']

[samoh001.091]

---

1 The usual vowel coalescence occurs (see §3.1.1). See §5.5.7 for why, before vowel-initial stems, the prefix vowel is assumed to be **i** even though it is always subsumed by the stem vowel.

(4) m-úu   ǹ-dénêu
     NC5-child   AG5-small
     small children

All targets beginning with short consonants seem to offer both choices, although **mi-** is the more common.

6.1.3 AG8 allomorphs

The distribution of the AG8 allomorphs is perhaps even more complex than that of the corresponding NC8 allomorphs (§5.5.7). Recall that the NC8 noun prefix has three allomorphs: either **Ø-** or **C-** before short consonants, and **vi-** before a vowel (NC8 roots do not seem to begin with long consonants). In the case of the agreement prefix, the same three allomorphs occur, but this time **vi-** has a wider distribution. Not only does it occur before vowels and all long consonants, it may also be found before short consonants. Thus before short consonants all three allomorphs are possible, and the main problem here is to describe the conditions under which each of them surfaces in this environment.

First of all, it should be admitted that it is frequently hard to decide whether a particular instance of a prefix is **Ø-** or **C-**. Textual analysis is complicated by the fact that in fast speech, the contrast between long and short consonants is often unclear. Nevertheless speakers have a metalinguistic awareness of these two possibilities, just as for the noun prefix, and this means it is possible to have fruitful elicitation sessions on the topic. The lengthening option (**C-**) seems to be considered to be the 'best' style of Cicipu, but I could find no evidence of a diachronic shift from **C-** to **Ø-**.

Apart from stylistic factors, the choice of prefix is influenced both by the type of agreement target, and by the semantics of the noun triggering the agreement. For most targets all three possibilities **vi-**, **Ø-** and **C-** can occur, as shown below for numerals:

(5) (a) mŭ-uwà   'o'   **n**-nósì
          1s-hear\RLS [sound] AG8-four

     (b) mŭ-uwà   'o'   **Ø**-nósì
          1s-hear\RLS [sound] AG8-four

     (c) mŭ-uwà   'o'   **vì**-nósì
          1s-hear\RLS [sound] AG8-four
          I heard four 'o's

Agreement on numerals is in fact optional, so it is not clear whether (5b) shows a null prefix or a lack of agreement altogether. However there is less doubt about the associative construction (6); since agreement is obligatory for every other noun class, there is no reason to analyse it as optional for AG8, especially given the independent existence of the **C-/Ø-** noun prefix alternation.

(6)   (a)   z-zá   **k=**kú-yímà
             NC8-person   AG8=NC9-smith
             blacksmith [lit. 'person of smithing']

       (b)   z-zá   **Ø=**kà-gàskíyà
             NC8-person   AG8=NC1-truth
             truthful person [lit. 'person of truth']

       (c)   l-lámá   **ví↓=**má-wáa
             NC8-sound   AG8=NC4-dog
             the sound of a dog
                                   [eamd016.399, tats005.001.127, tats002.001.068]

The article/relativiser **-nà** on the other hand, allows only **C-** (7a) or **vi-** (7b), and cannot occur without an overt prefix (7c):

(7)   (a)   'b'   **n**-nà   Ø-líbà-nà   kù-gíyà
             [letter]   AG8-REL   AG8-lack\RLS-PFV   NC9-hook
             'b' without a hook [i.e. not ɓ]

       (b)   'b'   **ví**-nà   Ø-líbà-nà   kù-gíyà
             [letter]   AG8-REL   AG8-lack\RLS-PFV   NC9-hook
             'b' without a hook [i.e. not ɓ]

       (c)   *'b'   **Ø**-nà   Ø-líbà-nà   kù-gíyà
             [letter]   AG8-REL   AG8-lack\RLS-PFV   NC9-hook
                                         [2007-01-10.010]

The distribution of **n-nà** and **ví-nà** in the corpus is not random, and they are in more or less complementary distribution. **N-nà** almost always occurs immediately after lexical heads (80 tokens, one exception), while **ví-nà** mostly (12 tokens, one exception) occurs as a headless relative, as in (8). This situation has a parallel in Hausa, where headless relatives require the longer relativiser *wanda,* and headed relative clauses are more likely to have the short form *da* (Newman 2000:534-536).

(8) ví-nà Ø-sí-zìnò   vó-ɔmɔ̀,   sée ú-cìyò-nò
    AG8-REL AG8-HAB-change NC8-monkey then 3S-collect\IRR-VENT
the one [k-káa 'woman, NC8'] *who changed into a monkey, she collected*

[saat001.004.014]

A similar distributional pattern is found with subject agreement, although with a different pair of allomorphs. For subject agreement there is no evidence for **C-** – only **vi-** (9a) or **Ø-** (9b) can occur. **Vi-** is unrestricted in distribution, but **Ø-** may only occur utterance-medially.

(9) (a) Ø-hìtílà   vì-ríɓà
        NC8-light   AG8-sink\RLS
        the light faded

    (b) Ø-hìtílà   Ø-ríɓà
        NC8-light   AG8-sink\RLS
        the light faded

[eaim007.1490]

Table 42 summarises what is known about the distribution of the AG8 allomorphs with respect to target type (recall that a choice is only possible for targets whose stems begin with a single short consonant).

Table 42: Possible AG8 allomorphs according to target type

Target	Section	C-	Ø-	vi-
Numerals	§6.2.2	✓	✓	✓
Adjectives	§6.2.3	✓	✓	✓
Negative copula	§6.2.6	✓	✓	✓
Associative construction	§6.2.8	✓	✓	✓
Article/relativiser	§6.2.10	✓	✗	✓
Subject marker	§6.2.13	✗	✓	✓
Demonstrative adverbs	§6.2.7	?	?	✓
nínnîi 'only'	§6.2.4	?	?	?
wh-words	§6.2.5	?	?	?

As well as the target type, the animacy of the noun triggering the agreement is also a factor in determining the form of the prefix. There is a correlation between nouns denoting humans and **C-** or **Ø-**, and between those denoting animals and inanimates and **vi-**. This has been investigated more fully for subject prefixes and numerals, but the correlation is believed to hold for the other target types which allow a choice as well.

Nouns for which **C-** or **Ø-** are preferred tend to have human referents: **z-zá** 'person', **vá-arì** 'man', **m-mácì** 'friend'. Nouns for which the **vi-** agreement prefix is preferred are mostly animals or inanimates, such as **c-cɔ́'ɔ** 'sheep', **vɔ́-ɔmɔ̀** 'monkey', and **Ø-rìkódà** 'audio recorder'.

   K-káa 'woman' seems to be somewhere in between. The following doublet, from a text about dogs, consists of two consecutive intonation units with a striking parallel between the two different AG8 allomorphs used in the second associative construction. In (10) the noun **k-káa** 'woman' triggers the **vi-** AG8 allomorph, while in (11) **z-zá** 'person/man' triggers **C-**[2].

(10) ù-lápà   kù-ɗóo   kú↓=k-káa   ví↓=kw-á'à   kú-llè
      3s-know\RLS NC9-voice   AG9=**NC8-woman**   AG8=NC9-house AG9-that
      *it knows the voice of the **woman** of that house*

                                                    [tats001.001.081]

(11) ù-lápà   kù-ɗóo   kú↓=z-zá   k↓=kw-á'à   kú-llè
      3s-know\RLS NC9-voice   AG9=**NC8-person** AG8=NC9-house AG9-that
      *it knows the voice of the **man** of that house*

                                                    [tats001.001.082]

NC8 agreement is a complex topic, and given the variation found even for individual nouns, as with the noun class prefixes it does not seem appropriate to set up separate agreement subclasses to deal with the separate prefixes.

6.2 *Agreement targets*

There is a very large number of agreement targets within the Cicipu noun phrase, and predicates also agree in gender with the subject. This agreement takes the form of prefixes or proclitics attached to the target. Other than for class 3 (§5.5.4), we generally find a one-to-one correspondence between noun class prefixes and agreement prefixes, although targets with vowel-initial stems are under-differentiating (i.e. having fewer than expected distinctions) due to the phonological similarity between classes 4 and 5. Overall, however, the agreement system is highly regular, with only one set of agreement markers shared by all targets. One general complication is that the tones of the various targets' prefixes vary (see below). In addition the agreement prefix vowel for class 6 may be either **ti-** or **tu-**, as was the case for the NC6 noun prefix (§5.5.6). The allomorphs of the AG5 and AG8 agreement prefixes have already been mentioned (§6.1).

---

2   For more on the effect of natural gender on variation in agreement see §8.4.5.3.

In the subsequent discussion on specific targets, all the prefixes follow the basic pattern set out below.

Table 43: Agreement prefixes found before consonant-initial and vowel-initial stems[3]

Shape	Before	1	2	3	4	5	6	7	8	9
CV-	C-initial stem	kA-	hA-	yi-	mA-	mi-/N-	ti-/tu-	wu-	vi-/C-/Ø-	ku-
C-	V-initial stem	k-	h-	y-	m-	m-	t-	w-	v-	kw-

6.2.1 Prefix tone

The various agreement targets, summarised below, can be divided into two main groups according to the tone on their agreement prefixes. Unlike some Bantu languages (§5.5.3), there are no class-dependent effects. Adjectives, numerals, the quantifier **-nínîi** 'only', wh-words, the negative copula, and the demonstrative adverbs all take low-tone prefixes. The associative construction, possessive pronouns (with the exception of 1PS), demonstrative modifiers, the article/relativiser, and the modifier **-mbɔ̂** 'another' all take high-tone prefixes.

Subject prefixes can appear with either H or L tone, according to the verb's mood (see §4.6.2).

There is no tone associated with the prefix on pronouns or the copula, since they are vowel-initial and take the C- series of prefixes.

---

3 Anderson's (1980a) description of the noun class system of the East Kainji language Amo refers to 'fused' vs. 'non-fused' prefixes. While the C- prefixes in Cicipu are indeed fused with the roots, the resultant forms are entirely regular (with the exception of the length of the fused vowel, which can be short or long), and so there is no need for the abstract representations and complex derivations proposed by Anderson for Amo.

Table 44: Agreement targets

Target	Prefix can be omitted?	Prefix tone	Section	
Numerals	Y, frequently	L	§6.2.2	
Adjectives	N	L	§6.2.3	
-nínnîi 'only'	Y	L	§6.2.4	**Set 1**
wh-words	N	L	§6.2.5	
Negative copula	N	L	§6.2.6	
Demonstrative adverbs	N	L	§6.2.7	
Associative construction	N	H	§6.2.8	
Possessive pronouns	Y, but rarely	H	§6.2.9	
Article/relativiser	N	H	§6.2.10	**Set 2**
Demonstrative modifiers	N	H	§6.2.11	
-mbɔ̃ 'another'	N	H	§6.2.12	
Subject marker	N	H or L	§6.2.13	
Copula	N	None (C-)	§6.2.14	
Pronouns	N	None (C-)	§6.2.15	

These target groupings are essentially the same as in Central Kambari (Crozier 1984:65-67), although the actual tones associated with the groups are different. In Central Kambari the tone on the prefix of set 1 targets is polar (as opposed to L in Cicipu), and for set 2 targets it is L (H in Cicipu).

6.2.2 Numerals

Attributive numeral phrases (§4.8.2) are formed by the singular noun followed by a numeral with an optional low-tone agreement prefix, which may occur up to but not beyond the number 9[4]. The following example shows each numeral with AG1 agreement:

(12)	No.			
	1	**kà-bárá**	**kò-tò**	*one old man*
	2	**kà-bárá**	**kà-yápù**	*two old men*
	3	**kà-bárá**	**kà-tàatù**	*three old men*
	4	**kà-bárá**	**kò-nósì**	*four old men*
	5	**kà-bárá**	**kà-tâu**	*five old men*
	6	**kà-bárá**	**kà-tórìhî**	*six old men*
	7	**kà-bárá**	**kà-tíndàyà**	*seven old men*
	8	**kà-bárá**	**kà-kùríllò**	*eight old men*
	9	**kà-bárá**	**kà-kùtíttí**	*nine old men*
	10	**kà-bárá**	**kúppá**	*ten old men*

[eamy005, eamy006]

---

4   This differs from the Central Kambari system, where only the numbers 1-7 take agreement prefixes (Hoffmann 1963:166). The Northwest group of languages, like Cicipu, have agreement on 1-9 (Smith 2007:71 for U̱t-ma'in, Dettweiler n.d. for C'Lela) – no information is available on the Kamuku or Reshe branches of West Kainji.

In the following list the numeral is kept constant (**yápù** 'two') and the class of the head noun is varied:

(13) Class
1	kà-bárá	kà-yápù	*two elders*
2	à-zá	hà-yápù	*two people* (plural form of noun)
3	ì-námà	yì-yápù	*two meats*
4	mà-díyá	mà-yápù	*two hares*
5	ǹ-'yɔ́'yɔ́	mì-yápù	*two fish* (plural form of noun)
6	cì-kóotò	tì-yápù	*two drums*
7	ù-yáa	wù-yápù	*two roads*
8	c-cɔ́'ɔ̀	vì-yápù	*two sheep*
9	kù-lácí	kù-yápù	*two girls*

Classes 2 and 5 do not occur with singular nouns and so are not normally involved in numeral phrases. Some speakers, however, optionally use the plural form of the noun in numeral phrases, as in the examples above. This is especially true for nouns with human referents. In this respect Cicipu patterns with Hausa (Jaggar 2001:359), although according to Bickel and Nichols (2007:213) "many languages draw the line between animate or human referents and the rest, requiring [number] marking only for nouns referring to human beings" (see also Payne 1997:96, Corbett 2000:70).

Cicipu differs from most Kainji languages, for which the *plural* form of the noun is required with numbers > 1 (e.g. Crozier 1984:44 for Central Kambari, Smith 2007:71 for U̱t-ma'in, Dettweiler n.d. for C'Lela, MacDonell 2007:55 for Pongu, Anderson 1980a:162 for the East Kainji language Amo). This seems to be the general pattern in Bantu/Grassfields Bantu (e.g. Welmers 1973:291 for Kiswahili, and several papers in Hyman 1980), and it is reasonable to suppose that the Cicipu requirement for singular nouns in numeral constructions is an effect of contact with Hausa.

When counting, the prefix may be omitted. Alternatively, if the speaker has a particular object or set of objects in mind, she may use the appropriate agreement prefix, regardless of whether the agreement controller (i.e. the noun denoting that object) is present or not. For example when counting days (**kwã́-a'ã̀**, NC9), the AG9 prefix **ku-** may be used. For numbers from 11-19, 21-29, 31-39 and so on, it is obligatory to use an agreement prefix on the 'units' figure:

(14) ù-kúppá ǹ   vì-tò
      NC7-ten   and AG8-one
      eleven

Sometimes there is no obvious referent, for example when giving a demonstration of the

counting system for the linguist. In such cases the AG8 prefix **vi-** is used with the units figure from 11 onwards. This is a case of 'neutral' agreement (§6.4).

6.2.3 Adjectives

Adjectives take an obligatory low tone prefix.

(15)
1	**kò-zúvù**	**kè-pénêu**	*big finger* [i.e. thumb]
2	**ó-ggòmbò**	**hè-pénêu**	*big bats*
3	**ì-námà**	**yì-pénêu**	*big meat*
4	**mà-díyá**	**mè-pénêu**	*big hare*
5	**ǹ-díyá**	**m̀-pénêu**	*big hares*
6	**cì-kóotò**	**tì-pénêu**	*big drum*
7	**ù-yáa**	**wù-pénêu**	*big road*
8	**c-cɔ́'ɔ̀**	**vì-pénêu**	*big sheep*
9	**kù-lácí**	**kù-pénêu**	*big girl*

[eamy005, eamy006]

6.2.4 nínnîi **'only'**

The word **nínnîi** 'only' is another gender agreement target[5], taking an optional low-tone prefix. I have only collected examples for classes 1, 4, 6, and 9 so far, but there is no reason to think it will not be as regular as the other targets.

(16) kù-yǜyû̂ **kù**-nìnnîi
NC9-sand   AG**9**-only
only sand

[Tidipo, saat002.004.014]

6.2.5 wh-words

The wh-words **yìní** 'what' (17), **-èné** 'which' (18), and **yàanú** 'how many' (19) can all function as noun modifiers, in which case they take a low-tone agreement prefix. Although **yàanú** means 'who' by itself (§4.3.6.2), when an agreement prefix is attached it becomes a noun modifier meaning 'how many'.

(17) kà-'ázzìkí   **kè**-yìní
NC1-prosperity AG**1**-what
what prosperity?

[saat002.002.392]

(18) á↓=ká-káasùwà **kè**-ené
LOC=NC1-market   AG**1**-which
to which market?

[sayb001.721]

---

5 Unless it is considered to be an adjective. There are no formal reasons not to do this, only semantic ones.

(19) kà-hí'î̀   **kà**-yàanú
NC1-night   AG1-how_many

how many nights

The full paradigm is given below for **-èné**:

(20) 1   **kè-ené**   ká-bárá   *which elder*
2   **hè-ené**   ó-ggòmbò̀   *which bat*
3   **yè-ené**   í-námà   *which meat*
4   **mè-ené**   má-díyá   *which hare*
5   **mè-ené**   ń-díyá   *which hares*
6   **tè-ené**   cí-kóotò   *which drum*
7   **wè-ené**   ú-yáa   *which road*
8   **vè-ené**   c-có'ò̀   *which sheep*
9   **kwè-ené**   kú-lácí   *which girl*

[eamy005, eamy006]

6.2.6 Negative copula

The negative copula consists of the negation particle **cé** (also found in negated verbal clauses, §4.3.5), with a low-tone agreement prefix.

(21) 1   **kà-rìmâi**   **kà-cé**   *it's not pleasure*   [tats002.008.015]
2   **àsâ hò-tò**   **hè-cé**   *it's not one grave*   [Tikula, sagb001.704]
3   **ì-náa**   **yì-cé**   *it's not cows*   [sayb001.749]
4   **mà-wáa**   **mè-cé**   *it's not a dog*   [eamd002.030]
5[6]
6   **tíirímpà**   **tì-cé**   *it's not like this thing*   [sayb001.430]
7   **ù-séi**   **wù-cé**   *it's not pain*   [tats002.008.016]
8   **Ø-'àgóogó**   **vì-cé**   *it's not a watch*   [saat001.005.099]
9   **kù-róonò**   **kù-cé**   *its not a loincloth*   [sayb001.356]

6.2.7 Demonstrative adverbs

The demonstrative adverbs agree in gender with the subject NP when used as predicates (§4.3.3.2). The prefixes are low-tone.

(22) 1   **kà-táarí**   **kà-pâa**   *the stone is here*
2   **à-táarí**   **hà-pâa**   *the stones are here*
3   **ì-námà**   **yì-pâa**   *the meat is here*
4   **mà-cìijíi**   **mà-pâa**   *the calabash is here*
5   **ǹ-gấi**   **mì-pâa**   *the swords are here*
6   **cì-kó'ǔ̀**   **tì-pâa**   *the chest is here* [i.e. body part]
7   **ù-rée**   **wù-pâa**   *the town is here*
8   **Ø-vóotò**   **vì-pâa**   *the goat is here*
9   **kù-sáyú**   **kù-pâa**   *the spear is here*

[2007-01-14.001]

---

6 **Mì-cé** has not been observed or elicited yet but there is no reason to think it is missing from the paradigm.

Each of the five adverbs follows the same pattern. This is shown below for class 3, but the pattern is regular across all the classes.

(23) ì-táatù   yì-pâa   *the mat is here* (near speaker)
      ì-táatù   yì-lêe   *the mat is there* (near hearer)
      ì-táatù   yì-'índè   *the mat is over there*
      ì-táatù   yì-'û̀   *the mat is there* (far away or out of sight)
      ì-táatù   yì-ɗôo   *the mat is here* (permanent place)

[2007-01-14.001]

6.2.8   Associative construction

The associative construction (§4.4.5.1) takes the form NP₁ AG_NP1-NP₂, where NP₁ is the 'possessed' noun and NP₂ the 'possessor' noun. NP₁ occurs in its citation tone pattern, the agreement proclitic AG_NP1 is high tone, while NP₂ undergoes a complex tonal perturbation (§3.4.7).

(24) 1   kà-dámá   ká↓=k-káa   *the word of the woman*
      2   à-dámá   há↓=k-káa   *the words of the woman*
      3   ì-námà   yí↓=k-káa   *the meat of the woman*
      4   mà-díyá   má↓=k-káa   *the hare of the woman*
      5   ǹ-díyá   mí↓=k-káa   *the hares of the woman*
      6   cì-kóotò   tí↓=k-káa   *the drum of the woman*
      7   ù-yáa   wú↓=k-káa   *the road of the woman*
      8   c-có'ɔ̀   ví↓=k-káa   *the sheep of the woman*
      9   kù-dávù   kú↓=k-káa   *the mortar of the woman*

[eamd032]

When the associative construction is considered as a whole, it is usually the first NP (the 'head' in the sense that it is the semantic determinant of the whole) that triggers agreement on external targets, but this is not always the case (see §6.5).

6.2.9   Possessive pronouns

The possessive pronouns (§4.4.5.1.1) take agreement prefixes, usually with H tone.

(25) kà-'íngùwà   kó-ttù
      NC1-village   AG1-1P.POSS
      our village

[svbg001.030]

The full paradigm is shown below, repeated from §4.4.5.1.1.

Table 45: Possessive pronoun paradigm with gender agreement prefixes

	1s	2s	3s	1p	2p	3p
1	ko-vôo	ká-avù	ké-evì	kó-ttù	kó-ɗɗò	kí-ivè
2	ho-vôo	há-avù	hé-evì	hó-ttù	hó-ɗɗò	hí-ivè
3	yi-vôo	yá-avù	yé-evì	yí-ttù	yí-ɗɗò	yí-ivè
4	mo-vôo	má-avù	mé-evì	mó-ttù	mó-ɗɗò	mí-ivè
5	mi-vôo or n-vôo	má-avù	mé-evì	mí-ttù	mí-ɗɗò	mí-ivè
6	ti-vôo	tá-avù	té-evì	tí-ttù	tí-ɗɗò	tí-ivè
7	wu-voô	wá-avù	wé-evì	wú-ttù	wú-ɗɗò	wí-ivè
8	vi-vôo or v-vôo	vá-avù	vé-evì	ví-ttù	ví-ɗɗò	ví-ivè
9	ku-vôo	kwá-avù	kwé-evì	kú-ttù	kú-ɗɗò	kwí-ivè

The agreement markers have been observed omitted, but this is rare.

(26) á-sɨ̀ɨ́   há-mpà **vôo**
      NC2-feather AG2-this **1s.POSS**

these my feathers

[saat001.006.156]

6.2.10 Article/relativiser

As discussed in §4.4.5.3, the same form **-nà** is used for the article and the relativiser. The article **-nà** can occur either before or after the noun head – in both cases it takes an obligatory high-tone agreement prefix.

(27) 1 **kà-dámá**   **ká-nà**   *the elder*
      2 **ó-ggòmbò**   **há-nà**   *the bats*
      3 **ì-námà**   **yí-nà**   *the meat*
      4 **mà-díyá**   **má-nà**   *the hare*
      5 **ǹ-díyá**   **mí-nà / n-nà**   *the hares*
      6 **cì-kóotò**   **tí-nà**   *the drum*
      7 **ù-yáa**   **wú-nà**   *the road*
      8 **c-cɔ́'ɔ̀**   **ví-nà / n-nà**   *the sheep*
      9 **kù-lácí**   **kú-nà**   *the man*

[eamd032]

As can be seen in the list above, the variation in AG5 and AG8 prefixes can lead to these classes being undifferentiated. See §6.1.3 for the distribution of the relativiser AG8 allomorphs.

The following example illustrates agreement on the relativiser, which is again obligatory.

(28) ì-rí   **yí**-nà   yì-lábà-nà   yì-'étếi
      NC3-thing   AG3-REL   AG3-lack\RLS-PFV   AG3-fine

bad things [lit 'things which lack fineness']

[oamy001.128]

6.2.11   Demonstrative modifier

Demonstratives modifiers (§4.4.5.2) take a high-tone prefix, which is generally obligatory[7].

(29) ká-ayá   ká-mpà   *this hut* (near to speaker)
      ká-ayá   ké-llè   *that hut* (near to hearer)
      ká-ayá   ké-"índè   *yonder hut* (far from both)
      kò-dóndó   kó-"û̀   *that garden* (very far away or out of vision)
      kwá'à   kú-ɗɗô   *this house* (permanent place)

The agreement paradigm is fully regular but is presented below for completeness.

Table 46: Demonstrative modifiers with gender agreement prefixes

	this	that	yonder	out of sight	this (perm.)
1	ká-mpà	ké-llè	ke-̀"índè	kó-"û̀	kó-ɗɗô
2	há-mpà	hé-llè	hé-"índè	hó-"û̀	hó-ɗɗô
3	yí-mpà	yí-llè	yí-"índè	yí-"û̀	yí-ɗɗô
4	má-mpà	mé-llè	mé-"índè	mó-"û̀	mó-ɗɗô
5	mí-mpà	mí-llè	mí-"índè	mí-"û̀	mí-ɗɗô
6	tí-mpà	tí-llè	tí-"índè	tí-"û̀	tí-ɗɗô
7	wú-mpà	wú-llè	wú-"índè	wú-"û̀	wú-ɗɗô
8	ví-mpà	ví-llè	ví-"índè	ví-"û̀	ví-ɗɗô
9	kú-mpà	kú-llè	kú-"índè	kú-"û̀	kú-ɗɗô

6.2.12   -mbɔ̃̀ 'another'

The modifier **-mbɔ̃̀** (§4.4.5.3) may be glossed as 'another'. It takes a high-tone prefix.

(30) 1   kà-bárá   kó-mbɔ̃̀   *another elder*
     2   à-bárá   hó-mbɔ̃̀   *other elders*
     3   ì-námà   yí-mbɔ̃̀   *another meat*
     4   mà-díyá   mó-mbɔ̃̀   *another hare*
     5   ǹ-díyá   mí-mbɔ̃̀   *other hares*
     6   cì-kóotò   tí-mbɔ̃̀   *another drum*
     7   ù-yáa   wú-mbɔ̃̀   *another road*
     8   c-cɔ́'ɔ   ví-mbɔ̃̀   *another sheep*
     9   kù-lácí   kú-mbɔ̃̀   *another girl*

                           [eamy005, eamy006]

6.2.13   Subject

Verbs are obligatorily prefixed by either a gender or person marker. The gender agreement prefixes are of CV- shape, unless the verb stem is vowel-initial in which case they have the C- forms, with the usual prefix/root vowel coalescence (§3.1.1). The tone depends on the mood of the verb (§4.6.2). Subject agreement is dealt with in more detail

---

7   The AG8 'near speaker' form is sometimes **llè** rather than **ví-llè**. This could be analysed as the **Ø-**allomorph of the AG8 prefix, although normally that is only possible before short consonants. Alternatively, it could be considered as lacking agreement.

in §7.4-7.6.

(31) kà-bárá   **kà-'wâa-nà**
      NC1-old_man AG1-pass\RLS-VENT
      an old man passed by

[tapf002.001.030]

6.2.14 Copula

The copula consists of either an [e] or an [i] vowel, with a gender agreement prefix of the C- form, but without the long vowel that usually results from coalescence between prefix and stem vowels. The [e] vowel is found in classes 1, 2, and 4, the prefixes of which contain the harmonising /A/ vowel (§6.1.1). It may therefore be analysed as the product of assimilation between the prefix vowel /A/ and an underlying /i/ in the copula (cf. §3.7.1). The tone on the copula itself is usually polar, but as discussed in §3.4.4 there are exceptions.

(32) 1   **kó-ggòmbò**   **k-é**   *it's a bat*
      2   **ó-ggòmbò**   **h-é**   *it's bats*
      3   **ì-námà**   **y-í**   *it's meat*
      4   **mà-díyá**   **m-è**   *it's a hare*
      5   **ǹ-díyá**   **m-ì**   *it's hares*
      6   **cì-kóotò**   **t-í**   *it's a drum*
      7   **ù-yáa**   **w-ì**   *it's a road*
      8   **c-cɔ́'ɔ̀**   **v-í**   *it's a sheep*
      9   **kù-lácí**   **kw-ì**   *it's a girl*

[eamy005, eamy006]

As noted in §4.3.3.1, in predicate nominal constructions the copula agrees with the predicate rather than the subject, an instance of 'back' agreement or 'attraction'.

6.2.15 Noun class pronouns

Noun class pronouns take the same form as the copula[8], but this time the tone depends on the syntactic position of the target rather than the phonological environment. In the complement position following verbs or the locative proclitic **Á** they occur with low tone as in (33); otherwise they take high tone as in (34).

(33) ǹ-zúwɛ̀̃   ó-gínó,   m-ǐndà   **h-è**   cé
      1s-roast\RLS NC2-groundnut 1s-see\RLS **AG2-PRO** NEG
      *I roasted groundnuts, I didn't see **them***

[saat001.002.003]

---

8 See Stassen (1997:77-85) on 'pro-copulas'.

292

(34) ń n=à-kábà-nà à-ràkúmí à-dúkwà ǹ **h-é**
    if and=3P-take\RLS-VENT NC2-camel 3P-go\RLS with **AG2-PRO**
    *if they took camels they went with **them***

<div align="right">[samoh001.198]</div>

Predicate locative constructions (§4.3.3.2) involve what appears to be a gender-marked pronoun followed by the copula.

(35) **k-í** **k-è** pâa
    AG1-PRO **AG1-COP** here
    it [e.g. **kà-táarí** 'the stone, NC1'] *is here*

<div align="right">[2007-01-14.001]</div>

Despite the PRO gloss just given, the category of the first word is not totally clear, since I have not found these precise forms in any other context. They differ from straightforward noun class pronouns by retaining the **i** vowel throughout the paradigm. The reason for glossing them as pronouns is that the structure is analogous to the personal pronoun predicate locative (e.g. **ìvɔ́ vì pâa** 'you(sg.) are here' – §4.3.3.2), with respect to both meaning and the H L tone pattern.

(36) 1 **k-í** **k-è** pâa *it is here* [e.g. bat]
    2 **h-í** **h-è** pâa *they are here* [e.g. bats]
    3 **y-í** **y-ì** pâa *it is here* [e.g. meat]
    4 **m-í** **m-è** pâa *it is here* [e.g. hare]
    5 **m-í** **m-ì** pâa *they are here* [e.g. hares]
    6 **t-í** **t-ì** pâa *it is here* [e.g. drum]
    7 **w-í** **w-ì** pâa *it is here* [e.g. road]
    8 **v-í** **v-ì** pâa *it is here* [e.g. sheep]
    9 **kw-í** **kw-ì** pâa *it is here* [e.g. girl][9]

<div align="right">[2007-01-14.001]</div>

6.3 Antecedentless agreement morphology and semantic agreement

In §2.2.1 I discussed the possibility of the use of agreement morphology in the absence of an agreement controller, and noted that this does not qualify as 'agreement' under Steele's definition of agreement adopted by Corbett and others. Instead the speaker is relying on the hearer making an inference as to the appropriate referent, based on the use of a particular agreement prefix in a particular context. The relevant context may be text-*external*, for example pointing to a stone while saying **ká-mpà** 'AG1-this'. Alternatively the referent may be inferred from the text (an 'indirect anaphor' according to Schwarz-Friesel 2007). Example (37) below is of this kind. Speaker B assumes that,

---

9 The AG9 construction is sometimes pronounced **kú kù**.

given the conversation is about grave sites and burial, speaker A will be able to make the inference that the AG1 indirect anaphor **k-é** denotes the corpse (**kà-kkwái**, NC1) of Wan Moolu.

(37)　　[Context: Speaker A: *The one who had the title 'Wan Moolu', Ukula hill hasn't seen his grave yet.*]

　　B:　tò　à-sì-kábà-kábà　ó-gìtù　　　ǹ　　**k-é**?
　　　　OK　3P-HAB-take-REDUP　3P-go_back\IRR　with　**AG1-PRO**
　　　　*OK so they are taking **it** [Wan Moolu's corpse] back?*

<div align="right">[Tikula, sagb001.763]</div>

The existence of the NC1 noun **kà-kkwái** 'corpse' in the lexicon was sufficient, in that particular context, to ensure coherence.

It is important to distinguish this usage of agreement morphology from semantic agreement. Semantic agreement (§2.2.3) involves the target agreeing with semantic properties of the controller (e.g. humanness). The form of the agreement morphology used in (37) was, in contrast, determined by reference to purely formal properties of another word, one that was entirely absent from the discourse. If the semantic structure of gender 1/2 was more coherent, it might be possible to talk of the AG1 morphology displayed by **k-é** in (37) being determined by the semantics of the gender. As things are, it seems more appropriate to assume that the antecedentless agreement is, in this case, made possible because of the prominence of the word **kà-kkwái** in the particular context, rather than any semantic associations of gender 1/2.

There are times, however, when semantics does seem to play a role, particularly in the case of the less incoherent genders, and when the concept involved is more abstract than, say, CORPSE. For example AG6 agreement morphology is frequently used to express manner, but there is no specific NC6 noun that can be appealed to as an explanation for this. Instead, it seems that the gender itself has certain semantic associations that can be invoked by the use of the corresponding agreement morphology. There are also intermediate cases – in the case of AG7 agreement indicating a place, it is not clear whether this is possible because of the existence of the single salient gender 7 noun **'ásù** 'place', or because of the fact that gender 7/8 contains a number of words for places and therefore, to some extent, is semantically-integrated by the notion of PLACE (see §5.3.4 and §5.4.1 for NC7 locative derivations). It seems likely that both these facts are relevant – after all, they are not independent, and a cognitive linguistics approach to gender structure (see §5.2) might see the association with PLACE as derived from a semantic

network integrated to a strong degree by the specific noun **'ásù**.

The gender with the most semantic coherence[10] is 8/2, which contains only nouns denoting humans or spirits. There exists at least one 'hybrid noun' **má-gàjì** 'priest, 4/5' which triggers either 4/5 or 8/2 agreement. This can be seen as exemplifying the extreme[11] end of this continuum, where the semantic structure of the gender 8/2 is of the most importance in allowing the reference to be made.

The examples that follow are arranged roughly along this continuum, starting with examples like (37) which rely heavily on a single word, and ending with a discussion of the hybrid noun **má-gàjì**, which relies on the coherent semantic structure of the 8/2 gender.

The first example is similar to (37), but instead of an antecedentless pronoun it involves an antecedentless demonstrative **mí-mpà**. The topic of the discussion from which (38) was taken was clothing worn in the past. The only textual antecedent that would fit **mí-mpà** semantically is the generic word for clothing **ì-hyácí** (NC3), but this does not fit grammatically, and there is no semantic reason to use AG5 agreement here. Instead reference by means of the antecedentless AG5 demonstrative **mí-mpà** is made possible by the existence in the lexicon of the word **mè-ttégù/ǹ-tí-téggù** 'shirt/s, 4/5', just as in the case of **kà-kkwái** 'corpse' and the antecedentless AG1 pronoun **k-é**.

(38) [Context: discussion of clothing – *but it wasn't these sort of clothes / they were something /*]

mí-mpà   mí↓=ó-yóyú   mí-mpà
AG**5**-this   AG5=NC2-shirt_k.o.   AG5-this
these [**ǹ-tí-téggù** 'shirts, 4/5'] *of 'oyoyu' these*

[sayb001.728]

Perhaps more interesting are cases of antecedentless agreement involving more abstract concepts, which consequently can be used in less restricted contexts, despite the fact that they still depend on the existence of individual words rather than the semantic coherence of a gender. One such case is the use of AG2 morphology when names are involved. Recall from §5.2.9 that the only noun in single class gender 2 is **à-húlá** 'name', which of course cuts down potential referents and makes antecedentless agreement less problematic[12]. The following extract is especially interesting because of

---

10 Barring, of course, inquorate genders containing only a single noun.
11 With respect to Cicipu at least. Strict semantic agreement as defined by Corbett (1991) does not occur – see the discussion at the end of this section.
12 There are of course many other NC2 word forms, but they are all plurals, and often the context will

the variety of expressions (shown in **bold**) that are used to refer to the item of clothing.

(39)   [Context: A family are trying to remember the name of a kind of clothing.]

Y: **ì-rí**   yí-nà   à-kámùkwà Ø-sá'à   ví-llè   à-vádì
    NC3-thing AG3-REL   3P-be.PST\RLS NC8-time AG8-that NC2-Vadì

    à-sí-yũ̂u   w-à   kà-kàɓàlí   kǒo?
    3P-HAB-wear 3S-REPRT NC1-leaf_covering   or

    the things *that they were that time the Avadì are wearing they are called kakaɓali or?*

T: ò'íi
    yes

    yes

M: àmáa àsì'ɔ́mì**vì**   yìní ǹ   Cìcípù?
    àmáa à-sì-'ɔ́mɔ̀=**vì**   yìní ǹ   Cì-cípù
    but   3P-HAB-call=**3S.PRO** what with NC6-Cipu

    *but what do they call **it** in Cicipu?*

T: **y-í**   y-ì   célè   "kà-kàɓàlí"
    AG3-PRO   AG3-COP   that   NC1-leaf_covering

    *that's **it** "kakaɓali"*

Y: **h-í**   h-è   lêe   hwá   á↓=cí-cípù "kà-kàɓàlí"
    AG2-PRO   AG2-COP   there definitely LOC=NC6-Cipu NC1-leaf_covering

    *that's **it** definitely in Cicipu "kakaɓali"*

[sayb001.345]

The excerpt is part of a wider discourse on clothing, and at this moment the item of clothing in question is the main discourse topic. This explains M's use of the person-marked object clitic **vì** in his contribution[13]. T's AG3 pronoun **y-í** may be considered to be straightforward (albeit cross-turn) syntactic[14] agreement with the **ì-rí** 'thing' in Y's first utterance. The most interesting form with respect to the present discussion is the AG2 pronoun **h-í** in the final utterance. According to my language consultant, the AG2 form was used because the reference is to a *name*.

While the AG3 pronoun **y-í** in the above example can be analysed as syntactic agreement, it is also possible to view it as deictic reference (see §2.2.1 on the blurring of the boundary between the two phenomena). In support of the latter analysis, there are also unambiguous cases of antecedentless AG3 agreement, which rely on the existence of

---

straightforwardly rule them out.
13 See §8.5.1.3 for discussion of this example with respect to discourse topicality.
14 In Corbett's sense of being determined by morphosyntactic properties of the controller.

the 8/3 noun **ì-rí** 'thing'[15]:

(40) [Context: discussion of methods for trapping fish. *there is hunting of* **à-sídǎ**
trees [**kà-sídǎ** 'tree, k.o., 1/2']. *In the past they would split split split* [i.e. the
branches]*, then they threw in the water*]

ée,	tò	n＝à-tũ̂u	y-ì,
yes	OK	when=3P-pour\RLS	AG3-PRO

*yes, OK when they poured **them***

[tats005.002.115]

Here the AG3 pronoun **y-ì** is the first explicit mention of the broken pieces of branches,
although their existence can easily be inferred from the previous sentence. When
questioned, my consultant said that AG3 agreement was used because the broken pieces
were **ì-rí** 'things'.

The most common exploitation of antecedentless agreement morphology in the
corpus is the use of AG1 agreement to refer to words, propositions, speech, and
languages, and this will be treated at some length here. While the word for 'word',
kà-dámá (1/2) is no doubt strongly responsible for this possibility (it is the word most
frequently cited in explanations by consultants), the association is likely to be re-
inforced by the existence of other 1/2 words with related meanings, such as **kà-nàbáyì**
'news, account', **kò-mísòonî** 'story', and **kè-ré'è** 'language'. Nevertheless the size of
gender 1/2 is so big and its structure so incoherent that it is probably still better to view
the links as links to a set of words, rather than to some abstract WORD concept. In what
follows I give a number of examples starting from reference to a single word and
building up to reference to entire languages or ways of speaking.

In (41) below the speaker uses the AG1 numeral **kò-tò** with the meaning 'one
word'. Nowhere in the surrounding discourse is there a suitable antecedent (or even a
'postcedent') for this indirect anaphor – however the selectional restrictions of the verb
hyãa 'say' are a strong constraint on possible referents, and there seems little room for
confusion. **Kà-dámá** 'word' was offered without hesitation when I asked for the reason
for using an AG1 prefix.

(41)
ùhyȋi	cé	sáa	**kòtò**
ù-hyãa＝ì	cé	sáa	**kò-tò**
3S-say\RLS=3S.PRO	NEG	even	**AG1-one**

*he didn't say **one** [word] to him*

[saat002.002.245]

---

15 While it is possible to elicit the NC8 'singular' **r-rí**, the NC3 'plural' **ì-rí** is much more common, even
for single entities.

As well as actual words, AG1 morphology can be used to refer to conceptual structures, in particular propositions. Again speakers consistently offer the existence of **kà-dámá** 'word' as the explanation for the use of AG1 morphology in such examples. In the following example taken from a song, the demonstrative **ké-llè** 'that' cataphorically links to the proposition expressed in the following sentence.

(42)   **ké-llè**   mà-hũu   m-è:
       AG1-that NC4-truth AG4-COP

       Wán ví↓=i-dɔ̂ɔ   náhà   ú-rée   ù-yẫa   ǹ-háalú
       Lord AG8=NC3-horse leave\RLS NC7-town 3S-do\RLS NC5-journey

       ***this** is the truth: Wan Viidɔɔ left the town he went on a journey*
                                                 [ovkz002.020.001]

The following three examples are similar, except the reference is anaphoric and by means of a pronoun. The collocation **k-í k-è n-nà** shown in (45) is a fairly common conversational response, and is unusual because of the AG8 agreement on the article. This seems to be a form of neutral agreement (§6.4) with the non-verbal clause **k-í k-è**, although why this should require the article is unclear – perhaps because the speaker is stressing that the proposition which has just been voiced is one familiar to him.

(43)   [Context: *news of the Acipu has reached the white man's land. He followed the traces and thought about the land from far off...*]

       k-í   **k-è**   w-ăyà-nà   pâa
       AG1-PRO **AG1-COP** 3S-come\RLS-PFV here

       ***that's** why he came here*
                                                 [Tikula, sagb001.099]

(44)   kádà á-pàndà   Cì-cípù / kà-tĩ̀i[16]   ké-'èesù   wú-utò-nò /
       PROH 3P-forget\IRR NC6-Cipu NC1-head AG1-add\IRR NC7-go_out-VENT

       k-í   **k-è**   **ká**-ayà-wà-nà=tù   páa-nì
       AG1-PRO **AG1-COP** AG1-come\RLS-APPL-PFV=1P.PRO here-NMLZ

       may Cicipu not be forgotten, let knowledge increase [lit. 'let heads increase coming out'], ***that's** what brought us here*
                                                 [Tikula, svsdt001.104]

---

16 The NC1 noun **kà-tĩ̀i** 'head' is unrelated to the AG1 agreement markers in the following intonation unit.

(45) A: sáa hínà àrîi   sáa mòdóo mè
        sáa hínà à-râa = ì   sáa mò-dóo m-è
        or   TOP 3P-eat\RLS=3S.PRO   or   NC4-slave AG4-COP
        either they defeated him or he's a slave

    B: **k-í**   **k-è**   **n**-nà
       AG1-PRO **AG1**-COP **AG8**-ART
       that's right!

                                                    [Tikula, sami001.455]

The next example involves the use of the AG1 relativiser. The speaker is emphasising his family's status within Mataari village by explaining that no decision can be made without their consent. The Hausa translation I was given for (46) was *see **abin** da mun yadda da shi* 'only **a thing** with which we agree'. The nine choices of agreement prefix on the Cicipu relativiser give the speaker the opportunity to be more specific about the referent, but without having to mention the actual word **kà-dámá** 'word'.

(46) ká'à lêe Mà-táarí sée **ká-nà** tì-yáddà-nà /
      now there NC4-[place] unless **AG1-REL** 1P-agree\RLS-PFV
      *now there Mataari, only **that** [i.e. a word] with which we agree /*

      [*If we speak, and the others don't agree, they won't be able to deprive us of our proper share.*]
                                                    [svbg001.032]

The final 'propositional' example is notable because unlike the previous examples, the immediate linguistic context seems to be of little use for determining the reference of the AG1 morphology. It comes from a folktale involving five superheroes, each of whom claimed to have the best powers. After two of them had demonstrated their skills, the assembled throng inform them that they have lived up to their word:

(47) tò   ìdó   h-índà   **kó-ɗɗò**
      OK   2P.PRO 3P-see\RLS   **AG1-2P.POSS**
      *OK you (pl.) **yours** [word] has been seen [lit 'you (pl.) they saw yours']*
                                                    [Tidipo, saat002.005.043]

The point is that this sentence comes after a 90-second account of feverish superpower activity during which no 'words' were mentioned – the crowd is referring back to the heroes' original claims. This is an argument for analysing AG1 morphology as having a 'default interpretation' (in the absence of an antecedent) as referring to something covered by the basic-level term **kà-dámá**.

    As well as pointing to specific propositions, AG1 morphology can also have a more

general interpretation as 'knowledge about', and in the following examples the AG1-marked associative constructions have been translated using *about* in each case. There is no specific word, statement, or even proposition relevant to these examples.

(48)　[Context: discussion of **ù-láa** 'fire, NC7']

sáa　yàanú　Ø-lápà　　　ǹ　　kɛ́=w-ì
or　who　AG8-know\RLS　with　AG1=NC7-PRO
*everyone knows **about** it*

[tats002.002.011]

(49)　[Context: speaker A says he doesn't know about the founding of Korisino. Speaker B checks...]

vú-u-làpà　　　cé　　kɛ́-llè?
2S-FUT-know\IRR　NEG　AG1-that
*you wouldn't know **about** that?*

[sayb001.145]

As usual, the existence of **kà-dámá** supplies the explanation – and this analysis is supported by the parallel construction in (50) which has an explicit antecedent and apparently identical meaning.

(50)　[Context: the speaker is 'buttering up' the hearer, whom he wants to start talking about the traditional religion]

'ínà　z-zá　　　ɗàa　　　'ínà　z-zá　　　ù-gúyá
some　NC8-person　surpass\RLS　some　NC8-person　NC7-can

kà-dámá　kó↓=kó-rínnò
NC1-word　AG1=NC1-traditional_religion
*some people know more than others **about** the traditional religion*

[Tikula, sagb001.303]

Antecedentless AG1 agreement can also be employed for the reification of an entire discourse during its closure. Example (51) came from the end of a retelling of the Pear Story (§1.4.1), and the referent of **k-í** seems to be the entire textual span of the narrative, or possibly the corresponding conceptual 'story'. This time consultants were less clear about the reason for the AG1 markers – **kà-nàbáyì** 'account', **kò-mísòonî** 'story', and **kà-dámá** 'word' were all offered. It is quite possible that the multiple possibilities reinforce rather than conflict with each other as potential referents.

(51)　**k-í**　　k-è　　lêe　k-áyà　　　kɔ̀-kɔ́tɔ̀
　　AG1-PRO　AG1-COP　there　AG1-come\RLS　AG1-finish\RLS
that's that it's finished

[tapf001.004.025]

Moving away from specific speech acts or events, antecedentless AG1 morphology can also be used to refer to actual languages or ways of speaking (i.e. dialects/idiolects). **Kà-dámá** can mean 'speech' or 'dialect' as well as 'word' and may be what the AG1 morphology is pointing to in the following two examples. The first comes from a discussion of the different dialects of Cicipu.

(52) Tì-zòoríyò dàbɔ̂, Tì-dɔ́dìmɔ̀ɔ kúmá **kí**-ivè dàbɔ̂
NC6-Zooriyo different NC6-Dodimo and AG1-3P.POSS different
*Mazarko dialect is different, Kadedan also **theirs** ['speech'] is different*
[eadt001.232]

The second example is given mostly in English for the reader's convenience. The important point to note here is the progression from AG6 agreement to AG1 in the final intonation unit.

(53) a. Grandson: and **Cicipu**(NC6) nowadays, people are not doing well /
b. Grandmother: yes /
c. AG6-not the same as that AG6-of-before /
d. you (sg.) now the **Cicipu**(NC6) people /
e. AG6-of-before AG6-that they would speak /
f. if you spoke AG6-it now /
g. children wouldn't know /
h. the meaning /
i. they wouldn't know the meaning /

j. ká↓=k-ká'à èrè sée hû-uwà **k-è** bíríbírí
AG1=NC8-now 3P.PRO until 3P-hear\IRR AG1-PRO dimly
[**language**] *of now they can hardly speak **it***
[svtmg001.087]

To the extent that we can consider the AG1 agreement in (53) to be 'semantic' agreement controlled by the word **Cì-cípù**, the progression from NP to AG6 to AG1 (on both pronouns and the associative construction) fits in with Corbett's (1991:240) prediction that for a given target, semantic agreement will not be found closer to the controller than syntactic agreement[17].

The examples up to now have all been strongly linked to particular words, even if it is arguable that there is some semantic connection between genders 1 and 1/2 and the idea of words or language. For the remaining examples the semantic link is stronger. They involve AG6, AG7, and AG9 morphology, and we will consider each of these in turn.

---

17 Immediately after the utterances in this extract, however, a third speaker asks the old woman **hûuwà** t-ì **ánà tìkɔ́gɔ́ kóo?** 'they speak AG6-it like Hausa?'. So Corbett's prediction only holds within the single discourse turn of the grandmother in this instance.

We saw in §5.3.3 and §5.3.5 that there is an association between nouns from gender 6 and 'manner'. This association is also apparent in the use of antecedentless AG6 agreement morphology. In particular the wh-word **-ènɛ́** (§4.3.6.2) and the relativiser **-nà** (§4.3.4) are both found with AG6 prefixes in words meaning 'how':

(54)   **t-ènɛ́**   t-ì   à-sì-yɛ̃́a   kà-bíkì   ká↓='á-kkwíi
        AG6-which   AG6-COP   3P-HAB-do   NC1-festival   AG1=NC2-dead_person
        ***how** do they do the festival of dead people?*

                                                 [Tikula, sagb001.412]

(55)   **tí-nà**   Ø-máasĩ̀i   sì-súmá   á↓=í-ɗáa
        AG6-REL   NC8-motorbike   HAB-run   LOC=NC3-ground
        ***how** [i.e. the way in which] a motorbike runs along the ground*
                                                  [tats002.006.039]

Similarly, the 'preposition' **tí** (§4.5) may actually just be an AG6 associative agreement prefix. There is no gender 6 word meaning 'manner' that has been omitted in these examples; instead, it seems to be the gender exponents themselves that are contributing that meaning.

Recall from chapter 5 (§5.2.4, §5.3.4, §5.4.1) that genders 7 and 7/8 are associated with places. This association has also been observed in the case of antecedentless AG7 agreement morphology. In the following example the use of AG7 marking on the pronoun evokes a place – in particular, the place where Caari celebrates the festival.

(56)   [Context: discussion of the groupings in which people gather for a festival]
        Cáarì,   tò   é   bésì   **w**-í   w-ì   lêe   'yɛ̃̀u   gàbáɗáyá   gàngàmé
        [name]   OK   3S.PRO   all   AG7-PRO   AG7-COP   there   1S.PRO   together   together
        *Caari, OK him all **it** [i.e. Caari's place] is there together with me*
                                                  [Tikula, sami001.192]

In the following example the subject **Mà-ppáyá** is clearly the antecedent, and so this is perhaps better thought of as semantic subject agreement[18].

(57)   lêe   Mà-ppáyá   **wù**-yô   ánà   mà-ppátá
        there   NC4-[name]   AG7-be\RLS   like   NC4-cave
        *there Mappaya **is** like a small cave*
                                               [Tikula[19], sagb001.648]

We saw in §5.4.2 that NC9 deverbal nominalisations denote repeated action, and just as

---

18 As the only example of semantic AG7 agreement observed, it is suggestive to find predicate agreement given Corbett's (1991: chp. 8) Agreement Hierarchy. But until more data is collected nothing more definite can be said.

19 It is not yet known whether this phenomenon is found in Tirisino.

for NC6 and NC7, we find a reflex in the agreement system. Two different kinds have been observed. First, when the AG9 prefix **kù-** occurs on a numeral in the absence of a controller it usually means 'times' (e.g. *he fell over three times*), as observed in §4.8.4[20]. Secondly, it seems to be possible to evoke a reified repetition using AG9 morphology:

(58) [Context: discussion of slavery. *A person keeps doing this suffering, doing this suffering, doing this suffering...*]

    tò   **kw-í**   kw-ì   tí-módó
    OK   **AG9-PRO**   AG9-COP   NC6-slavery
    OK this is slavery

                                  [tats002.005.052]

Finally we turn to the 'hybrid noun' mentioned at the start of this section. Corbett (1991:225) describes hybrid nouns as nouns which "take agreement in more than one gender, depending on the target". A well-known example is the German noun *Mädchen* 'girl' (Corbett 1991:227-228), which may be referred to using a choice of personal pronouns, either the neuter *es* ('syntactic' agreement) or the feminine *sie* ('semantic' agreement).

    Only one such noun has been discovered so far in Cicipu, **má-gàjì** 'priest' (originally from Hausa *magaji* 'heir'). When it was borrowed this word was presumably assigned to gender 4/5 by phonological assignment rules (see §5.6), the first syllable **ma-** having been reinterpreted as an NC4 prefix, as evidenced by the NC5 plural **ń-gàjì** 'priests' and the NC6 derivation **tí-gàjì** 'priesthood'. While **má-gàjì** seems to trigger AG4 agreement consistently in the associative construction (59-61), both AG4 and AG8 agreement can be found on the copula (61-62) and on verbs (63-64).

(59) má-gàjì   **má↓**=ká-ngú
    NC4-priest   **AG4**=NC1-[place]
    the priest of Kangu

                                    [svmk001.001]

(60) má-gàjì   **má↓**='yũ̂u-nì,   **má**-nà   **mò**-kôo-nò
    NC4-priest   **AG4**=there-NMLZ   **AG4**-REL   **AG4**-die\RLS-PFV
    the priest of up there, the one who died

                                [Tikula, sami001.022]

---

20 In Chadic languages there is often a link between the word for 'foot' or 'footprint' and the word for times – cf. Hausa *sau* 'footprint/times'. I do not know whether this is the case for Benue-Congo in general. In Cicipu the word for 'foot' is **kù-náa** (NC9), which may be reason for the AG9 morphology here.

(61) má-gàjì   **mú**=u-ɗàyà   **m**-é
NC4-priest **AG4**=NC7-[place] **AG4**-COP
it's the priest of Uɗaya

[Tikula, svsdt001.066]

(62) má-gàjì   **v**-ì   wú-u-yãa   'ínì
NC4-priest **AG8**-COP 3S-FUT-do\IRR like_that
it's the priest he will do thus

[tats001.003.028]

(63) má-gàjì   **mà**-yâa-nà
NC4-priest **AG4**-arrive\RLS-VENT
the priest arrived

[oar002.010.001]

(64) má-gàjì   **Ø**-sì-hyáa   ó-kíisó
NC4-priest **AG8**-HAB-say NC2-maigiro_spirit
the priest speaks to the Okiiso spirits

[tats001.003.040]

The AG4 agreement is consistent with the overt expression of NC4 morphology on the controller, and is therefore syntactic agreement. AG8 agreement, on the other hand, could be argued to be semantic agreement – while there is no 'semantic assignment rule' (Corbett 1991:225-226) assigning human nouns to gender 8/2[21], we saw in §5.2.5 that 8/2 is the most coherent of all the genders, consisting only of human/spirit nouns. Therefore if there are going to be hybrid nouns in Cicipu at all, 8/2 is the most likely gender for semantic agreement.

Corbett has shown how the possible agreement forms triggered by hybrid nouns is constrained by the Agreement Hierarchy (1991: chp. 8). The evidence from Cicipu is (currently) meagre, but it at least does not conflict with the Agreement Hierarchy – subject agreement allows either syntactic or semantic agreement, but noun modifiers such as the associative construction seem to be limited to syntactic agreement. Further research would be required to say anything beyond this.

The phenomenon of hybrid nouns anticipates the concerns of Part IV, which also involves an alternation between two different kinds of agreement. The difference is that in Part IV we are not dealing with an alternation between syntactic gender agreement and semantic gender agreement restricted to a few nouns, but with one between gender agreement and gender-independent *person* agreement, potentially affecting *all* nouns. Nevertheless certain of Corbett's predictions about the differences between syntactic and

---

21 There are also nouns denoting humans in genders 1/2, 4/5, 8/3, and 9/2.

semantic agreement are found to hold for the gender/person alternation as well (§8.8).

6.4 Neutral agreement

Neutral agreement in Cicipu (i.e. agreement with atypical controllers – Corbett 1991) involves AG8 as the 'exceptional case default' (Corbett 2007:267), regardless of the type of agreement target. This is not really surprising given that this class has a null prefix, amongst other allomorphs. Consequently controllers can function as a noun in this class without any structural change, and yet not appear 'out of place' either[22]. In the first subsection (§6.4.1) I will demonstrate that AG8 agreement is used for a variety of different kinds of atypical controller. In the second (§6.4.2) we will see that NC8 noun prefixes may also be assigned to atypical controllers.

6.4.1   AG8 neutral agreement

The names of people trigger AG8 agreement[23]. The following example shows the name of a person **Màjíijí** triggering AG8 agreement on the definite article, despite resembling an NC4 noun with a **mA-** prefix:

(65)   éví   Màjíijí   **n**-nà
        3s.PRO [name of person] AG**8**-ART
        him Majiiji [who we have just talked about]

[Tikula, sami001.032]

The controller in (66) below is atypical in that it is not actually a word, and therefore cannot have inherent gender. Here the sound 'o' triggers AG8 agreement on the numeral. As usual with AG8 agreement (§6.1.3), **vi-** or **Ø-** may be substituted for the lengthening prefix.

(66)   mǔ-uwà   'o'   **n**-nósì
        1s-hear\RLS [sound] AG**8**-four
        I heard four o's [i.e. the sound 'o']

[2007-01-10.010]

The next example demonstrates neutral subject agreement, again with a controller denoting a sound:

(67)   'a'   **ví**-ità=mù
        [sound] AG**8**-satisfy\RLS=1s.PRO
        an 'ah' satisfied me [i.e. that was the sound I wanted to hear]

[2007-02-09.002]

---

22 This is also the reason why loanwords enter NC8 (§5.6).

23 Non-human proper nouns often have inherent gender (§4.4.2).

The next example is similar, except that this time it is a letter ('b') triggering AG8 agreement on the relativiser **-nà** and the verb.

(68) 'b'   **n**-nà   Ø-líbà-nà   kù-gíyà
  [letter] AG8-REL   AG8-lack\RLS-PFV   NC9-hook
  'b' without a hook [i.e. not ɓ]

<div align="right">[2007-01-10.010]</div>

Individual words without inherent gender also trigger neutral agreement. Example (69) comes from a metalinguistic discussion about the verb form **haaya** 'they came', and shows AG8 agreement on the numeral.

(69) òkóo   "haaya"   **vì**-yápù
  there_is [word]   AG8-two
  there are two 'haaya's'

<div align="right">[samoh001.227]</div>

Prepositional phrases (70) or even entire clauses (71) can also trigger AG8 agreement, at least on the copula:

(70) dègè   'ũ̂   **v**-ì
  from   there_far_off   AG8-COP
  it was from far off

<div align="right">[sayb001.452]</div>

(71) [kàmí ì-yũ̂u   ò-róonò]   **v**-ì,
  before 2P-wear\RLS   NC2-loincloth AG8-COP

  ì-rí-mpàa-nì   yí↓ = e-bèntêe   y-áyà-nà   sáa   gò?
  NC3-thing-this-NMLZ   AG3=NC2_loincloth_k.o.   AG3-come\RLS-PFV   or   TOP
  it was [before you wore loincloths], this thing of 'bante' came, or?

<div align="right">[sayb001.360]</div>

6.4.2 Neutral gender assignment

Some atypical controllers not only trigger AG8 agreement, but also occur with a non-null NC8 prefix allomorph (§5.5.7), as in the name of the town **K-káinwà** in (72):

(72) Máamá   ví-nà Ø-yó-nò   **K-káinwà**
  [name]   AG8-REL AG8-be\RLS-PFV   **NC8-[town]**
  *Maama who is at **Kainwa***

<div align="right">[sayb001.665]</div>

Adverbs too are assigned to gender 8 when functioning as nominals. In (73), where **dòorí** 'formerly' and **ká'à** 'now' both function as 'possessor' NPs, they occur with the NC8 lengthening prefix. At least one adverb, **géi** 'much' (§4.9), is always pronounced with a long **g** by some speakers, even when functioning adverbally (74).

(73) à-bíkì   há↓=d-dòorí,   ǹ   há↓=k-ká'à
NC2-festival   AG2=NC8-formerly   and   AG2=NC8-now

the festivals of before, and of now

[svtmg001.287]

(74) ń nì   Ø-yúu   sɔ̀ɔ   t-ì   g-géi
if when   2S-CONT   drink   AG6-PRO   NC8-much

if you're smoking it a lot

[tats007.002.045]

Another potential case of neutral 'noun class' assignment, at least historically, involves the possessive pronouns (§4.4.5.1.1). The forms of three of the six possessive pronouns (1P, 2P, and 3S) are exactly what one would expect if they were straightforward associative constructions (§4.4.5.1), with the object enclitics (§7.3) as possessors. Example (76) has been reanalysed from (4.137, repeated as (75)) accordingly:

(75)   (a) c-cɔ́'ɔ̀   ví-**ttù**   (b) kù-lácí   kú-**ɗɗò**
       NC8-sheep   AG8-**1P.POSS**   NC9-girl   AG9-**2P.POSS**
       our sheep (sg.)   *your (pl.) girl*

[repeated from 4.137]

(76)   (a) c-cɔ́'ɔ̀   ví=**t**=**tù**   (b) kù-lácí   kú=**ɗ**=**ɗò**
       NC8-sheep   AG8=**NC8**=**1P.PRO**   NC9-girl   AG9=**NC8**=**2P.PRO**
       our sheep (sg.)   *your (pl.) girl*

The similarity between the tone patterns here and those found in the standard associative construction has already been observed (§4.4.5.1.1), but thinking of the possessive pronouns in this way can also explain the long consonants in the 1P and 2P forms. Here they can be viewed as the result of the **C-** NC8 allomorph attaching to the pronominal forms **tù** and **ɗò**, analogous to the situation with regular nominal 'possessors' e.g. **c-cɔ́'ɔ̀ ví↓-k-káa** 'the sheep of the woman'. The same analysis can also be applied to the demonstrative modifiers (§4.4.5.2) and ordinal numerals (§4.8.3), which also have a H L tone pattern and lengthened consonants. See also the diminutive and augmentative pre-prefixes in §5.3.6[24].

We noted in §5.6 that Corbett assumes gender in Bantu languages to be largely morphologically-assigned. If we adopt this view, then it is perhaps better to consider the examples just mentioned as cases of neutral *noun class* assignment to NC8, with the AG8 agreement following straightforwardly from this.

---

24 Synchronically, the patterns observed here can be thought of as a kind of sub-morphemic 'eidemic resonance' (Bickel and Nichols 2007:209-210).

6.5 Complex NPs and variation in agreement

Textbooks usually illustrate agreement by demonstrating that varying the controller results in systematic changes on the target, e.g. *the boy **runs**, the boys **run***. However it is also possible for the features on the target to vary, even though the controller remains fixed. This indeterminacy can arise in two ways. The first is for a morphosyntactically-simple noun to alternately trigger agreement morphology encoding different values for gender, as we saw for the hybrid noun **má-gàjì** in §6.3. The second way, which will be discussed in this section, involves morphosyntactically-complex nouns or NPs which offer more than one choice of controlling gender, according to the features of their subparts.

Two sorts of complex nominal words have been observed. The first is pre-prefixed nouns (§5.3.6). Although we might expect the outer prefix to determine the gender of the noun and hence the agreements that it triggers, this is not always the case; example (77) shows agreement with the outer prefix, and (78) with the inner.

(77) **mà**-kù-náa **má**↓=k-káa   vé-evì
NC4-NC9-leg AG4=NC8-woman AG8-3S.POSS
the little leg of his wife

[saat002.002.664]

(78) kò-**cì**-kóotò **tí**-mpà **tì**-pénénêu
NC1-NC6-drum AG6-this AG6-small
this big drum

[saat001.008.073]

Secondly, some compound nouns may trigger different agreements. Either the compound is treated as internally analysable as in (79a, 80a), or as opaque as in (79b, 80b).

(79) (a) kwákúllè   kw-í
that_day   AG9-COP
it's that day [cf. full form **kw-ắa'ắ kú-llè** 'NC9-day AG9-that']

[saat001.008.008]

(b) kwắakúllè[25] v-í
that_day   AG8-COP
it's that day

[Tikula, sami001.283]

---

25 The first vowel of **kwắakúllè** varies in length and nasalisation – this is not thought to be related to the choice of gender.

(80) (a) ì-rí-mpàa-nì   **yé**-evì   **yì**-tò
NC3-thing-this-NMLZ AG3-3S.POSS AG3-one
his one thingummy

(b) ì-rí-mpàa-nì   **vì**-yápù
NC3-thing-this-NMLZ AG8-two
two thingummys

As well as morphologically-complex nouns, some NPs offer a choice of controlling gender. In the case of associative NPs (§4.4.5.1), agreement can be triggered either by the 'possessed' head noun or by the modifying 'possessor' NP. The phrases in (81a) and (81b) both contain the same associative NP **à-zá há↓-n-kácí** 'hunters', lit. 'people of hunting', but in the former the copula agrees with the NC2 head noun **à-zá** 'people', while in the latter it agrees with the internal NP **ǹ-kácí** 'hunting'.

(81) (a) **à-zá**   há↓=n-kácí   **h**-è
NC2-person AG2=NC5-hunting AG2-COP
it was hunters

(b) à-zá   há↓=**n-kácí**   **m**-ì
NC2-person AG2=**NC5-hunting** AG5-COP
it was hunters

Similarly in the conversational exchange in (82) the first speaker's copula agrees with the NC8 head noun of the associative construction, **Ø-lóokàcí** 'time', but the second speaker's copula agrees with the NC3 'possessed' noun **rú-umá** 'war'.

(82) A: **Ø-lóokàcí** Ø-rú-umá **v**-ì
NC8-time   AG8-NC3-war AG8-COP
it was the time of war

B: ée, Ø-lóokàcí **Ø-rú-umá** **y**-ì
yes NC8-time   AG8-**NC3-war** AG3-COP
yes, it was the time of war

The following example of an associative construction is presented as it was spoken. However the last two words in (83) were actually transcribed by a native speaker consultant as *tina tile*, agreeing with the inner 'possessor' noun **tì-wɔ́mɔ́** 'chieftancy'[26].

---

26 When queried, consultants most often say that agreement should be with the outer 'possessed' NP. While this is more frequent in the corpus, the exceptions are too common for me to be happy classing them all as speech errors. See Crozier (1984:94) for an example of 'inner agreement' in Central Kambari.

(83) **mà-hwáarì**   má↓＝tí-wɔ́mɔ́ **má-nà** **mè**-lêe
    NC4-start   AG4=NC6-chief AG4-ART AG4-there

that's the start of the chieftancy

[sayb001.297]

These examples appear to show 'trigger-happy' agreement (Comrie 2003), where conflicting controllers compete for a single target. Another potential case involves agreement 'out of' prepositional phrases. Consider the relationship between the NP **kù-jénè** 'river' and the agreement target **cé** in the following example:

(84) á↓＝kú-jénè **kù**-cé
    LOC=NC9-river AG9-NEG

it's not in the river

[Tikula, taff002.013]

We might have expected AG8 neutral agreement here with the PP as an atypical controller, but instead the negator seems to seek out the more prototypical NC9 controller **kù-jénè** 'river', regardless of the fact that the scope of negation is the PP, not the NP. It is not known whether neutral agreement would be a possibility in (84) – if so, then this is another case of trigger-happy agreement.

The next two examples are similar in that the agreement controllers are embedded within a PP. Again it is not known whether AG8 agreement (i.e. with the PP as controller) could occur in (86), but these constructions are at least candidates for trigger-happy agreement.

(85) á↓＝Ø-'áttáurá **vì**-hyẫa z-zá   n-nà   pɔ̂ à-ráatè-nè
    LOC=NC8-**Torah** AG8-say\RLS NC8-person AG8-REL all 3P-hang\RLS-PFV

in the Torah it says the one that they hanged

[oamy001.219]

(86) á↓＝**mɔ́-kkɔ̀'ɔ̀** **mà**-hyẫa...
    LOC=NC4-**cup** AG4-say\RLS

on the cup it said...

[2008-04-01.005]

6.6 Gender resolution

'Gender resolution' is a term used by Corbett (1991:261ff) to describe what happens when two agreement controllers are conjoined, and then together trigger agreement on a target. As discussed in McGill (2007:84-85) Cicipu speakers tend to find ways of avoiding this structural configuration altogether. In summary, two constructions are marginally acceptable, illustrated by (87) showing agreement with the first and more

distant conjunct, and (88) showing syntactic resolution (the verb agrees with the plural form of the first conjunct), although the former may involve a comitative construction *e.g. the sword [with the spear]*, and the latter was not accepted by all consultants. Further research would be required to rule out the comitative analysis (see Haspelmath 2004 for tests).

(87) **mà**-gɑ̃́i ǹ kù-sáyù **mò**-yúwò-nò
      NC**4**-sword and NC9-spear AG**4**-fell\RLS-PFV
      the sword and the spear fell

[2007-01-31.001]

(88) ?**mà**-gɑ̃́i ǹ kù-sáyù **mì**-yúwò-nò
      NC**4**-sword and NC9-spear AG**5**-fell\RLS-PFV
      the sword and the spear fell

[2007-01-31.001]

Part IV – Gender and person agreement in Cicipu

Chapter 7 – Gender and person agreement on person markers

The previous chapter gave a high-level overview of the different agreement targets found in Cicipu. This chapter will now focus in on a subset of these, specifically the different paradigms of person markers. I use the phrase 'person marker' in the same way as Siewierska (2004), as a cover term for independent pronouns, pronominal clitics and affixes, and grammatical agreement markers. Here I consider the Cicipu gender-marked affixes and pronouns to be person markers which inflect for the grammatical category of GENDER (as in Siewierska's (2004:104) analysis of Kiswahili noun class pronouns). It is important to stress that in what follows the term 'person marker' is used to cover both 'person-marked' person markers (which inflect for PERSON) and 'gender-marked' person markers (which inflect for GENDER).

A distinction is made here between the pronouns, clitics, and agreement markers just mentioned, and other word classes such as the article and the demonstratives. Although the latter may substitute for nouns as the head of a noun phrase (and therefore can be referential), they have additional syntactic functions and are treated separately in §8.9. This leaves five distinct paradigms of person markers:

Table 47: Summary of Cicipu person markers

	Subject prefix	**Post-verbal object**	**Elsewhere**
Person	person subject prefixes	object enclitics	personal pronouns
Gender	gender subject prefixes	gender pronouns	

There are three different PERSON-marked paradigms depending on the morphosyntactic position: subject prefix, post-verbal object enclitic, and elsewhere. For GENDER-marked forms there is a two-way distinction between subject prefix and other environments. The first six sections of this chapter describe the properties of these five paradigms. For each of the pronoun/clitic paradigms (§7.1-7.3) I will discuss its morphological and syntactic status according to the typology of agreement markers outlined in §2.2.5. Sections 7.4-7.5 consider the phonological and morphological properties of the two paradigms of subject agreement markers. The arguments concerning the syntactic status of these two paradigms are complex and are discussed in §7.6. After an interim summary in §7.7, I argue (§7.8) that there really are separate paradigms of gender and

person agreement markers, rather than a single complex paradigm for each syntactic position.

Together with chapter 8, this chapter forms Part IV of the thesis. The division of labour with respect to the characterisation of the gender/person alternation on person markers can be summarised as follows. This chapter (especially §7.6) is concerned with the AGREEMENT PREREQUISITES (§2.2.4.1) necessary for the two different kinds of agreement. There turn out to be a number of prerequisites involving lexical, phonological, and morphosyntactic factors. They are summarised in flowchart form in §7.6.1.3. Chapter 8, on the other hand, is concerned with the AGREEMENT CONDITIONS which bear on the *choice* of gender or person agreement in those contexts where the prerequisites for both are met. These conditions involve the higher levels of linguistic structure: morphosyntax, semantics, and pragmatics.

7.1 Independent personal pronouns

The independent personal pronouns show a straightforward six-way person/number distinction. There is no inclusive/exclusive 1PP distinction.

Table 48: Independent personal pronouns

Person	Singular	Plural
1	ámú	ótú
2	ìvɔ́	ìdó
3	é *or* éví	éré

The tone melody on these pronouns varies according to both the phonological and the syntactic environment in a way which is not yet well-understood. In isolation the tone pattern seems to either L H or H H for second person, and H H for the others.

7.1.1 Morphology

There are at least three reasons to classify the forms given above as 'independent' person markers[1]. First, they can occur as complete utterances:

(1)   àmú?!
       1S.PRO
       me?!

[saat001.008.120]

Second, they can occur in co-ordinated constructions, as in (2-3). If the first conjunct is

---

1   See Siewierska (2004:18) and Dixon and Aikhenvald (2002) for the tests used here.

a pronoun, then it occurs in the plural form, although the meaning is ambiguous between singular and plural (see §4.4.5.4).

(2)   **èrè**   ǹ   Ø-lɔ́wɔ́lì   ò-tôonò
   3P.PRO   and   NC8-large_spider   3P-go_home\RLS
   he and Big Spider went home

<div align="right">[saat001.006.203]</div>

(3)   **òtù**   n=ìvɔ́
   1P.PRO   and=2S.PRO
   me and you (sg.)

<div align="right">[2006-12-08.005]</div>

A third indication that this group of person markers are independent is that they can occur as the head of noun phrases, and be modified by other words. The third-person forms (at least) can co-occur with the article, in which case they trigger AG8 agreement as in (4). This is best seen as a form of neutral agreement (§6.4).

(4)   [Context:The speaker has been asked which ones are the slaves]
   áa   Ìsháyà   ǹ   Múusá,   **érè**   **n-nà**   ò-módó   há-nà
   PL   Ishaya   and   Musa   **3P.PRO**   **AG8-ART**   NC2-slave   AG2-ART
   *Ishaya and Musa, **they** are the slaves*

<div align="right">[2008-03-07.003]</div>

The 3PS variant **ɛ́** is more often found in morphologically-complex words such as the demonstrative pronouns **ɛ́-llè** and **ɛ́-"índè** (§4.4.3.4), but it may also occur as a free form, in apparent free variation with the longer form **evi**:

(5)   **ɛ́**   dá'à   w-áyà   ù-tʉ́wɔ̃̀
   3S.PRO   moreover   3s-come\RLS   3s-pour\RLS
   and then he poured

<div align="right">[tapf002.004.047]</div>

(6)   **èvì**   dá'à   ù-kóndò
   3S.PRO   moreover   3s-enter\RLS
   and he entered

<div align="right">[saat002.002.431]</div>

7.1.2 Syntax

The independent personal pronouns are the only person-marked person forms that can occur in non-argument positions:

(7)   dùkwá   n=àmú!
   go\IMP   with=1S.PRO
   go with me!

<div align="right">[saat001.006.203]</div>

Occasionally independent pronouns occur in subject (i.e. pre-verbal) position, but as with many 'pro-drop' languages, such constructions are not pragmatically-neutral. Instead they mark some kind of pragmatic relation involving the subject referent, as in (8), where **ìvɔ́** is in argument focus.

(8)   [Context:The king discovered who killed the crocodile]

      ìvɔ́   Ø-húnà-nà   v-ì
      2s.PRO 2s-kill\RLS-PFV AG8-PRO
      YOU killed it

                                                    [saim001.034]

Very rarely, the complement position of the verb phrase may be filled by an independent personal pronoun at the expense of the usual object enclitic (§7.3). I have only found two examples of this, and in at least one of them, shown below, the sentence is pragmatically-marked, with no fewer than three left-dislocated topic NPs. The anaphoric chains in (9) are indicated with subscripts.

(9)   [z-zá   Ø-rùmónò n-nà]$_x$
      NC8-person AG8-black   AG8-ART

      [*the black one*]$_x$

      [éví n-nà   yó-nò   n=ù-úsũ̀]$_x$
      3s.PRO AG8-REL be\RLS-PFV with=NC7-power

      [*it which has power*]$_x$

      [**vésì mɔ̀-'yɔ́'yɔ́ má-nà   mǎ-ayà-nà**]$_y$ sée   [ú]$_x$-sɔ̀dù   [**èvì**]$_y$
      every NC4-fish   AG4-REL AG4-come\RLS-PFV then 3s-swallow\IRR **3s.PRO**
      [***every fish that came***]$_y$ [*it*]$_x$ *would swallow* [***it***]$_y$

                                              [Tikula, taff002.020]

7.2 Independent noun class pronouns

The independent noun class pronouns are listed below:

Table 49: Independent noun class pronouns in Cicipu

Class	1	2	3	4	5	6	7	8	9
Pronoun	ke	he	yi	me	mi	ti	wi	vi	kwi

7.2.1 Morphology

As with the personal pronouns, the noun class pronouns[2] are best analysed as independent words. The conjunction test is not very useful here since there are no examples in the corpus of co-ordinated noun class pronouns, and neither are they

---

2 See §6.2.14-6.2.15 for some remarks on the phonological properties of these pronouns.

attested as complete utterances. However two other tests are relevant. Firstly, noun class pronouns can occur as one part of an equational construction (10) and in appositional noun phrases (11).

(10) **w-í**   ù-láa
      AG7-PRO NC7-fire
      it is fire

[tats002.002.008]

(11) kúmá **kw-í**   kù-yúpù   kù-zámùkwà   ì-rí   yí↓=wɔ-ɔ́vɔ́ɔ
      and   **AG9-PRO** NC9-crocodile AG9-be\RLS   NC3-thing   AG3=NC7-fear
      and it the crocodile is a fearful thing

[tats002.006.012]

Secondly, noun class pronouns occasionally head complex noun phrases, as in (12), where the pronoun **ví** is followed by a modifying demonstrative:

(12) n-náa   n-nà   ù-yó-nò   v-ì,   [**v-í**   **ví-llè**]   ù-wîinà-nà
      NC8-cow   AG8-REL   3S-be\RLS-PFV   AG8-PRO   **AG8-PRO** **AG8-that** 3S-sell\RLS-PFV
      *the cow that he'd had, **that one** he sold*

[saim001.058]

7.2.2 Syntax

Noun class pronouns can occur in both argument and non-argument positions, although it is hard to find clear examples of the pronouns in subject position, even though they are accepted in elicitation sessions (13). From the corpus, (14) shows a noun class pronoun functioning as subject in a non-verbal clause, while (15) is a possible example of a verbal subject[3].

(13) **k-é**   kò-yúwò-nù   ó↓=kó-ocì
      AG1-PRO AG1-fall\RLS-RES   LOC=NC9-hole
      it [**kà-mángá** 'the rope, NC1'] *fell into the hole*

[2007-02-05.007]

(14) **v-í**   ù-láa   w-ì
      AG8-PRO NC7-fire   AG7-COP
      it [**Ø-gógóró** 'gin, NC8'] *is fire*

[tats002.008.027]

---

3 It would be better to have an example that ruled out inter-clausal interference.

(15) [Context: *that word* [**kà-dáa**, NC1] *is not hearsay...*]

á-yẫa   **k-ɛ́**   ká-wùtò-nò   á↓=Ø-rìkóodà
3P-do\IRR   **AG1-PRO**   AG1-go_out\IRR-VENT   LOC=NC8-recorder

*let them make **it** [the word] come out of the recorder*

<div align="right">[Tikula, sagb001.719]</div>

Noun class pronouns also occur in complement position, as shown below. In (16) the pronoun occurs in the complement position of the VP, while in (17) the pronoun is in the complement of the locative proclitic **A** (§4.5, §3.4.7). Example (18) shows a noun class pronoun functioning as the object of the preposition **ǹ** 'with' (§4.5).

(16) mà-gẫi   mé-llè,   ò-gútù   **m-è**   cé   á↓=kú-pé'ɪ́
NC4-sword   AG1-that   3P-put_back\RLS   **AG4-PRO**   NEG   LOC=NC9-scabbard

*that sword, they didn't put **it** back in the scabbard*

<div align="right">[samy001.067]</div>

(17) kà-táarí   ké-llè,   ǹ-tá'à   ń-kàbà-nà,   ǹ-yôo   é=**k-è**
NC1-stone   AG1-that   1S-want\RLS   1S-take\IRR-VENT   1S-go\RLS   LOC=**AG1-PRO**

*that stone, I wanted to bring it, I went to **it***

<div align="right">[2008-02-01.002]</div>

(18) n=à-kábà-nà   à-ràkúmí   à-dúkwà   ǹ   **h-ɛ́**
when=3P-take\RLS-PFV   NC2-camel   3P-go\RLS   with   **AG2-PRO**

*when they took camels they went with **them** [the camels]*

<div align="right">[samoh001.198]</div>

7.3 Object clitics

As well as the independent personal pronouns (§7.1), there is a corresponding set of six person-marked clitics found immediately after the verb, as in (19-20). These always express the object grammatical function.

(19) íg-gòonù=**mù**
2S.IRR-help\IRR=**1S.PRO**

*please help **me***

<div align="right">[saat002.001.097]</div>

(20) vàyí**vì**!
vayu-LH=**vì**
throw_away-IMP=**3S.PRO**

*throw **him** away!*

<div align="right">[saat001.006.203]</div>

The full paradigm is given in Table 50:

Table 50: Object clitics

Person	Singular	Plural
1	mù	tù
2	vù	ɗò
3	vì	rè

7.3.1 Phonology

The resemblance between these CV object clitics and the VCV independent personal pronouns (ámú 1ᴘꜱ, ìvɔ́ 2ᴘꜱ, éví 3ᴘꜱ, ótú 1ᴘᴘ, ìɗó 2ᴘᴘ, éré 3ᴘᴘ) is clear, and suggests that the former derive historically from the latter, having undergone a process of phonological reduction. Segmentally, the clitics are virtually identical to the second syllable of the independent pronouns, the only difference being a change of vowel in the 2ᴘꜱ form **vù**. Cross-linguistically such relationships between intra-language person paradigms are common (Siewierska 2004:251-255).

The tone on the object clitics is always L. This is even the case in the imperative mood, where otherwise the final H of the verbal word spreads on to the following syllable (§3.4.3), and so the L tone on the clitic should be regarded as lexically-assigned.

The 3ᴘꜱ form **vì** is of particular interest because of its phonological similarity to the ᴀɢ8 pronoun **vi**. They are formally indistinguishable in post-verbal position, apart from the fact that only the person-marked object clitic triggers the leftward *i*-spreading process described in §3.7.3. Example (21) is repeated from §3.7.3 (3.106).

(21) (a) A: hán   Aúdù?
          where   Audu
          where's Audu?

  B: míndìvì
      m-índà = vì
      1s-see\ʀʟꜱ=**3ꜱ.ᴘʀᴏ**
      I saw him

(b) A: hán   Ø-vóotò
          where   ɴᴄ8-goat
          where's the goat?

  B: míndà   vì
      m-índà   v-ì
      1s-see\ʀʟꜱ   ᴀɢ**8-ᴘʀᴏ**
      I saw it

7.3.2 Morphology

The object clitics share properties with both pronouns and canonical agreement affixes. Like pronouns they may not co-occur with a lexical NP object, and yet like affixes they have a fixed position with respect to the verb stem; they can only occur immediately to the right of the verbal complex. Although certain discourse markers may come between the verb and lexical NP objects, as in (22), this has not been observed for object clitics,

despite over a thousand tokens of the latter in the corpus.

(22) ù-dɔ́nɔ̀-nɔ̀ **dá'à** ú-yáa
 3s-follow\RLS-VENT **moreover** NC7-road
 and he followed the road

The object clitics can be distinguished not only from lexical NPs but also from other types of person marker, according to their position relative to the clausal negator **cé** (§4.3.5.1). As illustrated below, lexical NPs and independent personal pronouns always follow the negator (23-24), object clitics always precede it (25), and for noun class pronouns both possibilities occur (26-27)[4].

(23) à-gúyà **cé** **cí-cípù**
 3P-can\RLS **NEG** **NC6-Cipu**
 they don't know Cicipu

(24) m-ĭndà **cé** **èvì**
 1s-see\RLS **NEG** **3s.PRO**
 I didn't see HIM

(25) hú-u-gɔ̀ɔnɔ̀=**rè** **cé**
 3P-FUT-overcome\IRR=**3P.PRO** **NEG**
 they wouldn't overcome them

(26) mà-gɛ́i mé-llè, ò-gútù **m-è** **cé** á=kú-pé'í̵
 NC4-sword AG1-that 3P-put_back\RLS **AG4-PRO** **NEG** LOC=NC9-scabbard
 that sword, they didn't put it back in the scabbard

(27) Ø-'íttɔ̀ vé-evì Ø-dâa **cé** **w-ì**
 NC8-neighbour AG8-3s.POSS AG8-have **NEG** **AG7-PRO**
 his neighbour doesn't have it [**ù-pépí** (NC7) 'spirit']

Unlike the independent personal pronouns (§7.1), the object clitics cannot be separated from the right edge of the verbal word. For example they cannot be fronted (§4.3.2.1) – instead an independent personal pronoun must be used:

---

4 Example (27) is the only post-negation example involving a noun class pronoun in the corpus. Both structures can be elicited, although the pre-negation structure seems to be preferred.

(28) (a) ìvɔ́   vì   m-índà-nà
         2s.PRO   2s.COP   1s-see\RLS-PFV

         it was you I saw

   (b) *vù   (vì)   m-índà-nà
        2s.PRO   2s.COP   1s-see\RLS-PFV

[2008-02-01.002]

Further evidence comes from the various double-object constructions, including those made possible by the applicative suffix **-wA** (§4.6.4.2). As pointed out in §4.3.1.1 the object clitic can only express the first object – the second object slot must be filled by an independent pronoun, as in (29).

(29) àcîivì   èrè
      à-câa = vì   èrè
      3P-give\RLS=3s.PRO   3P.PRO

      they gave them to him

[2008-02-01.002]

As well as this positional restriction, the bound morpheme analysis is supported by the fact that object clitics cannot occur as complete utterances, even if followed by the copula.

   The evidence above suggests that the morphemes under discussion are bound to the verb. However they also differ from the more obvious verbal affixes discussed in §4.6, in at least three ways. First, object clitics are excluded from the domain of certain tonological process. Recall from §4.6.3.3 that for verbs 'heavier' than CVCV, the habitual tone pattern is L L H (L)*. For CVCV the tone pattern is L L H H. When (undisputed) suffixes are added to a CVCV root they contribute to the weight of the verbal word, and consequently the tone pattern changes from the light pattern to the heavy one. However when there is an object clitic after the verb, as in (30), the tone pattern is not affected, and we can conclude that the clitic does not affect the weight of the verbal word.

(30) ùsìwṍsívì   L L H H (light)
      u-si-wõsõ = vì
      3s-HAB-bark=3s.PRO

      it barks at him

[tats001.001.076]

This does not actually force us to decide between affixes, clitics and words, although we do know that if the "object clitics" *are* affixes then they are not in the first 'stratum' of

321

lexical phonology, since they must be ordered after the process which assigns tonal melodies to verbal words (cf. Kiparsky 1982 on English stress assignment and Level I/Level II affixes).

A second piece of evidence singling out the object clitics is the fact that they occur to the right of everything else that might be considered affixal, as illustrated below.

(31)   zzá   nnà   ùtóbìlìsìsùwòwònò**mù**   sháyì
         z-zá   n-nà   ù-tób<ìl><ìs><ìs>o-wò-wò-nò=**mù**   Ø-sháyì
         NC8-person AG8-REL 3S-cool\RLS<PLAC><CAUS><CAUS>-PASS-APPL-PFV=**1S.PRO** NC8-tea
        *the person who has caused tea to become cooled down in a forceful and iterative fashion for **me***

[Repeated from (4.191)]

Thirdly, there is evidence from the domain of vowel harmony. Recall from §3.5.3 that the vowel in Cicipu affixes is either neutral {i, u}, as in the resultative suffix **-nu**, or comes from the harmonising set {a, e, o, ɔ} like the applicative suffix **-wA**, in which case its phonetic value is determined according to the vowels of the verb root. In contrast, the vowels in the 2PP and 3PP object clitics **dò** and **rè** are invariant[5].

There *is* one argument in favour of the free pronoun analysis, and this is the fact that these markers can apparently function as NP heads, as in (32). Although such constructions were absent from the corpus, they were accepted without reservation in elicitation sessions.

(32)   m-ĭndì**vì**   **nínnîi**
         m-inda\LHL=**vì**   **nínnîi**
         1S-see\RLS=**3S.PRO**   **alone**
         *I saw **just him/her***

[2008-02-01.002]

On balance, there seem to be more reasons to view these person markers as clitics, rather than as affixes or pronouns.

Reduced object clitics

This is an appropriate point to introduce one further set of person markers, related to the object clitics just discussed, but with a more limited distribution. These three alternative forms are restricted to the singular, and each of them has unusual characteristics. In contrast to the markers just discussed, they are severely phonologically-reduced, and they are most often found before the negator **cé**. Following the intransitive example in

---

5   It should be noted that other clitics *do* harmonise (§3.5.3).

(33) the three reduced object clitics are presented in an elicited paradigm (34-36), and then in examples from the corpus (37-39).

(33) w-ǐndà   cé   INTRANSITIVE
3s-see\RLS NEG
He/she didn't see

[2008-02-11.003]

(34) w-ǐndà=**n**   cé   1PS
3s-see\RLS=**1s.PRO** NEG
*He/she didn't see **me***

[2008-02-11.003]

(35) w-ǐndà   **c**=cé   2PS
3s-see\RLS **2s.PRO**=NEG
*He/she didn't see **you (sg.)***

[2008-02-11.003]

(36) w-ǐndì   **c**cé   3PS
w-ǐndà=**ìC**   cé
3s-see\RLS=**3s.PRO** NEG
*He/she didn't see **him/her***

[2008-02-11.003]

(37) vú-u-zɔ̀ɔsɔ̀-wɔ̀=**n**   cé?
2s-FUT-laugh\IRR-APPL=**1s.PRO** NEG
*won't you flirt with **me**?*

[saat002.001.019]

(38) mú-u-càa   **c**=cé   k-káa   v-vôo
1s-FUT-give\IRR **2s.PRO**=NEG NC8-woman AG8-1s.POSS
*I won't give **you (sg.)** my wife*

[saat002.002.642]

(39) ǹgwáanùkwì   **c**cé   mázámázá
ǹ-gwáanùkwà=**ìC**   cé   mázámázá
1s-see\RLS=**3s.PRO**   NEG   quickly
*I didn't see **it** quickly*

[tats004.001.037]

The normal object clitics do sometimes occur before **cé**, as in (40), but this is rare, and speakers have stated a preference for the reduced forms in this environment.

(40) tú-u-hyàᷓa=**vù**   cé
1P-FUT-say\IRR=**2s.PRO** NEG
*we won't tell **you (sg.)***

[saat002.002.098]

323

Each of these reduced object clitics has special phonetic properties. The 1PS form **N** attaches to the end of the verbal complex resulting in a word-final consonant[6], otherwise unattested in Cicipu outside of ideophones. The 3PS form **i** also attaches to the end of the verbal word, replacing the final vowel rather than coalescing, since there is no increase in length. On the contrary, there is actually a decrease in length resulting from the gemination of the following consonant – the **c** of the negator **cé**. The 2PS form is odder still, being realised solely as the lengthening of the consonant immediately following the verb. Therefore the only difference between an intransitive sentence such as (33) and the corresponding sentence with a 2PS clitic object (35) is that the **c** consonant of the negator is lengthened. The difference in length can be seen in the following diagrams:

---

6   Unless **cé** is also considered to be a clitic. If so, it would be of the non-harmonising kind, like the non-reduced object clitics.
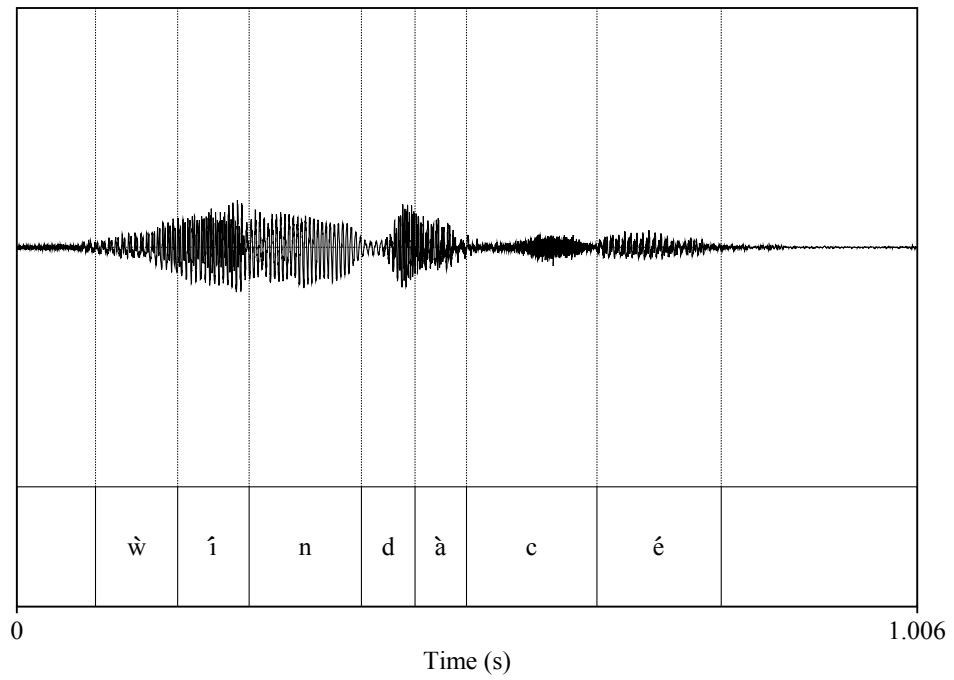
Figure 32: Waveform of example (33): **wíndà cé** *– I didn't see*
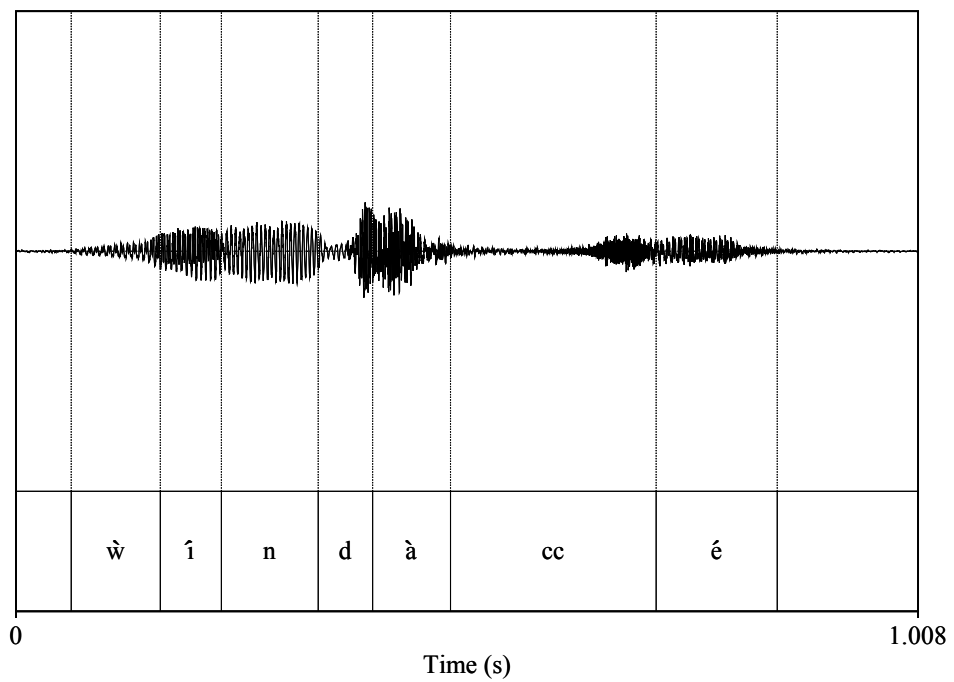
Figure 33: Waveform of example (35): **wíndà ccé** *– I didn't see you*

325

This kind of morpheme is cross-linguistically rare[7], although in Cicipu the same 'form' is also found as an allomorph of the 2PS subject agreement prefix (§7.4.1), and of the NC8 and AG8 prefixes discussed in §5.5.7 and §6.1.3. Since the 2PS reduced object clitic is mostly found before the negator **cé**, it usually appears as part of a long consonant [t:ʃ]. However the consonant-lengthening 2PS subject agreement prefix **C-** comes immediately before the verb, and its application can therefore result in the lengthening of *any* consonant. Therefore it seems better to also view the 2PS reduced object clitic as an underlyingly-unspecified weight unit **C**, which just happens to always occur before /**c**/.

The 3PS form (at least) is also found as the first object in double-object constructions, either with (41) or without (42) the applicative suffix **-wA** (§4.6.4.2):

(41)　ùyẫawì　　　　　cìgẫa
　　　u-yãa-wA = ì　　cì-gẫa
　　　3S-do\RLS-APPL=**3S.PRO**　NC6-quarrel
　　　*he quarrelled with **her*** [lit. 'did her quarrel']

　　　　　　　　　　　　　　　　　　　　　　　[saat001.004.021]

(42)　àcîi　　　　　màsùuríyà
　　　à-câa = ì　　mà-sùuríyà
　　　3P-give\RLS=**3S.PRO**　NC4-flute
　　　*they gave **him** a little flute*

　　　　　　　　　　　　　　　　　　　　　　　[saat001.003.049]

While these reduced forms are most likely to occur between the verbal word and the negator **cé**, or in double-object constructions, they are also sometimes found before the adverb **pâa** 'here' (43-44). Here, however, the reduced forms are the exception rather than the rule, and it is more usual to find the normal object clitics.

(43)　ǹ mò-ní　mà-kábà = m　　pâa mà-yâa　ǹ　Túngán Káɗé
　　　if NC4-water　AG4-take\RLS=**1S.PRO**　here AG4-arrive\RLS with [town]
　　　*if water took **me** from here to Tungan Kaɗe*

　　　　　　　　　　　　　　　　　　　　　　　[tats002.006.034]

7　Grammatical morphemes consisting of just a consonantal weight morpheme *are* found (e.g. the Arabic causative – Haspelmath (2002:22), see also Blevins 2004 chp. 7), but I am not aware of any agreement markers or incorporated pronouns expressed in this way. Morphemes consisting of a vowel weight morpheme seem to be more common; for example, one of Cicipu's West Kainji relatives Pongu uses a vowel weight morpheme as a 2PS object agreement prefix (James MacDonell p.c. 2008). In Hausa most members of the independent series of pronouns differ from their object pronoun counterparts simply by an increase in vowel length (Newman 2000:476), although this should probably not be viewed as the result of applying a vowel weight morpheme.

(44) ń-nàhà   p=pâa
      1s-leave\IRR **2s.PRO**=here

   *let me leave **you (sg.)** here*

| ńn | à | h | à | pp | âa | |

0   0.5218

Time (s)
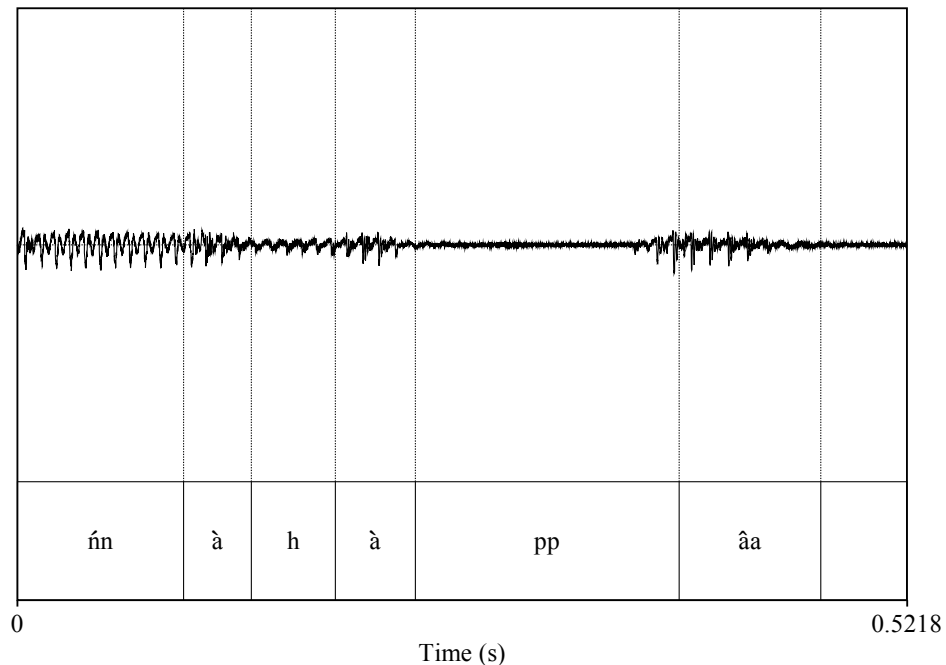
Figure 34: Waveform of example (44): **ńnàhà ppâa** *'let me leave you (sg.) here'*

The frequency with which these reduced clitics are found as opposed to the normal object clitics seems to be correlated with the syntactic relationship between the marker and the word that follows, in that the more tightly bound the environment is syntactically to the following word, the more likely it is to be filled by the reduced forms. The following table may clarify matters:

Table 51: Frequency of reduced forms as opposed to normal object clitics

Following word	**Frequency[8] of reduced forms**
Negator	Almost always
Secondary object	Frequent
Adverb	Possible
[*utterance final*]	Almost never

It seems that the reduced forms occur in environments where there is likely to be pressure exerted due to the compression of speech within intonation groups, the

8 These frequencies are currently impressionistic.

327

boundaries of which are influenced to a large extent by syntactic structure (e.g. Cruttenden 1997: chp. 4). It should be noted that it *is* possible for the reduced 3PS clitic to occur utterance-finally, as in (45), although it is extremely rare. The reduced 1PS clitic has not yet been observed here, and of course it would be logically impossible to find the reduced 2PS proclitic **C** utterance-finally.

(45) víndà   mòní   múur̀i
     v-índà   mò-ní   mú-u-ràa = **vi**
     2S-see\RLS   NC4-water   AG4-FUT-eat\IRR=**3S.PRO**
     *you see the water will overcome **him***

[Tikula, sagb001.165]

7.3.3   Syntax

Crozier (1984:123) makes two observations about the Central Kambari "object agreement suffixes[9]": (i) they are sufficient by themselves to refer to the object referent, and (ii) they cannot co-occur with a lexical object NP. The Cicipu object clitics are identical in these respects. We have already seen that they are sufficient to refer to the object referent (e.g. 19-20), and like Central Kambari they cannot co-occur with a lexical NP, i.e. 'clitic-doubling' is impossible. There is no reason to think that they are not referential, nor that they occur with anything other than the object grammatical function. Not all objects, however, are expressed by these object clitics. As we saw in (28-29) above, focused and secondary objects are expressed by independent pronouns instead.

7.4   *Person subject prefixes*

The gender/person alternation is perhaps most obvious for the subject prefixes, since the gender and person subject prefixes are more directly comparable than, say, the gender-marked independent pronouns vs. the person-marked object clitics. The person subject prefix marks the standard six person/number distinctions, with the 2PS paradigm cell having separate prefixes for realis and irrealis moods.

---

9   Crozier (1984:123) analyses the Central Kambari equivalents of the Cicipu object clitics as suffixes, although explicit reasons for this analysis are not given.

Table 52: Person subject prefixes (before consonants/before vowels)

	Singular	Plural
1	N-/m-	ti-/t-
2	RLS [C-, Ø-]/v- IRR [iC-, Ø-]/v-	i-/y-
3	u-/w-	A-/h-

7.4.1 Phonology

Before vowel-initial verb stems the usual coalescence applies (§3.7.1): the prefix vowel merges with the root vowel, resulting in a long vowel with the same quality as the root vowel. Recall from §3.4.6 that the tone on the prefix depends on the mood of the clause, the sonority of the prefix consonant, and whether or not the verb stem is consonant-initial or vowel-initial. The possibilities for 3PS subjects are set out below:

Table 53: Verb prefix tone patterns with a 3PS subject

	C-initial verb stem **dukwa** 'go'		V-initial verb stem **aya** 'come'	
Realis	ùdúkwà	L H L	wǎːjà	LH L
Irrealis	údùkwà	H L L	wâːjà	HL L

Perhaps the most interesting thing to be said about the phonology of the person subject prefixes concerns the 2PS forms before consonant-initial stems. In the realis mood the prefix is either **Ø-** or the consonant-lengthening **C-** familiar from the discussion of NC8/AG8 prefixes (§5.5.7, §6.1.3), as well as the reduced 2PS and 3PS clitics discussed in the previous section. The variation between the **Ø-** and **C-** 2PS subject allomorphs is complex, as was the case for the NC8/AG8 prefixes. Again the lengthened consonant is most clear word-medially, and in normal, fast speech word-initial long consonants are often undifferentiated from short ones. However the difference in length between examples such as (46) and (47) comes out clearly in careful speech.

(46) léllé **t-**tá'à
     truly 2S-want\RLS
     truly you (sg.) wanted

[eamy036.001.002]

(47) ìdó ì-**t**á'à
     2P.PRO 2P-want\RLS
     truly you (pl.) wanted

[sayb001.367]

The following examples show a lengthened voiced affricate (48), liquid (49), voiceless fricative (50), and glottal stop (51).

(48)   **j-j**ântà
     2s-crush\RLS
     you (sg.) crushed

                                             [eamy036.001.020]

(49)   **l-l**áttà
     2s-sleep\RLS
     you (sg.) slept

                                             [eamy036.001.015]

(50)   **s-s**âabà
     2s-used_to\RLS
     you (sg.) are used to it

                                             [eamy036.001.010]

(51)   nì = **'-'ɔ**pɔ̀
     when=**2s**-hold\RLS
     *when **you(sg.)** hold*

                                             [tats005.001.030]

In the irrealis mood, the 2ᴘꜱ prefix can be written **iC-**: the root consonant is lengthened just as in the realis mood, but there is also an **i** vowel preceding the geminate. Compare the geminate **d** in the singular in (52) with the short **d** in the plural in (53):

(52)   **íd**-dàmà-wà
     2s.IRR-speak\IRR-APPL
     you (sg.) should tell

                                             [tapf001.003.002]

(53)   **í**-dàmà-wà
     2ᴘ-speak\IRR-APPL
     you (pl.) should tell

                                             [sagb001.254]

There is dialectal and even idiolectal variation here. In Tikula the irrealis 2ᴘꜱ prefix is **vi-** rather than the **C-** lengthening prefix, and for some Tirisino speakers the 2ᴘꜱ prefix is **viC-**, combining elements of what seems to be the usual Tirisino pronunciation and of the Tikula form.

     The existence of these consonant-lengthening prefixes in both nominal and verbal morphology allows us to re-evaluate De Wolf's (1971) hypothesis about the historical derivation of the Kambari long consonants. This has been done in McGill (n.d.), where I

conclude that the lengthening prefix **C-** is derived from **vi-** by the loss of a vowel and the total assimilation of **v** to the following consonant.

7.4.2 Morphology

The morphological status of the person subject prefixes is more straightforward than that of the object clitics (§7.3). They show none of the expected properties of free words, since they cannot occur independently of the verb stem, cannot be conjoined with other words, and never occur with modifiers. They always appear in the same position: either prefixed to the verb root, or to the future (**u-**) or habitual (**si-**) prefixes if they are present. In addition, the 3PP prefix **A-** harmonises with the verb root in the same way as other affixes.

7.5 Gender subject prefixes

The nine gender subject prefixes are given below:

Table 54: Gender subject prefixes

Class	1	2	3	4	5	6	7	8	9
Prefix	kA-	hA-	yi-	mA-	mi-	ti-	wu-	Ø-/vi-	ku-

7.5.1 Phonology

Before vowel-initial verbs the usual rules of vowel coalescence apply, just as for person subject prefixes. In the realis mood the prefix tone is usually L, but sometimes a preceding H will cause it to rise. This behaviour contrasts with the person agreement prefixes which seem to be consistently L (§3.4.3).

Six of the nine gender agreement prefixes are remarkably similar in form to the six person prefixes just discussed, particularly before vowel-initial verbs. The ambiguous forms are highlighted in the tables below.

Table 55: Partial ambiguity of person prefixes with respect to gender prefixes, before consonant-initial stems

Prefix	Category
kA-	AG1
hA-/A-	AG2/3P
yi-/i-	AG3/2P
mA-	AG4
mi-	AG5
N-	AG5/1S
ti-	AG6/1P
wu-/u-	AG7/3S
vi-	AG8
Ø-	AG8/2S
ku-	AG9

Table 56: Total ambiguity of person prefixes with respect to gender prefixes, before vowel-initial stems

Prefix	Category
kV-	AG1
hV-	AG2/3P
yi-	AG3/2P
mV-	AG4/AG5/1S
ti-	AG6/1P
wV-	AG7/3S
vV-	AG8/2S
ku-	AG9

It is in the ambiguity between the AG2 and 3PP prefixes that the real difficulty for this study lies, since speech act participants are not normally referred to using gender agreement prefixes, and the semantic properties of the referents of NC7 nouns mean that they rarely trigger 3PS agreement (see §8.4.3). NC2, on the other hand, is by far the most common plural class for all kinds of nouns. In normal speech the AG2 prefix **hA-** can be very hard to distinguish from the 3PP agreement prefix **A-**, even before consonant-initial stems, since word-initial /**h**/ often elides intervocalically (§3.7.1). The problem is made worse by the fact that some of the most commonly occurring verbs have vowel-initial roots, for example **aya** 'come', **inda** 'see', and **uto** 'go out', as well as the future tense prefix **u-**, and here the gender-marked and person-marked verbs are identical.

332

Ambiguous cases of plural subject agreement are omitted from the analysis presented in chapter 8.

7.5.2 Morphology

The gender subject prefixes behave identically to the person subject prefixes in tests for morphological status, and so they too are considered to be affixes.

7.6 Syntactic status of the subject agreement prefixes

The main aim of this section is to categorise the two paradigms of Cicipu subject prefixes according to the three-way typology of agreement markers which was set out in §2.2.5. Corbett (2003b:184-192) provides several tests to distinguish between grammatical agreement and pronominal incorporation, building on earlier work by Bresnan and Mchombo (1987), Siewierska (1999), Evans (1999, 2002) and others. In the first five subsections of this section I will take each of Corbett's tests in turn, applying them to both of the subject prefix paradigms. In §7.6.6 I will briefly consider the results of two diagnostics offered by Bresnan and Mchombo. Subsection 7.6.7 is concerned with non-overt agreement, and asks whether it is better to posit null prefixes or to assume the absence of agreement altogether. Finally §7.6.8 compares the status of the Cicipu agreement markers to their counterparts in the related language Central Kambari.

7.6.1 'Multi-representation'

'Multi-representation[10]' refers to the possibility of a referent being indexed more than once in the clause. If a verbal marker can co-occur with a pronominal or lexical subject, this suggests it is a grammatical agreement marker rather than an incorporated pronoun. Similarly, according to Bresnan and Mchombo's (1987:752) "locality principle", grammatical agreement must be local to the predicate (i.e. the agreement relation holds between elements of the same clause), and therefore only anaphoric agreement can be non-local.

As will be seen in the discussion below, multi-representation is possible with both kinds of subject prefix in Cicipu, although it is more common with gender subject marking than with person subject marking. By this measure both gender and person agreement markers in Cicipu come out as *ambiguous* agreement markers according to

---

10 Corbett (2003b:185).

Bresnan and Mchombo's typology, although they are located at different points on this intermediate scale between pure grammatical agreement and pure anaphoric agreement. Gender agreement in the corpus is usually (but not always) a local relation, while person agreement is usually (but not always) non-local.

In the case of an explicit subject NP (i.e. grammatical agreement) there is usually no choice of agreement feature, and the next two subsections will make it clear which types of subjects can co-occur with each type of agreement.

7.6.1.1 Person subject prefixes

Determining whether a particular NP is a grammatical subject is not always straightforward for 'pro-drop' languages like Cicipu. The ability to trigger subject agreement is not pertinent since both gender and person agreement can be anaphoric, and so one cannot be sure that the 'subject' NP is within the clause – it might be a clause-external adjunct, as in Baker's (1996) analysis of polysynthetic languages[11]. However, evidence for the subjecthood of the NP triggering the verb agreement marker comes from the existence of an infinitive (§5.4.1) in Cicipu. Consider the following complex sentence:

(54) éré   hú-u-gùyà   cé   ù-zártà
      3P.PRO 3P-FUT-can\IRR   NEG   NC7-deprive_of_share
      they won't be able to deprive us of our proper share

[svbg001.035]

Here only the main clause verb **guya** takes a subject agreement prefix; the second verb **zarta** 'deprive s.o. of their proper share' appears in the invariant infinitive form **ù-zártà**, the subject agreement prefix having been replaced by an NC7 prefix. I assume, with Baker (1996:25, see also Kroeger 2004:104-106), that all infinitival clauses have an obligatory subject – since there is no agreement prefix to carry the subject grammatical function, the only other possibility is for there to be a null subject ('PRO') in the subordinate clause. The subject of the infinitival clause in Cicipu is always understood as co-referential with that of the main clause, and so PRO must be controlled by a main clause subject, in this case **éré** '3P.PRO'.

Further evidence that true subject NPs are possible in Cicipu comes from the impossibility of subject agreement being triggered by post-verbal NPs, unless they are offset from the clause by a pause. If the NP is a true subject then we would expect it to

---

11 According to Baker these have no true subject NPs, only adjuncts.

occur in the specifier position of the clause, i.e. preverbally, and this is precisely what we find. If, on the other hand, the 'subject' NP were an adjunct then we might expect to find greater variability in word order (see Baker 1996:101-102 for this argument applied in reverse to Mohawk).

The above tests are useful in demonstrating that the person subject prefixes have the *possibility* of taking part in grammatical agreement, but they are of limited use when it comes to encoding a corpus and deciding whether individual prefix tokens are instances of grammatical or anaphoric agreement. Therefore when encoding the corpus[12] I used a purely phonological diagnostic – if there was a pause between the pre-verbal NP and the verb then I assumed that it did not bear the subject grammatical function; if there was no pause then I assumed that it did, unless the NP was identifiable as a fronted constituent (§4.3.2.1). In certain cases (e.g. when both arguments have the same gender) fronted objects may be syntactically indistinguishable from focused subjects, in which case one must rely on the semantics to identify the grammatical relations.

The resulting counts show a strong tendency for person agreement to be anaphoric rather than grammatical. There are 3,042 occurrences of the 3PS[13] prefix in the corpus, and only 321 of these (11%) have an overt subject NP, according to the phonological diagnostic just described. In other words, the 3PS prefix is largely in complementary distribution with lexical NPs. Only certain kinds of NPs can be involved in grammatical agreement, as can be seen from Table 57; in the majority of the 321 cases the subject NP was headed by either an independent personal pronoun (§7.1), a demonstrative personal pronoun (§4.4.3.4), or the word **z-zá** 'person'.

---

12 See §1.4 for details of the corpus.

13 The figures exclude the Tikula texts (approx. 15% of the total corpus), which are yet to be interlinearised. Analysis of the distribution of the 3PP prefix is complicated, since the distinction between it and the AG2 prefix (the most common gender agreement prefix for plural controllers) is often neutralised – see §7.5.1. Accurate figures have not been compiled for first- and second-person prefixes, but a brief inspection of the texts suggest that co-occurrence with a subject NP is even rarer.

Table 57: Distribution of subject NPs co-occurring with 3PS subject agreement

Head noun of subject NP	Tokens in corpus
3PS personal pronoun (**éví**)	150
3PS demonstrative pronoun (**é-mpè**, **é-llè**, etc...)	29
z-zá 'person'	100
NC8 subject with V-initial verb	38
NC8 subject with C-initial verb	4
NC4 subject	1
Total	**321**

The independent personal pronoun **éví** accounts for around half of the cases of 3PS grammatical agreement. Rather than being set off from the clause by a pause, as we might expect for extra-clausal topicalised NPs, the pronoun actually coalesces with the person agreement prefix:

(55) évùudâanà   ù'úsǜ   sáa   yàanú
   évì   **ù-dâa-nà**   ù-'úsǜ   sáa   yàanú
   3S.PRO   **3s-surpass\RLS-PFV**   NC7-power   or   who
   he is greater than everyone

                   [eabg001.008]

The same is true for the 1PS and 1PP prefixes.

(56) àmúngúyà   cé
   àmú   **ǹ-gúyà**   cé
   1S.PRO   **1s-can\RLS**   NEG
   me I can't

                   [eaim003.1394A]

(57) òttúutònò   lêe   'ásù   wíivè
   òtù   **tì-útò-nò**   lêe   'ásù   wí-ivè
   1P.PRO   **1P-go_out\RLS-VENT**   there   place(NC7)   AG7-3P.POSS
   we came from their place there

                   [svbg001.025]

The 3PS demonstrative pronouns also trigger person subject agreement:

(58) éllùkôo
   é-llè   **ù-kôo**
   3S.PRO-that   **3s-die\RLS**
   that one died

                   [samoh001.259]

In addition to the personal and demonstrative pronouns, the lexical item **z-zá** also

triggers person subject agreement:[14].

(59) w-áa   á↓=d-dòorí   **z-zá**   ù-gúyà   ù-híyà   kwáanù   góomá   sű?
     3S-REPRT LOC=NC8-formerly **NC8-person** 3S-can\RLS NC7-grind bowl   ten   Q
     *it's said that in the olden days **one** could grind ten measures?*

[svtmg001.018]

(60) t-èné   t-ì   **z-zá**   w-ú-yằa-wà=vù   kà-líipì?
     AG6-which AG6-COP **NC8-person** 3S-FUT-do\IRR-APPL=2S.PRO NC1-wrong
     how can anyone do you wrong?

[oamy001.060]

It could be argued that what have been classed as subject NPs in (55-58) are actually topicalised and extra-clausal, despite the lack of a pause between the NP and the verbal word. However the subject referents in (59-60) are very bad candidates for topichood, since in both cases the referent is non-specific. The fact that person subject prefixes occur here is strong evidence that they can take part in grammatical as well as anaphoric agreement.

Turning back to Table 57, 38 of the remaining 43 cases of grammatical person agreement involve the combination of an NC8 controller[15] with a vowel-initial verb stem, as in (61-62):

(61) Ø-wɔ́mɔ́   vú↓=u-rée   **w**-ăyà
     NC8-chief AG8=NC7-town 3S-come\RLS
     The chief of the town came

[saim001.022]

(62) Ø-kùrì'áanì **wú**-u-'ɔ̀pɔ̀ z-zá
     NC8-Koran   3S-FUT-hold   NC8-person
     the Koran will hold a person

[tats005.001.134]

These **w-** prefixes certainly look like person agreement, but there is one other possibility that should be considered, if only to rule it out. Since vowel-initial syllables do not readily occur on their own in Cicipu, could it be that the **w-** prefix has simply been added as a kind of default? As noted in §4.6.2.3, a dummy approximant is added when the imperative of vowel-initial verbs is formed. So for example the root **uwa** 'feel' becomes **wùuwá!** in the imperative. Similarly before **i**-initial verb roots, a **y-** is added:

---

14 Despite its lexical meaning of 'person', **z-zá** is more often found as a grammaticalised morpheme meaning something like 'one' or 'the one'.

15 These examples include both inherent NC8 nouns, and atypical controllers (e.g. names of people) that have been assigned to this gender (see §6.4).

inda 'see' becomes **yìndá!**. Actually it is simple to show that the **w-** glossed as subject agreement in (61-62) is not there by default, since it is consistently **w-**, even before [i]. So instead of *kúngwá y-índà in (63) we have **kúngwá w-índà**:

(63) Ø-kúngwá   **w-**ĭndà
      NC8-God   3s-see\RLS
      God considered

[tats001.004.073]

This association between vowel-initial verbs and person agreement may be motivated by the fact that there are fewer agreement choices before vowel-initial verbs compared to before consonant-initial verbs. Recall from §6.1.3 that before consonant-initial verbs there are two AG8 allomorphs for the subject prefix: **vi-** or **Ø-**. As is the case across the agreement targets in general, the likelihood of **vi-** occurring increases as the animacy of the controller referent decreases. This is reflected in Table 58, which shows the frequency[16] with which the **vi-** and **Ø-** AG8 subject agreement prefixes take part in grammatical agreement before consonant-initial verb stems.

*Table 58: Distribution of the **vi-** and **Ø-** allomorphs of the AG8 subject prefix before C-initial verb stems, according to animacy of referent*

Type of referent	No. of occurrences of vi-	No. of occurrences of Ø-
Human/spirit	2 (3%)	66 (97%)
Animal	3 (17%)	15 (83%)
Inanimate	11 (20%)	44 (80%)

This distinction between **vi-** and **Ø-** is no longer available when the verb stem begins with a vowel – the **vi-** prefix remains, but the null prefix cannot occur since Cicipu nominal and verbal words do not in general begin with a vowel. Instead it seems that nouns which would have triggered the **Ø-** AG8 allomorph before a consonant-initial verb actually trigger the person agreement prefix **w-** before a vowel-initial verb. So the function (that of distinguishing animate[17] from inanimate controller referents) of the agreement is preserved regardless of the phonological structure of the verb, but the feature reflected on the target is not. The situation is diagrammed below:

---

16 Again based on the corpus minus the Tikula texts.

17 It is likely that discourse topicality is involved as well – see chapter 8 for the interaction of animacy and discourse topicality with respect to the alternation between gender and person agreement in general.

Table 59: Subject agreement prefixes triggered by NC8 controllers

	Lower in animacy	Higher in animacy
C-initial verb stems	**vi-** (AG8)	**Ø-** (AG8)
V-initial verb stems	**v-** (AG8)	**w-** (3PS)

In summary, person agreement seems to have been pressed into service in order to preserve a distinction in animacy/topicality in an environment in which normally only gender agreement is found.

This still leaves five cases of 3PS subject agreement unaccounted for. Given that there are 3,042 instances in the corpus, it seems reasonable to put them down to speech errors. However there is a certain amount of regularity in these exceptions. Four of them involve NC8 subjects with consonant-initial verb stems – all of which have unique referents – **kúngwá** 'God', **Yéesù** 'Jesus', **d-dáa** 'the king', and **bâammí** 'palm-wine'[18]. The fifth example is shown in (64) below. Here the subject is the NP **mɔ́ɔ méllè** 'that child, NC4', and this unexpectedly triggers person rather than gender agreement on the auxiliary verb **aya**.

(64)   éllè,   mɔ́ɔ   méllè **wǎyà**   ùcîivì
           é-llè,   m-ɔ́ɔ   mé-llè **w-ǎyà**   ù-câa = vì
           3S.PRO-that NC4-child AG4-that **3s-come\RLS** 3s-give\RLS=3S.PRO

           ì-rí-mpàa-nì   vì-yápù /
           NC3-thing-this-NMLZ AG8-two

           that one, then that child gave him two thingummys /

                                        [tapf001.001.030]

It seems likely that the speaker's choice of person subject agreement here was influenced by the left-dislocated NP **éllè**, which is co-referential with what appears to be the true subject NP **mɔ́ɔ méllè**, since sentences with non-NC8 gender-marked subject NPs triggering person agreement are rejected in elicitation work, as (65) illustrates. The 'subject' NP **kà-bárá** 'old man, NC1' cannot trigger person agreement on the verb (65b), unless there is an intervening pause (65c).

---

18 In the sense of palm-wine in general e.g. *palm-wine is good for you.*

(65) (a) kà-bárá   **kà**-tá'à   ká-yàntà   kà-yíiví
           NC1-old_man **AG1**-want\RLS AG1-squeeze\IRR NC1-leftovers

           the old man wants to squeeze the leftovers

      (b) *kà-bárá   **ù**-tá'à   ú-yàntà   kà-yíiví
           NC1-old_man **3S**-want\RLS 3S-squeeze\IRR NC1-leftovers

           [Intended meaning: *the old man wants to squeeze the leftovers*]

      (c) kà-bárá,   **ù**-tá'à   ú-yàntà   kà-yíiví
           NC1-old_man **3S**-want\RLS 3S-squeeze\IRR NC1-leftovers

           the old man, he wants to squeeze the leftovers

                                                         [2007-02-12.002]

Before leaving the person agreement prefixes, it is worth mentioning a special kind of
construction involving grammatical person agreement, not counted in Table 57 above. It
is possible to find lexical relativised subjects triggering *first-person* subject agreement
on the relative clause predicate. In the relative clause in the appositional NP in (66) we
might have expected 3PP agreement, but instead the 1PP subject prefix occurs.

(66) òtù   à-zá   há-nà   **tì**-kúsùu-nà
      1P.PRO NC2-person AG2-REL **1P**-remain\RLS-PFV

      us the people that (we-)*remain*

                                                         [Tikula, sagb001.284]

It could be argued that even though the relativiser **-nà** agrees in gender with **à-zá**, the
true head of the relativised clause is the first element of the apposition, the personal
pronoun. But this cannot be the case in (67), where instead of an appositional NP we
have a predicate nominal construction (§4.3.3.1.2) consisting of two NPs and then the
copula. I have indicated the structure by bracketing the two NPs.

(67) hábà   lêe   [ótú]   [à-zá   há-nà   **tì**-yó-nò,   m-úu
      come_on there 1P.PRO NC2-person AG2-REL **1P**-be\RLS-PFV NC5-child

      mí↓=z-zá   n-nà   tá'à-nà   à-zá]   h-è
      AG5=NC8-person AG8-REL like\RLS-PFV NC2-person AG2-COP

      Come on! [*we*] *are* [*people who* (we-)*are children of the one who loved
      people*]

                                                         [oamy001.188]

These examples are similar in structure to the second-person agreement triggered by
vocative nouns in other languages (e.g. Czech – Corbett 2006:132), but it is not known
whether this construction can be used with the other first- and second-person prefixes in
Cicipu, nor indeed whether it can occur with any other head noun apart from **z-zá**

'person'.

Overall, then, we can see that person subject markers are closer to 'anaphoric' agreement markers than to 'grammatical' agreement markers. It is important to stress that the difficulties involved in determining exactly when the person subject marker can co-occur with an NP subject are far from atypical cross-linguistically. Corbett (2003b:187) and Siewierska (2004:125) both make this point, and we should not be too surprised by the complexity of the Cicipu data.

7.6.1.2   Gender subject prefixes

Like their person-marked counterparts, gender subject prefixes are also ambiguous agreement markers in Bresnan and Mchombo's typology. The difference is that while the person prefixes occur as anaphoric pronominals 90% of the time, gender prefixes are much more likely to co-occur with an explicit subject NP. Out of the 843 instances of overt gender subject agreement in the corpus, 486 (58%) are immediately preceded by a subject NP, as in (68)[19].

(68) kà-bárá   **kà**-'wâa-nà
      NC1-old_man   AG1-pass\RLS-PFV
      an old man passed by

[tapf002.001.030]

Just as with the person subject prefixes, there is evidence from infinitival control to show that NPs like **kà-bárá** in (68) are true grammatical subjects. Example (69) below is analogous to the person-marked example (54) above, since the subject NP **rú'ú yéevì** 'his body' controls the subject of the infinitive **ù-hé'wè** 'to dry'. Just as before, this construction provides evidence that gender subject prefixes can co-occur with true subject NPs in the same clause.

(69) rú-'ú   yé-evì   **yì**-hwáarà   ù-hé'wè
      NC3-body   AG3-3S.POSS AG3-start\RLS NC7-dry_up
      his body starts to dry up

[tats007.001.037]

While gender subject agreement prefixes often co-occur with a subject NP, they also take part in anaphoric agreement. They are not licensed syntactically by any kind of

---

19 Again, the Tikula texts have not been counted. A significant proportion of these tokens come from texts with non-humans as topics, such as folktales and the topic-stimulation texts discussed in §1.4. It is likely that the incidence of anaphoric gender subject agreement is lower in everyday linguistic practice.

'binding theory', since they may occur arbitrarily far away from their antecedent. The following examples show antecedents in increasingly distant syntactic configurations, culminating in (73) where the expected antecedent is completely absent from the discourse.

The first example shows gender agreement in a subordinate clause with the subject of the matrix clause.

(70) **kù-lóngí**   kù-tá'à   **kú**-yàntà   kà-yíví
     NC9-monitor_lizard   AG9-want-RLS   AG9-squeeze\IRR   NC1-leftovers
     the monitor lizard wanted to squeeze the leftovers

                                                                                    [2007-02-12.002]

The following example shows three juxtaposed clauses. The subject NP only occurs in the first clause; in the others, the agreement marker is anaphoric.

(71) **ù-láa**   **wù**-yô   n = ù-'úsȕ /   **wù**-yô   kúmá   ǹ   tì -lípȁi /
     NC7-fire   **AG7**-be\RLS   and=NC7-power   **AG7**-be\RLS   also   and   NC6-goodness

     wù-yô   kúmá   n = ù-pácí
     AG7-be\RLS   also   and=NC7-difficulty
     fire has power / it is also good / it is also difficult

                                                                                     [tats002.002.059]

The next example is an instantiation of a common greeting pattern involving cross-turn anaphora – the referent **kw-á'à** 'house, NC9' is introduced by the questioner, and the gender subject agreement prefix **ku-** is used anaphorically in the answer. The same greeting pattern can be found with nouns of any class (e.g. **k-káa** 'wife, NC8', **m-úu** 'children, NC5').

(72)   A: t-èné   **kw-á'à?**
            AG6-which   **NC9-house**
            How's the house?

       B: **kú**-yô   sèré
            AG9-be\RLS   fine
            It's fine.

                                                                               [eamd003.003]

Finally, gender agreement morphology can be used even when the 'controller' is completely absent from the discourse. Example (73) was said by someone recounting what he had seen on a video. The form of the AG1 agreement marker seems to be related to **kà-dámá** (NC1) 'word' (see §6.3), but this word does not occur in the text itself, nor

does any other potential antecedent.

(73) lêe   v-ì   kɔ̀-kɔ́tɔ̀-nɔ̀
      there AG8-COP   AG1-finish\RLS-PFV

*it's there that **it** [the story] finished*

[tapf002.004.057]

It is clear then that the gender subject prefixes do not have to be bound in any particular syntactic configuration involving their antecedent. The antecedent may be the subject NP (in which case we have grammatical agreement), it may be elsewhere in the clause, extra-clausal, in a different sentence, a different turn, or even absent altogether[20].

7.6.1.3   Summary of prerequisites for subject agreement

Both gender and person subject prefixes take part in anaphoric as well as grammatical agreement, and so according to the typology discussed in §2.2.5 they are both ambiguous agreement markers. However we have seen that the agreement prerequisites differ for the two kinds of subject agreement, involving lexical, phonological and morphosyntactic factors. The flowchart in Figure 35 is intended as an elucidatory aid[21] to make it clear how the type of subject NP affects the kind of agreement found on the verb. It can be seen that the only environments in which there is a *choice* between gender and person agreement are (i) when there is no subject NP at all – in other words anaphoric agreement – and (ii) in the case of a NC8 subject before a vowel-initial verb stem.

---

20 It is odd, in the light of these examples and the frequency data given above, that language consultants often reject constructed examples of anaphoric gender agreement. This kind of elicitation seems to require careful negotiation of the context in order to supply the right felicity conditions for such utterances.

21 In particular, it is not meant to have any formal status in a processing or text-generation model.
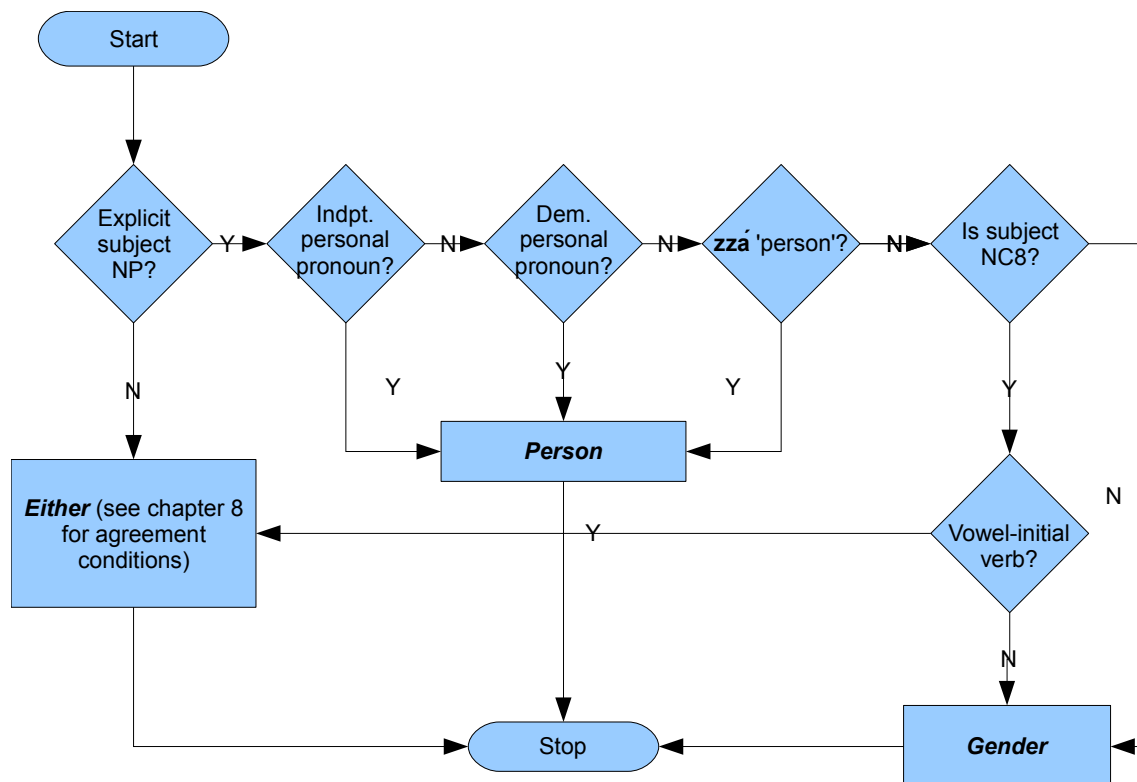
Figure 35: Flowchart showing which subject NPs trigger gender and person agreement

The prohibition on gender agreement for NPs which are headed either by an independent personal pronoun, a demonstrative personal pronoun, or the lexical item **z-zá** 'person' may seem a little arbitrary. A way round this problem is to claim that these expressions have in common the feature PERSON but not GENDER. This is easy to show for the personal and person-marked demonstrative pronouns. As will be apparent in chapter 8, person-marked independent pronouns may have any kind of NP as its antecedent, regardless of the gender of the noun. Similarly the person-marked demonstrative pronouns can readily be used deictically to indicate any kind of object, regardless of which gender the object's "basic-level" noun belongs to (§2.2.1). Neither show any signs of being marked for the feature GENDER, either through overt morphological marking or agreement.

To suggest that **z-zá** 'person' is not marked for GENDER is a more radical step. For a start, the word is morphologically marked with the characteristic NC8 lengthening prefix **C-**, and it triggers AG8 agreement on most of the agreement targets listed in §6.2. Nevertheless there are two other properties that sets this word apart from other 8/2 or 8/3 nouns. The first is its extreme flexibility of reference. While it has a 'default' meaning of 'man/person' (Hausa *mutum*), so that in the identificational clause **z-zá v-ì**

'it's a person' it can only denote a human, in other constructions (e.g. 74 below) it has the potential to refer to *any* concrete noun. Secondly, in addition to being able to co-occur with a person agreement prefix, **z-zá** is also special when it comes to gender agreement. Apparently uniquely, it cannot co-occur with the **vi-** allomorph of the AG8 gender subject prefix. So only (74a) and (74b) are possible, and not (74c)[22].

(74) (a) z-zá   n-nà   ù-yúwò-nò
           NC8-person   AG8-REL   3s-fall\RLS-PFV
           the one that fell

      (b) z-zá   n-nà   yúwò-nò
           NC8-person   AG8-REL   fall\RLS-PFV
           the one that fell

      (c) *z-zá   n-nà   vì-yúwò-nò
           NC8-person   AG8-REL   AG8-fall\RLS-PFV

[2008-04-15.001]

Compare this with the behaviour of other NC8 nouns such as **vá-arì** 'man' (75) and **k-káa** 'woman' (76), which allow Ø- or vi- but not u-.

(75) (a) vá-arì   n-nà   Ø-yúwò-nò
           NC8-man   AG8-REL   AG8-fall\RLS-PFV
           the man that fell

      (b) vá-arì   n-nà   vì-yúwò-nò
           NC8-man   AG8-REL   AG8-fall\RLS-PFV
           the man that fell

[2008-04-15.001]

(76) (a) k-káa   n-nà   Ø-yúwò-nò
           NC8-woman   AG8-REL   AG8-fall\RLS-PFV
           the woman that fell

      (b) k-káa   n-nà   vì-yúwò-nò
           NC8-woman   AG8-REL   AG8-fall\RLS-PFV
           the woman that fell

[2008-04-15.001]

Before vowel-initial stems **vi-** agreement (77a) is not possible with **z-zá** as subject either, leaving person agreement as the only option (77b), since Ø- is ruled out by phonotactic constraints (§3.1.1).

---

22 Recall from §6.1.3 that the **C-** AG8 allomorph found on other agreement targets does not occur on verbs in Cicipu.

345

(77) (a) *z-zá   n-nà   vǎ-ayà-nà
          NC8-person   AG8-REL   AG8-come\RLS-PFV

      (b) z-zá   n-nà   wǎ-ayà-nà
          NC8-person   AG8-REL   3s-come\RLS-PFV
          the one that came

Note that this is a lexical constraint on **z-zá**, rather than anything to do with animacy – the possibility of person agreement and the prohibition on the **vi-** allomorph hold whether **z-zá** is being used to refer to humans, animals, or even inanimate objects.

In this respect **z-zá** patterns with the independent personal pronouns, demonstrative personal pronouns, and personal names, rather than other NC8 nouns. Compare the grammaticality judgements in (74) with those in (78) involving the demonstrative pronoun **é-mpè** 'this'.

(78) (a) é-mpè   ù-ɗâa   é-mpè   ǹ   tì-lípâi
          3s.PRO-this   3s-surpass\RLS   3s.PRO-this   with   NC6-goodness
          this one is better than this one [lit. 'surpasses with goodness']

      (b) é-mpè   ɗâa   é-mpè   ǹ   tì-lípâi
          3s.PRO-this   surpass\RLS   3s.PRO-this   with   NC6-goodness
          this one is better than this one

      (c) *é-mpè   vì-ɗâa   é-mpè   ǹ   tì-lípâi
          3s.PRO-this   AG8-surpass\RLS 3s.PRO-this   with   NC6-goodness

So in the same way that person subject agreement is possible with the independent and demonstrative personal pronouns, personal names, and the noun **z-zá** as subject NP, *gender* subject agreement involving the **vi-** allomorph of the AG8 prefix is *impossible* with precisely the same grouping of expressions as subject.

This can be accounted for if we assume that, despite belonging to the morphological class 8/2, **z-zá** does not in fact have a specification for the morphosyntactic feature GENDER. This gender-deficient noun would, by default, trigger AG8 agreement on agreement targets that necessitate some kind of gender agreement, in just the same way as other atypical controllers such as personal names and personal pronouns (see §6.4 on neutral agreement) – compare **z-zá n-nà** 'the person/the one' with **Màjíijí n-nà** 'the Majiiji who we've been talking about' in (79), or **érè n-nà** 'them we've been talking about' in (80).

(79) Màjíijí **n-nà**
[name]   AG**8**-ART
the aforementioned Majiiji [a person's name]

[Tikula, sami001.032]

(80) érè   **n-nà**
3P.PRO   AG**8**-ART
them [i.e. those ones previously mentioned]

[sayb001.415]

This argument is strengthened by the apparent impossibility of the plural noun **à-zá** 'people/ones' triggering AG2 gender subject agreement. Instead it seems from the many examples in the corpus[23] that only 3PP person subject agreement is found, as in (81).

(81) à-zá   hé-llè   **à**-sì -súmá
NC2-person   AG2-that   **3P**-HAB-run
those ones were running

[samy001.021]

This example also raises a question, however: why, if it is gender-deficient, does **à-zá** trigger AG2 gender agreement on other targets such as the demonstrative modifier **-llè**? We have already used the neutral agreement to account for the *AG8* agreement triggered by **z-zá**, and from (80) above we can see that even plural referents such as the 3PP pronoun **érè** trigger AG8 agreement.

In fact there is evidence that atypical controllers with plural referents can also trigger AG2 agreement. Recall the associative plural marker **áa** (§4.4.2), which occurs before names to indicate that person and their family or associates. In the identificational clause (§4.3.3.1.1) in (82), AG2 agreement is triggered on the copula by the associative plural of the 8/3 noun[24] **Ø-Lɔ́wɔ́lì** 'large spider'.

(82) áa Ø-Lɔ́wɔ́lì   **h**-è
PL   NC8-large_spider   AG**2**-COP
There was Big Spider and associates

[saat001.008.005]

Additional evidence is provided by the 3PS personal pronoun, which can occur with the AG2 copula **h-è** if it has a plural referent (§8.9.4).

If we accept the analysis of **z-zá** as a gender-deficient noun which relies on the operation of neutral agreement for the AG8 morphology that it does trigger, then the

---

23 No relevant elicitation has been carried out, unfortunately.

24 In the context of the folktale from which this example is taken **Lɔ́wɔ́lì** should probably be thought of as a proper name rather a common noun – either way there is no lexical motivation for AG2 agreement.

flowchart presented in Figure 35 above can be simplified and the analysis made more elegant, since there is no need to explicitly mention the exceptional classes of subject NPs that allow person subject agreement. Figure 36 shows the revised analysis.
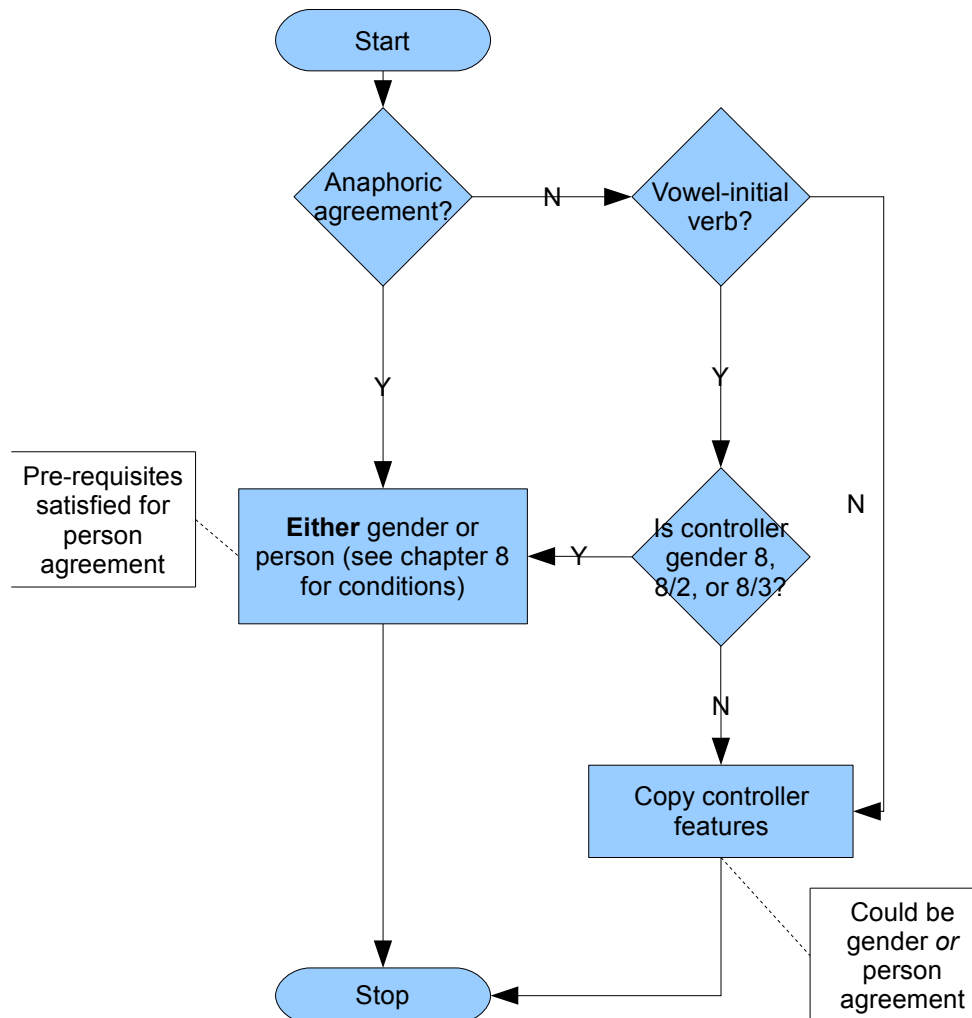
Figure 36: Revised flowchart modelling gender/person alternation

The fundamental rule here can be found towards the bottom-right of the chart. The rule simply states that the controller features are to be copied onto the target. So if the controller is marked for gender, as are most lexical NPs, then so will the target be. If the controller is *not* marked for gender, then only person will be marked on the target. This is what happens when the controller is a personal pronoun, a demonstrative personal pronoun (e.g. **é-mpè** 'this one', **é-llè** 'that one'), or, as has been argued above, the gender-deficient lexical item **z-zá** 'person'. The only way that this rule can be circumvented is if both (i) the prerequisites for person agreement are met (either

because the agreement is anaphoric, or because the agreement is grammatical with an NC8 subject and a vowel-initial verb stem), *and* (ii) the agreement conditions to be discussed in chapter 8 turn out in favour of person agreement. A final version of this flowchart will be presented in §8.7.

7.6.1.4 Agreement conditions

Now that the agreement PREREQUISITES have been established, we are equipped to deal with the agreement CONDITIONS which are relevant to the gender/person alternation in Cicipu. This will be the subject of chapter 8, and here I will only comment on the relevance of agreement conditions in Cicipu to the agreement/pronominal affix distinction. Concerning this topic, Corbett notes (2003b:190) that "If...conditions affect the presence or absence of a marker, then this suggests that the marker is a pronominal affix". The situation for Cicipu is complicated, of course, in that the two subject markers are not independent. If one is not present, the other must be, and *vice versa* (although see §7.6.7 on optional agreement). In §8.7 I will argue that gender agreement should be considered the default, and person agreement subject to conditions. This heuristic then places gender subject markers further along the scale towards grammatical agreement than person subject markers.

7.6.2 Number of arguments encoded on the verb

In Cicipu only one argument can be encoded on the verb, regardless of agreement feature, and this outcome favours neither the grammatical agreement nor the incorporated pronominal interpretation.

7.6.3 Referentiality

As mentioned in §2.2, Evans (2002) takes issue with Siewierska's (1999) use of the term 'grammatical agreement' to mean verb agreement which requires the presence in the clause of a lexical NP argument. Instead he argues that an agreement marker should be called 'grammatical' when it loses its referentiality (i.e. it can be co-indexed with non-referential NPs), independently of whether it can occur in the absence of a nominal argument. Both sets of Cicipu subject prefixes are able to co-occur with definite, indefinite, or non-specific[25] NPs. The following examples show gender subject agreement with a specific indefinite NP (83) and a non-specific NP (84).

---

25 See Haspelmath (1997:106-110) for the definition of non-specificity assumed here.

(83) hárì **kàkúyá** káyà kàká6ìvì
      hárì **kà-kúyá** k-áyà kà-ká6à = vì
      until NC1-termite AG1-come\RLS AG1-catch\RLS=3S.PRO

until a termite [mound] took hold of him

[Tidipo, saat002.004.018]

(84) dòorí, nì m-ɔ́ɔ **mò**-yôo á↓=ká-káasùwà /
      formerly when NC4-child AG4-go\RLS LOC=NC1-market

      ù-sì-cíyó cé Ø-kèeké /
      3S-HAB-get NEG NC8-bicycle

in the past, when a child went to the market, he wouldn't get a bicycle

[2008-02-01.001]

Example (85) shows gender agreement with a D-quantified (Austin and Bresnan 1996:237-239) subject NP **sáa kwèné kúlácí** 'every girl'.

(85) sáa kw-ènè kú-lácí **kú**-dúkwà
      or AG9-which NC9-young_girl **AG9**-go\RLS

every girl went

[eamd032.184]

Person agreement is equally compatible with indefinite and non-specific NPs. Example (86) involves an indefinite specific NP.

(86) wú-nà z-zá **wǎ**-ayà ǹ Ø-kèeké
      3S-ART NC8-person **3S**-come\RLS with NC8-bicycle

a certain person came along with a bike

[tapf002.003.013]

Examples (87-92) show person subject agreement with various kinds of non-specific referents. The first example (87) occurred in the corpus as an impersonal construction (§8.4.8) with a generic interpretation.

(87) ì-láɗámú kúmá / **hú**-u-yìndà y-ì cé ǹ kw-ándái /
      NC3-lightning and **3P**-FUT-see\IRR AG3-PRO NEG with NC9-dry_season

and lightning / it is not seen in the dry season /

[tats002.007.025]

The following examples all show person agreement with non-specific NPs, either **é-mpè** 'this one' (88) or **z-zá** 'person' (89-91).

(88) é-mpè   **ú**-yằa   tì-kɔ́gɔ́ /   é-mpè   **ú**-yằa   cé /   é-mpè   **ú**-sɔ̀ɔ
3s.PRO-this   **3s-do\IRR**   NC6-Hausa   3s.PRO-this   **3s-do\IRR**   NEG   3s.PRO-this   **3s-drink\IRR**

mó-yóo /   **ú**-yùwò /   é-mpè   **ú**-sɔ̀ɔ   cé   **ú**-yùwò /
NC4-beer   **3s-fall\IRR**   3s.PRO-this   **3s-drink\IRR**   NEG   **3s-fall\IRR**

one would follow Islam / another wouldn't /
one would drink beer / until he fell / another wouldn't drink until he fell /

                                                      [samoh001.231]

(89) z-zá   **wú**-u-yằa   Ø-kwáanikà   á=ú-yáa /
NC8-person   **3s-FUT-do\IRR**   NC8-overnight_stay   LOC=NC7-road

'ásù   wú-nà   gó   ù-'ɔ̀ɔ-nɔ̀ /   ú-hùngwà /
place(NC7)   AG7-REL   TOP   3s-tire\RLS-PFV   3s-rest\IRR

a person would do an overnight stay on the road /
where he got tired / he would rest /

                                                        [sayb001.755]

(90) w-áa   á↓=d-dòorí   **z-zá**   ù-gúyà   ù-híyà   kwáanù góomá sṹ?
3S-REPRT   LOC=NC8-formerly   **NC8-person**   3s-can\RLS   NC7-grind   bowl   ten   Q

it's said that in the olden days one could grind ten measures?

                                                      [svtmg001.018]

(91) t-èné   t-ì   **z-zá**   **w**-ú-yằa-wà=vù   kà-líipì?
AG6-which   AG6-COP   **NC8-person**   **3s-FUT-do\IRR-APPL=2s.PRO**   NC1-wrong

how can anyone do you wrong?

                                                      [oamy001.060]

The final example shows person agreement with a D-quantified subject:

(92) sáa w-èné z-zá   n-nà   **ù**-yó-nò   kà-ráyì
or   3s-which   NC8-person   AG8-REL   **3s-be\RLS-PFV**   NC1-life

every person who has life

                                                      [tats001.004.066]

It is clear then that both subject agreement paradigms can have referential and non-referential uses. However even the free pronouns of English have various indefinite readings (e.g. generic uses such as *they always get you in the end*) – the question is whether the indefinite readings are limited to unusual syntactic and semantic contexts (Evans 2002:17). Given the variety of examples presented above, it seems to me that with respect to referentiality both the Cicipu subject prefixes have more in common with English subject agreement than English free pronouns. Thus they are non-committal with respect to referentiality, and this points towards the grammatical agreement analysis.

7.6.4 Descriptive content

Regarding descriptive content, neither person nor gender subject affixes seem to have any semantic restrictions as to their use, and therefore we cannot say that there is any lexical meaning attached to the markers themselves. The most that can be said is that there is a strong correlation between the *use* of person agreement and animacy (especially humanness). However virtually any kind of referent can be indexed with person agreement if it is sufficiently discourse-topical (see §8.4), and thus while the 'descriptive content' criterion points towards grammatical agreement rather than incorporated pronouns, it does not distinguish between the two markers.

7.6.5 Balance of information

Considering 'balance of information' Corbett notes that for agreement markers there are features reflected on the controller that may be absent on the target, and vice versa. However this is not usually the case for incorporated pronominals, and Corbett remarks (2006:105) "It is rare to find (morphosyntactic) feature distinctions marked on the noun phrase which are not also marked on the pronominal affixes...". As far as this criterion is concerned, we find the opposite of what we might have expected. In common with other Benue-Congo languages (see §2.1.2), the noun class (and hence the morphosyntactic features GENDER, NUMBER and PERSON[26]) is indicated on the noun. Gender subject affixes carry the same three features. Person subject affixes, on the other hand, are marked only for PERSON and NUMBER, but not GENDER, and therefore the balance of information is shifted towards the NP, which also marks gender.

Thus this criterion places person agreement closer to 'grammatical agreement' than gender agreement, in contrast to the multi-representational criterion discussed in §7.6.1.

7.6.6 Other tests

As well as multi-representation (or the 'locality' principle), Bresnan and Mchombo (1987:759-764) discuss a number of other diagnostics for grammatical vs. anaphoric agreement, based on the assumed impossibility of anaphoric agreement being triggered by non-topical controller referents – if anaphoric agreement is topical agreement, then it cannot occur in contexts where we know that the controller referent is not a sentence topic. So if we find an agreement marker in such a context, then it is either a

---

26 PERSON is derivative of GENDER for all NPs other than first- and second-person pronouns.

grammatical or ambiguous agreement marker. Conversely, if an agreement marker *cannot* occur in such an environment, this suggests it is an anaphoric agreement marker.

It is important to be clear that these predictions are about different *types of agreement* rather than agreement markers. It is the type of agreement (anaphoric or grammatical) that is associated or disassociated with topicality. Individual agreement markers are then secondarily classified as grammatical, anaphoric, or ambiguous. The theory does not directly predict an association between, say, Cicipu person agreement markers and topicality, and Cicipu gender agreement markers and non-topicality. So even if we find (as we did in §7.6.1) that gender agreement is more common when there is an explicit subject NP and that person agreement is more commonly anaphoric, we should not expect Bresnan and Mchombo's subsequent diagnostics to pick out any asymmetry between the two paradigms. Nevertheless there is some merit in discussing a couple of their predictions, since even if they do not distinguish between the two paradigms, they at least serve to confirm that Cicipu gender and person agreement subject prefixes are ambiguous agreement markers, rather than pure anaphoric agreement markers.

7.6.6.1   Questioned subjects

Bresnan and Mchombo's first prediction concerns wh-questions. According to their theory, if a subject agreement marker is an incorporated pronoun (and hence indexed to a sentence-topical referent), then it should not be able to co-occur with questioned subject NPs, assuming that (i) a sentence-topic NP cannot also be in focus, and (ii) questioned constituents are in focus.

Assume for a moment that gender subject agreement prefixes were pure anaphoric agreement markers, and so in sentences like (93), repeated from (68), the apparent subject was actually a topicalised adjunct.

(93)   kà-bárá   kà-'wâa-nà
        NC1-old_man   AG1-pass\RLS-PFV
        an old man passed by

If that were the case, then by the logic outlined in the previous paragraph we would expect gender subject agreement to be impossible with questioned subject NPs. However this is not the case, and both **yìní** 'what, NC3' and **yàanú** 'who, NC8' behave just like straightforward lexical nouns, triggering gender agreement in (94) with **yìní** 'what,

NC3' and (95a) with **yàaní/yàanú** 'who, NC8'.

(94) yìnîihwáarùwànà   kò-rísìnôo?
     y-ìní   **yì**-hwáarà-wà-nà   kò-rísìnôo
     NC3-what AG3-start\RLS-APPL-PFV NC1-Karishen
     what started Korisino?

[sayb001.793]

(95)   (a) **yàaní** Ø-ɗávà-nà   ì-táatú yí-mpà?
           who   AG8-weave\RLS-PFV NC3-mat AG3-this
           who wove this mat?

(b) ***yàaní** **ù**-ɗávà-nà   ì-táatú yí-mpà?
           who   3S-weave\RLS-PFV NC3-mat AG3-this

[2007-01-22.004]

In Cicipu there is nothing wrong with the co-occurrence of a focused wh-word and a co-indexed gender subject agreement prefix, and this adds to the evidence that gender subject prefixes are ambiguous agreement markers.

The test also shows that wh-words behave in the same way as lexical nouns with respect to *person* agreement. Example (95b) above was judged ungrammatical, just like the analogous (65b) above with an ordinary, non-focused subject. Questioned subjects *can* co-occur with person agreement, but only under the same circumstances as ordinary non-questioned NPs, i.e. NC8 nouns before vowel-initial verb stems. In this environment person agreement is possible with **yàaní/yàanú** 'who, NC8' as the wh-word, but not with **yìní** 'what, NC3'. Compare (96) with (62) above, repeated as (97) below for convenience.

(96) yàanú **wú**-u-'ìngò   n=àmú?
     who   3S-FUT-go_home\IRR with=1S.PRO
     who will go home with me?

[saat001.008.060]

(97) Ø-kùrì'áanì **wú**-u-'ɔ̀pɔ̀ z-zá
     NC8-Koran   3S-FUT-hold NC8-person
     the Koran will hold a person

The fact that person agreement is possible in (96) suggests that, like gender agreement, person agreement prefixes are ambiguous agreement markers rather than pure anaphoric agreement markers.

7.6.6.2 Relativised constituents

Another of Bresnan and Mchombo's predictions concerns relativised constituents, which they take to be universally topics[27]. In Cicipu the same situation obtains for relativised subjects as for questioned subjects. Mostly we find gender agreement:

(98) kà-bárá   ká-mpà ká-nà   **kà**-'wâa-nà   pâa
     NC1-old_man   AG1-this AG1-REL   **AG1**-pass\RLS-PFV   here
     this old man who passed away here

[sayb001.477]

If the relativised subject is not NC8, then person agreement is impossible:

(99) (a) m-ɔ́ɔ   má-nà   **mɔ̀**-dóohò-nò
         NC4-child   AG4-REL   **AG4**-disappear\RLS-PFV
         the child who disappeared

     (b) *m-ɔ́ɔ   má-nà   **ù**-dóohò-nò
         NC4-child   AG4-REL   **3s**-disappear\RLS-PFV

[2007-02-07.002]

Person agreement *is* however possible, but apparently only in the same restricted environments as before, e.g. with the head noun **z-zá** 'person':

(100) (a) z-zá   n-nà   dóohò-nò
          NC8-person   AG8-REL   disappear\RLS-PFV
          the one who disappeared

      (b) z-zá   n-nà   **ù**-dóohò-nò
          NC8-person   AG8-REL   **3s**-disappear\RLS-PFV
          the one who disappeared

[2007-02-07.002]

If we accept Bresnan and Mchombo's assumption that relativised subjects are sentence topics, then the data further confirms that both gender and person agreement markers in Cicipu are ambiguous rather than pure anaphoric, but again it does not serve to distinguish the two kinds of agreement. Unlike the difference in Chicheŵa between ambiguous subject agreement markers and anaphoric object agreement markers, the asymmetry between Cicipu gender and person subject agreement cannot be uncovered by considering sentence topics. In chapter 8 we will see that topicality *is* an important factor in the distribution of gender and person agreement across *all* person markers, not just subject prefixes. However it is the notion of *discourse topicality* (§2.3.2) that turns out to be crucial, rather than sentence topicality.

---

27 Again sentence topic is meant rather than discourse topic (§2.3).

7.6.7 Are subject prefixes obligatory?

Before attempting to tackle the semantic and discourse factors governing the alternation between gender and person agreement, we should consider a third possibility: the absence of agreement altogether. We saw in §7.6.1 that overt subject NPs are not obligatory with either the gender or the person subject prefixes. But what about the prefixes themselves? Are they obligatory, or can the subject NP sometimes be followed directly by a bare verb stem?

It has already been mentioned (§7.6.1) that overt subject prefixes are obligatory before vowel-initial verb stems. With *consonant*-initial verb overt subject prefixes may be absent with either NC8 or 2PS subjects; recall that the NC8 noun prefix (§5.5.7), the AG8 agreement prefix (§6.1.3), and the 2PS agreement prefix (§7.4.1) were all analysed as having Ø- as one of their allomorphs, in order to fill the gaps in the paradigms of otherwise overt prefixes[28]. Example (101) shows an NC8 subject with no overt subject agreement.

(101) ánà   k-káa   ví-llè   Ø-mátà
       when   NC8-woman   AG8-that AG8-give_birth\RLS
       when that woman gave birth

[sayb001.168]

As well as nouns belonging to NC8, members of another group of nouns can occur as subject without triggering an overt subject prefix. This is the same group of nouns for which person agreement is possible before consonant-initial nouns (§7.6.1.1): the independent personal pronouns, the demonstrative personal pronouns, and the word **z-zá** 'person'. We saw in (74) and (78) above (the latter is repeated below for convenience) that these nouns either trigger the **ù-** 3PS subject prefix, or they occur without any overt prefix.

(102) (a) é-mpè   **ù-**ɗãa   é-mpè   ǹ   tì-lípâi
          3S.PRO-this **3S**-surpass\RLS 3S.PRO-this with   NC6-goodness
          this one is better than this one [lit. 'surpasses with goodness']

      (b) é-mpè   ɗãa   é-mpè   ǹ   tì-lípâi
          3S.PRO-this surpass\RLS   3S.PRO-this with   NC6-goodness
          this one is better than this one

      (c) *é-mpè   **vì-**ɗãa   é-mpè   ǹ   tì-lípâi
          3S.PRO-this **AG8**-surpass\RLS 3S.PRO-this with   NC6-goodness

---

28 See Baker (1996:21) and Siewierska (2004:22-24) for justification of this approach.

Anticipating the analysis in the rest of this subsection, the overall situation can be summarised as follows:

Table 60: Agreement possibilities before C-initial verb stems

Type of subject NP	**Possible prefixes before a C-initial stem**
Non-NC8 noun	Gender agreement only
NC8 noun	**vi-** (AG8) *or* **Ø-** (AG8)
Independent personal pronoun, person-marked demonstrative pronoun, **z-zá** 'person'	**ù-** (3PS) *or* lack of agreement

The problem here is how to handle the lack of overt agreement shown by the last two categories of nouns in Table 60. We saw in §6.1.3 that many Cicipu agreement targets offer either **vi-**, **C-** or **Ø-** agreement prefixes with NC8 nouns as controllers. Bearing in mind that many NC8 nouns also have **Ø-** *noun* prefixes (§5.5.7), the most straightforward analysis seems to be that the **Ø-** prefixes before verbs and other targets with NC8 controllers are true null AG8 agreement prefixes, as indicated in the table.

For the person-marked NPs in the third row it would be convenient if the **Ø-** prefix could be glossed as a null allomorph of the 3PS person prefix **u-**, since this would allow us to maintain that subject agreement is always obligatory. However there is no independent evidence for this from the other targets that allow person agreement[29]. Furthermore, plural personal pronouns (103-104) and other plural NPs lacking a gender specification (105) may also co-occur with verbs lacking overt morphology, and so to maintain this argument there would have to be a proliferation of null allomorphs.

(103) èrè   pándà
     3P.PRO   forget\RLS

    ***they** forgot*

                      [eamd004.F]

(104) òtù   gáanùkwà   cé
     1P.PRO   understand\RLS   NEG

    ***we** don't understand*

                      [Tikula, sagb001.337]

---

29 i.e. the demonstratives and the article. The copula and the wh-word **-ènέ** have vowel-initial stems, and so some kind of agreement is obligatory. See §8.9 for details.

(105) tí-nà **áa Ágrìc** hyẫa-nà "heating"
AG6-REL **PL agriculture** say\RLS-PFV heating

*how **the Agriculture people** say "heating"*

[samoh001.097]

Overall it seems best to analyse the lack of overt person agreement as the lack of agreement altogether. It has not been possible to determine when genderless-NPs such as **é-mpè** and **z-zá** will trigger person agreement and when they will not, and so no more will be said about this alternation with respect to grammatical agreement.

In clauses *lacking* a subject NP, there is almost always an overt subject agreement marker on the verb. So, for example in (106) below, the only possible allomorph of the NC8 gender agreement prefix is the AG8 allomorph **vi-**[30]. The corresponding sentence without overt agreement was judged ungrammatical.

(106) *(**vì**-)zámùkwà ì-rí yí↓=Ø-ɓáhánná
AG8-be\RLS NC3-thing AG3=NC8-destruction

[**vɔ́-ɔmɔ̀** 'monkey, NC8'] *it is a thing of destruction*

[tats002.001.024]

There are a handful of instances of this ungrammatical structure in the corpus, but they are each either sentence fragments repeating part of a previous clause, as in (107), or hesitations with an immediately preceding subject NP, as in (108).

(107) Ø-kùrì'áanì vì-sì↑-pɔ̀lɔ̀-pɔ́lɔ̀ z-zá.
NC8-Koran AG8-HAB-knock_down-REDUP NC8-person

Ø-pɔ̀lɔ̀-pɔ́lɔ̀ z-zá.
AG8-knock_down-REDUP NC8-person

The Koran knocks down people. Knocks down people.

[tats005.001.032]

(108) Ø-gógóró (.) **Ø**-hé'w<ìs>è-hé'wìsè z-zá
NC8-gin **AG8**-dry_up\RLS<CAUS>-REDUP NC8-person

gin (.) dries people up

[tats008.002.003]

If the non-overt agreement in these examples reflects, as has been assumed here, a true null prefix rather than the absence of agreement, then this AG8 **Ø**- allomorph seems to be a true grammatical agreement prefix, since it cannot occur without an explicit subject NP.

---

30 The 3PS person subject prefix **u**- would also be a possibility, as long as the monkey was a discourse topic (see §8.4).

7.6.8 Comparison with Central Kambari

Before leaving this section, it is worth taking a moment to look again at the gender/person alternation in the related language, Central Kambari (§2.2.4.2.3). There are two major differences between the alternation in Cicipu and Central Kambari. First, in Central Kambari gender subject agreement is grammatical rather than ambiguous, and is not found without an explicit NP (Crozier 1984:218)[31]. Secondly, in Central Kambari person subject agreement is more readily found after lexical subject NPs than in Cicipu.

It seems that in Cicipu neither kind of agreement has progressed as far along the grammaticalisation cline as their Central Kambari equivalents. Gender agreement in Cicipu is sometimes anaphoric (although usually it is not) and so it can be classified as ambiguous agreement, and placed further to the left of Central Kambari gender agreement on Siewierska's (1999) scale. Likewise person agreement in Cicipu is usually anaphoric, and should be placed close to the far left of the cline. The situation in the two languages can be visualised as in Figure 37:

	Anaphoric >	Ambiguous >	Grammatical
Central Kambari	**Person**		**Gender**
Cicipu	**Person**	**Gender**	

Figure 37: Relative positions of person and gender agreement in Central Kambari and Cicipu on Siewierska's grammaticalisation scale

So while in both languages there is an environment in which the alternation between the two agreement patterns is found, it is a different environment in each language. In Cicipu the alternation is mainly found when there is no subject NP, in other words when there is anaphoric agreement. In Central Kambari, on the other hand, the contrast is after lexical NPs, as in Hoffmann's (1963) examples quoted in §2.2.4.2.3 (2.23) and repeated below:[32]

---

31 Siewierska (1999:238-239) notes that this is actually very rare outside Indo-European.

32 I discussed the Cicipu equivalents of Hoffmann's examples with one of my language consultants and he judged them unacceptable, unless there was a pause inserted between the NP and the verb. In that case, the NP is most likely extra-clausal rather than a true subject.

(109) (a) ma-nun   u-kuwete
NC4-bird   3s-die

the bird died

(b) ma-nune   me-kuwete
NC4-bird   AG4-die

a bird died

7.7   Person marker summary

We have now examined the five different person marker paradigms in Cicipu. I concluded that there are two paradigms of free pronouns – one person-marked, and one gender-marked. The independent personal pronouns contrast with a paradigm of person-marked object clitics, which attach to the right-hand edge of the verbal word. 'Clitic-doubling' is not possible. The singular members of the object clitic paradigm have reduced counterparts with particularly interesting phonological properties. There are also two paradigms of subject agreement prefixes, one gender-marked and one person-marked. Both sets of prefixes are ambiguous between anaphoric agreement and grammatical agreement according to Bresnan and Mchombo's (1987) typology, but the gender-marked prefixes are closer to the 'grammatical agreement' end of the grammaticalisation cline. Table 61 summarises the morphological and syntactic status of the five person marker paradigms.

Table 61: Summary of properties of Cicipu person markers

Paradigm	Morph. status	Grammatical function	Features	Grammatical agreement?	Anaphoric agreement?
Independent PERSON	Word	Any	Person	No	Yes
Independent GENDER	Word	Any	Gender	No	Yes
Object clitics	Clitic	Object	Person	No	Yes
Subject PERSON	Prefix	Subject	Person	Sometimes, depends on subject NP	Yes
Subject GENDER	Prefix	Subject	Gender	Usually, depends on subject NP	Yes

I also determined the agreement PREREQUISITES for the two kinds of subject agreement (Figure 36 above). There are two distinct environments where prefixes from either paradigm can occur: (i) before vowel-initial verb stems with NC8 controllers, and (ii) in the absence of a subject NP (i.e. anaphoric agreement). Otherwise the features encoded

on the target are a straightforward match for the features encoded on the subject.

The syntactic status tests discussed in §7.6 were inconclusive and sometimes pointed in conflicting directions, as summarised in the following table:

Table 62: Summary of tests for syntactic status of the subject prefixes

Section	Test	Outcome
§7.6.1	Multi-representation	GENDER closer to grammatical agreement than PERSON
§7.6.2	Number of argument roles indexed on verb	Inconclusive
§7.6.3	Referentiality	Points towards grammatical agreement for both
§7.6.4	Descriptive content	Points towards grammatical agreement for both
§7.6.5	Balance of information	PERSON closer to grammatical agreement than GENDER
§7.6.6.1	Questioned subjects	Neither GENDER nor PERSON is pure anaphoric
§7.6.6.2	Relativised subjects	Neither GENDER nor PERSON is pure anaphoric

Other than multi-representation, the tests suggest that, if anything, person agreement markers are closer to true agreement markers than incorporated pronominals. However as we saw in §7.6.1.1, when we look at the distributional data in the corpus, person prefixes rarely take part in grammatical agreement. Perhaps the only thing we can say for sure is that neither of the subject prefixes is a pure anaphoric agreement marker.

The analysis in the following chapter will prove to be more fruitful when it comes to characterising the differences between these two paradigms. There I will consider the factors that influence the gender/person alternation in the environments which offer a choice of feature, in other words the agreement CONDITIONS. Before doing this, however, it is important to be clear that we really are dealing with separate paradigms of gender and person markers, rather than one complex paradigm.

7.8 Features involved

It is not uncommon for the various agreement targets of a language to differ with respect to their overt features (Corbett 2006:79). What does seem to be rather rarer is for a *single* target to vary in its agreement features. In the case of Cicipu, it is important to make sure that what we have so far been calling 'gender' and 'person' agreement really do contrast with each other in regard to the agreement features GENDER and PERSON. As an alternative analysis we might posit a single complex paradigm involving gender and

person, with gender as a subdivision of third person. This is assumed by Siewierska (2004:104) in her representation of Kiswahili person markers, and also by Welmers (1973:176) with respect to Bantu as a whole. Welmers is worth quoting at length because of the implicit contrast with the state of affairs in Cicipu.

> In a number of Bantu language grammars, there are references to "personal pronouns" in connection with verbs or with expressions of possession (which represent only one of many uses of the associative construction). The implication seems to be that morphemes or sets of morphemes meaning 'he/she, him/her, his/her' and 'they, them, their (personal)' belong with first and second person pronoun forms, and that referents to non-personal nouns have some kind of a secondary status. *Actually, of course, the referents for personal nouns, singular and plural, are nothing more than the concords for nouns of class 1 and 2. If there is to be any dichotomy, it should be between first and second person morphemes on the one hand and all class concords, including those for classes 1 and 2, on the other.* Actually, first and second person morphemes can also be treated with the noun class and concord system. There are no first and second person noun prefixes, to be sure, and first and second person concepts are hardly expected with attributives or demonstratives. But there are subject and object forms, and forms used after the associative morpheme (indicating possession). Further, there are independent referents for first and second person singular and plural, and also for nouns of all classes, which must be treated in a uniform way...In the light of these similarities, although the parallelism is not perfect, it would seem more elegant to include first and second person forms in the concord system [*my italics – S.M.*].

The contrasting distributions of the gender and person subject prefixes (§7.6) argues against this approach for Cicipu, as does the fact that H-tone spreading distinguishes between them, affecting only the gender prefixes (§3.4.3). Nevertheless, if we were to ignore this evidence and try to construct a composite paradigm for Cicipu, then we might start as in Table 63, which presents all the possible realis subject[33] agreement forms that occur before a consonant-initial verb, e.g. **dukwa** 'go'.

Table 63: Subject agreement prefixes for verbs in the realis mood (single flat paradigm)

1s	2s	3s	1p	2p	3p	AG1	AG2	AG3	AG4	AG5	AG6	AG7	AG8	AG9
n-	C-/Ø-	u-	ti-	i-	a-	ka-	ha-	yi-	ma-	mi-	ti-	wu-	vi-/Ø-	ku-

Theoretically it would be possible to treat these target forms as showing agreement in a single feature with fifteen values. However, as in most languages, there are certain relationships between Cicipu pronouns of a given person that suggest we need at least the NUMBER and PERSON features, in particular the anticipatory plurals discussed in

---

33 Any person marker could have been chosen here, since they all have an identical paradigm structure. Similarly the irrealis mood makes the same distinctions as the realis.

§4.4.5.4.

Once we add NUMBER and PERSON as separate features, then the paradigm becomes more complex if we still try to include the cells for the gender-marked prefixes. This is because the class markers differ with respect to number: some are singular, some are plural, and some can occur with both singular and plural nouns:

Table 64: Single paradigm with PERSON and CLASS as a subset of PERSON

PERSON	1	2	3										
NUMBER			CLASS	DEFAULT	1	2	3	4	5	6	7	8	9
SG	N-	C-/Ø-		u-	ka-		yi-	ma-		ti-	wu-	vi-	ku-
PL	ti-	i-		a-		ha-	yi-		mi-			vi-	

Note that the **u-**/**a-** cells have been included along with the noun classes, and are being considered as the morphological realisations of two 'minority target genders' in Corbett's (1991:160) terminology. They qualify as such because there are no nouns that *must* trigger these agreement forms, but every noun *may* trigger them.

In fact, the paradigm in Table 64 is still not complete, since we have not taken into account singular classes, which typically contain abstract or mass nouns. These may show the same morphology and agreement as the plural classes 2, 3, 5, and 8, and yet still trigger **u-** agreement under special circumstances (see §8.4.3). Thus a further revision of the paradigm is required (Table 65), and it should be clear that the interaction between person, gender and number is too complex for all the agreement forms to be insightfully represented in the same paradigm.

Table 65: Revised single paradigm with person and number features but not gender

PERSON	1	2	3										
NUMBER			CLASS	Default	1	2	3	4	5	6	7	8	9
SG	N-	C-/Ø-		u-	ka-		yi-	ma-		ti-	wu-	vi-	ku-
PL	ti-	i-		a-		ha-	yi-		mi-			vi-	
MASS				u-	ka-	ha-	yi-	ma-	mi-	ti-	wu-	vi-	ku-

Instead it is much simpler to construct two paradigms, one involving the features [PERSON] and [NUMBER], and the other involving the feature [CLASS], corresponding to Corbett's target gender:

Table 66: Person/number paradigm

	S	P
1	N-	ti-
2	C-/Ø-	i-
3	u-	a-

Table 67: Noun class paradigm

1	2	3	4	5	6	7	8	9
ka-	ha-	yi-	ma-	mi-	ti-	wu-	vi-/Ø-	ku-

In summary, we have two distinct paradigms of agreement markers affecting the same agreement target. Person agreement involves PERSON and NUMBER:

> PERSON: 1, 2, 3
>
> NUMBER: singular, plural

Gender agreement involves GENDER (or CLASS).

> Target gender: 1, 2, 3, 4, 5, 6, 7, 8, 9

7.9   Chapter summary

The bulk of this chapter (§7.1-7.6) gave an account of the phonological, morphological, and syntactic properties of the five paradigms of Cicipu person markers, and was summarised in §7.7. In particular, §7.6 set out the prerequisites necessary for the two different kinds of subject agreement, which turned out to involve lexical, phonological, and morphosyntactic factors. These were summarised in flowchart form in Figure 36. In the following chapter I turn to the conditions which bear on the choice of gender or person agreement, in those contexts where the prerequisites for both are met.

Chapter 8 – Gender and person agreement in Cicipu discourse

8.1 Introduction

The main aim of this chapter is to show that both the semantic notion of animacy and the discourse-pragmatic notion of topicality are necessary in order to fully describe the alternation between gender and person agreement introduced in the previous chapter. Recall that chapter 7 identified the agreement *prerequisites* (lexical, phonological, and morphosyntactic) relevant to this alternation on subject prefixes. Here I will show that animacy and discourse topicality are agreement *conditions* (§2.2.4.1), in other words factors which help determine the relative likelihood of the two competing agreement patterns occurring, given that the prerequisites for them both have been met. The agreement conditions discussed in this chapter apply to all three of the morphosyntactic positions discussed in chapter 7: subject prefixes, post-verbal object, and elsewhere. In the case of the subject prefixes, we will in this chapter only be concerned with those that meet the prerequisites for both kinds of agreement: (i) anaphoric agreement prefixes, and (ii) grammatical agreement prefixes with NC8 controllers before vowel-initial verb stems (recall Figure 36 from chapter 7). If the prerequisites for either gender or person agreement are not met, then there can be no competition in agreement, and hence no agreement conditions.

As a preliminary observation, it should be stressed that regardless of potential conditions on agreement such as animacy or discourse topicality, the initial referring expression has an important role to play in determining the features indexed on the subsequent anaphors. If the initial referring expression is marked for gender, then this increases the likelihood of subsequent person markers agreeing in gender. Conversely if the initial referring expression is unmarked for gender (e.g. a personal pronoun), then subsequent anaphors are far less likely to be marked for gender. If a speaker wishes to refer exophorically to some object, a tape-recorder, say, then they have two choices of 'near-speaker' demonstrative: (i) é-mpè, which is the 3PS form appropriate for any (singular) object, regardless of animacy or of the gender of the noun most commonly used to refer to it, or (ii) ví-mpà, the AG8 demonstrative which agrees in gender with the noun Ø-rìkóodà 'recorder'. However once the recorder has been introduced using one or

the other of these terms, subsequent mentions are likely to match the features and feature values of the original form. So while either of the 'consistent' conversations (1a) or (1b) would be natural, a 'mix' of differently-marked antecedent and anaphor would be less likely[1].

(1)   (a) A: **é-mpè**   hínà?   B: **ù**-ɗâa   é-mpè   ǹ   tì-lípâi
             3s.PRO-this TOP   3s-surpass\RLS 3s.PRO-this with NC6-goodness
             *What about **this one**?*   *It's better than this one.*

      (b) A: **ví-mpà**   hínà?   B: **vì**-ɗâa   ví-mpà   ǹ   tì-lípâi
             AG8-this TOP   AG8-surpass\RLS AG8-this with NC6-goodness
             *What about **this one**?*   *It's better than this one.*

[2007-02-13.006]

In such cases the subject marking is nothing to do with animacy or discourse topicality, but it is influenced to a strong degree by the form of the initial referring expression. Gender-marking *could* occur in (1a), in the sense that it would be grammatical, but would probably be regarded as 'clumsily-put' (see §2.3.3).

Nevertheless, however influential the initial referring expression may be in guiding the form of future anaphora, it is not the only factor. Although in the scenario just presented a change of feature would be unlikely, there are other circumstances where a switch from gender-marking to person-marking is more probable. To anticipate the later analysis, we will find that animacy and discourse topicality are both positively correlated with person agreement. Neither of the two notions is sufficient on its own, and moreover the two conditions cannot be 'ranked' in the manner of traditional Optimality Theory, since neither 'constraint' consistently wins out[2]. The most we can say is that subject agreement marking is conditioned on both *inherent* topicality and *discourse* topicality (see §2.3). Referents that are good inherent topics are almost (but not quite) always coded using person agreement morphology. Referents that are bad inherent topics almost (but not quite) never achieve the status of being person-marked. The situation can be visualised as in Figure 38, where lighter shading indicates a higher probability that the verb will agree in person rather than in gender.

---

1 The example given involves a turn-transition, but the effect is just as strong in monologues.
2 Although see Aissen (2003) for a discussion of variation in agreement and 'weak bidirectional OT'.
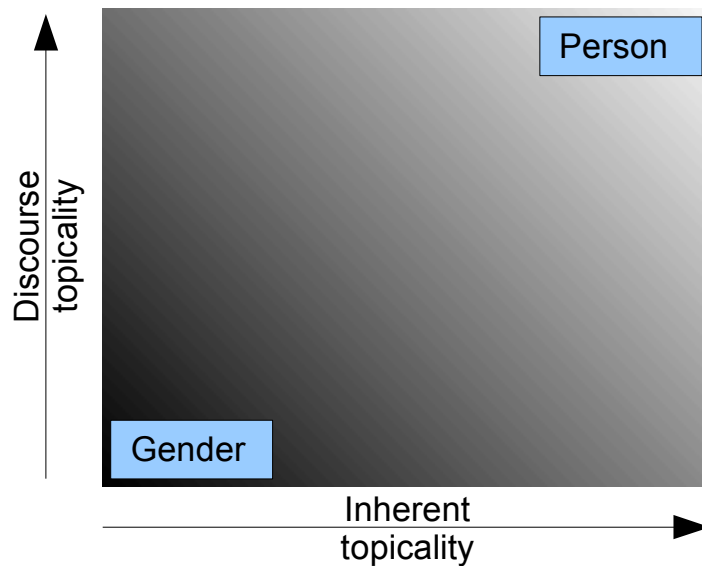
Figure 38: Variation in agreement according to inherent and discourse topicality

Referents situated at the bottom-left and top-right corners do not admit a choice – those that are neither inherently- nor discourse-topical are always coded with gender agreement morphology, while those with both kinds of topicality always trigger person agreement. Referents with only one kind of topicality, on the other hand, may be found with either kind of agreement.

In addition to animacy and discourse topicality, there is one further agreement condition on the gender/person alternation, this time involving the noun class of the controller. If the controller's gender is 8/2 (the default gender for humans), 8/3 (the default gender for animals and inanimates), or the single class gender 8, then person agreement is more likely than if the controller belongs to any other gender.

Finally, there is also an effect of 'natural' gender. Although Cicipu, as is typical for Niger-Congo languages, does not distinguish male and female referents *within* its gender system, this distinction does play a role in determining the likelihood of person vs. gender agreement. Female referents appear to favour gender agreement to a greater extent than males. It is tentatively suggested here that this difference should be incorporated into the Cicipu animacy hierarchy, rather than treated as an independent parameter.

The rest of this chapter is organised as follows. In §8.2 I explain the methodology used for coding and analysing the corpus. Section 8.3 provides a preliminary overview of participant reference in Cicipu, including a brief discussion on marked and

contrastive topics. The key parts of the chapter are §8.4, in which individual examples from the corpus are chosen to demonstrate the combined effect of animacy and discourse topicality, and then §8.5, where it is shown that there is a coding progression within paragraphs from lexical NPs > gender agreement > person agreement. Section 8.6 considers two possible alternative explanations for the distribution of gender and person agreement. The final two sections take a step back from the data – §8.7 asks which of the two kinds of agreement is 'marked' and which is 'unmarked', while §8.8 considers how the two agreement paradigms can be understood in terms of coding weight and anaphoric vs. deictic reference. Finally §8.9 turns to the remaining agreement targets that alternate between gender and person agreement, and considers to what extent the analysis developed for person markers also applies to them.

8.2 Data analysis

The methodology involved in collecting the data used for this study was set out in §1.4. In summary, the corpus consists of approximately six hours of transcribed and interlinearised audio and video recordings, supplemented by elicitation sessions with native speaker consultants. In addition to more traditional genres such as folktales and historical narratives, I also recorded people talking about a number of set topics of varying animacy, in order to stimulate texts with a higher density of agreement markers with non-human controller referents. This section describes how the corpus was analysed and coded.

The corpus is small according to the standards of corpus linguistics, with only about twelve thousand clauses[3]. Nevertheless, an exhaustive coding of even a small proportion of this corpus, covering the identification of discourse units and paragraph topics, and the enumeration and tracking of all referents and their agreement markers was beyond the scope of this thesis. The crucial examples which most clearly demonstrate the intra-paragraph progressions from gender to person agreement are spread thinly through the corpus, and such an enterprise, while it would have increased the methodological rigour of this study, would have consisted mainly of the encoding of thousands of the less-interesting discourse-topical human referents, which always trigger person agreement and therefore admit no choice. Likewise, the majority of non-human referents are non-topical and do not persist in the discourse. Only rarely are they

---

3   Although see Lüpke (2006) on 'small is beautiful'.

encoded as lexical NP subjects, and when they do they always trigger grammatical gender agreement, again offering no choice. It was considered to be a better use of the time available to take a more qualitative approach by concentrating on the more interesting examples, and these are discussed at length in §8.4 and §8.5.

Thus my approach to the data was to consider all cases of anaphoric agreement, apart from the majority cases where the referent was human and triggered only person agreement, in which case it was ignored. In particular I was interested in the way in which a referent was coded between lexical mentions. The following properties were coded for each person marker:

Table 68: Properties coded for person markers

Property	Categories
Animacy	human, spirit, anthropomorphised animal, animal, inanimate
Linguistic form	subject prefix, object clitic, free pronoun, possessive pronoun
Agreement features	gender, person, null
Discourse topical	yes, no

The first three properties are simple to determine[4]. The last property, discourse topicality, is much harder to code independently of the gender and person markers that (it is to be argued) provide part of the linguistic evidence for it. Identifying the statuses of referents in the minds of interlocutors can be difficult even when it is the linguist's native language (see Newmeyer 1998:133 for a pessimistic view), and an undeniable disadvantage of the methodology used here is the fact that the analysis has been carried out by a non-native speaker, whose intuitions with respect to conceptual structure stem from his understanding of the English translations that he himself has prepared. Despite such difficulties, as Tomlin et al. (1997:101) have observed, "The most conventional method of linguistic analysis – the introspective examination of discourse data – remains a central and important strategy in discourse studies"[5], and some of the major problems associated with introspection as a research technique are eliminated by the reliance in this study on relatively 'authentic' discourse data, rather than examples constructed in elicitation sessions.

Moreover, there are not too many viable alternative research methodologies when

---

4   The absence of overt agreement was interpreted according to the analysis in §7.6.7.
5   'Introspective' in the sense that the linguist reconstructs the discourse structure of the text through introspection, rather than the data itself being created through introspection.

it comes to the analysis of topics. Text-counting strategies (Tomlin et al. 1997:101-102, Myhill 2001:165-168) such as those employed by Givón (1983) were mentioned in §2.3.2.1. The work of Givón and his collaborators has been very influential and it is undeniable that this kind of methodology can lead to a better understanding of the pragmatic function of referential expressions. Hopper and Thompson (1984:711) talk about the 'manipulability' (or 'deployability') of forms, noting that the coding of referents often depends on how *important* they are going to be in the discourse subsequent to their introduction (i.e. their "discourse referentiality" as opposed to the semantic notion of "objective referentiality" – Payne 1997:264-266). A straightforward measure of discourse importance is the number of times that referents go on to be mentioned following their initial introduction. For example, Jaggar's (1983, 1988) research on Hausa narratives shows that if the initial NP is marked with the specific indefinite determiner *wani*, then the referent is likely to be mentioned more often in the subsequent discourse[6], a correlation which holds regardless of either the animacy of the referent or the grammatical function of the NP. The use of quantitative methodology thus provides hard evidence allowing Jaggar to improve upon the earlier analyses of *wani* described in Jaggar (1988:46).

So why not take a similar approach to the present problem and test the correlation between the use of gender/person agreement and the number of subsequent mentions? *A priori* we can predict that such an analysis would indeed show that person agreement is associated with a higher number of subsequent mentions than gender agreement, since there is of course overlap between Givón and Dooley's different approaches to discourse topic (§2.3). It is usually the case that what we explicitly refer to in texts are those things which hold intrinsic interest for us. The high degree to which discourse topics (in Dooley's sense) integrate texts *conceptually* makes them prime candidates for continued linguistic expression, and conversely the more the speaker mentions a referent, the more likely it is that the hearer will be encouraged to construe the unfolding discourse as relevant to that particular referent (i.e. it will become for him a discourse topic). Nevertheless this is by no means always the case, and while a text-counting approach would almost certainly reveal generalities at the *corpus* level, it is not fine-grained enough to account for the individual examples to be discussed in §8.4. As was stressed in §2.3, discourse topicality involves more than just referential density

---

6   As noted in §4.4.5.3 the pre-nominal article in Cicipu may well function in a similar way.

– the interlocutors must also have an intrinsic interest in the referent, and it is impossible to determine this solely by the inspection of formal properties of the text.

The other major methodological strategy mentioned by Tomlin et al. is experimental studies. I have already discussed the problems I experienced in this area in §1.4. Now that the alternation between gender and person agreement is better understood, the possibility is open for more targeted experimental work in the future (see §9.3).

In addition to the properties coded for individual markers, I also recorded the nature of the series of agreement markers (i.e. anaphoric chain) that they took part in. The following *a priori* possibilities exist:

- Consistent gender-marking
- Consistent person-marking
- Gender-marking followed by person-marking
- Person-marking followed by gender-marking
- More complex mixes of gender- and person-marking

We will see in §8.5 that the last two possibilities are strongly disfavoured in Cicipu. In other words, series of person markers are either consistent with respect to the agreement features they encode, or they show a progression from gender-marking to person-marking.

8.3 Participant reference and marked topics in Cicipu

8.3.1 Participant reference

As in many languages (Dooley and Levinsohn 2001:119-123), the way in which participants are introduced into a text in Cicipu depends on their "referential importance" (Chafe 1994:88), or to put it another way, the degree to which they will go on to integrate the text, in the sense that the remainder of the text is construed to be about them (see §2.3.2.2). The first-mentioned discourse participant in a narrative (which often turns out to be the most important) may be introduced with a special presentational construction, particularly in folktales (examples 3-6 below), which have a largely formulaic beginning in Cicipu. Example (2) is taken from a Cicipu Pear Story narrative (see §1.4).

(2) **wú-nà z-zá v-ì,** ánà ù-kúmbà ú-ɗángà
3s-ART NC**8-person** AG8-COP when 3s-climb\RLS NC7-tree

***there was a man**, when he climbed a tree*

[tapf002.004.008]

(3) **áa Ø-lɔ́wɔ́lì h-è,** èrè ǹ Ø-náatà
PL NC**8-large_spider** AG2-COP 3P.PRO and NC8-small_spider

***Big Spider and associates**, him and Small Spider*

[saat001.008.005]

(4) **wú-nà z-zá gò,** éllè ù-yô ǹ cì-níyáa
3s-ART NC**8-person** TOP that 3s-be\RLS and NC6-farm

á↓=u-nóo wú↓=kú-jénè
LOC=NC7-mouth AG7=NC9-river

***once there was a man**, he had a farm by the side of a river*

[saat001.003.005]

(5) **Ø-náatà gó** éré ǹ kù-lóngí,
NC**8-spider** TOP 3P.PRO and NC9-monitor_lizard

Ø-náatà, ù-yâa tù-mócì
NC8-spider 3s-do\RLS NC6-friendship

***Spider** him and Lizard, Spider, he was friends*

[saat001.006.005]

(6) **wú-nà mà-gɛ́i m-è gó**
3s-ART NC**4-sword** AG4-COP TOP

once there was a sword

[samy002.005]

The initial construction in each of these examples is a verbless existential clause, consisting of an NP, often prefaced with the person-marked indefinite article (§8.9.1) as in (2), (4) and (6), and always followed by either the copula (2-3), the topicaliser **gó** (4-5), or both the copula *and* the topicaliser (6).

Other important participants may be introduced in a marked construction such as an extra-clausal left-dislocated NP, but more commonly they occur in the focal domain of a clause as the object of a verb or instrumental preposition (e.g. *they saw a..., he was with a...*). Marginal participants or props are mainly introduced as the objects of verbs or prepositions.

New participants may also be introduced using subject NPs, contrary to Chafe's (1994:82-92) 'light subject constraint' hypothesis. Chafe found that subjects in conversational English *always* express either referents that are active or semi-active, or that are trivial in importance. Important referents are never introduced by NPs

372

functioning as subject. However in the nine Pear Story narratives that I recorded in Cicipu, in every single one the main character[7] was introduced by the subject of an intransitive motion verb. In more traditional Cicipu narratives the use of subject NPs to introduce main characters is less frequent, but it does happen. Clearly the light subject constraint does not apply as strongly to Cicipu as it does to English. As has been pointed out (e.g. Khorounjaia and Tolchinsky 2004), Chafe's hypothesis does not distinguish between transitive and intransitive subjects, and Du Bois' (1987) Preferred Argument Structure is a better fit with the Cicipu data, since it places no restriction on intransitive subjects (S), only on transitive subjects (A), which speakers prefer to reserve for 'given' referents.

Once a participant has been introduced it tends, unsurprisingly, to be referred to using minimal coding (§2.3.2.5), especially if it is human. For subjects the minimal coding is an agreement prefix, for objects it is a clitic, and for non-arguments it is a free pronoun. Anaphors may agree with their antecedent in either gender or person, as discussed in chapter 7. Introduction with a full NP and then minimal subject coding is illustrated in (7). Note how gender subject agreement occurs at first, followed by person subject agreement in the last clause of the extract. This foreshadows the progression from gender to person agreement that will be discussed in §8.5.

(7) ánà   ù-kúmbà-nà /   sée   **'ínà**   **z-zá**   yâa-nà /
     when   3s-climb\RLS-PFV   then   **certain**   NC**8-person**   arrive\RLS-VENT

     'ínà   **m-ɔ́ɔ**   mà-yâa-nà   é↓=k-kèeké /
     certain   NC**4-child**   AG4-arrive\RLS-VENT   LOC=NC8-bicycle

     *when he had climbed / then **a certain person** arrived / **a certain child** arrived on a bike /*

     ánà   **mà**-yâa-nà-nà /   **m**-áyà   **mɔ̀**-pɔ́lù   Ø-kèeké
     when   **AG4**-arrive\RLS-VENT-PFV   **AG4**-come\RLS   **AG4**-put_down\RLS   NC8-bicycle

     ù-kábà-nà /   mò-kúlòotû   mò-tò /
     3s-take\RLS-VENT   NC4-basket   AG4-one

     *when **he** arrived / then **he** came and **he** put down the bike and **he** took / one basket /*

                                                                        [tapf002.004.014]

Most discourse topics are human or at least animate. It is in our nature to talk about, and to be interested in, things that are most like ourselves. Nevertheless, sometimes we do have an interest in inanimate entities, in which case they may become discourse topics.

---

7   The boy on the bicycle.

Nevertheless the kind of topic chain we can see in (7) above is rarely seen for inanimate entities, regardless of how interested in them the interlocutors might be. In part this is because inanimates cannot generally be agentive and hence are rarely found as subjects in any case. But even in places where an inanimate referent *is* repeatedly encoded as subject, lengthy unbroken anaphor chains are not the norm, and instead the lexical NP is repeated more often. It is as if, despite the discourse topicality of the referent, its lack of inherent topicality requires it to be more frequently-mentioned than would be the case for a human or animate referent. This is illustrated in (8) below[8]. References to the Koran are in **bold**. Note how frequently the lexical NP occurs, despite the obvious discourse topicality of the referent and the lack of any competing referents.

(8)   OK **the Koran**, AG**8**-has a verse for example of (.) of this thing (.) of 'permissions' (.) of God (.) **the Koran** AG**8**-has in a permission of God (.) if you do, if you hold **the Koran** (.) OK forever (.) OK I know for sure, **the Koran** AG**8**-knocks down people (.) knocks down people (.) **the Koran** (.) AG**8**-kills people (.) **this big Koran** (.) when they place AG**8**-PRO, and then you touch AG**8**-PRO (.) OK AG**8**-kills people

[tats005.001.024]

This property of inanimate referents should be borne in mind throughout this chapter, especially in §8.5 on agreement progressions. Long agreement progressions starting with a lexical NP are rare for inanimate referents, in large part because the lexical NP is often repeated.

8.3.2   Marked and contrastive topics

In §2.3.2.6 we noted the difference between 'unmarked' and 'marked' topics. According to Dooley's (2007) use of the terms, unmarked topics, which are expressed using minimal coding, carry out the *integration* function of topics, whereas marked topics, which have more than minimal coding, often carry out the *access* function. Marked topics may also be used to set up a contrast between two or more different topics.

Cicipu has two overt topicalisers which may be used to mark topics, the particles **gó**[9] and **hínà**. It is not easy to pinpoint the difference between the two. Both can be found after NPs, in which case they are commonly left-dislocated topics, although they

---

8   For reasons of readability and space the English translation is given here and for other examples in this chapter. See the front matter for the key to the grammatical markup.

9   The tone on **go** is variable and not well-understood. High-tone seems to predominate, however, and in some cases the tone even seems to be extra-high.

may also occur at clause-level. **Hínà**, however, is more likely to be found after NPs, and it is most often used for contrastive topics as in the metalinguistic statement in (9).

(9) v-áarì vì-tò, á-arí **hínà** ò-yô gɛ̃́i
  NC8-man AG8-one NC2-man **TOP** 3P-be\RLS many
  'man' is one, as for 'men' they are many

[eabg001.048]

Gó often occurs when introducing major discourse participants, as was noted in §8.3.1. Naturally such participants are also discourse topics for large portions of the text. **Gó** also regularly occurs in the protasis of conditionals (10), or after sentence adverbials functioning as 'space-builders' (Fauconnier 1994) or 'points of departure' (Dooley and Levinsohn 2001:68-69), as in (11) and (12).

(10) ń **gò** v-ú-nɔ̀ɔ=mù, tò íd-dòonù
  if **TOP** 2S-FUT-give\IRR=1S.PRO OK 2S-sit\IRR
  if you'll give me [meat], then sit down

[Tidipo, saat002.003.057]

(11) ánnà **gó** h-ú-yìndà z-zá n-nà w-ú-ràa
  today **TOP** 3P-FUT-see\IRR NC8-person AG8-REL 3S-FUT-eat\IRR
  today they'll see the person who'll win

[Tidipo, saat002.005.042]

(12) kíilà **gó** sáa kò-tôo tì-sáa tíyó
  perhaps **TOP** or AG1-one 1P-might get
  perhaps we might have got every one

[saat001.006.054]

The topic-like nature of conditionals and other space-builders has been observed by Haiman (1978, see also Dancygier and Sweetser 2005:125), and the fact that **gó** in Cicipu marks both discourse topics and space-builders provides further evidence for this connection.

 Although the particles **gó** and **hínà** are the most explicit indicators of topics in their access function, the rest of this chapter will not be particularly concerned with them. Most topics are not in fact marked with either of these particles, and in any case, this chapter focuses on the integration function of anaphoric agreement, since that is principally where the gender/person alternation is found.

 Before closing the discussion on contrastive topics, it should be noted that they may also be expressed through a contrast in deixis. The "near-hearer" demonstrative pronoun **é-llè** (§4.4.3.4) seems to imply a continued use of a current topic, whereas the

"yonder" demonstrative **é-"índè** reactivates an earlier topic in contrast to the current one. This can be seen in the following example:

(13) [Context: This excerpt is taken from a discourse on tobacco and alcohol. After talking about the dangers of alcohol, the speaker then moved on to tobacco. The discussion was summarised as follows:]

v-índà **é-llè** ù-kôo, Ø-sánàdíyà ví↓=tá-abà, **é-"índè** kúmá,
2s-see\RLS **3s.PRO-that** 3s-die\RLS NC8-cause AG8=NC6-tobacco **3s.PRO-yonder** and

w-áyà ù-kôo, sɔ́bɔ̀ Ø-gógóró
3s-come\RLS 3s-die\RLS because NC8-gin

*you see, **this one** died, because of tobacco, and **that other one**, he died, because of gin*

[samoh001.259]

8.4 Animacy and discourse topicality as agreement conditions

The main aim of this section is to show that both animacy and discourse topicality are conditions on agreement for Cicipu person markers. For didactic purposes it seems best to hold the context-independent property of animacy constant while varying the context-dependent property of discourse topicality. So for each category of animacy starting from the inanimates and building up to human referents[10], I will consider the expression of first non-topical referents and then topical referents. If the hypothesis put forward at the beginning of this chapter (which was diagrammed in Figure 38) is accurate, then we expect that for each level of animacy topical referents will be more likely to be person-marked than non-topical referents. And as we ascend the animacy hierarchy, we also expect that the 'cut-off point' between gender-marking and person-marking should become lower – in other words the more animate the referent, the less topical it should have to be in order to achieve person-marking.

In addition to animacy and topicality, the noun class of the controlling referent is also a condition for agreement. Everything else being equal, nouns from class 8 (gender 8/2 or 8/3, or the single class gender 8[11]) are more likely to trigger person agreement than nouns from other classes. This condition is treated first (§8.4.1) since it is relevant for parts of the later discussion.

Clear examples of the alternation where all relevant factors are held constant except discourse topicality are very rare, only occurring in the corpus 'accidentally'

---

10 Only third-person referents are relevant here – there is no gender agreement for first- and second-person referents.

11 I have no data with respect to the plurals of gender 7/8.

rather than as a result of the topic-stimulation sessions. Nevertheless, when such contrasts do occur they are striking, and provide the clearest evidence for the account being put forward in this chapter. Therefore I will begin the discussion of animacy and discourse topicality by considering a 'minimal pair' of this kind (§8.4.2). The rest of the evidence for the relevance of discourse topicality comes from generalisations across texts, and the remaining observations in this subsection (§8.4.3) are of this kind[12].

8.4.1   Noun class 8

ɴᴄ8 nouns are a special case, in that although there are perfectly good ᴀɢ8 agreement forms for the person markers, in comparison to the other noun classes they are less often used. Instead we often find person agreement where we might have predicted gender agreement. This is especially obvious in texts where the same referent is referred to using more than one noun. In one topic-stimulation text about the crocodile, for example, the crocodile is denoted both by the 9/2 noun **kù-yúpù** 'crocodile', and by its 8/3 hypernym **kwáarò** 'creature' (from Hausa *kwaro*), as can be seen in (14), taken from the start of the text. In the part of the excerpt following the 8/3 noun (b-k), the speaker progresses to person agreement much more readily than in the part following the 9/2 noun (l-t).

---

12 Although see §8.5.2 for a text-level progression in discourse topicality involving a single referent.

(14) a. **the crocodile (kù-yúpù, NC9)** /

b. it's **a creature (kwáarò, NC8)** /

c which AG8-is..., 3S.PRO /

d. 3S-is in the water all the time /

e. living place-3S.POSS is there, in the water /

f. when 3S-comes to the bank /

g. then 3S-wants 3S-take some air /

h. 3S-will be able to come outside, 3S-lies down and 3S-takes air /

i. after a little time, 3S-goes back to the water /

j. whenever not in the water /

k. OK, 3S.PRO, 3S-doesn't understand /

l. and AG9-PRO **the crocodile (kù-yúpù, NC9)** /

m. AG9-is a fearful thing for people /

n. any water where there is said to be **a crocodile (kù-yúpù, NC9)** inside /

o. if a **crocodile (kù-yúpù, NC9)** is sensed to be there inside /

p. OK, people won't..., people will be afraid to enter that water /

q. since, **the crocodile (kù-yúpù, NC9)** /

r. AG9-grabs people /

s. AG9-can grab a person /

t. AG9-will kill /

u. in the water /

v. because a person in the water doesn't have weight /

w. when someone enters /

x. 3S.PRO when 3S-comes and 3S-takes someone and 3S-pulls

[tats002.006.001]

A similar contrast is found across two separate topic-stimulation sessions involving the gender 8 noun **gógóró** 'gin' (again borrowed from Hausa) and the gender 4 noun **mò-yóo** 'beer' respectively. The anaphors of **gógóró** progress much more quickly to person agreement than those of **mò-yóo**, even though there is little or no difference between the two referents with respect to animacy. This condition applies to both human nouns from gender 8/2 as well as lower animates and inanimates from 8/3.

8.4.2   A 'minimal pair' showing contrast in topicality

Consider the following two examples, which come from separate texts. The first

example (15) is from a topic-stimulation session about monkeys, the second (16) from one about dogs. Common to both texts is that they include a section about a dog chasing another animal – the monkey in the first instance, and a hare in the second. The crucial difference is that the dog is not a discourse topic in the first excerpt, but it is in the second. The 'chasees' also differ with respect to topicality – in (15) the monkey is a discourse topic and the central participant of the text. By contrast, in (16) the hare is incidental to the discourse and is only mentioned in order to demonstrate a particular characteristic of the dog – any other suitable animal could have taken its place. Observe the contrasting means by which the dog is referred to in each of the examples (**bold** references are to the **dog**, underlined references are to the monkey and the hare):

(15) ǹ w-ǐndà **mà-wáa** mò-wốsồ / hálì **m**-ǔuwà kù-sṹu
    and 3s-see\RLS **NC4-dog** AG4-bark\RLS even AG4-feel\RLS NC9-smell

    kwéevì mò'úngònò **m**ồdónì vì nìiládǐ /
    kw-éevì mò-'úngò-nò **m**ồ-dónồ = vì n = ì-ládǐ /
    AG9-3s.POSS AG4-get_up\RLS-VENT AG4-follow\RLS=3s.PRO and=NC3-speed

    *if it sees **the dog** bark / **it** feels its smell **it** gets up and **it** follows it fast*

    tò, n = ù-súmà / ù-náhà ù-gítà-nà mázámázá
    OK and=3s-run\RLS 3s-leave\RLS NC7-return-VENT hurriedly

    OK, when it runs / it won't come back in a hurry

                                   [tats002.001.053]

(16) tò, lóokàcí **wú-nà mà-wáa,** ǹ **m**ồ-dónồ mà-díyá /
    OK time **3s-ART NC4-dog** and AG4-follow\RLS NC4-hare

    n = **ù**-húnà mà-díyá á↓ = ká-dábá /
    and=**3s**-kill\RLS NC4-hare LOC=NC1-bush

    *OK, **some dogs** [lit. 'a certain dog'], when **it** follows a hare / when **it** kills a hare in the bush /*

    ù-sì-ràa cé 'ũ̂ / **ù**-sì-ràa cé /
    3s-HAB-eat NEG there **3s**-HAB-eat NEG

    sée **ú**-tòonò ǹ m-é /
    unless **3s**-come_back_home\IRR with AG4-PRO

    ***it** doesn't eat [it] there / **it** doesn't eat [it] / then **it** comes back home with it /*
                                     [tats001.001.039]

In (15), where the monkey is discourse-topical and the dog is not, the dog is consistently indexed by subject gender agreement (**mA-**), while the monkey is consistently indexed with person agreement (in a variety of morphosyntactic positions). In (16), where the dog is topical, it is at first marked with gender agreement, but crucially, it graduates to person agreement and doesn't revert to gender agreement, at least until after the next

lexical mention following this excerpt. The hare, on the other hand, is either marked with gender agreement or not at all[13]. All references to the dog across both the examples are subjects, and in (15) the referent of a subject indexed for gender acts on the referent of an object marked for person (**mɔ̀-dɔ́nì=vì**), illustrating that it is not the semantic role of the referent that determines the target agreement features. It is also important to note that the propositions encoded in several of the individual clauses in (15) are *about* the dog, and so the dog is a sentence topic according to Lambrecht's approach (see discussion of the 'grandma' example (2.36) in §2.3.2.2). It is only by considering the role of the participants in the wider discourse (i.e. discourse topicality) that the differences in morphosyntactic coding can be accounted for.

The third intonation unit in (15) (beginning **n=ù-súmà**) provides an interesting parallel to the second intonation unit in (16) (beginning **n=ù-húnà**). In both cases the previous clause encoded the dog as subject by means of anaphoric gender agreement. In (16) the subject remains the same going into the second intonation unit, seamlessly transitioning from gender agreement to person agreement. However in (15), the subject of the verb **ù-súmà** is now the monkey, no longer the dog. Note that apart from the association of person agreement with topic that I am arguing for here, everything else points to *topic continuity*. The subject is expressed using minimal coding and there is no discontinuity of action: the dog is following the monkey with speed in one clause, and then something is running off in the next and neglecting to come back. Had the speaker been making a point that dogs often disappeared off on their own while chasing monkeys, then no doubt this sentence could have been used with the dog remaining as the subject referent. However because the hearer understands that in (15) the monkey and not the dog is topical, the speaker can use the 3PS marker **u-** without fear of ambiguity. Animal referents can only be encoded using person agreement if they are discourse-topical, and so **u-** can only refer to the monkey in this text.

This pair of examples also brings home the fact that discourse topicality is about more than just referential density (recall the 'football' example from §2.3.2.2). In (15) four references to the dog are made in the space of two intonation units, but it nowhere triggers person agreement. Simply talking about a referent is not enough to make it a discourse topic. Instead, it is the speaker's *intrinsic interest* that elevates frequently-mentioned referents to the status of discourse topic, and, as far as non-human referents

---

13 See §4.6.4 for object omission.

in Cicipu are concerned, allows them to graduate from anaphoric gender agreement to anaphoric person agreement.

We now move on to look more methodically at the anaphors used to refer to non-topical and topical referents, moving up the animacy hierarchy from inanimates through animals to humans. We will find that in principle, most referents can be found with either gender or person agreement provided they are discourse-topical enough[14]. Nevertheless human referents are far more *likely* to be found triggering person agreement than animal referents, which in turn are more likely to trigger person agreement than inanimate referents. This is true for each of the morphosyntactic positions covered in chapter 7: subject prefix, post-verbal object, and elsewhere.

8.4.3   Inanimates

8.4.3.1   Inanimates, non-topical

Referents which are both inanimate and non-topical rarely function as the subjects of clauses, and in clauses where they might be expected to occur as objects (i.e. because they are a central participant in the verb's semantic structure) they are often omitted altogether (§4.6.4). On occasion, however, anaphoric chains of inanimate non-topical referents do occur. As predicted by the diagram in Figure 38 such referents are without exception coded using gender agreement.

Two examples are given below. The first describes how a monkey escapes from being tied up. The discourse topics are the monkey, and also the owner of the monkey. The rope is only of interest because when it slips off the monkey can escape; by itself it holds no intrinsic interest for the interlocutors. Similarly in the second example, the hoe is only the means by which one of the main discourse participants is made to wake up. It is of no interest by itself.

(17) **kà-mángá** kò-sírò-nò   **kɔ̀**-pɔ́'ɔ̃   **kò**-yúwò-nò,
     NC**1-rope** AG1-slip_off\RLS-VENT AG1-remove\RLS AG1-fall\RLS-VENT
     kà-náhà   vɔ́-ɔmɔ̀   'ũ̀
     AG**1**-leave\RLS NC8-monkey way_over_there
     the rope *slipped down **it** came off **it** fell, **it** left the monkey up there*
                                                              [saat002.002.536]

---

14 A set of examples from the corpus can be inspected in Appendix C, covering the 18 combinations of agreement feature (x2 – person, gender), animacy level (x3 – human, animal, inanimate), and morphosyntactic position (x3 – subject prefix, post-verbal object, and 'elsewhere').

(18) sée **kù-béyí** kù-sá'ū̀-nà z-zá '-'índè /
until **NC9-hoe** AG9-touch\RLS-VENT NC8-person AG8-yonder

ánà **kù**sá'ū̀nìvì **kù**gáatànìvì címmā̀i /
ánà **kù**-sá'ū̀-nà=vì **kù**-gáatà-nà=vì címmā̀i /
when AG9-touch\RLS-PFV=3S.PRO AG9-cut\RLS-VENT=3S.PRO little

*then **the hoe** touched that one / when **it** touched him, **it** cut him slightly /*

[Tidipo, saat002.004.053]

8.4.3.2 Inanimates, topical

NPs referring to inanimate referents rarely trigger person agreement, even when they are topical. In the whole corpus there are only seventeen inanimate referents cross-referenced by person agreement markers, ten of which surfaced in topic-stimulation texts deliberately designed to stimulate this kind of infrequent data. Nevertheless, despite the rarity of the combination, a fairly wide variety of inanimate referents have been observed indexed by person agreement markers. Examples from topic-stimulation texts include beer, gin, tobacco, the Bible, and the Koran. Examples from other kinds of texts include palm-wine, fish trapping, trees, a lake, a fence, and clothing. In each of the seventeen cases the referent is a discourse topic, and in the majority of cases there is no reason to think personification or metonymy might be involved. As observed in the previous subsection, person agreement never occurs with non-topical referents. Examples (19) and (20) show person agreement with NPs denoting the Koran and tobacco respectively, taken from topic-stimulation sessions. The topicality of the referents should be self-evident in both examples.

(19) a. **the Koran (kùrì'áanì, NC8)** AG8-forbids all kinds of things /

b. the things **3s**-forbids, are two places, you see **3s**-forbids beer /

c **3s**-forbids beer /

d. **3s**-forbids, evil eating /

e. you will do eating, you will eat you will eat you will eat you are full to bursting! /

f. OK **3s**-forbids, **3s**-forbids that thing /

g. yes, like that xxx it is forbidden like that /

h. but **3s.PRO the Koran**, **3s**-stands between us and God

[tats005.001.189]

382

(20) a. OK **3s.pro tobacco(tá-abà, nc6)** when you are smoking **ag6-pro** /

b. if you are healthy /

c you will find your heart blackening, heart blackening /

d. until **3s-does, 3s-spoils** for a person the lungs that are inside /

e. OK heart, near to the heart /

f. **3s-will block** the holes of the lungs

<div align="right">[tats007.002.033]</div>

The information given in (21) below was unsolicited, and comes from a passage about various methods for trapping fish. The speaker lists several different techniques and describes each one. The different kinds of trapping therefore form a series of contrastive discourse topics. The AG1 agreement markers derive from a general word for trapping, **kà-yímá** (NC1). Note how in the second line the person-marked article **wú-nà** (§8.9.1) is used to introduce a new kind of trapping, in contrast to the gender-marked **ká-nà** in the first line. Reference to this **má-rúwó** trapping continues with person agreement markers in the third line.

(21) àkwái **ká-nà** à-sí-hyằa ká↓=má-gítá / ée, á-kùmbà gíihwá /
there_is AG1-ART 3P-HAB-say AG1=NC4-trapping_k.o. yes 3P-climb\IRR side

ée, òkóo, **wú-nà** à-sí-hyằa, kí↓=i-rí-mpà / má-rúwó /
yes there_is 3s-ART 3P-HAB-say AG1=NC3-thing-this NC4-trapping_k.o.

ée, **éví**, ù-yô ǹ=ì-rí-mpàa-nì, à-ggájì
yes 3s.PRO 3s-be\RLS and=NC3-thing-this-NMLZ NC2-shrub_k.o.

*there is **another** [trapping technique] called of 'magita' / yes, they climb the side / yes, there is, **another** [trapping technique] called, of this thing / 'maruwo' / yes, **it**, **it** is with this thing, 'aggaji' /*

<div align="right">[tats005.002.103]</div>

So there are indisputable occurrences of person agreement with inanimate referents, as long as they are discourse topics. Nevertheless being topical is not always sufficient to ensure that anaphors are person-marked, and in fact topical inanimate referents are more often encoded by gender agreement, as demonstrated by the remaining examples in this subsection.

Example (22) is a conversational aside recorded during a session on the history of the Akula division of the Acipu. The speaker had just noted a sacrificial pot which had been knocked over. The example shows a nice contrast between the inanimate pot, which is encoded with gender agreement even though it is a discourse topic (and clearly marked as such by its introduction using a presentational construction), and the human,

who is marked with person agreement despite being non-topical and unknown to the speaker.

(22) v-îndà   **má-tíllú mé-llè,** 'ínà z-zá   'úng‹òs›ò-nò   **m-è** /
2s-see\IRR NC4-pot   AG4-that ART NC8-person rise‹CAUS›\RLS-PFV AG4-PRO

mà-kámà   ù-'ísánúu-ní, ù-ɓóngùlò-wò   **m-è**   ɓóngùlò-wò
AG4-be.PST\RLS   NC7-stand-NMLZ 3s-fall_down\RLS-APPL AG4-PRO fall_down\RLS-APPL

*you see **that pot**, someone caused **it** to rise up / **it** was standing, he/she pushed **it** over*

[Tikula, svsdt001.070]

The next example comes from the summing up of a topic-stimulation text concerning 'fire'. Like other non-individuated nouns such as water or lightning, fire proves to be highly resistant to person-marking.

(23) sɔ́bɔ̀   ǹ   'íní,   **ù-láa** wù-yô   ǹ=ù-'úsṹ /
because and like_that NC7-fire AG7-be\RLS and=NC7-power

wù-yô   kúmá ǹ   tì-lípằi / **wù**-yô   kúmá   ǹ=ù-pácì /
AG7-be\RLS   also   and NC6-good   AG7-be\RLS   also   and=NC7-difficulty

*because of that, **fire** has power / **it** is also good / **it** is also difficult*

[tats002.002.059]

The final two examples show relatively long anaphoric chains. In the second example (25) the referent ***that meat of that day*** is introduced using a marked topic construction with a resumptive pronoun, but nevertheless does not achieve person agreement.

(24) a. A: there Mappaya it is like a small cave, **the water (mò-ní, NC4)** AG4-stays clean /

b. B: hmm /

c   A: clean! /

d.   only goodness /

e.   they are scrambling to get AG4-PRO /

f. B: and when you fetched AG4-PRO, AG4-would finish all at once or AG4-wouldn't finish?

g. A: AG4-finished but AG4-didn't finish quickly

h.   they did digging

i.   AG4-spurted sweetness!

j.   AG4-surpassed that of Kokko'o in sweetness

[Tikula, sagb001.648]

384

(25) a. "you **that meat (ì-námà, NC3) of that day** even now you haven't eaten AG3-PRO yet?"

b. he said "no! I ate Ø long ago"

c he said "aha!"

d. I knew things would turn out like this

e. really AG3-lasted a long time

f. AG3-lasted a long time

g. AG3-lasted

h. I know now for sure you've eaten AG3-PRO

i. he said "OK

j. since you've eaten AG3-PRO

k. you've got to give me one wife"

[saat002.002.628]

It is not clear why a speaker will sometimes choose person agreement to refer to a topical inanimate referent and other times gender agreement, and it may not be possible to give a deterministic account. As mentioned in §2.2.4.1 some linguists have treated topicality as more of an agreement 'prerequisite' than a 'condition', such that the presence or absence of agreement can be precisely determined according to the topicality of the controlling referent. Concerning zero anaphors in Chinese, for example, Li and Thompson (1979:312) maintain that their distribution is in principle predictable, but that "the principle contains variables dependent on the speaker's perception of the pragmatic situation". This position does not seem tenable for the Cicipu alternation, since in the examples above the referents concerned are indisputably discourse topics, and yet they are still not marked with person agreement. In Cicipu topicality is necessary (but not sufficient) for person agreement to occur with an inanimate referent.

8.4.4 Animals

In the texts I have collected, and of course more generally in discourse, animal referents are rarely topics. Although they have greater potential than inanimates to take part in anaphoric chains, and are more often the subjects of clauses, they are rarely of intrinsic interest to the interlocutors. Instead the situations in which they take part reveal something about or of interest to some other topical discourse participant, usually one of the speech participants or another human.

Animal referents fall in-between inanimates and humans in the various versions of

the cross-linguistic animacy hierarchy (e.g. Silverstein 1976, Frawley 1992:95). Similarly in Cicipu animals can be distinguished from inanimates on the one hand and humans on the other with respect to the gender/person alternation. As we saw above, inanimate referents are rarely encoded using person agreement markers, even when topical. Humans, on the other hand, are rarely encoded using gender agreement markers, even when non-topical. The alternation has a more balanced character as far as animal referents are concerned, and thus animals reflect the importance of discourse topicality more clearly than other kinds of referents.

8.4.4.1 Animals, non-topical

Non-topical animal referents are without fail encoded using gender agreement. The dog in (15) above was one such referent. Another example is given in (26) below, taken from a folktale about a hunter stuck to a rock. The hare is a minor character in this text, only being introduced right at the end, in unit 120 of 132. The sole purpose of the hare is to provide a means of freeing the hunter, who is the main character and a discourse topic. Observe that the hare is able to speak in this story, and so has been anthropomorphised, and is therefore higher up the animacy hierarchy than normal animals. Nevertheless it is not encoded using person marking, despite ample opportunity for this. This is because it is not a discourse topic.

(26) a. then AG4-came **the hare (mà-díyá, NC4)** AG4-came AG4-said "OK get ready", AG4-came **the hare(NC4)** AG4-came AG4-went yonder far off /

b. AG4-came AG4-ran towards AG4-came AG4-collided with him crash! /

[Tidipo, saat002.003.127]

8.4.4.2 Animals, topical

In contrast to non-topical animals, topical animals are very often indexed using person agreement, and in fact the corpus does not contain any examples of topical animal referents which fail to progress from gender-marking to person-marking. Example (27) comes from a topic-stimulation text about the tortoise, which is easily seen to be a discourse topic in the excerpt presented.

(27) a. **the tortoise (cì-kúlú, NC6)** /

     b. God created **3s.OBJ** well /

     c when God created **3s.OBJ** /

     d. he beat **3s.OBJ** together with an iron bowl /

     e. because whenever /

     f. **3s**-travels /

     g. the sun doesn't bother **3s.OBJ** /

     h. if **3s**-feels the sun bothering **3s.OBJ** /

     i. then **3s**-enters self-**3s.POSS** into house-**3s.POSS** /

     j. then **3s**-hides...

<div align="right">[tats004.002.006]</div>

The next example comes from a folk history of agriculture. The donkeys are initially marked with gender agreement in (a) but as the discourse progresses they are referred to using person agreement markers.

(28) a. you see **the donkey (mà-jàkíi, 4/5)** if you do guineacorn, it's **AG5-them** that **AG5**-will carry the guineacorn for you /

     b. **3s**-will take for you to the house, camels(NC2) formerly it is said /

     c when they took camels(NC2) they went with **AG2-them** /

     d. xxx[15]-were destroying the farm it was said it would be better to return the camels and come back to **donkeys(NC5)** /

     e. they (people) would do for **3P.OBJ** (donkeys) load standing /

     f. after they finished loading, you see they (people) took **3P.OBJ** (donkeys) /

<div align="right">[samoh001.196]</div>

One further example comes from an interview of an old man conducted by his son. The main topic of the passage below is a particular horse, introduced by the Hausa loanword **má-hàukàcí** 'lunatic, NC4'. The next anaphoric reference is a gender-marked pronoun (b), but thereafter the horse is referred to using only person agreement morphology. The use of the word **má-hàukàcí** seems to indicate a certain anthropomorphism, and it might be thought that this is the reason for the person-marking rather than topicality. However note the plural reference in (g), which refers to the (non-anthropomorphised) horses in general.

---

15 Unfortunately it is not clear from the recording whether the camels are denoted here using person (**a-**) or gender (**ha-**) agreement.

(29) a. Son: so you knew **horse (d-dɔ̀ɔ, NC8)** well? /

    b. Father: a! **horse** now back then one **lunatic (má-hàukàcí, NC4) of a horse** we had, if Kooki wasn't there /

    c.   there was no-one who could catch **AG4-PRO** /

    d. Son: OK /

    e. Father: OK going there, by Molloci /

    f.   by Maguji there, by Tizebi /

    g.   where they would find **3P.OBJ** in the marsh /

    h.   grazing /

    i.   with arriving /

    j.   the one who knew unless he called **3S.OBJ**, now there /

    k.   you would find **3s-**would come /

                                  [sayb001.500]

Finally example (30) is in the form of a riddle. Riddle-telling of course deviates from the normal co-operation that is assumed to occur between interlocutors (Grice 1975), in that the speaker does not want to be (too) easily understood. If the hearers do not know the riddle, then they do not have the same referents and potential topics available to them as the speaker. At one level the main discourse topic of a riddle is the answer, in this case 'a broom'. However on the surface level the broom is not mentioned at all, and it is the animate "cows of our father" that are the discourse topic, occurring left-dislocated in a marked topic construction, with a person-marked resumptive pronominal clitic.

(30) ì-náa   y=Ø-ɔ́sɔ́ɔ   ví-ttù   páttì kà-mángá kò-tò   k-é
       NC3-cow AG3=NC8-father AG8-1P.POSS **all**   NC1-rope   AG1-one AG1-COP

     à-sí-và'ã̀=rè
     3P-HAB-tie=**3P.PRO**

     all the cattle of our father *it's with one rope they tie* ***them***
                                     [saat001.002.059]

8.4.5 Humans

Humans are of course the archetypal topics (e.g Kuno 1976, Payne 1997:151), but it does not follow that *all* human referents encoded in a text are discourse topics. Humans can be of trivial importance in a text, despite their inherent topicality. If a referent fails to persist and therefore does not integrate the discourse schema for a text, or if the interlocutors have no intrinsic interest in that referent, then it cannot be said to be a

discourse topic. So we might expect to find some examples of non-topical human referents in the corpus, and thus it is a non-trivial question as to whether they will be marked with gender or person agreement.

In fact, the possibility of gender-marking for human referents is restricted by a lexical property of the word for 'person' **z-zá**, discussed in §7.6.1.1. The majority of male[16] referents in the corpus are introduced into the discourse either by name, or using an NP headed by **z-zá**, and neither of these cases allows the possibility of gender agreement. Moreover, most of the remaining NPs used to introduce human referents are headed by NC8 nouns – recall from §8.4.1 that these also favour person rather than gender agreement. Ruling out referents introduced by **z-zá** and other NC8 nouns leaves only eleven corpus examples of nouns used to introduce adult humans[17], eight of which have discourse topical referents. Those with non-topical referents will be discussed first.

8.4.5.1   Humans, non-topical

Two of the three non-topical cases involve gender subject marking which persists not only beyond the initial verb, but also into the next intonation unit. This can be seen in the following two examples. Example (31) is taken from a sermon about forgiveness. The point of the passage from which the excerpt is taken is to demonstrate certain facets of the character of Jesus, who is a discourse topic of this section (note the left-dislocated personal pronoun in the first intonation unit). The officials are of no intrinsic interest, and are only mentioned in order to reflect the nature of the discourse topic. They are introduced in this excerpt and they are not referred to again. Although they do actually progress to person marking at the end of the third intonation unit, gender agreement persists for longer than might have been expected for human referents, i.e. beyond the intonation unit in which the NP occurs.

---

16 For female referents the word for 'woman' **k-káa** is the 'basic-level' term, rather than **z-zá** (at least for male speech). See §8.4.5.3 on 'natural gender'.

17 Children in the corpus are usually introduced using the word **móɔ** 'child, NC4' (e.g. ex. 7 above). On the basis of the limited data available they seem to come in between animals and adult humans in the animacy hierachy.

(31) èví **ìrígáatí** yìlɔ́wìvì /
èví **ì-rígáatí** yì-lɔ́wɔ̀=vì /
3s.PRO NC3-**official_k.o.** AG3-drive_away\RLS=3s.PRO

yéewìvì / àmáa Yéesù tá'àrè /
yé-ewè=vì / àmáa Yéesù tá'à=rè /
AG3-hate\RLS=3s.PRO but Jesus love\RLS=3P.PRO

*him **the senior officials** drove him away / **they** hated him / but Jesus loved* **them** /

[oamy001.211]

In example (32), taken from a discussion about family history, the referent of *their elder* holds no intrinsic interest for the questioner. He is trying to discover when the interviewee first moved to his present village, and is using the childhood of the referent of *their elder* as a temporal standard for comparison. The referent has no other role and plays no further part in this discussion, and so cannot be said to be a discourse topic. As with the previous example, this may explain the persistence of the gender agreement marking across several intonation units.

(32) a. they said **their elder (kà-bárá, NC1)**, **AG1-**grew up here in this house /

b. OK when **AG1-**grew up here, were you here /

c. **AG1-**grew up here or, were you across there /

[sayb001.084]

8.4.5.2 Humans, topical

Of the eight topical adult human referents introduced using non-NC8 nouns, five are encoded exclusively using person agreement. The other three do admit anaphoric gender agreement, but only within the same intonation unit, in contrast to the non-topical referents discussed in (31) and (32) above. When the reference is taken up again in the next intonation unit, person agreement takes over, as can be seen in (33) and (34), taken from the same folktale. The teacher and the prostitute are two of the main characters in the story, and are discourse topics of the respective excerpts.

(33) **má-llû** m-áyà **mà**-rúbùtà má-tákàddá /
NC4-**teacher** AG4-come\RLS AG4-write\RLS NC4-book

w-áyà **éví** w-ú-nùunà ù-mátúwáa-ní wé-**evì** lêe
3s-come\RLS 3s.PRO 3s-FUT-show\IRR NC7-magic-NMLZ AG7-3s.POSS there

*then **the teacher** wrote a little piece of paper, **he** was going to demonstrate his magic there*

[sahs001.003.004]

(34) dègèlée **kà-káarùwà** k-áyà **kà**-yâa Ø-kɔ́llíyá
    then   **NC1-prostitute** AG1-come\RLS AG1-do\RLS NC8-makeup

    kà-yâa Ø-kɔ́llíyá / **w**-áyà **ù**-rívà mó-ní
    AG1-do\RLS NC8-makeup 3s-come\RLS 3s-look\RLS NC4-water

    *then **the prostitute** put on makeup **she** put on makeup / then **she** looked at the water*

<div align="right">[sahs001.003.010]</div>

8.4.5.3 'Natural' gender

In addition to the three-way division between humans, animals, and inanimates, there is also an argument for setting up two sub-categories for men and women, as diagrammed below:

Humans		Animals	Inanimates
Men	Women		
High probability of person agreement			High probability of gender agreement

Figure 39: How animacy affects the Cicipu gender/person agreement alternation

This subdivision between men and women must be regarded as tentative. It is based partly on my own inspection of the corpus, and partly on discussion with language consultants.

The main evidence from the corpus is the fact that although there are relatively few mentions of women, most of the examples of gender-marked human anaphors have female referents. This is true with respect to both pronouns/clitics and affixes. In general, post-verbal objects with human referents are almost always expressed by person-marked object clitics, rather than gender-marked pronouns. The few examples of gender-marked pronouns in this position all involve female referents[18]. In (35) the AG9 agreement markers derive from **kù-lácí** 'girl, 9/2', the singular form of **ɔ̀-lɔ́cí**.

(35) ká'à cì-mé'ì̃ tɔ́↓=ɔ-lɔ́cí há-nà hò-yó-nò pâa,
    now NC6-inside AG6=**NC2-girl** AG2-REL AG2-be\RLS-PFV here

    v-índà **kú-nà** Ø-yó-nò á↓=u-tá'á sɯ́? /
    2s-see\RLS **AG9-REL** 2s-be\RLS-PFV LOC=NC7-want Q

    ù-hyâa 'ii, w-índà **kw-ì** /
    3s-say\RLS yes 3s-see\RLS AG9-PRO

    *now among the **girls** that are here, do you see **the one** that you want? / he said yes, he saw **her** /*

<div align="right">[saat002.002.123]</div>

---

18 Gender-marked pronouns with male referents *can* be elicited however (2007-02-23.001).

Example (36) below shows a direct contrast between female and male referents. The pronoun referring to the woman is marked with gender agreement, while in the parallel construction afterwards, her husband is indexed with a person-marked object clitic[19].

(36) ǹ-yó ǹ kúmá wú-nà **kà-mɛ́̃yɛ̃̀** /
1s-be\RLS and more 3s-ART **NC1-elder_sibling**

*I also have another **elder sibling** /*

ká-nà kò-yó-nò ù-'ítà pâa / à-sì-hyɛ́a **k-è** /
AG1-REL AG1-be\RLS-PFV NC7-marry here 3P-HAB-say **AG1-PRO**

*who is married here / they call **her** /*

m̀-pándà / Sáahìyà /
1s-forget\RLS [name]

I forget / Sahiya /

ǹ	**váarì**	**véevì**	àsìhyîivì	Mámmà /
ǹ	**v-áarì**	**vé-evì**	à-sì-hyɛ̂a=**vì**	Mámmà /
and	**NC8-husband**	**AG8-3s.POSS**	3P-HAB-say=**3s.PRO**	[name]

*and **her husband** they call **him** Mamma /*

[saat002.008.001]

In addition to pronouns/clitics, subject agreement markers also seem more likely to be inflected for gender when they have female referents. The following three examples come from the same folktale, and all involve topical female referents triggering anaphoric gender subject agreement.

(37) ví-llè t-tò kúmá, **vì**-sì-zìnò-zínò vɔ́-ɔmɔ̀
AG8-that AG8-one more **AG8**-HAB-turn_into-REDUP NC8-monkey

*and that one [**k-káa** 'woman, NC8'], she habitually turned into a monkey*

[saat001.004.008]

(38) ǹ ví-"índè Ø-yôo **vì**-cídà-nà
when AG8-that AG8-go\RLS **AG8**-pluck\RLS-PFV

*when that one [**k-káa** 'woman, NC8'] went picking*

[saat001.004.013]

(39) k-káa ví-llè vá-avù, **vì**-sì-zínò-zíno vɔ́-ɔmɔ̀
NC8-woman AG8-that AG8-2s.POSS **AG8**-HAB-turn_into-REDUP NC8-monkey

that wife of yours, she habitually turns into a monkey

[saat001.004.029]

It should be stressed that gender subject agreement is still not the norm for the women

---

19 It should be mentioned that there is an alternative explanation for this example, since it could be argued that the difference in feature is due to the class of the controlling noun – recall from §8.4.1 that NC8 nouns are more likely to trigger person agreement than nouns of other classes.

in this story – these are the only three occurrences. However it does seem that for NC8 nouns, female referents have a greater potential than male referents to trigger gender subject agreement. Despite the far greater number of nouns with male referents in the corpus, there is only a single example of gender subject agreement involving an NC8 noun with a male referent, taken from a song[20]:

(40) d-dáa   **vì**-hyầa-nà   ń-yầa-wà = ɗò
      NC8-king   AG8-say\RLS-PFV   1S-do\IRR-APPL=2P.PRO
      the king told me to tell you (pl.) [lit. 'do for you']

[sagj001.010.001]

Although these examples may not by themselves be enough to convince the sceptic, more than one language consultant has independently suggested that gender agreement is more likely to occur with female controllers than with male[21]. On one occasion, when I was investigating the phrases in (41), I was told that using the gender-marked pronoun as in (41b) implied that it was a *woman* that was told, not a man.

(41) (a) ǹdámùw**ìvì**   (b) ǹdámùwà   **vì**
          ǹ-dámà-wà = **vì**   ǹ-dámà-wà   **v-ì**
          1S-tell\RLS-APPL=**3S.PRO**   1S-tell\RLS-APPL   **AG8-PRO**
          I told him/her/it   *I told her/it*

[2008-03-28.001]

8.4.6 Sentence and discourse topic

We have now seen the effect of topicality on the indexing of referents at all levels of animacy. The effect is most obvious for inanimates and animals, which are never indexed by person agreement unless they are topics. The relevant data for humans is limited to a small number of examples involving non-NC8 nouns, but there is still a discernible trend for gender agreement to persist longer with non-topical humans at the expense of person agreement, as in (31-32) above. Such examples recall the discussion of Lambrecht's 'grandma' example (2.36) in §2.3.2.2, where transient discourse participants were classed as topical by Lambrecht, but non-topical according to Dooley's theory of discourse topic. Other examples such as (15), (17-18) and (26) are similar in that either (i) the referent is only transient in the discourse, or (ii) even if it

---

20 The song is of the 'call-and-response' type common in sub-Saharan Africa. The female caller consistently used gender agreement for this line, whereas the male responder consistently used person agreement.

21 These consultants were all men, and the question remains as to whether this pattern is repeated or reversed in women's speech.

semantically integrates a paragraph it does not *thematically* integrate it, since it holds no intrinsic interest for the speaker. When one considers individual clauses and propositions in isolation, then the referents encoded by gender agreement in the above-mentioned examples qualify as sentence topics, just like the discourse-topical referents encoded by person agreement. Thus the notion of sentence topicality does not help us to distinguish between the wholly gender-marked referents discussed in the 'non-topical' sections above, and the person-marked referents in the 'topical' sections. If we are to look to topicality to help account for the distinction between gender and person agreement marking, then it is discourse topicality we require rather than sentence topicality.

8.4.7   Conflict with speakers' intuition

Despite the existence of textual examples of unambiguous person agreement occurring with non-human or inanimate subject referents, native speakers, when questioned, maintain that person agreement can only occur with human referents, just as they also often maintain (again, contrary to the textual evidence) that gender agreement cannot occur anaphorically. It is interesting that neither Hoffmann nor Crozier mention such difficulties with respect to Central Kambari, and it seems likely that person agreement is more common with non-human referents in Central Kambari than in Cicipu. This partial loss of a semantic restriction ('bleaching') fits in with the suggestion made in §7.6.8 that the Central Kambari markers are further down the grammaticalisation cline than the Cicipu ones.

It was mentioned in §7.6.1.2 (fn. 20) that with careful negotiation of the context in elicitation sessions, native speakers will accept instances of anaphoric gender agreement as grammatical. It proved much harder to achieve this for constructed examples of person agreement with non-human controller referents. For example, for the following exchange my consultant adamantly refused to accept person agreement (o-) in the answer:

(42) A: hán   **à-mángá**?
          where NC2-rope

     where are the ropes?

      B: (a) **hò**-yúwò-nù   kó-ocì
               AG2-fall\RLS-RES NC9-hole

          they fell down the hole

          (b) ***ò**-yúwò-nù   kó-ocì
               3P-fall\RLS-RES   NC9-hole

[2007-02-05.007]

This refusal to admit person agreement was presumably due to the difficulty, in an artificial environment, of persuading speakers to view referents low in inherent topicality as intrinsically interesting.

The example again illustrates the importance of recognising the notion of discourse topic (rather than just sentence topic). According to Lambrecht's theory of information structure (§2.3.1), the rope in (42) is the topic of the answer sentences. However non-human sentence topics never achieve the level of accessibility required for minimal coding unless they are also discourse topics. To explain the reluctance of native speakers to accept sentences such as (42b), we therefore need to appeal to the notion of discourse topic.

In contrast to my experience with constructed examples in elicitation sessions, when recorded examples of person agreement with non-human controllers were taken from less artificial texts and played back to speakers in context, they were never rejected as wrong.

8.4.8 Unexplained data

There is one major class of exceptions that remains unaccounted for, and this consists of the various kinds of constructions that involve non-referential subjects, including expletives, weather verbs, and impersonal constructions. Subject agreement is obligatory in Cicipu; it cannot simply be omitted. Instead third-person agreement prefixes are found in these constructions. The problem for the current account of Cicipu agreement is that these kind of subjects are neither animate nor topical.

Expletives make use of 3PS agreement, as in (43-44), while in impersonal constructions we find 3PP (45-46), a common technique in Niger-Congo and in other languages that lack true passives (Siewierska 2005).

(43) ù-kámàatà   cé   tí-sɔ̀ɔ   t-áabà
3s-be_supposed_to\RLS NEG 1P-drink\IRR NC6-tobacco
***it** is not proper for us to smoke tobacco*

[tats007.002.077]

(44) ù-'wâa   Ø-káa'ìdà
3s-pass\RLS NC8-principle
***it's** beyond reason*

[tats007.002.006]

(45) ì-láɗámú   kúmá / **hú**-u-yìndà y-ì   cé   ǹ   kw-ándái /
NC3-lightning and   3P-FUT-see\IRR AG3-PRO NEG with NC9-dry_season
and lightning / it is not seen in the dry season / [lit. '**they** will not see it']

[tats002.007.025]

(46) [Context: there is only a single protagonist in this narrative]
sée **há**-asà   kà-hí̃'ĩ̀   kà-yápù
then 3P-spend\IRR NC1-night AG1-two
then two nights passed [lit '**they** spent two nights']

[tats002.001.057]

Weather-verbs are slightly more complex. They are usually found with 3PS agreement e.g. **ù-yɔ́'ɔ̀** '3s-rained', which looks just like the impersonal agreement found in English *it rained*. However on asking a native speaker 'What rained?', the reply is invariably "**kúngwá**", which is ambiguous in meaning between 'sky' and 'God'. **Kúngwá** is also found as the explicit subject of weather verbs on eleven occasions in the corpus, and so it is hard to be sure in cases of anaphoric agreement that the subject is truly non-referential. Examples such as (47) below, which is taken from a topic-stimulation session on lightning, suggest that the subject is at least sometimes referential, since it has a marked topic subject referent, encoded by a left-dislocated NP.

(47) sáa kúmá **Ø-kúngwá,** dá'à   ù-yɔ́'ɔ̀
or   more **NC8-sky**   moreover 3s-rain\RLS
*or **the sky**, **it** rains*

[tats002.007.032]

Nevertheless, even if singular person agreement is understood to refer to the sky or to God, this interpretation does not seem possible with plural person agreement. **ɔ̀-yɔ́'ɔ̀** '3P-rained' is also acceptable, and in that case it was not possible to elicit a potential plural referent. Thus this construction is another example of person agreement with a non-referential subject.

This non-referential use of person agreement seems to be a 'last resort' in the

absence of a suitable gender to choose, although we may wonder why AG8 gender agreement does not occur in such examples, since that is what is normally found in environments where the controller has no specific gender (§6.4). In other Benue-Congo languages such as Sesotho (Demuth 1990), expletive subjects do occur with weather verbs, in which case the locative noun class subject prefix *ho-* occurs. If, as Demuth argues, this prefix is entirely void of semantic content, then weather-verb agreement in Sesotho is an instance of neutral gender agreement[22].

Almost all linguists would classify such subjects as non-topical (e.g. Dooley 2007:80, Lambrecht 1994:140). There is a parallel between the Cicipu data being considered here and expletive-*it* constructions in English. In both cases the minimal coding that occurs with topical referents (person subject agreement on the one hand, unstressed pronoun on the other) is being used as an apparently 'dummy' subject – about as far removed from a suitable topic as could be imagined.

An alternative approach is to say that the pronoun does in fact have some meaning, usually that it denotes some highly abstract setting, perhaps always at least semi-active in our consciousness. Examples of such analyses include Bolinger (1977), Lakoff (1987) on expletive-*there*, and Langacker (1991:365). However this explanation does not wash for Cicipu, since activated inanimate objects never trigger person agreement unless they are topics. It is probably best just to treat expletives and weather verbs as another kind of neutral agreement, available as an alternative to the usual NC8 agreement for a different kind of atypical controller.

8.5 Progressions

The previous section showed that both animacy and discourse topicality are relevant for the gender/person alternation across the corpus as a whole. Here I will look at what happens in individual anaphoric chains, and consider whether any patterns can be discerned with respect to agreement. Recall from §2.3.2.2 that the definition of discourse topic in Dooley's conception depends on the paragraph. A discourse topic is a referent which thematically integrates the discourse schema of a paragraph, in that each step of the discourse schema for the paragraph is understood to be about that referent. Discourse topics of one paragraph do not automatically carry over into the next, and as

---

22 In still other Benue-Congo languages such as Chicheŵa (Bresnan and Kanerva 1989:10-11) gender agreement does occur with weather verbs, although this is referential agreement e.g. 'outside is cold' (Bresnan and Kanerva argue that Chicheŵa lacks expletive subjects altogether).

we saw in §2.3.2.5 the activation level of referents is reset across paragraph boundaries – this was evident both from psycholinguistic experiments and from the increased coding weight required to refer to a previously-activated referent once a boundary is crossed. We also observed that in languages as diverse as Chinese, Japanese, Koine Greek, and Mbyá Guarani (as well as several Benue-Congo languages) there are within-paragraph progressions from greater to lesser coding weight which reflect the (re-)establishment of the referent as the topic of the new discourse unit. This section will show that in Cicipu we find a similar within-paragraph progression from lexical NP > gender-marked anaphors > person-marked anaphors.

8.5.1   Within-paragraph progressions

Recall from §8.1 that in general only instances of anaphoric agreement are being considered in this chapter, rather than grammatical agreement, since the latter does not generally permit a choice of feature. This decision limits the number of agreement progressions that will be found in the corpus. For example, the switch in (48) below was not counted as a 'progression' for the purposes of this analysis, since the gender agreement prefix **ma-** is an instance of grammatical agreement. This kind of switch is extremely common in the corpus.

(48) sée   mà-gɐ́i   **má-**'ìsànù   bè'í́   **ù-**párà=rè
      then NC4-sword AG4-stand\IRR until 3s-stalk\RLS=3P.PRO
      then the sword waited until it spied on them

[samy001.067]

Consequently the only candidates for progressions are cases where a lexical NP is followed by an anaphor chain of two or more person markers without interruption by another lexical NP. The vast majority of these are consistent in their marking, either sticking to gender-marking throughout, or to person-marking throughout. Of greater significance here are the remaining 65 'mixed' examples. There is a clear preference for mixed chains of anaphors to start off with gender-marking and then progress to person-marking. There are 54 chains of this kind in the corpus, 19 of which occur in the topic-stimulation texts, and 35 in the rest of the corpus. In contrast, there are only 2 chains in the reverse direction (PERSON > GENDER). There are 9 'double' chains, involving more than one change of feature (e.g. GENDER > PERSON > GENDER > PERSON).

The mean length of the 65 anaphor chains is seven person markers, with chains

ranging in size from two anaphors (the logical minimum) to 29. The median length is four markers. Of more interest for this study is the interaction between number of progressions and animacy. Table 69 shows that progressions from gender to person agreement can be found for referents at any point in the animacy hierarchy.

Table 69: Distribution of agreement progressions according to animacy

Animacy	no. of GENDER > PERSON progressions
Human	8
Spirit	3
Non-human folktale character	20
Animal	17
Inanimate	6

The high figures in the cells for folktale characters and animals reflect the intermediate inherent topicality of these referents. Readers are invited to inspect the folktale in Appendix A, which contains several progressions. Humans are at the top end of the inherent topicality scale, and so anaphoric gender-marking is very rare, despite the high number of references to humans in the corpus overall. As a result the possibility of progressions starting out from gender-marking is limited. Inanimate objects, on the other hand, are at the bottom end of the inherent topicality scale, which means that person-marking is rare, and similarly there is less chance of a progression ending up with person-marking.

There is also a correlation between the discourse topicality of the referent and its animacy, as shown in Table 70.

Table 70: Distribution of GENDER > PERSON agreement progressions according to animacy and discourse topicality

Animacy	Discourse-topical		Non-discourse-topical	
Human	5	(60%)	3	(40%)
Spirit	2	(67%)	1	(33%)
Non-human folktale character	18	(90%)	2	(10%)
Animal	17	(100%)	0	(0%)
Inanimate	6	(100%)	0	(0%)

We should not be surprised at the direction of this correlation, since we are only considering anaphor chains involving *both* gender and person agreement. The general

case is for human referents not to trigger anaphoric gender agreement at all. So if gender agreement does occur, then we might expect it to be for a special reason – and in the case of humans, being non-topical is exceptional[23]. Animals and inanimates, on the other hand, occur lower down the animacy hierarchy, and so for these, gender agreement is the norm. Person agreement with animals or inanimates is exceptional, indeed it is only possible AT all if the referent is discourse-topical. Folktale characters have properties in common with both humans and non-humans, and they are intermediate in terms of the percentage of gender/person agreement progressions that involve discourse topics.

Progressions have been observed for nouns of all classes except NC7. Recall from §5.2 that nouns from genders 7 and 7/8 are often abstract nouns, or kinds of trees. Such nouns are neither animate nor good candidates for discourse topics, and so they rarely trigger person agreement. NC7 nouns *can* trigger person agreement, but the corpus would probably have to be a lot bigger to find examples of anaphoric chains progressing from AG7 to 3PS.

The remainder of this section will proceed by considering five progressions (or series of progressions) in some detail. Then the nature of the 'boundary' at which gender-marking becomes person-marking will be examined, to see if there are any general statements that can be made with respect to where it might fall. Finally the exceptional anaphor chains will be considered, since they also show interesting properties.

8.5.1.1   Examples of progressions

The excerpt in (49), repeated from (7), provides a straightforward instance of a within-paragraph progression involving a human referent. The referent is first introduced into the discourse by the NC8 noun **z-zá** 'person', but then immediately 're-introduced' using the NC4 noun **m-ɔ́ɔ** 'child'. This noun is the subject of its clause, and so automatically triggers AG4 gender agreement on the verb **yaa** 'arrive'. The next three subject prefixes agree in gender, but the fourth, on the verb **kaba**, agrees only in person.

---

23 This is especially true since we are counting individual progressions, not referents, and typically a topical referent will give rise to many more anaphor chains in a text than a non-topical one.

(49) ánà   ù-kúmbà-nà / sée **'ínà**   **z-zá**   yâa-nà /
when   3s-climb\RLS-PFV   then   **certain**   **NC8-person**   arrive\RLS-PFV

'ínà   **m-ɔ́ɔ**   mà-yâa-nà   é↓=k-kèeké /
certain   **NC4-child**   AG4-arrive\RLS-VENT   LOC=NC8-bicycle

*when he had climbed / then **a certain person** arrived / **a certain child** arrived on a bike /*

ánà   **mà**-yâa-nà-nà /   **m**-áyà   **mɔ̀**-pɔ́lù   Ø-kèeké
when   **AG4**-arrive\RLS-VENT-PFV   **AG4**-come\RLS   **AG4**-put_down\RLS   NC8-bicycle

ù-kábà-nà /   mò-kúlòotû   mò-tò /
3s-take\RLS-VENT   NC4-basket   AG4-one

*when **he** arrived / then **he** came and **he** put down the bike and **he** took / one basket /*

The second example shows a plural referent progressing from gender to person agreement. Most plurals in Cicipu trigger AG2 agreement, which as mentioned in §7.5.1 can be hard to distinguish from 3PP person agreement. There is no such problem in the following example, since the relevant noun **m-ɔ́ɔ**/**m-úu** 'child/children' has 4/5 gender. Here we find anaphoric gender agreement within the same intonation unit as the initial NP, and then person agreement in the next intonation unit.

(50) **m-úu**   **m**=**mó-ní**   **mí-llè**   sée   **mí**-kàbà   kà-ráyì   ká↓=z-zá /
NC5-child   **AG5**=**NC4-water**   **AG5-that**   unless   **AG5**-take\IRR   NC1-life   AG1=NC8-person

sée   **á**-làhà   z-zá /
unless   **3P**-leave\IRR   NC8-person

***those water babies** then **they** take the life of a person / then **they** leave a person* [i.e. leave them lifeless] /

[tats005.002.144]

After two relatively simple illustrations, we now consider a longer stretch of text, this time taken from the end of a folktale about an anthropomorphised sword. The passage in (51) contains three progressions involving the sword, one from GENDER > PERSON (a-f[24]), and two full LEXICAL NP > GENDER > PERSON progressions (i-o and p-x[25]). These repeated progressions show that the three-stage progression is not confined to the introduction of the referent. Instead gender agreement may re-occur after a stretch of person-marked anaphors, provided a lexical NP intervenes. These multiple progressions suggest that it is not merely anthropomorphism that is governing the agreements (see §8.6.1).

---

24 In this paragraph the first reference to the sword is by means of gender agreement. See the discussion of exception 1 in §8.5.1.3.
25 Paragraph (g-h) contains only person-marked anaphors.

(51) a. ¶OK /

    b. when AG4-tired /

    c AG4-couldn't carry on /

    d. then AG4-wait a bit /

    e. **3s**-spied on them /

    f. if **3s**-stalked them, perhaps they would forget /

    g. ¶then **the sword (mà-gâi,** NC4) AG4-waited a bit **3s**-spied on them **3s**-spied on them **3s**-spied on them **3s**-spied on them /

    h. perhaps those people would forget /

    i. ¶OK those people, them, they weren't forgetting /

    j. because /

    k. if they forgot, **that sword(**NC4**)**, if they didn't put AG4-PRO back in the scabbard /

    l. perhaps when they were to stand together /

    m. **3s**-would come and **3s**-cut them up /

    n. OK they didn't want [this to happen] /

    o. they wanted to carry on going far off /

    p. ¶OK, some time [later] /

    q. they were looking for something /

    r. OK, perhaps they would throw something at **the sword(**NC4**)** /

    s. when they threw at AG4-PRO, AG4-grew in anger /

    t. **3s**-was jumping quickly **3s**-was following them quickly here OK /

    u. because when they didn't come together as two people /

    v. **3s**-was feeling anger **3s**-was feeling anger **3s**-was feeling anger /

    w. **3s**-had to wait /

    x. because, **3s**-didn't get two people in one place /

[samy001.061]

The following example again involves multiple progressions, this time with a generic topical referent, the crocodile.

(52) a. ¶and AG9-PRO **the crocodile (kù-yúpù,** NC9**)** /

    b. AG9-is a fearful thing for people /

    c any water where there is said to be **a crocodile(**NC9**)** inside /

d. if **a crocodile(NC9)** is sensed to be there inside /

e. OK, people won't (.) people will be afraid to enter that water /

f. since, **the crocodile(NC9)** /

g. **AG9**-grabs people /

h. **AG9**-can grab a person /

i. **AG9**-will kill /

j. in the water /

k. because a person in the water doesn't have weight /

l. when someone enters /

m. **3S.PRO** when **3S**-comes and **3S**-takes someone and **3S**-pulls /

n. **3S**-causes [the person] to sink in the water doing pulling the person there in the water sinking /

o. and a person, in the water /

p. he will not find it difficult to die /

q. since, water doesn't have a branch which a person will hold /

r. much less a person can get a foothold /

s. ¶OK **3S.PRO** /

t. how **3S**-rules all the water /

u. **3S**-eats a person fish, person, whenever **3S**-has the chance to do for person **3S**-takes the chance /[26]

v. and **3S.PRO** is in the water /

w. if water took me from here to Tungan Kaɗè /

x. inside one minute /

y. if **3S**-runs /

z. **3S**-arrives there /

aa. right away /

ab. like how a motorbike runs on the ground /

ac. it's like that its speed is inside the water /

ad. ¶OK, when it becomes evident that **the crocodile(NC9)** **AG9**-reaches exasperation in the water /

ae. **AG9-**starts taking people /

af. OK whenever someone enters inside the water **3S**-spots the person from far off /

---

26 The meaning of this line is obscure, but the indexing is accurate.

ag. in just a minute **3s**-arrives there /

ah. attempt-**3s.poss** when **3s**-arrives to hold the person /

ai. **3s**-causes [the person] to sink in the water /

aj. **3s**-kills /

ak. and when **3s**-kills **3s**-will eat /

[tats002.006.012]

In the first and third paragraphs there are straightforward progressions from LEXICAL NP > GENDER > PERSON. The persistence of lexical mentions from (a-f) is probably due to the generic nature of the referent being discussed, but beyond the greater density of lexical mentions, generic referents do not seem to behave any differently from specific individuals with respect to the feature alternation – both gender and person agreement are found, and the restrictions on the nature of possible progressions are the same.

The second paragraph has no instances of gender-marking at all. This is because the first mention of the crocodile in this paragraph is not a lexical NP, but is an independent personal pronoun, and given the unidirectional progression from GENDER > PERSON agreement being argued for here, there should be no opportunity for gender-marking. And indeed this is what we find here.

We can make three further observations about this example in passing. First, despite the fact there is no lexical NP denoting the crocodile in the second paragraph, there is still an increase in coding weight in line (s) compared to the previous mention in (n). The use of a marked topic structure (the left-dislocated independent personal pronoun éví '3s.PRO') is required because of the resetting of activation levels that occurs across a paragraph boundary. That a pronoun is possible here rather than a full NP derives from the fact that the crocodile is a topic of the discourse as a whole, and by virtue of this it retains a certain level of activation throughout the text (e.g. Anderson et al. 1983). The second observation is to note that the other instance of an independent personal pronoun in the second paragraph (line v) occurs for a different reason. In this case the activation level of the crocodile is high and no paragraph boundary has been crossed. Instead the form of the referring expression is determined by the syntactic construction, in this case a non-verbal clause:

(53)   **éví**   kúmá   cì-mé'ĭ̀   tí↓=mó-ní

       3s.ᴘʀᴏ   more   ɴᴄ6-inside   ᴀɢ6=ɴᴄ4-water

       *and **it** is in the water*

<div align="right">[tats002.006.033]</div>

This is not true "overcoding" since in this particular syntactic environment (locative non-verbal clauses) there is no predicate with a slot for agreement morphology. Therefore a free pronoun *is* the minimal coding in this case. Finally, going back to the first paragraph, there is one more example of an overcoded personal pronoun (line m). This occurs because of yet another reason: in this case, the full pronoun is a response to the competing referent *someone* introduced in the previous line.

     The fifth and final illustration in this subsection also involves two progressions, this time in consecutive paragraphs. The referent concerned is the mass noun **Ø-gógóró** 'gin, ɴᴄ8'. In both paragraphs the initial referring expression is an appositional series of NPs **éví Ø-gógóró** 'it gin'.

(54) a.   ¶OK /

     b.   but **3s.ᴘʀᴏ gin (Ø-gógóró, ɴᴄ8)** /

     c   ᴀɢ8-it has a lot of faults /

     d.   there are things that **3s-causes** /

     e.   **3s-causes** a lot of things /

     f.   ¶because you see **3s.ᴘʀᴏ gin(ɴᴄ8)**, if you are drinking **ᴀɢ8-ᴘʀᴏ** /

     g.   if you don't have good things while you are eating /

     h.   *inshaAllahu* there are consequences /

     i.   because **3s.ᴘʀᴏ** when you drink **3s.ᴏʙᴊ** too while drinking /

     j.   because **3s.ᴘʀᴏ 3s-**won't build you up

<div align="right">[tats007.001.014]</div>

In Cicipu, appositional controllers of this kind seem to be equivalent to the second NP (i.e. the lexical NP) with respect to agreement. For example, they trigger gender rather than person agreement when they occur in subject position, as demonstrated in (55).

(55) Ø-sá'à   ví-llè   t-èné   t-ì   **évì**   **kà-bárá**   **kó-ttù**
NC8-time AG8-that AG6-which AG6-COP **3S.PRO** **NC1-old_man** **AG1-1P.POSS**

ká-ayà-nà   **kò**-cíyò   Ø-màkáamì   lêe...?
AG1-come\RLS AG1-get\RLS NC8-position there

*that time how was it that **he our father** came and got an influential position there...?*

[svbg001.009]

In line with this observation, composite NPs of this kind can head a GENDER > PERSON progression as in (54f-j). Consequently it seems best to analyse them as carrying the feature GENDER as well as PERSON. Example (54) is also of interest because of the contrast between the near-homophonous AG8 pronoun and 3PS object clitic in lines (f) and (i) (see §7.3.1). Here are the relevant clauses in Cicipu:

(56)   ń   nì   Ø-yúu   sɔ̀ɔ   v-ì /
if   and   2S-CONT   **drink**   AG8-PRO

if you are drinking it /

[tats007.001.019]

(57)   nì   sîivì...
nì   Ø-sɔ̂ɔ = vì
and   2S-**drink**\RLS=3S.PRO

when you drink it...

[tats007.001.022]

Observe how the **i** vowel from the person-marked clitic spreads leftward onto the verb in (57), but that this fails to happen for the otherwise-identical gender pronoun in (56).

It should be noted that the multiple progressions discussed here are not the norm. For this section I chose three examples of multiple progressions in consecutive or near-consecutive paragraphs, since they could be usefully discussed without having to reproduce very large sections of texts. Normally agreement progressions are distributed more disparately (remember there are only 65 examples in the entire corpus). In total there were thirteen referents which took part in multiple progressions, although often these progressions would be separated by many other *consistent* anaphor chains (e.g. all gender, or all person).

8.5.1.2   The boundary between gender and person agreement

The point at which gender agreement progresses to person agreement in anaphor chains is not random, and two general observations can be made. First, the length of time taken to progress to person agreement depends on the conditions discussed in §8.4. So for

example for the lower animates, the same reluctance to display person agreement at all manifests itself in the fact that when person agreement *does* occur, it is often towards the end of a long chain of gender-marked anaphors. By the same token, anaphoric chains with human referents, if they display gender-marking at all, usually progress to person-marking early in the chain.

Secondly, even though any number of consistently-marked anaphors may occur in a single chain without any pause, the switch from gender to person marking rarely occurs *within* an intonation unit. Example (58) is typical:

(58) ánà **cì-kúndú** tì-yâa-nà,   tì-dúkwà-nà   **tì**-yôo
   when **NC6-hyena** AG6-arrive\RLS-PFV AG6-go\RLS-VENT AG6-go\RLS

   tì-jôolò-nò /   **ù**-zâa   cé   Ø-nàatànáatà /
   AG6-check\RLS-VENT 3s-find\RLS   NEG   NC8-small_spider

   when hyena arrived, he came and checked / he didn't see spider /

                                    [saat001.008.130]

Out of the 54 GENDER > PERSON switches, only 9 take place within an intonation unit, and in each of these cases there is a constituent separating the person-marked anaphor from the previous gender-marked one, as in (59).

(59) **mà-kkácì**   **má↓=má-wáa** m-áyà   **mà**-dúkwà **mà**-húɗò-no
   NC4-old_bitch AG4=NC4-dog   AG4-come\RLS AG4-go\RLS AG4-sweep\RLS-VENT

   à-hángáyáǔ   hé=t-îi   **w**-áyà   **ù**-gít<ìl>à-nù   ó↓=mó-ní /
   NC2-dried_piece AG2=NC6-shit 3s-come\RLS 3s-snap<PLAC>\RLS-RES LOC=NC4-water

   then the old bitch went and swept up dried pieces of shit, then she broke them up in the water /

                                    [saat002.002.252]

8.5.1.3   Exceptions to the GENDER > PERSON directional constraint

At the beginning of this section I noted that as well as the 54 anaphoric chains progressing from gender to person agreement, there were 11 anaphoric chains that did not fit that pattern. Explanations can be offered for some of these exceptions, but others remain mysterious. This subsection discusses the exceptions in turn, arranged with what are judged to be the more convincing explanations first.

Exception 1

This exception has a straightforward explanation. It involves a 'reversion' from person- to gender-marking midway through a folktale about a sword. The reversion took place after a one-minute interruption to the narrative, caused by a motorbike going past the compound where the recording was taking place. The interrupted chain of reference to

the sword starts with the lexical NP in line (b). Gender-marking in (c) is followed by person-marking (c, g) until the interruption at (h). When the speaker restarted, the sword was referred to using a gender subject prefix (j) rather than the expected lexical NP.

(60) a. ¶because of that,

b. **that sword (m̀à-g̃ẫi, NC4),**

c AG4-was carrying on with speed, 3s-was carrying on with speed hemming them in,

d. but they didn't want to come together.

e. because,

f. they wanted to frustrate AG4-PRO,[27]

g. but 3s-wouldn't get people to 3s-kill.

h. there's a motorbike.

[one-minute break]

i. ¶OK,

j. OK when AG4-tired AG4-tired AG4-tired,

k. AG4-couldn't,

l. AG4-had to wait a bit.

m. then 3s-stalked them,

[samy001.053]

The reason for resuming reference with a gender agreement marker rather than a nominal is not altogether clear, but certainly the absence of minimal coding (3PS agreement) can be explained by the passage of time and subsequent shift in attention since the previous reference. The activation of the referent was no doubt presumed to have decayed in the consciousness of the hearer beyond the point where minimal coding was appropriate, and a heavier kind of coding (i.e. gender-marking) was therefore required to re-activate it.

Exception 2

A further exception can be explained by appealing to the subjective nature of the notion of *intrinsic interest*. Referents may only be of intrinsic interest (and hence discourse topics) to *some* of the discourse participants. In the following conversation about clothing (61), an old man and his son are trying to remember the word for a particular

---

27 The gender-marking in (f) is of course also 'out-of-turn' – see the miscellaneous exceptions discussed after Exception 6.

item of clothing worn in the past. Since the item in question was a discourse topic for them, it is referred to using person agreement (see lines k, m, r, and t). When the father calls his wife to confirm the correct word, she uses a gender-marked pronoun (line s) – presumably because the referent held no particular interest for her at that time, and so from her point of view was not a topic.

(61) a. Son: Back then the old man told us there was **a thing(ì-rí, NC3)** what was **AG3-PRO** called,
what was **AG3-PRO** called,
mokuru?
kai! I forget.
there was **something(NC3)** that women wear, men too there was something women wea-, men wear.

     b. Father: what thing.

     c   Son: what did they say?

     d. Father: isn't it *iyayiba*? women wore them.

     e. Son: it's not *iyayiba*, what did they say?
what they called **AG3-PRO**,
AG3-was similar to the *koyoyu*, or did they call that *kokompo*?
kai!
I forget the name of **AG3-PRO**.

     f. Father: no! *kokompo* is not like *kebente*.

     g. Son: no like for women for women.

     h. Father: women?

     i. Son: yes.

     j. Father: ah! it's not these clothes that the Avaɗi are wearing here,
like these skirts.

     k. Son: what was the name-**3S.POSS**

     l. Father: hmm?

     m. Son: what was the name-**3S.POSS**

     n. Father: hey Tawi!

     o. Mother: yes?

     p. Father: **the thing(NC3)** that the Avaɗi wear they say *kakaɓali*?

     q. Mother: yes

     r. Son: but what do they call **3S.OBJ** in Cicipu.

     s. Mother: **AG3-PRO** is *kakaɓali*.

     t. Father: it's that in Cicipu for sure.
but that time there was no woven clothing, only **3P.PRO**.

[sayb001.326]

Exception 3

The next example is similar to Exception 1, in that it involves a double GENDER > PERSON progression without any intervening lexical NP. This time, however, there is no delay or paragraph boundary which might cause the activation level of the referent to decay. Instead the reason for the repetition seems to be poetic – the second time the progression occurs (d) it is simply a formal repetition of the first (c)[28], and it should not be too surprising that the normal rules of participant reference are suspended in such situations. Arguably this should be regarded as a single repeated progression, and should not really count as an exception. The referent in question is **kò-yòngólì** 'kind of large black ant, NC1'.

(62) a. then **the ant (kò-yòngólì, NC1)** came out xxx he brought out **the ant(NC1)** he put **Ø** inside /

    b. of the seeds /

    c **AG1**-kept on gathering, **AG1**-kept on gathering seeds while **3s**-transporting them to the side /

    d. **AG1**-kept on gathering seeds while **3s**-transporting them to the side until **3s**-had gathered them all /

    e. all /

    f. **3s**-left the sand there separate, **3s**-left the seeds separate /

[saat002.002.149]

Exception 4

The next example is more difficult to explain in terms of discourse topicality or paragraph structure. It comes from earlier on in the folktale from which (60) above was taken, a story about a sword chasing two people. The sword is a discourse topic throughout the text, and the passage quoted below forms a complete paragraph. The problem is that after several person-marked anaphors (b-c), the chain reverts to gender agreement (line e) and remains in that state until the end of the paragraph (f, j).

---

28 The intonational contours imposed over the two statements are almost identical, which presumably adds to the poetic effect.

(63) a. OK, **the sword (mà-gȃi, NC4)** AG4-was tiring quickly AG4-tiring quickly AG4-tiring quickly,

b. **3s**-was following the people because if it was just one person, **3s**-wouldn't kill him.

c OK, **3s**-wanted two people, OK **3s**-was following this one **3s**-would follow that one it **3s**-would follow this one,

d. they wouldn't come together,

e. AG4-was tiring,

f. AG4-didn't even get the people,

g. assembled as two,

h. far less killing them.

i. because,

j. AG4-couldn't assemble them,

k. as two people

[samy001.034]

One solution to this problem is suggested by the distribution of information within the paragraph, and that is to consider the above passage as consisting of two paragraphs rather than one. The first grouping (a-d) would then be part of the overall narrative schema, with the second (e-k) being simply a recapitulation of the first. This reanalysis would leave us with two anaphoric chains, the first progressing from gender to person agreement as expected for a discourse topic, and the second with consistent gender agreement. The functional motivation for the return to gender marking in (e) would be the processing shift that has been argued to take place across paragraph boundaries (§2.3.2.4). There are conceptual reasons to take this approach, since both the individual groupings express what seems to be the same discourse schema, comprising of (i) the sword tiring, (ii) its desire to assemble the two people into one place, and (iii) its failure to do so.

Unfortunately the formal signals do not support such a paragraph division. Line (d) has a non-terminal pitch contour (Chafe 1994:59-62), indicated by the comma. Secondly, there is only a short pause of 0.25s between (d) and (e), compared to 0.7s before (a) and 0.5s after (k). Finally, the increase in amplitude characteristic of the start of new paragraphs is lacking entirely from (e).

An alternative, and perhaps preferable, reason for the unexpected use of gender coding in line (e) might be the switch of subject that took place in the previous

411

intonation unit. It is possible that the increased competition arising from the occurrence of the two people as subject in line (d) led the speaker to augment the coding used for the subject of line (e). This coding technique has been observed experimentally by Anderson et al. (1983), who found that English subjects unexpectedly used lexical NPs rather than pronouns within paragraphs, just when there was a competing referent. Against this explanation it could be argued that 3PS coding would have been sufficient to unambiguously identify the sword, since neither of the two humans are ever referred to independently in this story. However the effect of competing referents on speakers' choices of referring expression is not limited to cases of *logical* ambiguity, as Anderson et al. observed. In their experiments the mere activation of an additional referent was sufficient to cause heavier coding than was 'logically' necessary for disambiguation. It may be that something similar is happening in this example, and that the use of slightly-more-than-minimal coding in the form of gender agreement in line (e) is a small price to pay to help the hearer determine the correct subject referent.

Exception 5

This example is similar to Exception 4, in that there is an unexpected occurrence of gender agreement midway through a paragraph (line d).

(64) a. **this sword (mà-gẫi, NC4)**,
      b. AG4-came and AG4-increased,
      c   **3s**-increased in shivering.
      d. when AG4-increased in shivering,
      e. **3s**-came and **3s**-left.
      f. when it **3s**-left,
      g. **3s**-went and **3s**-spotted someone.
      h. up on a horse.

[samy002.036]

This time, however, there is no competing referent, and the paragraph cannot be divided into two. Neither (a-c) nor (d-h) form a coherent paragraph on their own, either conceptually or formally. Nevertheless there is *some* kind of division between (c) and (d), as well as between (e) and (f). This can be seen in the form of an terminal intonational contour (represented by a full stop), and by the fact that (d) and (f) both package up the event-line information introduced in the previous few intonation units

and present it as now-presupposed information in a dependent clause. It may be that as well as being sensitive to complete discourse units (i.e. paragraphs), the Cicipu gender/person alternation is also, to some degree, sensitive to smaller structures. Dooley refers to the incomplete building-blocks of paragraph schemas as 'micro-level units' (2007:14, 59-61, see also p.46 fn. 42 on tail-head linkage), while Chafe (1994:139-145) uses the term 'sentence' to refer to the structural unit corresponding to Dooley's conceptual definition (see also Hinds 1979:150).

Dooley (2007:60-61) observes that *non*-topical referents usually have to have their activation status renewed after crossing the boundary of a micro-level unit, even within the same paragraph. Consequently speakers typically resort to heavier coding in such cases. Although the sword in (64) above is most definitely a discourse topic, it may be that the slight increasing in coding in line (d) is nevertheless related to the crossing of such a boundary.

Exception 6

This exception involves an anaphoric chain of two subject agreement markers, in the reverse order with respect to the GENDER > PERSON progression. The grasscutter in (65) is only an incidental participant in the folktale being described, not a discourse topic. The only explanation I can think of is based on the fact that grammatical person agreement occurs on the first verb **aya** 'come' (recall from §7.6.1.1 that grammatical person agreement is possible with NC8 controllers before vowel-initial verbs). It may be that this instance of phonologically-enabled person agreement influences the subsequent marking on the verb **yãa** 'do', and normal service (i.e. gender-marking) is only resumed in the next intonation unit.

(65) a. **a grasscutter (d-díyò, NC8)** 3S-came and **3S**-did a hole.
     b. a place that AG8-entered.

<div align="right">[Tidipo, saat002.004.025]</div>

The remaining five exceptions involve person-marking in one syntactic position, followed by unexpected gender-marking in a different syntactic position (e.g. a person-marked subject prefix followed by a gender-marked pronoun in the complement of VP). One solution to this would be to propose that syntactic positions vary with respect to the rate at which the GENDER > PERSON progression takes place, allowing for the co-

occurrence of multiple semi-independent (yet unidirectional) progressions. This is similar to what Corbett (1991:240) assumes when he restricts his predictions about progressions from syntactic to semantic agreement to *parallel targets* (see §8.8). A much bigger corpus would be required to test whether it is really necessary to weaken the claim about the nature of the progression in this way, by making it dependent on syntactic position.

In any case, the fact that there are unexplained exceptions should not be too disturbing. Given the complexity of the factors involved in the gender/person alternation, and the indeterminacy of the notion of discourse topic, it would be surprising if there had been no exceptions at all. As pointed out in §2.3.3, judgements about discourse topicality are quite different to grammaticality judgements. Speakers vary in eloquence and often do "put things badly", and so we should not be surprised if hypotheses relating to "well-put" language are not reflected 100% in the data. I have devoted a considerable amount of space to discussing the exceptions here, and the fact that most of them appear to some extent to be motivated gives further support to the hypothesis that lexical NPs, gender-marked anaphors and person-marked anaphors occur in an ordered progression within paragraphs. Discussion of the exceptions has also been useful in that it has demonstrated the relevance of some of the concepts discussed in §2.3 such as activation level and cost (Exception 1), intrinsic interest (Exception 2), and competing referents (Exception 4).

8.5.2   Text-level progressions

In addition to the within-paragraph progressions which have just been discussed, *a priori* it is at least possible that there might be discernible progressions in the way discourse participants are coded at a higher level. Dooley's (2007) conception of discourse topic is not limited to the paragraph but is applicable to any discourse unit. In particular there can be global topics whose span is the whole text. Consequently we might expect to find a higher-level progression superimposed on top of the intra-paragraph ones, and in fact we do. There are several candidates for text-level progressions in Cicipu.

One of the clearest such texts is a topic-stimulation text concerning witches (tats002.004). As well as the witch (the main participant), for most of this text (units 18-79 out of 89) a spirit (ù-pépí, also 'wind', NC7) is also on-stage. In the first part of the

textual span of the spirit (units 18-41), which is largely concerned with the witch's activities in securing its services, person markers denoting the spirit are exclusively gender-marked. However once the spirit has been set loose to trouble its victim, it is thereafter (units 42-79) gender-marked only in those environments in which gender-marking is obligatory – otherwise it is exclusively person-marked. In this text there are not actually any individual anaphoric chains showing the kind of progressions that were discussed in §8.5.1. However the text as a whole clearly shows a higher-level progression.

A number of other texts show similar effects. The reverse state of affairs, with person-marking concentrated towards the beginning of a text and gender-marking predominating towards the end does not occur in any of the texts in the corpus[29].

These observations are suggestive of a higher level of discourse organisation, and would merit investigation on a larger scale, not only in Cicipu but also in the other West Kainji languages that have this alternation between gender and person agreement. Further evidence that this might be a fruitful line of research comes from the single Tsureshe folktale included in Agamalafiya (2007). The two main characters are a frog and one another (unidentified) animal. There is a high incidence of subject markers[30] in this story, and both main participants are coded predominantly with gender subject agreement towards the beginning of the tale and person subject agreement towards the end.

8.6 Alternative explanations?

In the course of this chapter three agreement conditions have been proposed involving the noun class, animacy, and discourse topicality of the controller referent. The positing of three independent conditions further exacerbates an analytical problem which was complicated enough by the end of chapter 7. It is therefore tempting to find some means of reducing the number of conditions.

8.6.1 Animacy

One way of doing this is suggested by the fact that, as in many other cultures, animals in

---

29 There is an obvious reason why this should be the case: we do not tend to start talking about participants which are inherently interesting to us, lose interest in them, and yet at the same time carry on talking about them (which of course is necessary for there to be any anaphoric chains at all).

30 It is not clear from the accompanying analysis whether they should be regarded as affixes or pronouns – they are written as free forms by Agamalafiya, but as affixes by Boettger and Boettger (1967).

Cicipu folktales tend to be anthropomorphised if they are principal characters. Not only can they occur as the subject referents of predicates which usually require human referents (e.g. think, speak, cook etc...), they are also coded using person agreement more often than animals outside of this genre.

In the same way, it could be argued that when, say, a shrub is cross-referenced by person agreement in a more prosaic text, then this is also a case of anthropomorphism[31]. If this explanation were satisfactory, then we could do away with discourse topicality altogether, and claim that person agreement is being used because the referent is being conceptualised as especially animate. As an illustration of how that might work, consider the following extract from an English conversation provided by Hewings and Hewings (2005:217), with the addition of **bold** type for what I have assumed to be references to the dog:

> he's a funny old stick Nev is...he says you see **that dog** he's got a lit-, there's **a little terrier dog** mooching round and **it**'s a nice little dog actually **it**'s about this big...**Ø** belongs to his daughter Jackie...so he says see **that dog** down there and I said yeah...I couldn't see **it** like I'm sitting there having me hair cut like doing...and he's, and he's standing there in front of me and he says you see **that dog** down there and I say yeah...and he said uh...you wouldn't think **he** could get **his** head in that cup would you...he's got a mug, a tea mug on the thing...and I looked at the mug and you know it's an ordinary mug and I said...no he said **he** can't...he said all that hair he said sticks out he said when it's smoothed back and he says...**his**, **his** head's narrow he said

Note how earlier on in the excerpt the neuter pronoun *it* is used to refer to the dog, whereas in the latter part of the excerpt the masculine pronouns *he* and *his* occur. In this case the AGENT semantic role assigned by the verb complex *could get* encourages anthropomorphism, which then results in masculine coding. The earlier references using the neuter pronouns are the subjects of descriptive clauses.

However the Cicipu alternation seems to be quite different in nature. In particular, it is possible to get repeated progressions from lexical NP > gender agreement > person agreement for the same referent in a single text, as shown in §8.5.1.1, seemingly independent of the lexical semantics of the verb (see §8.6.2 below). Once we think of an animal or inanimate referent as having human characteristics, we tend not to go back to thinking of it as inanimate in subsequent sentences. Moreover changes from gender to

---

31 Alternatively it could be explained in terms of Kuno's "empathy" (e.g. Kuno 1976). See Siewierska (2004:208) for discussion. The arguments against anthropomorphism below apply equally well to empathy.

person agreement within the same sentence (rather than intonation unit) are common in Cicipu, but a change from *it* to *him* in the same sentence for English would be stylistically odd[32].

8.6.2   Transitivity

I have not yet coded the corpus used for this study for the lexical semantics of the verb or for any other aspect of transitivity (as characterised by Hopper and Thompson 1980), although this has been suggested to me as a potentially relevant factor by more than one linguist. An inspection of the two passages discussed in §8.4.2 shows that there is certainly no deterministic link between the choice of agreement features and transitivity. Nevertheless it remains an open question as to whether there is any significant effect of transitivity, beyond the trivial correlation which would result from the fact that discourse topical or animate referents are more likely to be engaged in transitive activities than non-topical or inanimate ones.

8.7   Revised flowchart

Having reviewed all the available facts about the alternation between person and gender agreement, we are now in a position to revise the flowchart that was presented in Figure 36 (§7.6.1.3) by adding the three agreement conditions discussed in this chapter. This has been done below[33].

---

32 Numerous examples like "**My dog** had **it**s muzzle on because **he** doesn't like other dogs" can be found using Google but they sound odd to me.
33 Recall that these flowcharts are intended purely as an elucidatory aid. In particular, they are not intended to represent processing nor text-generation models.
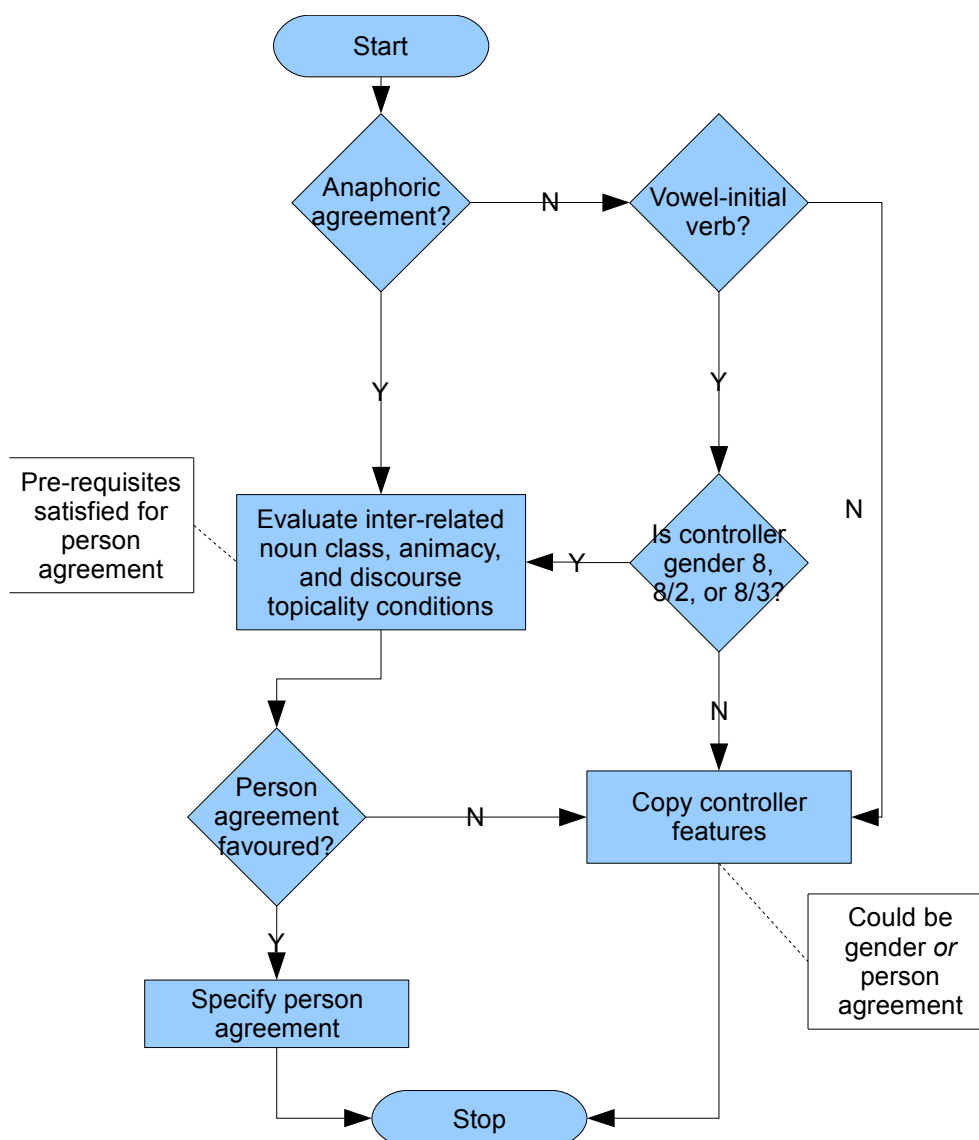
Figure 40: Final revision of flowchart modelling gender/person alternation

There is an asymmetry between gender and person agreement in the flowchart, in that only person agreement is explicitly 'hard-coded'. Gender agreement does not need to be specified at all, since it only occurs when there is a gender-marked antecedent, and so can be represented simply by the blind copying of features[34]. This asymmetry accounts for the direction of the agreement progressions. Since gender agreement is not explicitly specified, it cannot come from anywhere other than the controlling referent. Assuming that in a chain of anaphors each target goes on to function as the controller of the next anaphor in a chain, then once the chain has progressed to person agreement there is no way for it to return without re-mentioning the lexical item. Exceptional cases (especially

---

34 This account does not work for the instances of 1PP person agreement triggered by lexical NPs discussed in §7.6.1.1 (i.e. those in the form *we are the people who [we]-are...*).

exceptions 1, 4, and 5 in §8.5.1.3 above, where there is reason to believe the activation status of the referent might have decayed) are not considered to be agreement here, even though agreement morphology is involved. Instead it is suggested they are being used deictically (see §8.8 below).

The asymmetry in the flowchart also means that person agreement is being treated as marked, whereas gender agreement is unmarked. *A priori* the opposite position might have seemed more probable, for three reasons. Firstly, gender agreement is far rarer than person agreement in the Cicipu text corpus. This is not surprising given the cross-linguistic association of anaphora, animacy and topic. Secondly, all languages have the category person (Siewierska 2004:8-13), whereas not all have gender. Furthermore person agreement on the verb is fairly common in the world's languages, being found in 28 of Bybee's sample of 50 languages (Bybee 1985:30). By contrast, gender agreement as a verbal inflection is less common, occurring in only eight of these languages. Even when gender agreement does occur, it is more often optional and dependent upon pragmatic conditions (Stassen 1997:35-36). The third reason derives from the fact that gender subject agreement in Cicipu implies third person subject. So the change from gender agreement to person agreement involves an increase in the number of features marked on the verb from [PERSON] to both [PERSON] and [GENDER].

Nevertheless, the only practical approach is to assume that gender agreement is the default. If we restrict our attention to Benue-Congo languages, it is clear that there is nothing marked about gender agreement in the group as a whole. Moreover, it seems impossible to give a coherent account of gender agreement as marked in Cicipu, since in that case the persistence of gender agreement would be inversely correlated to the discourse salience of the controller referent. As we saw in §2.2.4.2, whenever there is variation cross-linguistically, it is the *presence* of agreement rather than its absence that correlates positively with animacy or topicality.

8.8   Gender, person and coding weight

We might ask why it should be person agreement that indexes referents high in inherent and discourse topicality, whereas gender agreement is the more straightforward indicator of grammatical agreement. Given that some linguists have suggested that the primary function of gender is to keep track of participant reference (e.g. Contini-Morava 2002, Corbett 2006, see also Dooley 2007:85), it might seem that gender agreement is

baling out just when we need it most – i.e. when there is no longer a nearby formal 'hint' in the form of a lexical NP. However we need to consider more carefully the process by which addressees search for possible referents of anaphors. It is not just a case of searching back for 'recent mentions' in the text, as pointed out by Schwarz-Friesel (2007). Sometimes a referent achieves a privileged position within the discourse unit, in the sense that the discourse unit is integrated around it and it holds intrinsic interest for the speaker, i.e. it is a discourse topic. In that case the use of third-person marking will be enough for the hearer to uniquely identify the intended referent, even when they are inanimate and have competition from non-topical human referents, provided that selectional restrictions and world knowledge permit this identification to be made. We should not be surprised by this; after all, languages which allow null anaphora like Chinese (Li and Thompson 1979) successfully make use of the notion of discourse topic to disentangle considerably more ambiguity than is present in Cicipu texts. Similarly a human referent can achieve this privileged status by virtue of its inherent topicality. In the case of competing referents, none of which have achieved the status of discourse topicality, gender agreement is useful to distinguish which one is intended (as noted by Heath 1983, Bosch (1983:57), Corbett (1991:320-322), Contini-Morava 2002, and others).

We can look at this phenomenon in two ways – one way is to think of it as a manifestation of the principle of economy. Concerning coding weight, Givón suggested that the following principle is at work: "Expend only as much energy on a task as is required for its performance" (1983:18). More generally, Grice's (1975:45) second maxim of quantity states "do not make your contribution more informative than is required". If person agreement is sufficient to identify the referent, then there is no need to supply increased coding weight in the form of gender agreement. As Dooley (2007:104) puts it, "the more the conceptual structure of a discourse unit 'points to' an element and makes it conceptually accessible to the addressee, the less need there will be to 'point to' it linguistically". Bosch (1983:57) refers to this conceptual accessibility as the "salience" of the referent (see also Ariel 1990)[35].

While "coding weight" can be understood phonologically in the case of the post-verbal alternation – the person-marked forms which occur in this environment are

---

35 The two scales of 'inherent topicality' and 'discourse topicality' in Figure 38 could have been combined into a single dimension of 'entity salience'. However this notion is impossible to operationalise without recourse to the independent factors that comprise it.

bound object clitics, while the gender-marked forms are free pronouns – this cannot be the case for the two paradigms of subject prefixes. As we observed in §7.5.1 they are very similar in phonological weight, to the point of considerable ambiguity. Givón (1983) observes that there are several independent scales involved in the concept of coding weight, and Levinsohn (2000:267ff) and Seifart (2005:249-250) stress the role of the semantic specificity of the anaphoric device. However it should be clear from the discussion so far (especially §6.3 on antecedentless agreement) that very little, if any, meaning can be attached to either gender or person markers, so they cannot be ranked in terms of semantics. Instead, as mentioned in §8.7 they can be ranked in terms of the number of features they expound – two for the gender markers, one for the person markers. I am not aware of any languages outside Benue-Congo showing this kind of 'feature' coding weight hierarchy – but if one were to be set up then the Cicipu markers would be the predicted way round, with the higher number of features correlating with lower accessibility. Inspection of the available texts in the other Kainji and Plateau languages which have this alternation suggests that the same LEXICAL NP > GENDER AGREEMENT > PERSON AGREEMENT progression holds for these languages too, although this requires further investigation (see chapter 9).

The other way to think of the phenomenon is that the use of an unexpected coding strategy actually functions as a signal to the hearer that, as well as selecting the most appropriate referent, they should also elevate this referent to the status of discourse topic i.e. to integrate their understanding of the current discourse unit around that referent. According to this second understanding then, the switch to person-marking *causes* rather than reflects an instance of discourse topicality. It should be possible to test this hypothesis experimentally (again see chapter 9).

This view of person agreement as picking out a contextually-salient referent calls to mind the distinction between anaphoric and deictic reference discussed in §2.2.1. Recall that agreement morphology may either be (i) triggered by an antecedent controller in the discourse, or (ii) used to encourage the hearer to infer an appropriate referent, which may or may not have been mentioned previously in the text. Would it make sense to think of gender agreement being 'controlled' by the antecedent NP, with person agreement morphology being a deictic signal to the addressee to encourage them to search for an appropriate referent?

There is evidence to suggest that gender-marked person markers in an anaphoric chain headed by a lexical NP show true agreement rather than deictic reference. As we have seen, gender agreement is rarely found far from an antecedent – for three reasons: (i) speakers do not tend to talk about inanimate or non-topical referents for a long time, (ii) when they do, they tend to repeat the NP, and (iii) in the few cases when lower animates actually do achieve the status of discourse topics, they may progress on to person agreement. Corbett (1991:244-245) comments that some linguists argue that there is no such thing as anaphoric agreement, and that all pronouns deictically refer to the referent of the antecedent, without mediation from their antecedents. However if this were the case for Cicipu then we might expect all gender-marked person markers to display the same freedom with respect to agreement morphology as we find for deictic reference (recall the 'stone' example in §2.2.1). Instead what we find is that chains of gender-marked anaphors remain consistent as to the gender they show. This is even the case if the initial referring expression is a superordinate, non-basic term such as ì-rí 'thing'. The lack of variation in this and the other person-marker chains in Cicipu supports Corbett's (1991:245) contention that anaphoric agreement cannot be reduced to deictic reference.

The relative influence of the controlling antecedent and the referent itself is presumably scalar rather than discrete in nature. The further removed the target is from the last occurrence of the antecedent the less strong we can expect the influence of the controlling expression to be. In contrast, the referent will remain in at least semi-activation and indeed its influence can be expected to grow with time if it continues to integrate the discourse[36] and becomes more entrenched as a discourse topic within the paragraph. As discussed above, at a certain point in the coding progression this prominence may enable the speaker to leave off using anaphoric gender agreement and start using person agreement, in which case we could argue that reference is no longer being mediated by the controller element in the text[37].

---

36 Recall from §2.3.2.2 that this can happen without the topic being mentioned – for example if events are reported from the topic referent's point of view.

37 The use of the controller/target terminology actually obscures what happens when agreement morphology is processed by the hearer. The speaker knows what the referent of the person marker is, and in the case of anaphoric agreement she chooses the appropriate form depending on syntactic properties of the controller. The hearer's task is the opposite. Starting from the features marked on the anaphor, the selectional restrictions imposed by the 'host predication' (Cornish 2007:22), and his own world knowledge, the hearer must choose the referent which will result in the highest amount of textual coherence. Whether a referent is brought to prominence by the decaying influence of a textual antecedent or by the increasing degree to which it integrates the discourse is essentially irrelevant to

However it would be wrong to simply equate gender-marked person markers with agreement and person-marked person markers with deictic reference. The most we can say is that gender-marked person markers are more strongly associated with agreement and less likely to be involved in deictic reference. There cannot be a deterministic link between the type of agreement and the type of reference, since we have seen examples in §6.3 of gender agreement morphology being used with indisputable deictic reference, and in §7.6.1.1 of person subject agreement prefixes co-occurring with true NP subjects.

Before leaving person markers and going on to other targets which share the gender/person alternation, it is worth comparing this alternation in Cicipu to Corbett's theoretical distinction between syntactic and semantic agreement. Corbett (1991:239-241) identifies certain constraints on their occurrence, in particular the following corpus-level constraint:

> For any particular target type, the further it is removed from its controller, the greater the likelihood of semantic agreement (Corbett 1991:240).

In support of this claim Corbett gives examples of progressions from lexical NP > syntactic agreement > semantic agreement in Old English and Chicheŵa (1991:241-242, 250), which of course evoke the Cicipu progressions from lexical NP > gender agreement > person agreement.

However Cicipu person agreement cannot be considered to be semantic agreement since it does not involve any semantic features – as we saw in §8.4, referents of all kinds can be indexed with person agreement markers. In §2.2.4.2.3 we saw that the East Kainji language Amo also has two paradigms of independent pronouns, one inflected for gender and the other for person. Corbett (1991:247) identifies the Amo personal pronoun series as displaying semantic agreement – implicitly treating the 3PS/3PP agreement pattern as additional "minor target genders" (Corbett 1991:160) only available for human referents. However if nouns with *any* kind of referent can trigger person agreement, then there is no longer any semantic basis for agreement. Corbett (2006:155) remarks that "The distinction between syntactic and semantic agreement links to Steele's definition in that the covariance involves a 'semantic or formal property' of the controller". With Cicipu person agreement no semantic property of the controller is involved. Therefore it should not be considered as a kind of gender

---

the hearer's main task.

agreement, and the only feature value that the antecedent and anaphor have in common is [PERSON 3]. The progressions discussed in §8.5 can be seen as following a path from canonical grammatical agreement, through non-canonical anaphoric agreement, ending up with person marking which sometimes may not even be 'agreement' at all, but rather deictic reference requiring pragmatic inference on the part of the hearer. This kind of inference can only be successful if the intended referent has sufficient inherent or discourse topicality.

8.9   The gender/person alternation on other agreement targets

This chapter and the last have been concerned with PERSON MARKERS – pronouns, pronominal clitics, and subject agreement prefixes. However in chapter 4 we saw that the alternation between gender and person agreement actually applies to four additional agreement targets: the article, the demonstratives, the interrogative quantifier **-èné** 'which', and the copula[38]. Siewierska (2004:145) notes that cross-linguistically person agreement is very rarely found on targets other than predicates, possessed nouns, and adpositions, and certainly the grammatical category of 'person' may not seem to make much sense when applied to the targets just identified. While person-marked **w-èné** 'which' or **wú-mpà** 'this' readily co-occur with a noun to form an NP with the feature value [PERSON 3], the meaning of 'which you?' or 'this me' is less clear. Furthermore, as we saw in §4.4.5.3 the 'person' forms of the article (**wú-nà** (sg.) and **á-nà** (pl.)) only occur in indefinite NPs, which are of course incompatible with first and second persons.

It could be argued that the formal identity of the 'person' agreement markers across targets is simply a coincidence, and that a semantically-empty term such as *u-* agreement should be used instead. Nevertheless the term 'person agreement' is retained here because (i) it is less clumsy, (ii) in the case of the article, plural controllers trigger *a-* prefixes just like 3PP verb agreement, and (iii) to a limited extent, the conditions identified above (animacy, topicality, NC8 controller) also influence the agreement alternation on these more exotic targets. In the rest of this section I will consider the four targets in turn. It should be stressed that there are far fewer relevant examples than for the person markers, and accordingly the analysis is sketchier and in need of refinement based on additional research.

---

38 There is dialectal variation here – in Tikula person agreement does not seem possible on the article, the demonstratives, or **-èné**.

8.9.1 The article

Recall from §4.4.5.3 that the article **nà** can occur either before or after the noun. When it appears pre-nominally (e.g. **ká-nà kà-bárá**[39] 'a certain old man') it indicates an indefinite but specific NP, whereas post-nominally (e.g. **kà-bárá ká-nà** 'the old man') it marks an anaphoric (and therefore also definite) NP. Agreement on the post-nominal article is straightforward in that it always takes a gender-marked prefix with high tone. However with the pre-nominal article there is a choice between gender (66) and person (67) agreement.

(66) òkóo **má-nà** mò-ní 'ṹ á↓=ká-ppátà
there_is AG4-ART NC4-water there_far_off LOC=NC1-pit
there is some water there in the pit

[tats005.002.254]

(67) Ø-lóokàcí ví-llè **wú-nà** má-hàukàcí mɔ́=d-dɔ̂ɔ m-é
NC8-time AG8-that 3S-ART NC4-lunatic AG4=NC8-horse AG4-COP
back them there was this lunatic of a horse

[sayb001.501]

With a plural noun as the controller, either the 3PP agreement prefix **a-** (68) or the 3PS prefix **wu-** may occur on the article (69). The latter example shows **wú-nà** marking both singular (**z-zá**) and plural (**ɔ̀-bɔ́wɔ́**) nouns.

(68) **á-nà à-zá** ɔ̀-sì-sɔ́ɔ ì-cṹ'ṹ yí=Ø-súkúdãi
3P-ART NC2-person 3P-HAB-drink NC3-seed AG3=NC8-k.o._plant
some people *drink the seeds of the sukudãi plant*

[tats005.001.184]

(69) kú-nà kw-ã́a'ã́, **wú-nà ɔ̀-bɔ́wɔ́,** h-úndà **wú-nà z-zá**
AG9-ART NC9-day 3S-ART AG2-thief 3P-see\RLS 3S-ART NC8-person
one day, ***some thieves****, they saw* ***someone***

[sahs001.001.001]

There is one further form of the pre-nominal article, and this is **í-nà**, which does not appear to be related to any of the person subject prefixes. Like **wú-nà** it occurs with either singulars or plurals. Other than the restriction of **á-nà** to plural nouns, I have not found any way to predict the distribution of these three articles. To some extent the difference is idiolectal, and even age-mates from the same village differ in their usage.

---

39 **Kà-bárá** 'old man' is written with its citation tone here. In actual fact there is downstep between pre-nominal articles and their head nouns [kánà↓ kábárá] – one further environment in which nouns are found with their 'complement' tone (§3.4.7). If the noun precedes the article then it occurs with citation tone.

If we group the three 'non-gender' articles together (**wú-nà**, **á-nà**, and **í-nà**), then the difference in distribution between the gender and person forms of the article is actually similar to what we have found for the pronouns and subject agreement markers. Here I will concentrate on **wú-nà** (the distribution of **í-nà** is similar, and **á-nà** is rare in the corpus).

Before the noun **z-zá/à-zá** 'person/people', only the three person-marked forms can occur, not the gender-marked AG8 article **ví-nà** or the corresponding AG2 plural **há-nà**. Recall that in §7.6.1.1 we saw the same prohibition against **z-zá** triggering gender subject agreement. In the corpus there are thirty-six occurrences of the article **wú-nà** with **z-zá** or **à-zá** as the head noun, and if we discount these (since the prerequisites for gender agreement have not been met, and therefore there is no choice) this leaves 64 cases of **wú-nà** plus noun head. It turns out that the same three conditions we identified earlier for the person-forms are also at work for agreement on the article: the class of the controller (NC8 or non-NC8), the animacy of the controller, and the discourse topicality of the controller.

Of the 64 controller referents, 36 belong to NC8, leaving the other eight noun classes accounting for only 28 of the examples. In contrast there is not a single instance in the corpus of a pre-nominal gender-marked article with an NC8 controller[40], although they can be elicited. The results for the other classes are given in Table 71, and although the figures are lower than we might wish, they do suggest that nouns from NC8 are more likely to trigger person agreement than those belonging to any other class.

---

40 These do occur in the Tikula texts, but as has already been mentioned Tikula does not offer the gender/person alternation for the article and the demonstratives, and so gender agreement is the only option. Therefore the counts given in this subsection and the next exclude texts from the Tikula dialect.

Table 71: No. of occurrences in the corpus of gender-marked and person-marked pre-nominal articles, according to the noun class of the controller

Class of controller NP	Gender-marked article	Person-marked article
1	1	3
2	0	1
3	2	17
4	4	5
5	0	1
6	0	0
7[41]		
8	0	36
9	14	1
Total	**21**	**64**

When dealing with such small numbers there is a danger of the counts being skewed by common conventionalised expressions, such as **kú-nà kwɛ́a'ɛ́** 'one day', which accounts for 12 out of the 14 NC9 nouns triggering gender agreement, and **wúnà ì-rí** 'a certain thing, something' which is responsible for all 17 NC3 nouns triggering person agreement. We can mitigate against this by excluding all nouns that account for more than a certain number of tokens (e.g. 5). The revised count is as follows:

Table 72: No. of occurrences in the corpus of gender-marked and person-marked pre-nominal articles according to the noun class of the controller (excluding nouns contributing more than five tokens)

Class of controller NP	Gender-marked article	Person-marked article
1	1	3
2	0	1
3	0	0
4	4	5
5	0	1
6	0	0
7		
8	0	10
9	2	1
Total	**6**	**22**

41 Figures cannot be given for NC7 because the gender-marked form is **wú-nà**, the same as the person-marked article.

While the low figures in general may not be terribly convincing, the high figure for NC8 person-marking is at least suggestive. The NC8 nouns were of varying animacy: **bákàaké** 'verse', **dábbà** 'animal', **lóokàcí** 'time', **sá'à** 'time', **síyà** 'side', **'írì** 'kind', **'álìjìní** 'spirit', and **vɔ́mɔ́cí** 'stranger'[42]. None of these nouns had referents which went on to become discourse topics.

Moving on to animacy, nouns with human referents are more likely to trigger person agreement on the pre-nominal article than gender agreement. Discounting the NC8 nouns, the remaining 18 occurrences of pre-nominal articles from Table 72 are distributed according to the animacy of the controller as shown in Table 73.

Table 73: No. of occurrences in the corpus of gender-marked and person-marked pre-nominal articles according to the animacy of the controller (excluding nouns contributing more than five tokens and NC8 nouns)

Animacy of controller NP	Gender-marked article	Person-marked article
Humans	0	5
Folktale character	0	2
Animals	3	2
Inanimates	3	3
Total	**6**	**12**

It can be seen that nouns with human referents favour person agreement, whereas those with animal and inanimate referents are split between the two kinds of agreement.

Lastly, if we consider the 13 non-human referents from Table 73 we also find a correlation with discourse topicality, shown in Table 74. Of course, a non-specific indefinite article cannot encode a topic referent in its *integration* function (§2.3.2.6), since the topic cannot be 'a matter of standing interest or concern' at the point at which it is introduced to the discourse. However discourse topics may also be signalled by special constructions when they are used in their *access* function (cf. 'VIP' strategies of reference – Dooley and Levinsohn 2001:119-123).

---

42 All but the last are Hausa loanwords – recall from §5.6 that genders 8, 8/2, and 8/3 have a high proportion of loans.

Table 74: No. of occurrences in the corpus of gender-marked and person-marked pre-nominal articles according to the animacy of the controller (excluding nouns contributing more than five tokens, NC8 nouns, and nouns with human referents)

Topicality of controller NP	Gender-marked article	Person-marked article
Discourse topic	3	7
Non-discourse topic	3	0
Total	**6**	**7**

This distribution mirrors what we saw for the person markers earlier in the chapter: for non-human referents being discourse topical is a necessary but not always sufficient condition for person agreement.

It seems then that the same three factors occur as conditions on agreement on the article as for the person markers. The figures given in the above tables are obviously lower than one would like, but they are wholly in line with what we would expect from the earlier sections in this chapter[43].

8.9.2 Demonstratives

Recall from §4.4.5.2 that demonstrative modifiers can occur either before or after the head noun. Just as with the article, the post-nominal demonstratives agree in gender while the pre-nominal ones offer a choice of gender or person. The difference in meaning between the pre-nominal and post-nominal modifiers is not as obvious as in the case of the article (indefinite specific vs. definite), but as we will see below there is a link with discourse topicality.

As one would expect in a language so strongly left-headed as Cicipu, in the majority of cases the demonstrative occurs after the noun. In the entire corpus there are only 34 pre-nominal tokens (27 person-marked and 7 gender-marked) compared with 491 post-nominal tokens (which are obligatorily gender-marked). As with the person markers and the article, a sizeable number of the person-marked demonstratives occur before the word **z-zá** 'person' (11 out of 27 tokens). **Z-zá** is never preceded by a gender-marked demonstrative, and it is only followed by a gender-marked demonstrative on one occasion, in contrast to the general overwhelming preference for post-nominal gender-marking. However this statistic apart, since there are so few pre-nominal demonstratives it is difficult to tell what distinguishes the gender-marked ones from the

---

43 It is not logically possible to find coding progressions from gender to person, since the restriction to indefinite NPs means that the same referent cannot be coded with the pre-nominal article twice.

person-marked ones. Therefore I will treat both together, and concentrate on what characterises their use as opposed to the post-nominal demonstratives.

The most obvious difference that I was able to elicit from speakers is that unlike the post-nominal demonstratives, the pre-nominals cannot be used deictically to point out referents in the speech situation. However this is largely irrelevant with respect to the corpus, since the majority of tokens are being used anaphorically (or cataphorically) instead. After discarding the 11 tokens involving **z-zá**, one of the most striking things about the remaining 23 pre-nominal demonstratives is that they almost all occur with abstract nouns. The distribution is also highly genre-dependent, with 6 tokens occurring in reported conversations or thoughts, 11 in prayers, and 1 in a song – recall from §1.5 that these genres comprise a very small part of the corpus. In contrast, the genres that make up the bulk of the corpus are largely devoid of pre-nominal demonstratives: folktales (1 token), interviews/historical narratives (3 tokens), and topic-stimulation texts (1 token).

However in addition to these observations, and more relevant for the present discussion, the referents encoded in this way are almost *all* discourse topics, in that the speaker is either introducing a major topic of the subsequent discourse, or recapitulating a major topic of the previous discourse. The examples that follow give a flavour of the kind of environments in which pre-nominal demonstratives are found.

Example (70) comes from the start of a conversation, where the speaker is explicitly setting out the topic for the subsequent discourse. This is deictic reference, but anchored to an abstract entity in the knowledge of the text-internal speakers, rather than to the external world.

(70)   [Context: one folktale character confronts another]
"**wú-mpà ì-rí**   yí-nà   Ø-yó-nò   á↓=ú-yǎa-wà=mù
3s-this   NC3-thing   AG3-REL   2S-be\RLS-PFV   LOC=NC7-do-APPL=1S.PRO
yì-dà   cé   yì-'étéi!"
AG3-be   NEG   AG3-fine
*"**this thing** you are doing to me is not a fine thing!"*

[saat001.007.035]

In (71) the initial reference to the topic NP is formally-marked as such by left-dislocation. The anaphoric reference to the topic is made using a pre-nominal demonstrative.

(71)   [Context: *the fine things* [**ì-rí** 'things, NC3'] *that you eat / after you eat it then you smoke / OK you see tobacco spoils /*]
wú-llè   **ì-rí**   yí↓=ká-ráa
3s-that   **NC3-thing**   AG3=NC1-eat
those thing of eating

[tats007.002.090]

Contrast (71) with (72), where the fire is the sole discourse topic, and the NP with the postnominal demonstrative is non-specific and non-topical.

(72)   [Context: a description of what fire does. *No matter how big the collection of cornstalk stacks, or of grass or whatever...*]
nì=d-dúwà-nù   ú-láa,   ù-láa   wú-u-ràa   **ì-rí**   **yí-llè**   pỗ
and=2s-put\RLS-RES   NC7-fire   NC7-fire   AG7-FUT-eat\IRR   **NC3-thing**   **AG3-that**   all
*when you set fire, fire will consume all of **that thing***

[tats002.002.058]

In (73) the speaker is recapitulating the topic of the current discussion, which was the fact that his children have rejected Cicipu and turned to Hausa instead. This time the NP with the pre-nominal demonstrative is formally marked as a discourse topic by left-dislocation from the main clause.

(73)   [Context: interview about the loss of Cicipu language and culture]
àmáa   **tí-mpà**   **Tì-kógó**   tí-nà   m-úu   mí-ttù   ǹ-kábà-n-nà /
but   **AG6-this**   **NC6-Hausa**   AG6-REL   AG5-child   AG5-1P.POSS   AG5-take\RLS-PFV-VENT
òtù   gáanùkwà   cé /
1P.PRO   understand\RLS   NEG
*but **this Hausa** that our children have brought / we don't understand /*

[Tikula, sagb001.337]

Example (74) is similar to (73) in that the reference is backward-looking. The speaker was asked to give the reason for the enmity existing between two divisions of the Acipu, and after explaining that marriage practices were behind the quarrel, he summarised his explanation as follows:

(74)   lêe   sóbò   nì   cì-'ítànî   t-ì   tì-yỗo-nù-nà
there   because   with   NC6-marriage   AG6-COP   AG6-cause\RLS-RES-PFV
ké-llè   **kà-gáabá**
AG1-that   **NC1-enmity**
*there it's because of marriage it caused **that enmity***

[Tikula, sami001.422]

Although the enmity had been mentioned in the discourse prior to this point, and so in some sense this is anaphoric reference, it is not of the same kind as the post-nominal example (72), where there is no discontinuity in the structure of the discourse. In (74)

431

the speaker is taking a step back from the main event-line of the narration to provide a summary of what he has just said[44]. When the discourse topic of the event-line, the enmity, is introduced into the new "mental space" (Fauconnier 1994) set up to process the summary in (74), it is encoded using a pre-nominal demonstrative.

The following example is similar to (74) in that it comes from a summing up:

(75)   [Context: Summing up an interview on the history of the Akula. *If it weren't for all you elders remaining here...*]
        hú-u-6àa   cé   **ké-llè**   **kà-nàbáyì**
        3P-FUT-have\IRR   NEG   AG1-that   NC1-story
        *they wouldn't have **that story***

                                                                    [Tikula, sagb001.682]

However instead of anaphoric reference to an NP such as **kà-gáabá** 'enmity', here the example involves the reification of the preceding discourse. This is background information rather than foreground information (an 'evaluation non-event' in Grimes' 1975 terminology). As was the case for (74), when the referent of **ké-llè kà-nàbáyì** 'that story' is introduced into the new mental space required for this evaluation of what has gone before, we find a pre-nominal demonstrative.

This kind of transfer of referents between mental spaces characterises most of the pre-nominal examples. All instances of pre-nominal demonstratives encode discourse topics, apart from six occasions when **wú-mpà** 'this' occurs with a time word, as in **wú-mpà kw-ǎa'ǎ** 'this day'. All of these are within Christian prayers and this may well be a calque from a formal Hausa style.

It was noted in §4.4.5.2 that the position of Kiswahili demonstratives relative to their head nouns is said to be dependent on the topicality of the referent. The passage is worth quoting (Lyons 1999:115):

> Demonstratives can also encode the fact that a referent is the current topic of the discourse. It is often difficult to distinguish such topic demonstratives from anaphoric ones, since a topic is likely to have been just mentioned...But a particularly clear case is provided by Swahili, in which *topic is expressed by the position of the demonstrative*, and anaphoric reference quite differently. We have seen above that Swahili makes a two-way distance contrast and has an anaphoric demonstrative. These forms occur post-nominally, but the deictic *h-* and *-le* forms *can appear pre-nominally, and then they indicate that the referent is the current topic* [My italics – S.M.].

---

44 i.e. he is encoding a "non-event" rather than an "event" in Grimes' (1975) terms.

It seems that the position of the demonstrative in Cicipu serves a similar purpose[45], although what the consequences are of using a person-marked pre-nominal demonstrative rather than a gender-marked one remains to be seen.

I will end this subsection with a short observation about the use of demonstrative *pronouns*. While the gender-marked demonstrative modifiers can also occur as pronouns (§4.4.5.2), the person-marked ones cannot – instead a dedicated demonstrative personal pronoun (**é-mpè**, **é-llè**, and so on – see §4.4.3.4) must be used. **É-mpè** in particular is interesting because it can readily be used deictically to refer to any kind of animate or inanimate object, much more so than the anaphoric pronoun **éví**. It may be that this is related to the fact that referents that are being explicitly pointed out are often of maximum current interest to the speaker, and therefore good candidates to become discourse topics.

8.9.3   -èné **'which'**

The wh-word **-èné** 'which' also offers a choice between gender and person agreement. **-èné** usually occurs before the head noun, but it may also appear afterwards. As with the article and demonstrative, person agreement cannot occur post-nominally. Most of the examples in the corpus are found with person agreement in the universal quantifier construction (§4.9) e.g. **sáa w-èné lóokàcí** 'all the time'. Once these examples are removed from consideration, there are only twelve examples (9 gender- and 3 person-marked) left in the corpus, and this is simply too small a number to deduce anything further.

8.9.4   Copula

The following tables illustrate the two paradigms involving the copula[46].

---

45 Hausa also has a contrast between pre-nominal and post-nominal demonstratives, but Hausa pre-nominals differ from Cicipu ones in that they are associated with both new information and exophoric reference (Newman 2000:150-152).

46 Recall from §4.3.3 that the copula occurs in both identificational and predicate nominal constructions. It is the former kind that is of interest here, since the copula only agrees in person if the NP controller is a personal pronoun, which being referential cannot function as a predicate.

Table 75: Noun class pronouns plus copula (repeated from §6.2.15)

Class	Pronoun	Copula	Translation	
1	**k-í**	**k-è**	*it's it*	[e.g. **kó-ggòmbò** 'bat']
2	**h-í**	**h-è**	*it's them*	[e.g. **ó-ggòmbò** 'bats']
3	**y-í**	**y-ì**	*it's it*	[e.g. **ì-námà** 'meat']
4	**m-í**	**m-è**	*it's it*	[e.g. **mà-díyá** 'hare']
5	**m-í**	**m-ì**	*it's them*	[e.g. **ǹ-díyá** 'hares']
6	**t-í**	**t-ì**	*it's it*	[e.g. **cì-kóotò** 'drum']
7	**w-í**	**w-ì**	*it's it*	[e.g. **ù-yáa** 'road']
8	**v-í**	**v-ì**	*it's it*	[e.g. **c-cɔ́'ɔ̀** 'sheep']
9	**kw-í**	**kw-ì**	*it's it*	[e.g. **kù-lácí** 'girl']

Table 76: Personal pronoun/copula complexes (repeated from Table 16)

	Singular		**Plural**	
1	**ámbì**	*it's me*	**óttì**	*it's us*
2	**ìvɔ́ vì**	*it's you (sg.)*	**ìdó yì**	*it's you (pl.)*
3	**évvì**	*it's him/her*	**érè**	*it's them*

The morphological structure of the first- and third-person entries in Table 76 is not altogether transparent, in contrast to the tidy gender-marked paradigm in Table 75. The person-marked forms were called 'focused independent personal pronouns' in §4.3.3.1, but they can also be thought of as showing the copula agreeing in person. This structure is still transparent in the second-person entries, and it seems likely that the other entries are also derived from a structure analogous to the gender-marked constructions in Table 75: an independent personal pronoun, plus the C- form of the corresponding person subject agreement prefix [in **bold**], plus the copular stem **ì**.

Table 77: Suggested historical derivation of the person-marked copula forms

	Singular		**Plural**	
1	ámú # **m** + ì	> ámbì	ótú # **t** + ì	> óttì
2	ìvɔ́ # **v** + ì	> ìvɔ́ vì	ìdó # **y** + ì	> ìdó yì
3	éví # **w** + ì	> évvì	éré # **h** + ì	> érè

The singular entries would perhaps be more convincing if **v** was substituted for **m** and **w**, which calls to mind ᴀɢ8 neutral gender agreement (§6.4) rather than person agreement, while the 3ᴘᴘ derivation would be entirely unmotivated were it not for the rest of paradigm. However there is no problem with the second-person cells, and the 1ᴘᴘ derivation looks reasonable.

The distinction between grammatical and anaphoric agreement which was applied to verbs in §7.6 is also relevant to the copula, at least in the third person. Gender-marked copulas can be thought of as taking part in grammatical agreement, since the subject NP is obligatory:

(76) *(m-í)   m-è   lêe
      AG4-PRO AG4-COP there
      it's there [Mɔ̀-kkɔ́'ɔ́ 'name of a lake, NC4']

[Tikula, sagb001.647]

The 3PS copulas, on the other hand, usually take part in anaphoric agreement, as in (77), although explicit subject NPs seem possible in restricted circumstances (e.g. (78) with a proper name), as was the case for the person subject prefix.

(77) 'ɔpívì!   évvì!
      'ɔpɔ́\LH = vì   évvì
      hold\IMP=3S.PRO   3S.COP
      stop him! ***it's him****!*

[saat001.008.132]

(78) Rùgúncí   évvì   lêe
      [name of lake] 3S.COP there
      Rugunci is there

[tats005.002.137]

With respect to *choice* of agreement feature, this usually depends on the choice of subject expression. If the speaker chooses a gender-marked NP or pronoun then the copula must agree in gender, as in (76). If not, then the copula almost always agrees in person. On rare occasions, in the case of plural person agreement the expected complex form **érè** is replaced by the **é** variant of the 3PS pronoun followed by an unexpected gender-marked copula, in which case the pronoun is interpreted as having plural meaning[47]. There is only one example in the corpus (cf. **h-í h-è** in Table 75):

(79) é   h-è
      3S.PRO   AG2-COP
      it's them [à-yúpù 'crocodiles, NC2']

[tats005.002.250]

The crocodiles were discourse-topical at this point in the text, and it may be that this (coupled with their relatively low animacy) accounts for the hybrid nature of (79). As with all the targets being discussed in this section, more research is required to improve

---

47 Otherwise **é** is always understood as singular.

on the current sketchy account.

Before leaving this section it is worth expanding on the similarities between the two variants of the 3PS pronoun, ɛ́ and ɛ́ví, and the person-marked copula ɛ́vvì. It is likely that the form ɛ́vvì is actually the result of the *renewal* of copular encliticisation, with ɛ́ví being the now fully-grammaticalised result of a previous derivation, involving the coalescence of the original 3PS pronoun ɛ́ and the copula.

Apart from the fact that the postulated original ɛ́ still exists, further evidence comes from the possessive pronouns (§4.4.5.1.1). In the Tirisino dialect the 3PS possessive pronoun is **-ɛ́vì**, as in (80a). In Tikula however, the corresponding form is just ɛ́, as in (80b). It seems likely that Tikula preserves the original form.

(80) (a) kà-táarí ké-evì
   NC1-stone AG1-3S.POSS
   his/her stone [Tirisino]

  (b) kà-táarí ké-e
   NC1-stone AG1-3S.POSS
   his/her stone [Tikula]

The suggested derivation has parallels in other African languages, where independent pronouns also show a fossilised pronoun plus copula construction (e.g. the Mande language Bokobaru and several Omotic languages – Siewierska 2004:255-257).

8.10   Chapter summary

After preliminary sections on methodology (§8.2) and participant reference in Cicipu (§8.3), we saw in §8.4 that the notions of both *inherent* topicality (specifically, animacy) and *discourse* topicality are required to account for the alternation between gender and person agreement in Cicipu. Neither on its own is sufficient, since unindividuated inanimate referents (e.g. tobacco) can trigger person agreement if they are sufficiently discourse-topical, and humans which are not discourse topics can trigger gender agreement where person agreement might otherwise be expected. As well as these semantic and pragmatic conditions on agreement, there is also a morphosyntactic condition – we saw that NC8 controllers are more likely to trigger person agreement than nouns from other classes. Furthermore, there is evidence for an effect of natural gender, and it may be that the animacy hierarchy in Cicipu should distinguish between male and female humans.

436

Concerning the properties that the theoretical notion of topic must have in order to account for the alternation, we found that the notion of *intrinsic interest* introduced in §2.3.2.2 is important. Sections of a discourse, and in some cases entire texts[48], may be integrated semantically by a referent in which the interlocutors have only a passing interest. However it does not follow from the repeated mentions of such referents that they will always integrate the text *thematically*. Just talking about something is not enough to make it topical, and if our conception of topic is restricted to the number and frequency of mentions within texts then we will not be able to make use of this notion to account for examples such as those discussed in §8.4.2. Similarly a theory of topic restricted to the sentence such as Lambrecht's (§2.3.1) cannot account for the fact that in several of the examples in this chapter (see §8.4.6), the thing that an individual utterance is about is indexed with 'non-topical' gender agreement rather than 'topical' person agreement. To understand this alternation in Cicipu we must consider the importance of the referent to the wider discourse, rather than just the presuppositions and assertion relevant at one snapshot in time.

In §8.5 we saw that gender and person exponents are distributed according to a coding progression from lexical NP > gender agreement > person agreement. In §8.6 we briefly considered possible alternative explanations for the two paradigms of agreement. Section 8.7 brought together the analyses in chapters 7 and 8 and presented the various factors involved in flowchart form. It was argued that gender agreement should be considered the default, with person agreement occurring in specific circumstances. In §8.8 it was suggested that the coding progression identified in §8.5 might be regarded as a progression in coding weight, and the relationship of the two kinds of agreement to anaphora and deixis was discussed. Finally in §8.9 we investigated the other targets which show the gender/person alternation. We found that the factors relevant for subject prefixes and pronouns also appeared, to some extent, to be relevant for the article. For demonstrative modifiers the position of the demonstrative seems to depend on these factors, but there was not enough data to investigate the difference between gender and person agreement.

---

48 For example the match report discussed in §2.3.2.2.

Chapter 9 – Conclusion

In this thesis I have described the Cicipu gender system, focusing on the alternation between gender and person agreement in discourse. Part I introduced the Cicipu language and the research methodology used here (chapter 1), before setting out the relevant theoretical context (chapter 2). Part II provided a sketch phonology (chapter 3) and grammar (chapter 4). The scope narrowed in Part III where I described the Cicipu noun class system, concentrating on noun classification and derived nominals in chapter 5, and on agreement in chapter 6. Part IV turned to the major research questions of this thesis: chapter 7 investigated the pre-requisites for both gender and person agreement, while chapter 8 looked at the relevant conditions when there was a choice between the two different kinds of agreement.

The conclusion is divided into three sections. In §9.1 I summarise the main findings of Part IV and revisit the research questions set out at the beginning of the thesis. In §9.2 I discuss the contribution the research makes to linguistic science. Finally in §9.3 I suggest topics for further study, arising both from Part IV and more generally.

9.1 Main findings

Part IV was devoted to the alternation between gender and person agreement on person markers (i.e. agreement prefixes, pronominal clitics, and pronouns), considered from the point of view of the research context set out in chapter 2. In chapter 7 I investigated the phonological, morphological, and syntactic properties of the five different paradigms of Cicipu person markers, and concluded that the INDEPENDENT GENDER (noun class) PRONOUNS and INDEPENDENT PERSONAL PRONOUNS are, as the names suggest, free pronouns rather than bound morphemes. There is a separate paradigm of post-verbal PERSON-MARKED OBJECT CLITICS, occurring in complementary distribution with the independent personal pronouns. These object clitics have reduced versions that predominate in certain syntactic environments (particularly before the clausal negator cé). In addition there are two paradigms of subject agreement prefixes – one marked for GENDER, and one marked for PERSON. Reasons were given in §7.8 for considering these as two separate paradigms rather than a single complex one. Both of these sets of prefixes must be considered ambiguous agreement markers according to the typology of Bresnan and Mchombo (1987) and Siewierska (1999), but the gender subject prefixes are closer to the

'grammatical' end of the grammaticalisation cline than the person subject prefixes, since person agreement only rarely occurs with an explicit subject NP. Neither type of agreement seems as far along the cline as the corresponding kinds in Central Kambari.

The properties of the five paradigms are summarised in Table 78, and Figure 41 represents the relative status of the two subject agreement paradigms in Central Kambari and Cicipu with respect to grammaticalisation.

Table 78: Summary of properties of Cicipu person markers (repeated from Table 61)

Paradigm	Morph. status	Grammatical function	Features	Grammatical agreement?	Anaphoric agreement?
Independent PERSON	Word	Any	Person	No	Yes
Independent GENDER	Word	Any	Gender	No	Yes
Object clitics	Clitic	Object	Person	No	Yes
Subject PERSON	Prefix	Subject	Person	Sometimes, depends on subject NP	Yes
Subject GENDER	Prefix	Subject	Gender	Usually, depends on subject NP	Yes

	Anaphoric	>	Ambiguous	>	Grammatical
Central Kambari		**Person**			**Gender**
Cicipu		**Person**		**Gender**	

Figure 41: Relative positions of person and gender agreement in Central Kambari and Cicipu on Bresnan and Mchombo/Siewierska's agreement marker typology (repeated from Figure 37)

In §2.2.4 I discussed cross-linguistic examples of variation in agreement and considered the kind of conditions that are typically said to be relevant when there is a choice. This led on to the discussion of topicality in §2.3, where evidence was given that cross-linguistically the concept of discourse topicality has a role to play in influencing grammatical structure. In chapter 8 we saw that discourse topicality is a condition on agreement on Cicipu person markers, together with the animacy and the noun class of the controller. Person agreement is more likely if the controller referent is discourse-topical or highly animate, or if the controller NP is from noun class 8. Gender

agreement is more likely in the reverse scenarios: if the controller referent is non-discourse-topical, low in animacy, or if the controller is from some other noun class. The situation with respect to discourse topicality and animacy can be diagrammed as in Figure 42.
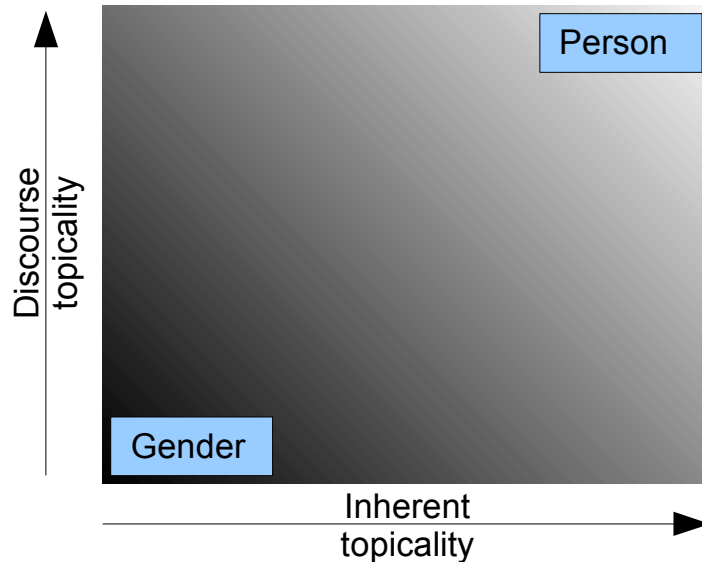

Figure 42: Variation in agreement according to inherent and discourse topicality, repeated from Figure 38

The examples in §8.4 showed that the relevant notion for the topicality agreement condition is *discourse* topic rather than sentence topic (§8.4.6), and that inanimate or animal referents with a high degree of referential density are not indexed with person agreement, unless they are also of intrinsic interest to the speaker (in other words they *thematically* integrate the text as well as semantically integrate it).

In §8.4.5.3 the effect of natural gender was discussed, with male referents seemingly more likely to trigger person agreement than females (at least for male speakers).

Section 8.5 showed that if, when the conditions mentioned above are applied to a particular referent, they work out so that either gender or person agreement is possible, then there is a very strong tendency for gender agreement to precede person agreement in mixed anaphoric chains. Thus there is a coding progression from LEXICAL NP > ANAPHORIC GENDER AGREEMENT > ANAPHORIC PERSON AGREEMENT, which recalls the other coding progressions mentioned in §2.3.2.5. Finally, preliminary investigation of the other agreement targets which show the gender/person alternation (§8.9) suggests that the same three conditions (animacy, discourse topicality, and noun class) also apply.

The research questions set out in §1.1.1 are repeated and summarily answered below:

- What are the agreement targets in Cicipu?
- Which of these agreement targets inflect for more than one agreement feature paradigm?
  - o Are the syntactic environments in which the different paradigms occur mutually-exclusive?
  - o If not, what are the factors that influence the agreement paradigm when there is a choice?
  - o If topicality is one of these factors, what kind of topicality is it? Is the data better explained by theories of 'sentence topic' or of 'discourse topic'?

The agreement targets in Cicipu were set out in §6.2. Inside the NP almost all modifiers agree in gender, including numerals, adjectives, wh-words, demonstratives, the associative construction/possessive pronouns, the article/relativiser, and the quantifiers nínnîi 'only' and -mbɔ̃ 'another'. Outside the NP, agreement is found on verbs, the copula and negative copula, pronouns/object clitics, and demonstrative adverbs when used as predicates.

Several different targets inflect for either GENDER or PERSON: verbs, pronouns/object clitics, the article, demonstratives, - èné 'which', and the copula. Each of these targets can occur in syntactic constructions where there is a choice of feature. In the case of the subject prefixes, object clitics, and pronouns, there are three conditions influencing the choice; the noun class of the controller NP, and the animacy and topicality of the controller referent. As mentioned above, the notion of discourse topic explains the data better than sentence topic.

9.2 Contribution to linguistics

The sketch phonology and grammar (Part II) and description of agreement (Parts III and IV) offered here is the first descriptive work on Cicipu, a member of the severely under-studied Kainji group. Similarly, the corpus submitted to the Endangered Languages Archive at SOAS contains the only extant audiovisual recordings of the language. As described in §1.4, this corpus consists of six hours of time-aligned interlinearised texts (approximately 12,000 clauses), together with recordings of almost two thousand

lexemes. Although I was compiling it in order to document a previously undescribed language of uncertain endangerment, this electronic corpus proved to be an indispensable tool for both the descriptive work in Part II and the study on agreement in Parts III and IV.

In addition to this documentary and descriptive work, the main theoretical contribution of this study is towards our understanding of (i) incorporated pronouns and (ii) agreement, in particular the factors relevant to variation in agreement. I will discuss both of these below, and also comment on the relationship between discourse analysis and typology.

It is not surprising that the various Kainji and Plateau languages mentioned in §2.2.4.2.3 have not made it into the literature on incorporated pronouns, since they are poorly known and very little has been published on them. A further complication is the persistent West African linguistic practice of writing prefixes as separate words, so that even when languages such as Amo (Anderson 1980a) make it into the typological literature, they are not recognised as having more than one series of incorporated pronouns. It is to be hoped that the analysis of another Kainji language presented in this thesis will lead to a re-evaluation of the automatic association that is so often made between anaphoric agreement and topicality. As was stressed in §2.2.5, not everything that is introduced into a discourse automatically becomes a topic (otherwise topicality loses its independent explanatory power). Some referents are transient in discourse, but nonetheless stay around long enough to require some sort of coding that neither sets them up as a new/contrastive referent, nor marks them out as a discourse topic. In Cicipu gender agreement markers serve this purpose, whereas person agreement markers function as indicators of discourse topics (particularly in the case of inanimates and animal referents). As we have seen both gender and person agreement markers can be involved in anaphoric agreement – in fact anaphoric reference to non-topical inanimates or animals is always by means of gender agreement. Consequently the Cicipu data supports Culy's (2000) contention that topicality is an independent dimension along which incorporated pronouns can be placed[1], and suggests that Siewierska's (2004) observation that dependent markers do not vary with respect to their discourse function should be re-evaluated (see §2.2.5.1).

---

1   It is not possible to go as far as Culy does for Takelma -*kʰwa* when he analyses it purely as a discourse topic marker. As we saw in chapter 8 person agreement can index non-topical referents if they are sufficiently animate.

If topicality is indeed an independent dimension to the grammatical vs. anaphoric agreement distinction, then this raises the question as to whether *grammatical* agreement markers can also vary in their discourse function. In Cicipu there is rarely a choice between grammatical gender agreement and grammatical person agreement. The only environment in which this choice materialises is when there is both an NC8 subject NP and a vowel-initial verb stem (§7.6.1.1). In that scenario it seems likely that both animacy and topicality play a role, although there are too few examples in the corpus to be certain. In Central Kambari, however, the two kinds of agreement are only in contrast when there *is* an explicit subject NP (§7.6.8). It seems likely, given Crozier's (1984) characterisation of person agreement as "topic" agreement, that grammatical agreement markers can vary as to their discourse function in just the same way as anaphoric agreement markers. The Tsureshe text mentioned in §2.2.4.2.3 suggests that the language patterns with Central Kambari in this respect.

As was observed in §2.2.4, recent studies on agreement have acknowledged the importance of topicality with respect to variation in agreement. The data from Cicipu presented here argue for an open mind as to the proper domain of 'topic'. Whatever its merits on other grounds, we saw in §8.4 that Lambrecht's theory of information structure cannot distinguish between inanimate and animal referents that trigger only gender agreement, and those indexed by anaphoric chains that progress to person agreement. Rather it is the property of being discourse-topical (§2.3.2) that enables inanimates or animals to trigger person agreement. Similarly we saw a difference between non-discourse-topical and discourse-topical adult humans, in that gender agreement appears to persist across intonation boundaries in the former but not in the latter. Therefore Cicipu provides an exception to Lambrecht's contention that discourse topic (as opposed to sentence topic) has little to do with the grammatical form of sentences (§2.3.1).

Lambrecht's dismissal of discourse topic with respect to morphosyntax may be a little premature, given there have been so few cross-linguistic studies on discourse topic. Ideally, of course, typologists interested in the morphosyntactic effect of discourse-pragmatic properties such as topicality would base their conclusions on comparative discourse analysis of a broad range of languages. However this is not currently possible, partly due to the practical limit on the number of languages a typologist can know well,

and partly due to the lack of interlinearised text collections in a wide range of languages (see Myhill (2001:163) for discussion). The texts collected and made available as part of this project are one contribution towards offsetting this dearth (which is particularly severe for Niger-Congo languages), but a vast amount of work still needs to be done in this area. Consequently typology is almost always of the "grammar-extraction" variety (Wälchli 2007). If the linguistic effects of topic and focus are mentioned at all in reference grammars, it is not usually done to the extent that would make cross-linguistic comparison possible for someone who did not already know the language (Myhill 2001:162-163). Often discussion is limited to elicited paradigms, and as Bearth (1999:141) has pointed out, linguistic devices which appear to be straightforward indicators of topic/focus in elicited paradigms may turn out to behave very differently in "natural text"[2].

Typologists of course do more than just browse reference grammars, and typology is also informed by syntactic theory, in particular typologically-oriented frameworks such as Lexical-Functional Grammar. Thus the effect of an inevitable reliance on elicited material at the expense of more authentic forms of communication is reinforced by the adoption of insights about the structure of the *sentence* gleaned from generative theories of syntax such as LFG. This would not be a problem if the domain of inquiry of typology was restricted to the sentence. However this is not the case, and instead mainstream typologists share a "general ideological compatibility" with discourse analysts, based on "the view that the study of language should be based upon analysis of empirical data" (Myhill 2001:161). Studies such as the present one, which involve a detailed textual analysis of agreement in a single language, serve to mitigate the inevitable skewing of typological theory towards the domain of the sentence.

I mentioned above that it is "not currently possible" for typologists to base their theories of topic-sensitive phenomena such as agreement on comparative discourse analysis, which of course begs the question what can be done to make it possible? The creation of interlinearised text collections of minority and endangered languages is an obvious and urgent priority. But is it enough just to provide an interlinearised corpus and leave the rest to the typologist of tomorrow (or five hundred years' time)? Or is there a way to furnish documentary corpora with annotation sufficiently "thick" (Nathan and Austin 2004) to, say, enable a typologist interested in variation in agreement to

---

2   See also McGill (2004) on focus in the Gur language Sisaala-Paasaal.

arrive at the analysis which was derived here?

Advances in technology have made the creation and access of language-specific corpora much easier than in the past, and recent developments in documentary linguistics have led to "best practice" standards for interlinearised annotation (e.g. Schultze-Berndt 2006, Bow et al. 2003) as well as stimulating interest in ontologies for linguistic description (e.g. Farrar and Langendoen 2003). However, despite these developments it remains the case that even the most comprehensive outputs of language documentation projects remain relatively inaccessible to typologists. Making corpora of interlinearised texts publicly available is a step in the right direction, but without a more fine-grained coding of the corpus than is usual the analysis of such corpora will be too time-consuming for most typologists. In general, it is unreasonable to expect language documenters to take into account specific areas of typology when building their corpora – why should they be particularly concerned with, say, the discourse variation of agreement, compared to any other kind of phenomenon? Language documenters face a difficult enough task without putting yet another burden on their shoulders.

However for documentary projects which focus on a particular area of linguistic theory (such as the present work), it is feasible and arguably desirable to have as one of their outputs a documentary corpus which has been explicitly marked up, in such a way that the patterning of linguistic structures relevant to their analyses can be replicated by others who have less language-specific knowledge and time to spare than the original researcher. To take the present study as an example, ideally it should be possible for a typologist to use the Cicipu corpus to electronically test the claims I have made about the association between animacy, topicality[3], noun class 8, and agreement features. As well as making the original linguist's work more accountable[4], this approach can of course be of help in the initial stages of developing and testing hypotheses about the language. Provided the corpus is coded to an appropriate level of detail, then it will also be of use in cross-linguistic investigations. For example in a cross-linguistic study on variation in agreement, variables such as definiteness, specificity, animacy, topicality,

---

3   This is controversial – see Myhill (2003) for the point of view that topicality is too vague a notion to be of any use for cross-linguistic discourse analysis. Dooley's definition of topic is cross-linguistically applicable but still subjective – topics cannot be read directly from the text, but instead require an introspective consideration of thematicity.

4   Just making one's corpus "transparent" to the philologist of the future (Woodbury 2003:47) does not ensure that the associated descriptive/analytical work is "accountable" (Bird and Simons 2003), since there may not be anyone sufficiently interested in both (i) your texts and (ii) your analysis to go to the trouble of testing your conclusions.

focus, and precedence (and perhaps others) would all have to be coded, and a means provided to match up individual controllers with targets. This would involve a signficant amount of coding beyond what would be immanent in even a well-structured interlinear corpus and accompanying lexicon, but would provide a means not only of finding out the factors that determine variation in individual languages, but also of testing more general claims, for examples the ones Corbett (1991:239-240, see §8.8) makes about the nature of agreement progressions involving syntactic and semantic agreement. The details of such an approach are a matter for further investigation.

9.3   Areas for further study

Several of the topics discussed in Part IV merit further research. First, the effect of natural gender on the choice of agreement feature (§8.4.5.3) is a promising topic, given that sex is generally held to be irrelevant with respect to Benue-Congo gender systems. In particular, it is necessary to find out whether male speakers' tendency to use person agreement for male referents and gender agreement for female referents is replicated or reversed in the speech of women. This line of research would provide a good opportunity for addressing the main weakness in the present Cicipu corpus, which is the lack of representation of female speakers.

Secondly, the text-level progressions discussed in §8.5.2 require further investigation. One possible technique would be to ask for retellings of a story structured in the same way as the 'witch' text (§8.5.2), with a participant of incidental importance in the first part progressing to a character more likely to function as a discourse topic in the second.

Thirdly, the investigation into the alternation on other targets in §8.9 was rather sketchy and needs development.

Fourthly, now that more is understood about the gender/person alternation (and about the kind of experiments likely to prove successful in Acipuland – see §1.4) it is possible to design experiments to provide corroborative evidence for the analysis proposed in chapter 8. For example, in a story with two characters the gender/person indexing of these referents could be manipulated to test the effect on subjects' perceptions about who the story is about, or how it will continue (i.e. a "discourse completion task", Bardovi-Harlig 2002:184). An experiment of this kind, if successful, would also demonstrate that person agreement (or the lack of it) functions as an active

indication to the hearer as to the discourse status of the indexed referent (thus causing him to re-evaluate the way he is construing the discourse), rather than merely passively reflecting what is already known (see §8.8).

Areas mentioned in Part III that require further research include the nature of the NC7 infinitive (§5.4.1, is it truly a mixed category?), the distinction in meaning between the various different deverbal nominalisations (§5.4), the NC8/AG8 geminating prefixes (§5.5.7, §6.1.3), hybrid nouns (§6.3, are there any others beside **má-gàjì** 'priest', and what are their agreement possibilities with respect to Corbett's Agreement Hierarchy?), trigger-happy agreement (§6.5), and the possibility of gender resolution with the most distant conjunct (§6.6).

In addition to improving the analyses set out in the present work, the study of agreement in other Kainji languages should be a high priority. It is well-known that despite the vast number of languages on the continent, African languages punch below their weight when it comes to their contributions to theory-building[5]. This is in part due to the bias in typological sampling methods[6], but also because of the dearth of detailed descriptions of African languages. This is especially true for the Kainji branch of Benue-Congo. Concerning gender and person agreement, we saw hints in §2.2.4.2.3 that the alternation found in Cicipu and Central Kambari may in fact be present in most Kainji languages. In order to test this, more detailed descriptions of agreement in these languages are necessary – and in fact are desirable for other reasons, since even the little that is known about them suggests they have interesting agreement systems. For example, in the Northwest languages (e.g. U̱t-ma'in, Smith 2007) the noun class prefix becomes a suffix when the noun functions as a subject, in which case the gender subject prefix on the verb does not occur. Thus it seems as if a former agreement prefix has been reinterpreted as a nominal suffix. The existence of a separate paradigm of person agreement provides an opportunity to test the current status of the 'flipped' marker. It was mentioned in §2.2.4.2.3 that in Pongu, in addition to the gender and person subject prefixes familiar from Cicipu, there are also what appear to be *portmanteau* subject prefixes, combining both features, thus leading to a three-way contrast. Studies on these agreement systems would dovetail nicely with much-needed documentary projects.

Many of the other topics touched on in this thesis warrant further research. Just a

---

5 Phonological theory being the major exception (especially with respect to tone and autosegmental phonology).

6 See Mous (2003) on the loss of linguistic diversity in Africa.

few that stand out are gemination and its implications for syllable structure (§3.1.8, §5.5.7), the extreme tonal downstep in complement position (§3.4.7), the behaviour of loanwords with respect to vowel harmony (§3.5.2), the number of nasalisation contrasts (§3.6.2), the neutralisation of lexical contrasts caused by the backwards spreading of the i vowel from the 3PS object clitic (§3.7.3), the ordering possibilities in the NP (§4.4.5.6), the morphological status of the pluractional and causative infixes (§4.6.1), and the connection between the (homophonous?) applicative and anticausative (§4.6.2-4.6.3).

Forty-four years ago Carl Hoffmann wrote in the introduction to one of the first modern works on a Kainji language "The publication of this wordlist hardly needs justification" (1965:7). Sadly, even now it remains true that a single published paper would be a step forward for most Kainji languages. Nevertheless there has been a recent resurgence of interest in the group, and I hope that this thesis will be a good foundation, and perhaps even a motivation, for other linguists as they take up the study of Cicipu and other Kainji languages.

Appendix A – Two Cicipu texts

The corpus used for this study is available as audio-enabled interlinearised text at www.cicipu.org. Two sample texts are presented here.

The bloodthirsty sword (samy002)

This folktale was told by Tirisino speaker Markus Mallam Yabani in Galadima village on 5th February 2008. The main character is a sword, references to which are in bold. Note the progressions from LEXICAL NP > GENDER > PERSON (§8.5). This text has a higher incidence of lexical NP mentions than is usual, which of course provides greater opportunity for multiple occurrences of such progressions within a single text. The minor character (the horseman) is consistently indexed by person agreement since the antecedent z-zá 'person' offers no other choice (see §7.6.1.1). Paragraph boundaries have been indicated with a ¶, based on the formal and conceptual evidence described in §2.3.2.4. Note the correlation with the discourse marker tò 'OK'.

mísòoní   mísòoní!
story   story
story story!

míi mìi! [formulaic reply]

àyá   tí-yòo-nò!
come\IMP   1P-go\IRR-VENT
Come let's go!

hói hòi! [formulaic reply]

¶ wú-nà **mà-gɛ̂i**   m-è   gó / **mà-gɛ̂i**   **mé-llè** /
3S-ART   **NC4-sword** AG4-COP TOP   **NC4-sword** **AG4-that**
*¶ Once there was a **sword** / **that sword** /*

éví   sáa w- èné   s-sá'à /   **ù**-sì-tá'á   mò-hɪ̂'ɪ̀ /
3S.PRO or   3S-which NC8-time **3S-HAB**-want   NC4-blood
***it** all the time / **it** needed blood /*

mà-gɛ̂i   **mé-llè**   kúmá /
NC4-sword **AG4-that** and
*and **that sword** /*

sáa w-èné s-sá'à **ù**-sì-tá'á **w**-îndà ù-'úgù mò-híˊì áↆ=í-ɗáa /
or 3s-which NC8-time **3s**-HAB-want **3s**-see\IRR NC7-spill NC4-blood LOC=NC3-ground
*all the time **it** needed to see the spilling of blood on the ground /*

¶ tò **mà-gâi** **má-nà** kúmá / **mò**-yô, ǹ kàrá gȅi /
OK **NC4-sword** **AG4-ART** and **AG4**-be\RLS with sharp much
¶ *OK and **the sword** / **it** was very sharp /*

sɔ́bɔ̀dà n=**ù**-cíyò z-zá, **wú**-u-kɔ̀dɔ̀ z-zá /
because when=**3s**-get\RLS NC8-person **3s**-FUT-cut\IRR NC8-person
*because when **it** got someone, **it** would cut them /*

kɔ̀-kkɔ́dǐ kà-yápù, wú-nà lóokàcí kɔ̀-kkɔ́dǐ kò-nósì /
NC1-slice AG1-two 3s-ART time(NC8) NC1-slice AG1-four
into two slices, sometimes four slices /

¶ tò **mà-gâi** **mé-llè** / káakúllè, **mò**-cíyò cé mò-híˊì /
OK **NC4-sword** **AG4-that** that_day **AG4**-get\RLS NEG NC4-blood
¶ *OK **that sword** / that day, **it** didn't get any blood /*

mà-gâi mà-hwâarà kù-zìzá'ù, **ù**-yúu bɔ̀lɔ̀ mò-híˊì /
NC4-sword AG4-start\RLS NC9-shiver **3s**-CONT look_for NC4-blood
***the sword** started shivering, **it** was looking for blood /*

ù-hwâarà ú-yǎa tùnàní /
3s-start\RLS NC7-do thinking(NC8)
it started thinking /

wú-u-dɔ̀nɔ̀ hàná-hàná **ú**-yòo **ú**-cìyò mò-híˊì ká'à /
3s-FUT-follow\IRR where-REDUP **3s**-go\IRR **3s**-get\IRR NC4-blood now
*where **it** could follow to go and get blood now /*

sɔ́bɔ̀dà **ù**-tá'à cé mò-híˊì míↆ=i-vóotò / **ù**-tá'à cé máↆ=má-wáa /
because **3s**-like\RLS NEG NC4-blood AG4=NC3-goat **3s**-like\RLS NEG AG4=NC4-dog
*because **it** didn't like the blood of goats / **it** didn't like [the blood] of the dog /*

ù-tá'à cé máↆ=n-náa / **ù**-tá'à cé mɔ́ↆ=kú-tɔ́ɔ /
3s-like\RLS NEG AG4=NC8-cow **3s**-like\RLS NEG AG4=NC9-chicken
***it** didn't like [the blood] of the cow / **it** didn't like [the blood] of the chicken /*

sée gɔ́ máↆ=z-zá / sɔ́bɔ̀dà máↆ=z-zá mà-ɗâa n=ù-jíˊi /
unless TOP AG4=NC8-person because AG4=NC8-person AG4-surpass\RLS with=NC7-value
only [the blood] of man / because [the blood] of man was more precious /

mà-ɗâa kúmá ǹ kà-rìmâi /
AG4-surpass\RLS and with NC1-pleasure
it was also more pleasurable /

¶ tò **mà-gâ̂i** **mé-llè** m-ǎyà **mà**-hwâarà / kù-zízá'ú /
OK NC4-sword AG4-that AG4-come\RLS AG4-start\RLS NC9-shiver
¶ *OK then **that sword** then **it** started / shivering /*

m-ǎyà **mò**-'úngò **mà**-hyẫa tò /
AG4-come\RLS AG4-rise\RLS AG4-say\RLS OK
*then **it** got up and **it** said OK /*

nàhá **ú**-yòo bɔ̀lɔ̀ 'ásù wú-nà z-zá ù-yó-nò /
let\IMP 3s-go\IRR look_for place(NC7) AG7-REL NC8-person 3s-be\RLS-PFV
*let **it** go and search for where people are /*

ánà **ù**-yôo-nò / 'ásù wú-nà **ù**-hyẫa-nà **ú**-yòo **ú**-zàa z-zá /
how 3s-go\RLS-PFV place(NC7) AG7-REL 3s-say\RLS-PFV 3s-go\IRR 3s-find\IRR NC8-person
*when **it** went / to the place **it** had thought **it** would go and find someone /*

à-zâa, kóngó / báa á-zá /
3P-find\RLS abandoned_habitation(NC8) NEG NC2-person
only an abandoned habitation was found / there was no-one there /

¶ **mà-gâ̂i** **mé-llè** / m-ǎyà **mò**-'óosù / **ù**-'ésènù kù-zìzá'ù /
NC4-sword AG4-that AG4-come\RLS AG4-again\RLS 3s-add\RLS NC9-shiver
¶ ***that sword** / **it** again / **it** shivered even more /*

lóokàcí n-nà **mè**-'ésènù kù-zìzá'ù / **w**-ǎyà **ù**-'wâa /
time(NC8) AG8-REL AG4-add\RLS NC9-shiver 3s-come\RLS 3s-pass\RLS
*when **it** shivered even more / then **it** left /*

ánà **ù**-'wâa-nà / **ù**-yôo **ù**-hángù wú-nà z-zá / g-gédù á↓=d-dɔ̂ɔ /
how 3s-pass\RLS-PFV 3s-go\RLS 3s-spot\RLS 3s-ART NC8-person NC8-up LOC=NC8-horse
*when **it** left / **it** went and spotted someone / up on a horse /*

¶ **mà-gâ̂i** **mé-llè** m-ǎyà **mà**-hyẫa tò ká'à / **wú**-u-cìyò mò-hî̃'ĩ /
NC4-sword AG4-that AG4-come\RLS AG4-say\RLS OK now 3s-FUT-get\IRR NC4-blood
¶ *then **that sword** then **it** said OK now / **it** would get blood /*

tò kúmá èsée **mà-gâ̂i** mà-lápà cé d-dɔ̂ɔ àkwái ì-ládì gễi /
OK and of_course NC4-sword AG4-know\RLS NEG NC8-horse there_is NC3-speed much
*OK and of course **the sword** didn't know the horse was very fast /*

mà-gâ̂i **mé-llè** / m-ǎyà **mà**-hyẫa kâi!
NC4-sword AG4-that AG4-come\RLS AG4-say\RLS EXCL
***that sword** / then **it** said kai!*

wú-u-cìyò mò-hî̃'ĩ sɔ́bɔ̀ **ù**-hángù z-zá / ká'à **wú**-u-cìyò mò-hî̃'ĩ /
3s-FUT-get\IRR NC4-blood because 3s-spot\RLS NC8-person now 3s-FUT-get\IRR NC4-blood
***it** would get blood because **it** spotted someone / now **it** would get blood /*

¶ ánà **ù-hángù-nà** z-zá / **mà-gâi** mò'ópò̀ ì-láɗì /
how 3s-spot\RLS-PFV NC8-person **NC4-sword** AG4-hold\RLS NC3-speed
¶ *when **it** spotted someone / **the sword** took off* /

z-zá d-dôɔ w-ĭndà **mà-gâi** / á↓=u-wáyá 'ásù wé-evì /
NC8-person NC8-horse 3s-see\RLS **NC4-sword** LOC=NC7-come place(NC7) AG7-3s.POSS
*the horseman saw **the sword** / coming towards him* /

sóbòdà **mà-gâi** **mé-llè,** z-zá n-nà pồ w-índà-nà **m-è** /
because **NC4-sword** **AG4-that** NC8-person AG8-REL all 3s-see\RLS-PFV **AG4-PRO**
*because **that sword**, whoever saw **it*** /

ń nì èvì z-zá t-tò báa sáa z-zá /
if when 3s.PRO NC8-person AG8-one NEG or NC8-person
if it was just one person and no-one else /

wú-u-làpà **mà-gâi** **mé-llè** mà-yúu tà'à mò-hî̀' î̀ /
3s-FUT-know\IRR **NC4-sword** **AG4-that** AG4-CONT like NC4-blood
*he would know **that sword** wanted blood* /

dègèlêe wǎ-ayà, ù-vásà d-dôɔ vé-evì / ù-'wâa n=ì-láɗì /
then 3s-come\RLS 3s-hit\RLS NC8-person AG8-3s.POSS 3s-pass\RLS with=NC3-speed
then, he slapped his horse / he went off quickly /

¶ **mà-gâi** mà-súmà **mà**-súmà **mà**-súmà **mà**-súmà **mà**-súmà **mà**-súmà!
NC4-sword AG4-ran\RLS **AG4**-ran\RLS **AG4**-ran\RLS **AG4**-ran\RLS **AG4**-ran\RLS **AG4**-ran\RLS
¶ ***the sword** ran **it** ran **it** ran **it** ran **it** ran **it** ran!*

ánà **mà-gâi** mà-súmà-nà / tò **ù**-lápà cé z-zá d-dôɔ kúmá /
how **NC4-sword** AG4-ran\RLS-PFV OK **3s**-know\RLS NEG NC8-person NC8-horse and
*when **the sword** ran / OK **it** didn't know that the horseman* /

ù-yôo, ánà d-dôɔ Ø-zárùwà-nà kà-táarí /
3s-go\RLS how NC8-horse AG8-step_over\RLS-PFV NC1-stone
he went, when the horse stepped over a rock /

mà-gâi ánà **mò**-yôo-nò, **èvì** w-áyà **wǎ**-ayà **w**-áyà **ù**-yôo /
NC4-sword how **AG4**-go\RLS-PFV **3s.PRO** 3s-come\RLS 3s-come\RLS 3s-come\RLS 3s-go\RLS
***the sword** when **it** went, **it** then came and went* /

ù-dú'ò kà-táarí / dègèlêe **w**-áyà **ù**-kɔ́t<ɔ̀l>ɔ̀-nù /
3s-collide\RLS NC1-stone then 3s-come\RLS **3s**-cut<PLAC>\RLS-RES
***it** collided with the rock / then **it** folded up* /

tò kàahí **ú**-'ùngò-wò / **ú**-gùtù ù-lápìlà /
OK before **3s**-rise-ANTIC\IRR **3s**-go_back\IRR NC7-prepare
*OK before **it** could get up / and recover* /

z-zá   d-dɔ̂ɔ   yã̂a   ù-'wîi /
NC8-person NC8-horse   do\RLS NC7-distance
the horseman was far off /

¶ tò **mà-gã̂i**   **mé-llè** /   **mǎ**-ayà   **mò**-dôonù lêe / **ù**-'ɔ́pɔ̀ /
OK **NC4-sword** **AG4-that** **AG4**-come\RLS **AG4**-sit\RLS   there   **3s**-hold\RLS
¶ *OK **that sword** / then **it** sat down there / **it** started /*

kú↓=u-yã́a   tùnàní   **ù**-'ɔ́pɔ̀   mí̃-isĩ̀ / tè-ené t-ì   **èvì**   wú-u-yã̀a /
AG9=NC7-do thinking(NC8) **3s**-hold\RLS NC5-crying AG6-which AG6-COP **3s.PRO** 3s-FUT-do\IRR
*thinking **it** started crying / what would **it** do?*

ìndúu sáa w-ènè   ù-lénjí / sée   n=**ù**-cíyò   mò-hí̃ĩ̀   má↓=z-zá /
here_is or 3s-which NC7-day unless and=**3s**-get\RLS NC4-blood AG4=NC8-person
*you see every day / **it** had to get human blood /*

tò ánnà **èvì**   wú-u-yã̀a tè-ené?
OK today **3s.PRO** 3s-FUT-do\IRR AG6-which
*OK today what would **it** do?*

¶ dègèlêe **mà-gã̂i**   **mé-llè** / **mǎ**-ayà   **mò**-dôonù lêe /
then   **NC4-sword** **AG4-that** **AG4**-come\RLS **AG4**-sit\RLS   there
¶ *then **that sword** / then **it** sat down there /*

mà-gã̂i   **mé-llè** / **mè**-'ésènù cé   kúmá ù-nú /
NC4-sword **AG4-that** **AG4**-add\RLS NEG and   NC7-work
that sword / **it** could no longer work /

sɔ́bɔ̀dà káakúllè / **ù**-cíyò   cé   mò-hí̃ĩ̀   má↓=z-zá /
because that_day   **3s**-get\RLS NEG NC4-blood AG4=NC8-person
*because that day / **it** didn't get human blood /*

káakúllè **mà-gã̂i**   **mé-llè** mò-kôo /
that_day   **NC4-sword** **AG4-that** AG4-die\RLS
*that day **that sword** died /*

mísòoní mísòoní jùngònû!
story   story   close\IMP
story story close! [formulaic ending]

The loss of Tikula (sagb001)

This text was recorded in Garkuwa village near Maburya on the 9th February 2008. The speaker is Garkuwa Bawa, a member of the Akula chieftaincy and speaker of the endangered Tikula dialect (§1.2.3). This unsolicited lament is an excerpt from a much longer recording about the history of the Akula.

Ø-nápà   òtù   à-cípù   h-é /   tì-yô   á↓-má-kū̃u /
2s-know\RLS 1P.PRO NC2-Cipu AG2-COP 1P-be\RLS LOC=NC4-mountain

you know we are Acipu / we are at the mountain /

tò   ká'à,   Cì-cípù   tù-kúsù   cé,   Tì-kɔ́gɔ́   tù-'úsà /
OK   now   NC6-Cipu   AG6-remain\RLS   NEG   NC6-Hausa   AG6-take\RLS

OK now, Cicipu no longer remains, Hausa has taken /

ɔ̀-kɔ́gɔ́   à-'úsà=tù   m-úu /   ɔ̀-kɔ́gɔ́   à-'úsà=tù   m-úu /
NC2-Hausa 3P-take\RLS=1P.PRO NC5-child NC2-Hausa 3P-take\RLS=1P.PRO NC5-child

The Hausas have taken our children from us / The Hausas have taken our children from us /

à-ɓàa   cé   á↓=u-yã̀a   Cì-cípù /
3P-be   NEG   LOC=NC7-do   NC6-Cipu

they [the children] *don't speak Cicipu /*

tò,   ìndúu   íɗò   yǎ-ayà   í-kòyùkò   Cì-cípù   kúmá /
OK   here_is   2P.PRO   2P-come\RLS   2P-teach\IRR   NC6-Cipu   and

OK, here you are, you've come to teach Cicipu /

tò,   Kúngwá   yã̀a   í-ràasà=tù /   Kúngwá   yã̀a   í-gùyà /
OK   God(NC8)   do\IRR   2P-serve\IRR=1P.PRO   God(NC8)   do\IRR   2P-can\IRR

OK, may God enable you to serve us / God make you able /

m-úu   mí-ttù   kúmá   n-nà   ò-yó-nò   á↓=ɔ-kɔ́gɔ́,   á-gùyà /   Cì-cípù /
NC5-child AG5-1P.POSS and   AG5-REL 3P-be\RLS-PFV LOC=NC2-Hausa 3P-can\IRR NC6-Cipu

and our children who are amongst the Hausas, may they be able to speak / Cicipu /

tò   mí-nà   mì-yó-nò   ɔ́↓=ɔ-kɔ́gɔ́,   à-ɓàa   cé   á↓=u-kóyùkó   yé   Cí-cípù /
OK   AG5-REL AG5-be\RLS-PFV LOC=NC2-Hausa 3P-be   NEG   LOC=NC7-teach 3P.PRO NC6-Cipu

OK those that are amongst the Hausas, they aren't teaching them Cicipu /

bèllé   á-gùyà /
far_less   3P-can\IRR

so how can they learn /

ízzìnì   ìɗò   yá-ayà   ì-kábà   Cì-cípù   m-úu   mí-ttù /
permission 2P.PRO 2P-come\RLS 2P-take\RLS NC6-Cipu NC5-child AG5-1P.POSS

even if you come and take Cicipu to our children /

á-ssà   hó-ttù   à-gúyà   cé /   tékè   v-í /
NC2-grandchild AG2-1P.POSS 3P-can\RLS NEG   useless   AG8-COP

our grandchildren can't [speak Cicipu] */ it's in vain /*

à-ɓàa   cé   á↓=u-dámú-wá   m-úu   mí-ttù   ǹ   Cì-cípù /
3P-be   NEG   LOC=NC7-speak\APPL NC5-child AG5-1P.POSS with   NC6-Cipu

they don't speak to our children in Cicipu /

à-ɓàa cé á↓=u-tépú-wé m-úu ǹ Cì-cípù /
3P-be NEG LOC=NC7-talk\APPL NC5-child with NC6-Cipu
they don't talk to children in Cicipu /

vésì z-zá n-nà ù-yó-nò ɔ́↓=ɔ-kɔ́gɔ́ /
all NC8-person AG8-REL 3S-be\RLS-PFV LOC=NC2-Hausa
everyone who is amongst the Hausas /

ù-ɓàa cé á↓=u-tá'á ú-kòyùkò m-úu Cì-cípù /
3S-be NEG LOC=NC7-want 3S-teach\IRR NC5-child NC6-Cipu
he doesn't want to teach the children Cicipu /

òtù ká'à kwáakúnà tù-kôo-nò Cì-cípù tì-kɔ́tɔ̀ /
1P.PRO now the_day 1P-die\RLS-PFV NC6-Cipu AG6-finish\RLS
us now, the day we die Cicipu is finished /

tò, v-índà, ótú tì-tá'à cé 'ínì /
OK 2S-see\RLS 1P.PRO 1P-like\RLS NEG like_that
OK, you see, us we don't want that to happen /

tò ánà à-gúyà-nà cé Cì-cípù, tú-u-dàmà cé k-kɔ́gɔ́ /
OK how 3P-can\RLS-PFV NEG NC6-Cipu 1P-FUT-speak\IRR NEG NC8-Hausa
since they don't know Cicipu, we can't have secrets from the Hausa /

tò dâa hínà ǹ tì-kúdɔ̀ / ǹ tì-tá'à tí-yùwwò kè-rêe /
OK formerly TOP when 1P-meet\RLS when 1P-like\RLS 1P-turn_around\IRR NC1-language
OK in the past when we met / when we wanted to switch language /

sée tí-dàmà ǹ Cì-cípù k-kɔ́gɔ́ Ø-nápà cé /
then 1P-speak\IRR with NC6-Cipu NC8-Hausa AG8-know\RLS NEG
then we spoke in Cicipu and the Hausa didn't know /

tò, ká'à à-náhà Cì-cípù, ɔ-'ɔ́pɔ̀ tì-kɔ́gɔ́ tì-tò /
OK now 3P-leave\RLS NC6-Cipu 3P-hold\RLS NC6-Hausa AG6-one
OK, now they left Cicipu, they've taken just Hausa /

tò, vú-u-yùwwò kè-rêe tè-ené? vú-u-yùwwò kè-rêe?
OK 2S-FUT-turn_around\IRR NC1-language AG6-which 2S-FUT-turn_around\IRR NC1-language
OK, how can you switch language? will you switch language?

vú-u-yùwwò cé kè-rêe /
2S-FUT-turn_around\IRR NEG NC1-language
you won't switch language /

m-úu mí-ttù ká'à, ò-dóow<ùs>ù kà-tíi kí-ivè /
NC5-child AG5-1P.POSS now 3P-disappear<CAUS>\RLS NC1-head AG1-3P.POSS
our children now, they're causing themselves to disappear /

v-índà,   à-sì-tá'á   cì-dámí   sɔ́bɔ̀   nì   Ø-yôo   á↓=ká-dáadà
2S-see\RLS   3P-HAB-need   NC6-whispering   because   when   2S-go\RLS   LOC=NC1-court
you see, they need whispering [i.e. secret talk] *because when you go to court*

nì = t-tɔ̀ɔ   vá-avù   Cípù /   v-índà   mò-ní   mú-u-rìi /
with=NC8-neighbour   AG8-2S.POSS   Cipu(NC8)   2S-see\RLS   NC4-water   AG4-FUT-eat;3S.PRO\IRR
with your neighbour a Cipu / you see water will eat him [i.e. he's in trouble!] /

sée   vú-yùwwì   ǹ   Cì-cípù /   "ɗáanà /   tí-yã̀a   kàzá /
unless   2S-turn_around;3S.PRO\IRR   with   NC6-Cipu   my_friend   1P-do\IRR   like_this
unless you turn to him in Cicipu / "my friend / let's do this /

kàzá /   wû-utò-wò = tù" /   à-náhà-nù   t-ì   dùmbúm!
like_this   3S-go_out-APPL=1P.PRO   3P-leave\RLS-RES   AG6-PRO   completely
this will get us out" / they've completely left the language!

Tì-kɔ́gɔ́   tì-tò,   Tì-kɔ́gɔ́   t-ì   tì-nínnîi /
NC6-Hausa   AG6-one   NC6-Hausa   AG6-COP   AG6-alone
just Hausa, it's just Hausa alone /

'ú-u-dàmà   cé   z-zá   ǹ   Tì-kɔ́gɔ́ /
tì-kɔ́gɔ́   NEG   NC8-person   with   NC6-Hausa
one can't speak against people in Hausa /

v-índà   ń   dà   à-yã̂a   Cì-cípù,   n = ì-ládì   y-í,
2S-see\RLS   if   when   3P-do\RLS   NC6-Cipu   with=NC3-speed   AG3-COP
you see if they could speak Cicipu, quickly,

sée   í-hyã̀a   t-tɔ̀ɔ   vá-avù   "kàrkàtá   lêe   'ùngó   tí-sùmà!"
then   2S-say\IRR   NC8-neighbour   AG8-2S.POSS   swerve\IMP   there   rise\IMP   1P-run\IRR
then you tell your neighbour "swerve there, get up and let's run!

'ásù   wú-mpà   wù-ɗãa   ù-'úsṹ   wú-ttù /   mà-sâa   ú-sùmà /
place(NC7)   AG7-this   AG7-surpass\RLS   NC7-power   AG7-1P.POSS   ?-should   NC7-run
this place is too much for us / we should run" /

tò   Cì-cípù   tì-tò   t-í   ká'à /   Cì-cípù   tì-tò   t-í /   bâa   tí-nà /
OK   NC6-Cipu   AG6-one   AG6-COP   now   NC6-Cipu   AG6-one   AG6-COP   NEG   AG6-ART
it's just one language now / it's just one language [i.e. Hausa] */ not any other /*

tú-uwà   cé   kà-rìmái   kó↓ = m-úu   mí-ttù /
1P-feel\RLS   NEG   NC1-pleasure   AG1=NC5-child   AG5-1P.POSS
we're not happy with our children /

m-úu   mí-ttù   ǹ-yáddà-nù = tù /
NC5-child   AG5-1P.POSS   AG5-abandon\RLS-RES=1P.PRO
our children abandoned us /

ìndúu yè 'yûu á↓=u-mátá m-úu / m-úu mì-gúyà cé Cì-cípù /
here_is 3P.PRO yonder LOC=NC7-bear NC5-child NC5-child AG5-can\RLS NEG NC6-Cipu

there they are outside giving birth to [other] *children / the children don't know Cicipu /*

tò hán ámfáaní ví-llè /
OK where benefit(NC8) AG8-that

OK where's the use in that /

à-'áasà cé kà-síllùm, á-yẽ̀a kà-síllùm /
3P-forbid\RLS NEG NC1-Islam 3P-do\IRR NC1-Islam

Islam was not forbidden / let them practice Islam /

àmáa sáa yàanú ú-kòyùkò m-ɔ́ɔ Cì-cípù tú↓-ú-rée wé-evì /
but or who 3S-teach\IRR NC4-child NC6-Cipu AG6-NC7-town AG7-3S.POSS

but everyone should teach his child the Cicipu of his own town /

kà-záamání ká-ayà tù-gwédè / àmáa tú-uwà cé kà-rìmái ká↓=k-è /
NC1-future AG1-come\RLS 1P-thank\RLS but 1P-feel\RLS NEG NC1-pleasure AG1=AG1-PRO

the future has come we're grateful / but we're not happy with it /

òtù à-zá há-nà tù-kúsù-nà /
1P.PRO NC2-person AG2-REL 1P-remain\RLS-PFV

us the people that remain /

óo, tú-uwà cé kà-rìmái ká↓=k-è, sáa címmẽ̀i /
yes 1P-feel\RLS NEG NC1-pleasure AG1=AG1-PRO even little

yes, we're not happy with it, not a bit /

Appendix B – Contributors

Texts from the following speakers were recorded, annotated, and used in this study. Each of them has kindly given their permission for the recordings and annotations to be distributed without restriction.

Name	Sex	Age[1]	Dialect	Contributions
Amos Bako Timothy	M	40	Tirisino	Folktales, riddles, topic-stimulation
Bigi Magaji	M	55	Tirisino	Wordlists, Historical account
Gajere Kuyama	M	50	Tirisino	Songs
Garba Damina	M	55	Tirisino	Introduction to the Cipu king
Garba Daniel	M	18	Tirisino	Topic-stimulation, pear story
Hussaini Sale	M	13	Tirisino	Folktales, news stories
Ibrahim Danjuma	M	30	Tirisino	Wordlists, news stories
Ibrahim Yabani	M	30	Tirisino	Wordlists, elicitation
Ige Gajere	F	35	Tirisino	Songs
Ishaya Audu	M	30	Tirisino	Prayers, riddles, pear story
Mallu Kuusu	M	60	Tirisino	Greetings, discussion of chieftancy
Markus Mallam Yabani	M	30	Tirisino	Wordlists, folktales, riddles, prayers, sermon, topic-stimulation, pear story, elicitation
Mohammadu Mallam	M	20	Tirisino	Historical account, topic-stimulation
Murna Musa	F	25	Tirisino	Songs, topic-stimulation
Musa Bigi	M	25	Tirisino	Riddles
Musa Danjuma	M	30	Tirisino	Wordlists, riddles, topic-stimulation, elicitation
Musa Mallam Magaji	M	60	Tirisino	Topic-stimulation
Rahi Umar	F	28	Tirisino	Songs
Sara Kwande Ibrahim	F	25	Tirisino	Songs
Tani Kuyama	F	30	Tirisino	Songs
Tenii Magaji	F	75	Tirisino	Historical account
Timothy Danjuma	M	40	Tirisino	Wordlists
Tontono Danjuma	F	55	Tirisino	Songs
Yabani Galadima	M	75	Tirisino	Historical account, pear story

---

1   All ages are approximate, and are relative to the commencement of my fieldwork in 2006.

Dogo Timbidii	M	40	Tikula	Wordlists
Garkuwa Maburya Bawa	M	80	Tikula	Historical account
Malli Idanbanya	M	75	Tikula	Historical account
Mohammed Musa	M	35	Tikula	Historical account, topic-elicitation
Saani Ɗantanyi	M	55	Tikula	Historical account, conversation
Ayuba Sani	M	30	Tidipo	Folktales, elicitation
Ibrahim Wakiso	M	60	Tidipo	Wordlists
Ishiaku Ibrahim	M	20	Tidipo	Elicitation
Tabitha Markus	F	25	Tidipo	Pear story
Umar Magaji	M	60	Tizoriyo	Wordlists

Appendix C – Person and gender agreement on person markers referring to humans, animals, and inanimates

In §8.4 it was stated that all person markers can agree anaphorically in either person or gender, irrespective of the animacy of the controller referent. The examples from the text corpus below demonstrate this, although it should be remembered that some combinations are rarer than others. The 18 combinations have different values for agreement feature (x2 – person, gender), animacy level (x 3 – human, animal, inanimate), and morphosyntactic position (x 3 – subject prefix, post-verbal object, and 'elsewhere' – see chapter 7).

Subject, human, person:

(1)   **ù**-yúwò
    3s-fall\RLS
    he *fell*

[tapf001.005.020]

Subject, human, gender:

(2)   Ø-lóokàcí   n-nà   **kò**-'úngò-nò
    NC8-time   AG8-REL   **AG1**-rise\RLS-PFV
    when ***he*** [**kà-bárá** 'old man, NC1'] *grew up*

[sayb001.085]

Subject, animal, person:

(3)   **ù**-yô   ǹ   Ø-ámfáaní sòsái
    3s-be\RLS with NC8-benefit very
    it [**kà-ràkúmí** 'camel, NC1'] *has many benefits*

[tats004.001.012]

Subject, animal, gender:

(4)   kà-ràkúmí ǹ   **kò**-dôohò,   **kò**-sì -tôonò   cé
    NC1-camel   when **AG1**-disappear\RLS **AG1**-HAB-come_home\RLS NEG
    the camel when ***it*** *goes missing,* ***it*** *doesn't come home*

[tats001.002.033]

Subject, inanimate, person:

(5)   **ú**-làngwà z-zà   Ø-hùuhú
    3s-spoil\IRR NC8-person NC8-lung
    it [**tá-abà** 'tobacco, NC6'] *spoils a person their lungs*

[tats006.002.036]

Subject, inanimate, gender:

(6) n = ù-lénjí **hè**-hé'wè
and=NC7-day **AG2**-dry\RLS

*in the afternoon **they*** [**ò-gínó** 'groundnuts, NC2'] *dry*

[eaim006.1444]

Post-verbal object, human, person:

(7) éllè ùhy**îivì**
é-llè ù-hyâa = **vì**
3s-that 3s-say\RLS=**3s.PRO**

*that one said to **him** [the man]*

[saat001.003.014]

Post-verbal object, human, gender:

(8) à-sì-hyá̋a **k-è**, m̀-pándà, Sáahìyà
3P-HAB-say **AG1-PRO** 1s-forget\RLS [name]

*they call **her*** [**kà-mǎ̋yầ** 'older sibling, NC1'], *I forget, Sahiya*

[saat002.008.003]

Post-verbal object, animal, person:

(9) sáa ccîi**vì** hàráawà
sáa c-câa = **vì** Ø-hàráawà
even 2s-give\RLS=**3s.PRO** NC8-fodder

*even though you give **it** [**kà-ràkúmí** 'camel, NC1']* fodder

[tats001.002.034]

Post-verbal object, animal, gender:

(10) kádà á-hùnà **v-ì**
PROH 3P-kill\IRR **AG8-PRO**

*they are not to kill **it** [**kwáarò** 'creature, NC8']*

[eaim009.033]

Post-verbal object, inanimate, person:

(11) àsìhy**îivì** kòpógó
à-sì-hyá̋a = **vì** kò-pógó
3s-HAB-say=**3s.PRO** NC1-cochlospermum_tinctorium

*they call **it** 'kopogo'* [a kind of tree]

[saat001.003.014]

Post-verbal object, inanimate, gender:

(12) n = ò-'ísò   **m-è**
      when=3P-cook\RLS **AG4-PRO**
      *when they drink **it*** [**mò-yóo** 'beer, NC4']

Elsewhere, human, person:

(13) **èré**   m-úu
      3P.PRO NC5-child
      ***they** were children*

Elsewhere, human, gender:

(14) m-ɔ́ɔ   mé-llè   dá'à **m-í**   m-è   mà-yáɗùkwà-nà
      NC4-child AG4-that PART **AG4-PRO** AG4-COP AG4-spread_out\RLS-PFV
      *that child it's **her** that multiplied*

Elsewhere, animal, person:

(15) à-sì-kábá   ì-tángì   n = **èví**
      3P-HAB-take   NC3-thing with=**3s.PRO**
      *they carry things with **it*** [**kà-ràkúmí** 'camel, NC1']

Elsewhere, animal, gender:

(16) ǹ-sì-yá̃a   ì-túmà   ǹ   **k-é**
      1s-HAB-do   NC3-farm with **NC1-PRO**
      *I farm with **it*** [**kà-ràkúmí** 'camel, NC1']

Elsewhere, inanimate, person AND Elsewhere, inanimate, gender:

(17) kó-osì   **éví**   pâa, kù-cíyè   **kw-í**   kw-ì   pâa
      NC1-eye   **3s.PRO**   here NC9-hand **AG9-PRO** AG9-COP   here
      *the eye **it** is here, the hand **it** is here*

462

References

Agamalafiya, Apollos. 2007. A preliminary sketch of the Tsureshe language. Unpublished manuscript.

Agamalafiya, Apollos. 2008. Development of the Reshe language. Paper presented at the 1ˢᵗ Workshop on West Kainji languages, Kontagora.

Aikhenvald, Alexandra Y. 2000. *Classifiers: a typology of noun categorization devices*. Oxford: Oxford University Press.

Aikhenvald, Alexandra Y. n.d. Gender and noun class in one language: Paumarí in typological perspective. Electronic ms. [Available online: http://www.aikhenvaldlinguistics.com/images/pdfs/publications/Paumari%20NCL.pdf]

Aissen, Judith. 2003. Differential coding, partial blocking, and bidirectional OT. In Pawel Nowak & Corey Yoquelet (eds.), *Berkeley Linguistics Society*, 29, 1-16. Berkeley Linguistics Society.

Anderson, A., S. C. Garrod & A. J. Sanford. 1983. The accessibility of pronominal antecedents as a function of episode shifts in narrative text. *Quarterly Journal of Experimental Psychology* 35, 427-440.

Anderson, Stephen C. 1980a. The noun class system of Amo. In Larry M. Hyman (ed.), *Noun classes in the Grassfields Bantu borderland* (Southern California Occasional Papers in Linguistics 8), 155-178. Los Angeles: University of Southern California.

Anderson, Stephen C. 1980b. The noun class system of Nygembɔɔn-Bamileke. In Larry Hyman (ed.), *Noun classes in the Grassfields Bantu borderland* (Southern California Occasional Papers in Linguistics 8), 38-56. Los Angeles: University of Southern California.

Anderson, Stephen R. 1976. On the notion of subject in ergative languages. In Charles N. Li (ed.), *Subject and topic*, 1-24. New York: Academic Press.

Ariel, Mira. 1990. *Accessing noun-phrase antecedents*. London: Routledge.

Asher, Nicholas. 2004. Discourse topic. *Theoretical linguistics* 30, 163-201.

Austin, Peter K. 2004. Clitics in Sasak, eastern Indonesia. Electronic ms. [Available online: https://eprints.soas.ac.uk/59/]

Austin, Peter K. & Joan Bresnan. 1996. Non-configurationality in Australian languages. *Natural Language and Linguistic Theory* 14(2), 215-268.

Baker, Mark C. 1996. *The Polysynthesis Parameter.* Oxford: Oxford University Press.

Bardovi-Harlig, Kathleen. 2002. Pragmatics and second language acquisition. In Kaplan, Robert B. (ed.) *The Oxford handbook of applied linguistics*, 182-192. Oxford, Oxford University Press.

Barlow, Michael. 1992. *A situated theory of agreement* (Outstanding Dissertations in Linguistics). New York: Garland.

Bearth, Thomas. 1999. The contribution of African linguistics towards a general theory of focus. Update and critical review. *Journal of African Languages and Linguistics* 20(2), 121-156.

Bearth, Thomas. 2003. Syntax. In Derek Nurse & Gérard Philippson (eds.), *The Bantu languages*, 121-142. London: Routledge.

Berman, Ruth A. & Dan I. Slobin. 1994. *Relating events in narrative: a crosslinguistic developmental study.* Hillsdale, NJ: Lawrence Erlbaum.

Bertho, J. 1952. Aperçu d'ensemble sur les dialectes de l'ouest de Nigéria. *Bulletin de l'IFAN (Institut Français de l'Afrique Noire)* 14, 259-271.

Biber, Douglas. 1984. Pragmatic roles in Central Somali narrative discourse. *Studies in African Linguistics* 15, 1-26.

Bickel, Balthasar. 2003. Referential density in discourse and syntactic typology. *Language* 79(4), 708-736.

Bickel, Balthasar, Bernard Comrie & Martin Haspelmath. 2004. *The Leipzig glossing rules. Conventions for interlinear morpheme by morpheme glosses.* Leipzig: Max-Planck Institute for Evolutionary Anthropology.

Bickel, Balthasar & Johanna Nichols. 2007. Inflectional morphology. In Timothy Shopen (ed.), *Language typology and syntactic description*, 2nd edn, 3 vols, vol. III, 169-240. Cambridge: Cambridge University Press.

Binnick, Robert I. 1991. *Time and the verb: a guide to tense and aspect*. Oxford: Oxford University Press.

Bird, Steven & Gary Simons. 2003. Seven dimensions of portability for language documentation and description. *Language* 79(3), 557-582.

Bleek, W. H. I. 1862. *A comparative grammar of South African languages. Part I: phonology.* London: Trübner.

Bleek, W. H. I. 1869. *A comparative grammar of South African languages. Part II: the*

concord; section I: the noun. London: Trübner.

Blench, Roger M. 1998. The status of the languages of Central Nigeria. In Mathias Brenzinger (ed.), *Endangered languages in Africa*, 187-206. Köln: Köppe Verlag.

Blench, Roger M. n.d. a. The West Kainji languages of northwestern Nigeria: Comparative wordlist. Electronic ms. [Available online:
http://www.rogerblench.info/Language%20data/Niger-Congo/Benue-Congo/West%20Kainji/West%20Kainji-subclassification.pdf]

Blench, Roger M. n.d. b. The Benue-Congo languages: a proposed internal classification. Electronic ms. [Available online:
http://www.rogerblench.info/Language%20data/Niger-Congo/Benue-Congo/General/Benue-Congo%20classification%20latest.pdf]

Blevins, Juliette. 2004. *Evolutionary Phonology*. Cambridge: Cambridge University Press.

Boettger, E. & V. Boettger. 1967. Tsureshe Grammar. Unpublished ms.

Bolinger, Dwight. 1977. *Meaning and form*. London: Longman.

Bolinger, Dwight. 1979. Pronouns in discourse. In Talmy Givón (ed.), *Discourse and syntax* (Syntax and semantics 12), 289-309. New York: Academic Press.

Bond, Oliver. 2009. Infrastructure requires a foundation: a base for the canons of negation. Paper presented at Creating Infrastructure for Canonical Typology, University of Surrey.

Bosch, Peter. 1983. *Agreement and anaphora* (Cognitive Science Series). London: Academic Press.

Bow, Catherine, Baden Hughes & Steven Bird. 2003. Towards a general model of interlinear text. *Proceedings of EMELD Workshop 2003: Digitizing and annotating texts and field recordings*, East Lansing, Michigan [Available online:
http://www.linguistlist.org/emeld/workshop/2003/bowbadenbird-paper.pdf].

Bresnan, Joan. 2001. *Lexical-Functional syntax*. Oxford: Blackwell.

Bresnan, Joan & Jonni M. Kanerva. 1989. Locative inversion in Chicheŵa: a case study of factorization in grammar. *Linguistic Inquiry* 20(1), 1-50.

Bresnan, Joan & Sam A. Mchombo. 1987. Topic, pronoun, and agreement in Chicheŵa. *Language* 63(4), 741-782.

Bresnan, Joan & Lioba Moshi. 1990. Object asymmetries in comparative Bantu syntax.

Linguistic Inquiry 21(2), 147-185.

Bybee, Joan. 1985. *Morphology: a study of the relation between meaning and form* (Typological Studies in Language 9). Amsterdam: John Benjamins.

Bybee, Joan, Revere Perkins & William Pagliuca. 1994. *The evolution of grammar: tense, aspect, and modality in the languages of the world*. Chicago: University of Chicago Press.

Cahill, Michael. 2000. Tonal associative morphemes in Optimality Theory. *OSU Working Papers in Linguistics* 53, 31-70.

CAPRO Research Office. 1995. *Kingdoms at war*. Jos: CAPRO Media.

Chafe, Wallace (ed.) 1980. *The Pear stories: cognitive, cultural, and linguistic aspects of narrative production*. Norwood, NJ: Ablex.

Chafe, Wallace. 1990. Some things that narratives tell us about the mind. In Bruce K. Britton & Anthony D. Pellegrini (eds.), *Narrative thought and narrative language*, 79-98. Hillsdale, NJ: Lawrence Erlbaum.

Chafe, Wallace. 1994. *Discourse, consciousness, and time*. Chicago: University of Chicago Press.

Chambers, Mary. 2009. *Which way is up? Motion verbs and paths of motion in Kubokota, an Austronesian language of the Solomon Islands*. PhD dissertation, School of Oriental and African Studies.

Chapman, Shirley & Desmond C. Derbyshire. 1991. Paumarí. In Desmond C. Derbyshire & Geoffrey K. Pullum (eds.), *Handbook of Amazonian languages*, vol. 3, 161-352. Berlin: Mouton de Gruyter.

Childs, G. Tucker. 1994. African ideophones. In Leanne Hinton, Johanna Nichols & John J. Ohala (eds.), *Sound symbolism*, 178-204. Cambridge: Cambridge University Press.

Chomsky, Noam. 1981. *Lectures on Government and Binding*. Dordrecht: Foris.

Chung, Sandra. 1998. *The design of agreement: evidence from Chamorro*. Chicago: University of Chicago Press.

Cinque, Guglielmo. 2005. Deriving Greenberg's Universal 20 and its exceptions. *Linguistic Inquiry* 36(3), 315-332.

Clements, G. N. 2000. Phonology. In Bernd Heine & Derek Nurse (eds.), *African languages: an introduction*, 123-160. Cambridge: Cambridge University Press.

Clements, G. N. & Samuel Jay Keyser. 1983. *CV Phonology: a generative theory of the syllable*. Cambridge, Mass.: MIT Press.

Comrie, Bernard. 1981. *Language universals and linguistic typology*. Oxford: Blackwell.

Comrie, Bernard. 1988. Topics, grammaticalized topics, and subjects. *Proceedings of the Fourteenth Annual Meeting of the Berkeley Linguistics Society*, 265-279. Berkeley, California.

Comrie, Bernard. 2003. When agreement gets trigger-happy. *Transactions of the Philological Society* 101(2), 313-337.

Connell, Bruce & D. Robert Ladd. 1990. Aspects of pitch realization in Yoruba. *Phonology* 7, 1-30.

Contini-Morava, Ellen. 2002. (What) do noun class markers mean? In Wallis Reid, Ricardo Otheguy & Nancy Stern (eds.), *Signal, meaning, and message*, 3-64. Amsterdam: John Benjamins.

Cook, Vivian J. & Mark Newson. 1996. *Chomsky's Universal Grammar: an introduction*, 2nd edn. Oxford: Blackwell.

Corbett, Greville G. 1991. *Gender*. Cambridge: Cambridge University Press.

Corbett, Greville G. 2000. *Number*. Cambridge: Cambridge University Press.

Corbett, Greville G. 2003a. Agreement: terms and boundaries. In William Griffin (ed.), *The Role of Agreement in Natural Language* (Texas Linguistics Forum 53), 109-122.

Corbett, Greville G. 2003b. Agreement: the range of the phenomenon and the principles of the Surrey Database of Agreement. *Transactions of the Philological Society* 101(2), 155-202.

Corbett, Greville G. 2005a. The canonical approach in typology. In Zygmunt Frajzyngier, Adam Hodges & David S. Rood (eds.), *Linguistic diversity and language theories*, 25-49. Amsterdam: John Benjamins.

Corbett, Greville G. 2005b. Number of genders. In Matthew Dryer, Martin Haspelmath, David Gil & Bernard Comrie (eds.), *The World Atlas of Language Structures*, 126-129. Oxford: Oxford University Press.

Corbett, Greville G. 2006. *Agreement*. Cambridge: Cambridge University Press.

Corbett, Greville G. 2007. Gender and noun classes. In Timothy Shopen (ed.),

Language typology and syntactic description, 2nd edn, 3 vols, vol. III, 241-279. Cambridge: Cambridge University Press.

Creissels, Denis. 2000. Typology. In Bernd Heine & Derek Nurse (eds.), *African languages: an introduction*, 231-258. Cambridge: Cambridge University Press.

Croft, William. 1991. *Syntactic categories and grammatical relations: the cognitive organization of information*. Chicago: University of Chicago Press.

Croft, William & D. Alan Cruse. 2004. *Cognitive Linguistics*. Cambridge: Cambridge University Press.

Crothers, John. 1978. Typology and universals of vowel systems. In Joseph H. Greenberg (ed.), *Universals of human language*, vol. 2, 93-152. Stanford: Stanford University Press.

Crozier, David H. 1984. *A study in the discourse grammar of Cishingini*. PhD dissertation, University of Ibadan.

Crozier, David H. & Roger M. Blench. 1992. *An Index of Nigerian Languages*, 2nd edn. Dallas: SIL.

Cruttenden, Alan. 1997. *Intonation*, 2nd edn. Cambridge: Cambridge University Press.

Culy, Christopher. 2000. An incorporated topic marker in Takelma. In Miriam Butt & Tracy Holloway King (eds.), *Proceedings of the LFG 00 Conference, University of California, Berkeley*, 63-81. Stanford: CSLI Publications.

Dahl, Östen & Viveka Velupillai. 2005. The Perfect. In Martin Haspelmath, Matthew Dryer, David Gil & Bernard Comrie (eds.), *The World Atlas of Language Structures*, 271-272. Oxford: Oxford University Press.

Dalby, David. 1964. Provisional identification of languages in the Polyglotta Africana. *Sierra Leone Language Review* 3, 83-90.

Dalrymple, Mary & Irina Nikolaeva. 2005. Nonsubject agreement and discourse roles. In *Oxford Working Papers in Linguistics, Philology, and Phonetics* 10, 71-92. [Available online: http://www.ling-phil.ox.ac.uk/download/OWP2005.pdf]

Dancygier, Barbara & Eve Sweetser. 2005. *Mental spaces in grammar: conditional constructions*. Cambridge: Cambridge University Press.

Daniel, Michael & Edith Moravcsik. 2005. The associative plural. In Martin Haspelmath, Matthew Dryer, David Gil & Bernard Comrie (eds.), *World Atlas of Language Structures*, 150-153. Oxford: Oxford University Press.

Danziger, Eve. 1993. *Manual for the Space Stimuli Kit 1.2*. Nijmegen: Max Planck Institute for Psycholinguistics.

Davis, Stuart. 1999. On the representation of initial geminates. *Phonology* 16(1), 93-104.

De Wolf, Paul Polydoor. 1971. *The noun class system of Proto-Benue-Congo*. The Hague: Mouton de Gruyter.

Demuth, Katherine. 1990. Locatives, impersonals and expletives in Sesotho. *The Linguistic Review*, vol. 7, 233-249. Dordrecht: Foris.

Demuth, Katherine & Mark Johnson. 1989. Interaction between discourse functions and agreement in Setawana. *Journal of African Languages and Linguistics* 11, 21-35.

Denny, J. Peter & Chet A. Creider. 1986. The semantics of noun classes in Proto Bantu. In Colette Craig (ed.), *Noun classes and categorization*, 217–239. Amsterdam: John Benjamins.

Dettweiler, Stephen H. 2000. Vowel harmony and neutral vowels in C'Lela. *Journal of West African Languages* 28(1), 3-18.

Dettweiler, Stephen H. n.d. C'Lela grammar sketch. Electronic ms.

Dettweiler, Stephen H. & Sonia G. Dettweiler. 2002. Sociolinguistic survey (level one) of the Kamuku language cluster. Electronic ms. http://www.sil.org/silesr/2003/silesr2003-003.pdf

Dingemanse, Mark. 2006. The semantics of Bantu noun classification. Electronic ms. http://www.mpi.nl/Members/MarkDingemanse/Publications

Dixon, Robert M. W. & Alexandra Y. Aikhenvald. 2000. Introduction. In Robert M. W. Dixon & Alexandra Y. Aikhenvald (eds.), *Changing valency: case studies in transitivity*, 1-29. Cambridge: Cambridge University Press.

Dixon, Robert M. W. & Alexandra Y. Aikhenvald. 2002. *Word: a cross-linguistic typology*. Cambridge: Cambridge University Press.

Dooley, Robert A. 2007. Explorations in discourse topicality. *SIL Electronic Working Papers* 2007-010. http://www.sil.org/silewp/2007/silewp2007-010.pdf

Dooley, Robert A. & Stephen H. Levinsohn. 2001. *Analyzing discourse: a manual of basic concepts*. Dallas: SIL International.

Dryer, Matthew. 2005. Position of interrogative phrases in content questions. In Martin Haspelmath, Matthew Dryer, David Gil & Bernard Comrie (eds.), *World Atlas of*

Language Structures, 378-381.

Du Bois, John W. 1987. The discourse basis of ergativity. *Language* 63(4), 805-855.

Duchan, Judith F., Gail A. Bruder & Lynne E. Hewitt. 1995. *Deixis in narrative: a cognitive science perspective*. Hillsdale, NJ: Lawrence Erlbaum.

Duff, E. C. 1972. Gazetteer of the Kontagora Province. In A. H. M. Kirk-Greene (ed.) *Gazetteers of the Northern provinces of Nigeria*, 1-72. London: Frank Cass. [Originally published in 1920 by Waterlow (London)].

Evans, Nicholas. 1999. Why argument affixes in polysynthetic languages are not pronouns: evidence from Bininj Gun-wok. *Sprachtypologie und Universalienforschung* 52, 255-281.

Evans, Nicholas. 2002. The true status of grammatical object affixes: evidence from Bininj Gun-wok. In Nicholas Evans & Hans-Jürgen Sasse (eds.), *Problems of polysynthesis*, 15-50. Berlin: Akademie Verlag.

Falk, Yehuda N. 2006. On the representation of case and agreement. In Miriam Butt & Tracy Holloway King (eds.), *The Proceedings of the LFG 06 Conference*. Stanford: CSLI Publications.

Farrar, Scott & D. Terence Langendoen. 2003. A linguistic ontology for the Semantic Web. *GLOT International* 7(3), 97-100.

Fauconnier, Gilles. 1994. *Mental spaces: aspects of meaning construction in natural languages*, 2nd edn. Cambridge: Cambridge University Press.

Fauconnier, Gilles. 1997. *Mappings in thought and language*. Cambridge: Cambridge University Press.

Ferguson, Charles A. & Michael Barlow. 1988. Introduction. In Michael Barlow & Charles A. Ferguson (eds.), *Agreement in natural language*, 1-22. Stanford: CSLI.

Foley, William A. 2003. Genre, register and language documentation in literate and preliterate communities. In Peter K. Austin (ed.), *Language documentation and description*, vol. 1, 85-98. London: Hans Rausing Endangered Languages Project.

Fortescue, Michael D. 1984. *West Greenlandic*. Beckenham, Kent: Croom Helm.

Fox, Barbara A. 1987. Anaphora in popular written English narratives. In Russell S. Tomlin (ed.), *Coherence and grounding in discourse*, 157-174. Amsterdam: John Benjamins.

Francis, Hartwell S., Michelle L. Gregory & Laura A. Michaelis. 1999. Are lexical

subjects deviant? *Papers from the 35th regional meeting of the Chicago Linguistics Society*, vol 1., 85-97. Chicago: Chicago Linguistics Society.

Frawley, William. 1992. *Linguistic semantics*. Hillsdale, NJ: Lawrence Erlbaum.

Gerhardt, Ludwig. 1989. Kainji and Platoid. In John Bendor-Samuel (ed.), *The Niger-Congo languages*, 359-376. Lanham, Md.: University Press of America.

Gernsbacher, Morton Ann. 1984. Cognitive responses to linguistic topic changes in discourse. In Morton Ann Gernsbacher (Chair), *Cross-disciplinary approaches to language processing. Proceedings of the sixth annual conference of the Cognitive Science Society*, 83-88.

Gernsbacher, Morton Ann. 1985. Surface information loss in comprehension. *Cognitive Psychology* 17, 324-363.

Givón, Talmy. 1976. Topic, pronoun, and grammatical agreement. In Charles N. Li (ed.), *Subject and topic*, 149-189. New York: Academic Press.

Givón, Talmy. 1983. Topic continuity in discourse: an introduction. In Talmy Givón (ed.), *Topic continuity in discourse: a quantitative cross-language study* (Typological Studies in Language 3), 5-41. Amsterdam: John Benjamins.

Gordon, Raymond G., Jr. 2005. *Ethnologue: languages of the world*, 15th edn. Dallas: SIL International.

Greenberg, Joseph H. 1955. *Studies in African linguistic classification*. New Haven: Compass Publishing.

Greenberg, Joseph H. 1963. Some universals of grammar with particular reference to the order of meaningful elements. In Joseph H. Greenberg (ed.), *Universals of language*, 73-113. Cambridge, MA: MIT Press.

Grice, H. Paul. 1975. Logic and conversation. In Peter Cole & Jerry L. Morgan (eds.), *Speech acts* (Syntax and semantics 3), 41-58. New York: Academic Press.

Grimes, Joseph E. 1975. *The thread of discourse*. The Hague: Mouton de Gruyter.

Grimes, Joseph E. 1978. *Papers on discourse*. Dallas: Summer Institute of Linguistics.

Grinevald, Colette & Frank Seifart. 2004. Noun classes in African and Amazonian languages: towards a comparison. *Linguistic Typology* 8(2), 243-285.

Gunn, H. D. & F. P. Conant. 1960. *Peoples of the middle Niger region, Northern Nigeria*. London: International African Institute.

Haiman, John. 1978. Conditionals are topics. *Language* 54, 564-589.

Halliday, M. A. K. & Ruqaiya Hasan. 1976. *Cohesion in English*. London: Longman.

Haspelmath, Martin. 1997. *Indefinite pronouns*. Oxford: Oxford University Press.

Haspelmath, Martin. 2002. *Understanding morphology*. London: Arnold.

Haspelmath, Martin. 2004. Coordinating constructions: an overview. In Martin Haspelmath (ed.), *Coordinating constructions*, 3-39. Amsterdam: John Benjamins.

Hayes, Bruce. 1989. Compensatory lengthening in moraic phonology. *Linguistic Inquiry* 20(2), 253-306.

Heath, Jeffrey. 1975. Some functional relationships in grammar. *Language* 51, 89-104.

Heath, Jeffrey. 1983. Referential tracking in Nunggubuyu (Australia). In John Haiman & Pamela Munro (eds.), *Switch reference and universal grammar: proceedings of a symposium on switch reference and universal grammar, Winnipeg, May 1981* (Typological Studies in Language 2), 129-149. Amsterdam: John Benjamins.

Hedinger, Robert. 1980. The noun classes of Akɔɔ̄sē. In Larry M. Hyman (ed.), *Noun classes in the Grassfields Bantu borderland* (Southern California Occasional Papers in Linguistics 8), 1-26. Los Angeles: University of Southern California.

Heine, Bernd. 1982. African noun class systems. In Hansjakob Seiler & Christian Lehmann (eds.), *Apprehension: Das sprachliche Erfassen von Gegenständen, Part I: Bereich und Ordnung der Phänomene*, 189–216. Tübingen: Gunter Narr.

Hendrikse, A. P. 1997. Systemic polysemy in the Southern Bantu noun class system. In Hubert Cuyckens & Britta E. Zawada (eds.), *Polysemy in Cognitive Linguistics*, 185-212. Amsterdam: John Benjamins.

Hewings, Ann & Martin Hewings. 2005. *Grammar and context: an advanced resource book*. London: Routledge.

Himmelmann, Nikolaus P. 1998. Documentary and descriptive linguistics. *Linguistics* 36, 161-195.

Hinds, John. 1979. Organizational patterns in discourse. In Talmy Givón (ed.), *Discourse and syntax* (Syntax and semantics 12), 135-158. New York: Academic Press.

Hinds, John. 1984. Topic maintenance in Japanese narratives and Japanese conversational interaction. *Discourse Processes* 7, 465-482.

Hockett, Charles F. 1958. *A course in modern linguistics*. New York: Macmillan.

Hoffmann, Carl. 1963. The noun-class system of Central Kambari. *Journal of African*

Languages 2(2), 160-169.

Hoffmann, Carl. 1965. A wordlist of Central Kambari. *Journal of West African Languages* 2(2), 7-31.

Hoffmann, Carl. 1967. An outline of the Dakarkari noun class system and the relation between prefix and suffix noun-class systems. In G. Manessy (ed.), *La classification nominale dans les langues négro-africaines*, 237-259. Paris: CNRS.

Hoffmann, Carl. 1972. A note on vowel contraction in Central Kambari. *Research Notes (Department of Linguistics and Nigerian Languages, Ibadan)* 5(2/3), 73-91.

Hopper, Paul J. & Sandra A. Thompson. 1980. Transitivity in grammar and discourse. *Language* 56, 251-299.

Hopper, Paul J. & Sandra A. Thompson. 1984. The discourse basis for lexical categories in universal grammar. *Language* 60, 703-752.

Hume, Elizabeth, Jennifer Muller & Aone van Engelenhoven. 1997. Nonmoraic geminates in Leti. *Phonology* 14(3), 371-402.

Hyman, Larry M. 1980. *Noun classes in the Grassfields Bantu borderland* (Southern California Occasional Papers in Linguistics 8). Los Angeles: University of South California.

Jaggar, Philip J. 1983. Some dimensions of topic-NP continuity in Hausa narrative. In Talmy Givón (ed.), *Topic continuity in discourse: a quantitative cross-language study* (Typological Studies in Language 3), 365-424. Amsterdam: John Benjamins.

Jaggar, Philip J. 1985. *Factors governing the morphological coding of referents in Hausa narrative discourse*. PhD dissertation, University of California, Los Angeles.

Jaggar, Philip J. 1988. Discourse-deployability and indefinite NP-marking in Hausa: a demonstration of the universal 'Categoriality Hypothesis'. In Graham Furniss & Philip J. Jaggar (eds.), *Studies in Hausa language and linguistics in honour of F. W. Parsons*, 45-61. London: Kegan Paul.

Jaggar, Philip J. 2001. *Hausa*. Amsterdam: John Benjamins.

Jelinek, Eloise. 1984. Empty categories, case, and configurationality. *Natural Language and Linguistic Theory* 2, 39-76.

Ji, Shaojun. 2002. Identifying episode transitions. *Journal of Pragmatics* 34, 1257-1271.

Johnston, Harry H. 1919. *A comparative study of the Bantu and Semi-Bantu languages*, vol. 1. Oxford: Clarendon Press.

Kameyama, Megumi. 1985. *Zero anaphora: the case of Japanese*. PhD dissertation, Stanford University.

Katamba, Francis. 2003. Bantu nominal morphology. In Derek Nurse & Gérard Philippson (eds.), *The Bantu languages*, 103-120. London: Routledge.

Kayne, Richard S. 1994. *The antisymmetry of syntax* (Linguistic Inquiry Monographs 25). Cambridge, MA: MIT Press.

Keenan, Edward L. & Bernard Comrie. 1977. NP accessibility and universal grammar. *Linguistic Inquiry* 8, 63-100.

Khorounjaia, Ekaterina & Liliana Tolchinsky. 2004. Discursive constraints on the lexical realization of arguments in Spanish. In Ruth A. Berman (ed.), *Language development across childhood and adolescence*, 83-109. Amsterdam: John Benjamins.

King, Tracy Holloway. 1997. Focus domains and information structure. In Miriam Butt & Tracy Holloway King (eds.), *The Proceedings of the LFG 97 Conference*. Stanford: CSLI Publications.

Kiparsky, Paul. 1982. From cyclic phonology to lexical phonology. In Harry van der Hulst & Norval Smith (eds.), *The structure of phonological representations (Part I)*, 131-176. Dordrecht: Foris.

Kisseberth, Charles & David Odden. 2003. Tone. In Derek Nurse & Gérard Philippson (eds.), *The Bantu languages*, 59-70. London: Routledge.

Klamer, Marian. 2008. Differential marking of intransitive subjects in Kambera (Austronesian). In Helen de Hoop & Peter de Swart (eds.), *Differential subject marking*, 281-299. Dordrecht: Springer.

Koelle, Sigismund Wilhelm. 1963. *Polyglotta Africana: photomechanic reprint of the original edition, Church Missionary Society, London 1854*. Freetown: Fourah Bay College.

König, Ekkehard & Peter Siemund (with Stephan Töpper). 2005. Intensifiers and reflexive pronouns. In Martin Haspelmath, Matthew Dryer, David Gil & Bernard Comrie (eds.), *World Atlas of Language Structures*, 194-197. Oxford: Oxford University Press.

Kreiman, J. 1982. Perception of sentence and paragraph boundaries in natural conversation. *Journal of Phonetics* 10, 163-175.

Kroeger, Paul R. 2004. *Analyzing syntax: a lexical-functional approach*. Cambridge: Cambridge University Press.

Kuno, Susumu. 1976. Subject, theme, and the speaker's empathy – a reexamination of relativization phenomena. In Charles N. Li (ed.), *Subject and topic*, 417-444. New York: Academic Press.

Ladefoged, Peter & Ian Maddieson. 1996. *The sounds of the world's languages*. Oxford: Blackwell.

Lakoff, George. 1987. *Women, fire and dangerous things: what categories reveal about the mind*. Chicago: University of Chicago Press.

Lambrecht, Knud. 1994. *Information structure and sentence form*. Cambridge: Cambridge University Press.

Langacker, Ronald W. 1987. *Foundations of Cognitive Grammar: theoretical prerequisites*, vol. 1. Stanford: Stanford University Press.

Langacker, Ronald W. 1991. *Foundations of Cognitive Grammar: descriptive application*, vol. 2. Stanford: Stanford University Press.

Langacker, Ronald W. 2000. *Grammar and conceptualization*. Berlin: Mouton de Gruyter.

Langacker, Ronald W. 2001. Discourse in Cognitive Grammar. *Cognitive Linguistics* 12(2), 143-188.

Laoust, E. 1928. *Cours de berbère marocain.* Paris: Hachette.

Lehiste, Ilse. 1979. Sentence boundaries and paragraph boundaries – perceptual evidence. In P. R. Clyne, W. F. Hanks & C. L. Hofbauer (eds.), *The elements: a parasession on linguistic units and levels*, 99-109. Chicago: Chicago Linguistics Society.

Levinson, Stephen C. 2004. Deixis. In Laurence R. Horn & Gregory L. Ward (eds.), *The handbook of pragmatics*, 97-121. Oxford: Blackwell.

Li, Charles N. & Sandra A. Thompson. 1979. Third-person pronouns and zero-anaphora in Chinese discourse. In Talmy Givón (ed.), *Discourse and syntax* (Syntax and semantics 12), 311-335. New York: Academic Press.

Lichtenberk, Frantisek. 2000. Inclusory pronominals. *Oceanic Linguistics* 39(1), 1-32.

Longacre, Robert E. 1996. *The grammar of discourse*, 2nd edn. New York: Plenum.

Lovelace, David. 1989. Computer printout on Tsuvaɗi.

Lüpke, Friederike. 2006. Small is beautiful – the contribution of small field-based corpora to different linguistic disciplines. In Peter K. Austin (ed.), *Language Documentation and Description*, vol. 3, 75-105. London: Hans Rausing Endangered Languages Project.

Lyons, Christopher. 1999. *Definiteness*. Cambridge: Cambridge University Press.

McCarthy, Michael. 2001. *Issues in applied linguistics*. Cambridge: Cambridge University Press.

MacDonell, James. 2007. A phonological and grammatical sketch of the Pongu language. Electronic ms.

McGill, Stuart. 2004. *Focus and activation in Paasaal: the particle rɛ*. MA dissertation, University of Reading.

McGill, Stuart. 2007. The Cicipu noun class system. *Journal of West African Languages* 34(2), 51-90.

McGill, Stuart. 2008. The Damakawa language. Electronic ms. [Available online: http://www.cicipu.org/papers/damakawa.pdf]

McGill, Stuart. n.d. The classification of Cicipu. Electronic ms. [Available online: http://www.cicipu.org/papers/2009-05-06_classification_cicipu_working.pdf]

McKinney, Norris P. 1978. Participant identification in Kaje narrative. In Joseph E. Grimes (ed.), *Papers on discourse*, 179-189. Dallas: Summer Institute of Linguistics.

Maho, Jouni F. 1999. *A comparative study of Bantu noun classes*. Göteborg: Acta Universitatis Gothoburgensis.

Maibaum, Anita. 1978. Participants in Jirel narrative. In Joseph E. Grimes (ed.), *Papers on discourse*, 203-207. Dallas: Summer Institute of Linguistics.

Marten, Lutz. 2005. Tone cases in Herero. Electronic ms. SOAS/University of Namibia.

Mathews, A. B. 1926a. Anthropological notes on the Busawa. National Archives, Kaduna: SNP 17/8/k.2068.

Mathews, A. B. 1926b. Historical and anthropological report on the Achifawa. National Archives, Kaduna, SNP 17/8/k.2068.

Mathews, A. B. 1950. The Kisra legend. *African Studies* 9, 144-147.

Mchombo, Sam A. 2004. *The syntax of Chichewa*. Cambridge: Cambridge University Press.

Meek, Charles Kingsley. 1925. *The northern tribes of Nigeria*, 2 vols. London: Oxford University Press.

Meinhof, Carl. 1932. *Introduction to the phonology of the Bantu languages*. Revised, enlarged and translated from German (Grundriss einer Lautlehre der Bantu-sprachen, 2nd edition, 1910) by N. J. van Warmelo in collaboration with the author and Dr. Alice Werner. Berlin: Dietrich Reimer.

Mel'čuk, Igor. 1993. Agreement, government, congruence. *Lingvisticæ Investigationes* 17(2), 307-373.

Mithun, Marianne. 1992. Is basic word order universal? In Doris Payne (ed.), *Pragmatics of word order flexibility*, 15-61. Amsterdam: John Benjamins.

Mithun, Marianne. 2003. Pronouns and agreement: the information status of pronominal affixes. *Transactions of the Philological Society* 101(2), 235-278.

Morimoto, Yukiko. 2002. Prominence mismatches and differential object marking in Bantu. In Miriam Butt & Tracy Holloway King (eds.), *Proceedings of the LFG 02 Conference, National Technical University of Athens*, 292-314. Stanford: CSLI Publications.

Morimoto, Yukiko. 2008. From topic to subject marking: implications for a typology of subject marking. In Helen de Hoop & Peter de Swart (eds.), *Differential subject marking*, 199-222. Dordrecht: Springer.

Mous, Maarten. 2003. Loss of linguistic diversity in Africa. In Mark Janse & Sijmen Tol (eds.), *Language death and language maintenance*, 157-170. Amsterdam: John Benjamins.

Mufwene, Salikoko S. 1980. Bantu class prefixes: inflectional or derivational? *Chicago Linguistic Society*, vol. 16, 246–258.

Munro, Pamela. 1997. The Garifuna gender system. In Jane H. Hill, P. J. Mistry & Lyle Campbell (eds.), *The life of language: papers in linguistics in honour of William Bright* (Trends in Linguistics 108), 443-462. Mouton de Gruyter.

Myhill, John. 2001. Typology and discourse analysis. In Deborah Schiffrin, Deborah Tannen & Heidi E. Hamilton (eds.), *The handbook of discourse analysis*, 161-174. Oxford: Blackwell.

Nathan, David & Peter K. Austin. 2004. Reconceiving metadata: language documentation through thick and thin. In Peter K. Austin (ed.), *Language Documentation and Description*, vol. 2, 179-187. London: Hans Rausing Endangered Languages Project.

Newman, John F. 1978. Participant orientation in Longuda folk tales. In Joseph E. Grimes (ed.), *Papers on discourse*, 91-104. Dallas: Summer Institute of Linguistics.

Newman, Paul. 1990. *Nominal and verbal plurality in Chadic* (Publications in African Languages and Linguistics 12). Dordrecht: Foris.

Newman, Paul. 2000. *The Hausa language: an encyclopedic reference grammar.* New Haven: Yale University Press.

Newmeyer, Frederick J. 1998. *Language form and language function*. Cambridge, MA: MIT Press.

Onishi, Masayuki. 1994. *A grammar of Motuna (Bougainville, Papua New Guinea)*. PhD dissertation, Australian National University.

Palmer, F. R. 2001. Mood and modality, 2nd edn. Cambridge: Cambridge University Press.

Payne, Thomas E. 1997. *Describing morphosyntax*. Cambridge: Cambridge University Press.

Perlmutter, David. 1995. Phonological quantity and multiple association. In John A. Goldsmith (ed.), *The handbook of phonological theory*, 307-317. Oxford: Blackwell.

Perrin, Mona. 1978. Who's who in Mambila folk stories. In Joseph E. Grimes (ed.), *Papers on discourse*, 105-118. Dallas: Summer Institute of Linguistics.

Polinsky, Maria & Bernard Comrie. 1999. Agreement in Tsez. *Folia Linguistica* 33(2), 109-130.

Postal, Paul M. 1970. On so-called pronouns in English. In Roderick A. Jacobs & Peter S. Rosenbaum (eds.), *Readings in English transformational grammar*, 56-82. Waltham, Mass.: Ginn.

Prazan, Ceslaus. 1977. *The Dukkawa of northwest Nigeria*. Pittsburgh: Duquesne University Press.

Quirk, Randolph, Sidney Greenbaum, Geoffrey Leech, and Jan Svartvik. 1985. *A*

comprehensive grammar of the English language. Longman: London.

Rialland, Annie. 2007. Question prosody: an African perspective. In C. Gussenhoven & T. Riad (eds.), *Tones and tunes: studies in word and sentence prosody*, 35-62. Berlin: Mouton de Gruyter.

Richardson, I. 1967. Linguistic evolution and Bantu noun class systems. In G. Manessy (ed.), *La classification nominale dans les langues négro-africaines*, 373-390. Paris: CNRS.

Rijkhoff, J. 2002. *The Noun Phrase* (Oxford Studies in Typology and Linguistic Theory). Oxford: Oxford University Press.

Rosch, Eleanor. 1978. Principles of categorization. In Eleanor Rosch & Barbara B. Lloyd (eds.), *Cognition and categorization*, 27-48. Hillsdale, NJ: Lawrence Erlbaum.

Rowlands, E. C. 1962. Notes on some class languages of northern Nigeria. *African Language Studies* 3, 71-83.

Sagna, Serge. 2008. *Formal and semantic properties of the Gújjolaay Eegimaa nominal classification system*. PhD dissertation, School of Oriental and African Studies.

Schadeberg, Thilo C. 2003. Derivation. In Derek Nurse & Gérard Philippson (eds.), *The Bantu languages*, 71-89. London: Routledge.

Schultze-Berndt, Eva. 2006. Linguistic annotation. In Jost Gippert, Nikolaus P. Himmelmann & Ulrike Mosel (eds.), *Essentials of language documentation*, 213-251. Berlin: Mouton de Gruyter.

Schwartz, Jean-Luc, Louis-Jean Boë, Nathalie Vallée & Christian Abry. 1997. Major trends in vowel system inventories. *Journal of Phonetics* 25, 233-253.

Schwarz-Friesel, Monika. 2007. Indirect anaphora in text: a cognitive account. In Monika Schwarz-Friesel, Manfred Consten & Mareile Knees (eds.), *Anaphors in text*, 3-20. Amsterdam: John Benjamins.

Seifart, Frank. 2005. *The structure and use of shape-based noun classes in Miraña (North West Amazon)*. PhD dissertation, Radboud University, Nijmegen.

Selvik, Kari-Anne. 1997. When a dance resembles a tree: a polysemy analysis of three Setswana noun classes. In Hubert Cuyckens & Britta E. Zawada (eds.), *Polysemy in cognitive linguistics*, 161-184. Amsterdam: John Benjamins.

Siewierska, Anna. 1984. *The passive: a comparative linguistic analysis*. London: Croom

Helm.

Siewierska, Anna. 1999. From anaphoric pronoun to grammatical agreement marker: why objects don't make it. In Greville G. Corbett (ed.), *Agreement (Special Issue of Folia Linguistica 33/2)*, 225-251. Berlin: Mouton de Gruyter.

Siewierska, Anna. 2004. *Person*. Cambridge: Cambridge University Press.

Siewierska, Anna. 2005. Passive constructions. In Martin Haspelmath, Matthew Dryer, David Gil & Bernard Comrie (eds.), *World Atlas of Language Structures*, 434-437. Oxford: Oxford University Press.

Silverstein, Michael. 1976. Hierarchy of features and ergativity. In Robert M. W. Dixon (ed.), *Grammatical categories in Australian languages*, 112-171. Canberra: Australian Institute of Aboriginal Studies.

Skopeteas, S., I. Fiedler, S. Hellmuth, A. Schwarz, R. Stoel, G. Fanselow, C. Féry & M. Krifka. 2006. *Questionnaire on Information Structure (QUIS)* (Interdisciplinary Studies on Information Structure (ISIS) 4). Potsdam: Potsdam University.

Smith, Rebecca Dow. 2007. *The noun class system of ut̲-Ma'in, a West Kainji language of Nigeria*. MA dissertation, University of North Dakota.

Snider, Keith & James Roberts. 2004. SIL comparative African wordlist. *Journal of West African Languages* 31(2), 73-122.

Souag, Lameen. 2007. The typology of number borrowing in Berber. Paper presented at the Fifth Cambridge Postgraduate Conference in Linguistics. [Available online: http://www.ling.cam.ac.uk/camling/Manuscripts/CamLing2007_Souag.pdf]

Spencer, Katharine. 2008. Kamuku languages survey report. Paper presented at the 1st Workshop on West Kainji languages, Kontagora.

Stallcup, Kenneth L. 1980. Noun classes in Esimbi. In Larry M. Hyman (ed.), *Noun classes in the Grassfields Bantu borderland* (Southern California Occasional Papers in Linguistics 8), 139-153. Los Angeles: University of Southern California.

Stark, Janie P. 2003. *The Tsikimba-Hausa-English dictionary*. Salka: Kambari Language Project.

Stark, Janie P. 2004a. *The Cishingini-Hausa-English dictionary*. Salka: Kambari Language Project.

Stark, Janie P. 2004b. *The Tsishingini-Hausa-English dictionary*. Salka: Kambari

Language Project.

Stassen, Leon. 1997. *Intransitive predication*. Oxford: Oxford University Press.

Steele, Susan. 1978. Word order variation: a typological study. In Joseph H. Greenberg, Charles A. Ferguson & Edith A. Moravcsik (eds.), *Universals of human language*, vol. 4: Syntax, 585-623. Stanford: Stanford University Press.

Stevens Jr., Phillips. 1975. The Kisra legend and the distortion of historical tradition. *The Journal of African History* 16(2), 185-200.

Stewart, Marjorie H. 1980. The Kisra legend as oral history. *International Journal of African Historical Studies* 13(1), 51-70.

Strawson, P. F. 1971. Identifying reference and truth values. In D. D. Steinberg & L. A. Jakobovits (eds.), *Semantics*, 75-95. Cambridge: Cambridge University Press.

Temple, O. 1922. *Notes on the tribes, provinces, emirates, and states of the northern provinces of Nigeria*. Compiled from official reports by O. Temple; edited by C. L. Temple. Lagos: CMS Bookshop.

Thompson, Sandra A. & Robert E. Longacre. 1985. Adverbial clauses. In Timothy Shopen (ed.), *Language typology and syntactic description*, vol. 2: Complex constructions, 171-234. Cambridge: Cambridge University Press.

Tomlin, Russell S. 1995. Focal attention, voice, and word order: an experimental, cross-linguistic study. In Pamela Downing & Michael Noonan (eds.), *Word order in discourse*, 517-554. Amsterdam: John Benjamins.

Tomlin, Russell S., Linda B. Forrest, Ming Ming Pu & Myung Hee Kim. 1997. Discourse semantics. In Teun A. van Dijk (ed.), *Discourse studies: a multidisciplinary introduction, vol. 1: Discourse as structure and process*, 63-111. London: Sage.

Uyechi, Linda. 1991. The functional structure of the Navajo third person alternation. *Papers from the 27th regional meeting of the Chicago Linguistics Society*, 434-446. Chicago: Chicago Linguistics Society.

van der Berg, Helma. 1999. Gender and person agreement in Akusha Dargi. In Greville G. Corbett (ed.), *Agreement (Special Issue of Folia Linguistica 33/2)*, 153-168. Berlin: Mouton de Gruyter.

van Dijk, Teun A. 1981. Episodes as units of discourse analysis. In Deborah Tannen (ed.), *Analyzing discourse: text and talk*, 177-195. Georgetown: Georgetown

University Press.

Van Valin, Robert D., Jr. 2005. *Exploring the syntax-semantics interface*. Cambridge: Cambridge University Press.

Vendler, Zeno. 1957. Verbs and times. *Philosophical Review* 66, 143-60, revised in Zeno Vendler (1967), *Linguistics in Philosophy*, 97–121. Cornell: Cornell University Press.

Visser, Marianna W. 1989. The syntax of the infinitive in Xhosa. *South African Journal of African Languages* 9(4), 154-185.

Voeltz, F. K. Erhard & Christa Kilian-Hatz. 2001. Introduction. In *Ideophones*, F. K. Erhard Voeltz & Christa Kilian-Hatz (eds.), 1–8. Amsterdam: John Benjamins.

Wälchli, Bernhard. 2007. Descriptive typology, or; the typologist's extended toolkit. Electronic ms. [Available online:
http://ling.uni-konstanz.de/pages/home/a20_11/waelchli/waelchli-desctyp.pdf]

Wald, Benji. 1979. The development of the Swahili object marker: a study of the interaction of syntax and discourse. In Talmy Givón (ed.), *Discourse and Syntax* (Syntax and Semantics 12), 505-524. New York: Academic Press.

Watters, John R. 1980. The Ejagam noun class system: Ekoid Bantu revisited. In Larry Hyman (ed.), *Noun classes in the Grassfields Bantu borderland* (Southern California Occasional Papers in Linguistics 8), 99-137. Los Angeles: University of Southern California.

Watters, John R. 2000. Syntax. In Bernd Heine & Derek Nurse (eds.), *African languages: an introduction*, 194-230. Cambridge: Cambridge University Press.

Watters, John R. 2003. Grassfields Bantu. In Derek Nurse & Gérard Philippson (eds.), *The Bantu languages*, 225-256. London: Routledge.

Welmers, William E. 1973. *African language structures*. Berkeley: University of California Press.

Westermann, Diedrich Hermann. & M. A. Bryan. 1952. *Languages of West Africa: Part II*. London: Oxford University Press.

Williamson, Kay. 1989. Niger-Congo overview. In John Bendor-Samuel (ed.), *The Niger-Congo languages*, 1-45. Lanham: University Press of America.

Williamson, Kay & Roger M. Blench. 2000. Niger-Congo. In Bernd Heine & Derek Nurse (eds.), *African languages: an introduction*, 11-42. Cambridge: Cambridge

University Press.

Wolff, H. Ekkehard. 1993. *Referenzgrammatik des Hausa* (Hamburger Beiträge zur Afrikanistik 2). Münster: LIT.

Woodbury, Tony. 2003. Defining documentary linguistics. In Peter K. Austin (ed.), *Language Documentation and Description*, vol. 1, 35-51. London: Hans Rausing Endangered Languages Project.

Yip, Moira. 2002. *Tone*. Cambridge: Cambridge University Press.

Zubin, David & Klaus-Michael Köpcke. 1986. Gender and folk taxonomy: the indexical relation between grammatical and lexical categorization. In Colette Craig (ed.), Noun classes and categorization, 139-180. Amsterdam: John Benjamins.