Gender and person markers in Cicipu discourse: ‘non-topical’ and ‘topical’ anaphoric agreement

Stuart McGill

Proceedings of Conference on
Language Documentation & Linguistic Theory 2

Edited by Peter K. Austin, Oliver Bond, Monik Charette, David Nathan & Peter Sells

13-14 November 2009 School of Oriental and African Studies, University of London

Hans Rausing Endangered Languages Project
Department of Linguistics
School of Oriental and African Studies
Thornhaugh Street, Russell Square
London WC1H 0XG
United Kingdom

Department of Linguistics:
Tel: +44-20-7898-4640
Fax: +44-20-7898-4679
linguistics@soas.ac.uk
http://www.soas.ac.uk/academics/departments/linguistics

Hans Rausing Endangered Languages Project:
Tel: +44-20-7898-4578
Fax: +44-20-7898-4349
elap@soas.ac.uk
http://www.hrelp.org

© 2009 Stuart McGill

No part of this publication may be reproduced, stored in a retrieval system, or transmitted, on any form or by any means, electronic, mechanical, photocopying, recording or otherwise, without the prior permission of the author(s) of that part of the publication, except as permitted by UK copyright law.

ISBN: 978-0-7286-0392-9

This publication can be cited as:
or:
Gender and person markers in Cicipu discourse; ‘non-topical’ and ‘topical’ anaphoric agreement

STUART MCGILL
School of Oriental and African Studies

1. INTRODUCTION

Over the last few decades linguists have made considerable progress in their understanding of agreement markers/ incorporated pronouns. One of the most important theoretical contributions has been Bresnan and Mchombo’s (1987) distinction between grammatical agreement and anaphoric agreement, and the association of the latter with topical referents. This association between anaphoric agreement and topicality has largely gone unchallenged, and is frequently repeated in the typological literature. The aim of this paper is to present data from the agreement system of Cicipu (Kainji, Nigeria) which shows that topicality is not necessarily derivative of the distinction between grammatical and anaphoric agreement, but is instead an independent dimension along which anaphoric agreement markers can be classified.

The paper is structured as follows: section 2 questions the assumption that anaphoric agreement markers must always have topical referents; section 3 then introduces the two paradigms of agreement in Cicipu, gender and person; while section 4 considers their distribution across a corpus of approximately 12,000 clauses. We find that both gender and person markers may take part in grammatical and anaphoric subject agreement, with the person markers associated with referents that are discourse-topical and/or animate.

2. AGREEMENT MARKERS AND TOPICALITY

For languages with optional agreement, the presence of agreement morphology is often correlated with the topicality of the controller referent. A well-known case is object agreement in the Bantu languages. Bresnan and Mchombo (1987) describe the asymmetries between (obligatory) subject and (optional) object agreement in the Bantu language Chichewa, and provide an analysis framed in Lexical-Functional Grammar (LFG). They make the distinction between GRAMMATICAL AGREEMENT, where the agreement marker is a redundant affix expressing the gender of a co-occurring subject NP (e.g. (1a) for a Cicipu example), and ANAPHORIC AGREEMENT, in which case the agreement marker is an

---

1 This paper is based on my PhD thesis, submitted to the School of Oriental and African Studies.
2 The controller is the constituent (usually an NP) ‘controlling’ the agreement morphology displayed on the agreement ‘target’ (Corbett 2006).
incorporated pronominal with full argument status (e.g. 1b). In other words, the agreement marker in (1b) is the subject.³

(1) (a) kà-rákùmí kò-dôohò
NC1-camel AG1-disappear\RLS 'the camel disappeared'

(b) kò-dôohò
AG1-disappear\RLS 'it [e.g. kà-rákùmí 'camel, NC1'] disappeared’ [Cicipu]

According to Bresnan and Mchombo’s analysis, subject agreement in Chicheŵa and other Bantu languages is ambiguous between grammatical and anaphoric agreement, whereas object agreement is pure anaphoric agreement. This explains a number of asymmetries between Chicheŵa subject NPs and ‘object’ NPs (which are really clausal adjuncts).

The relevance of the distinction between grammatical and anaphoric agreement for this paper is the association that the latter is generally held to have with topicality. Actually there is nothing in the theoretical introduction to Bresnan and Mchombo (1987: 743-752) that implies anaphoric agreement markers must have topical antecedents – the logic of their argument is the other way round, i.e. that floating topics in Chicheŵa can only be licensed by an incorporated pronoun in the clause. However later on in the paper it becomes clear that they also assume the reverse condition holds, in other words that incorporated pronouns must always anaphorically link to a topic. This entailment has often been repeated by both LFG researchers and typologists, but is rarely questioned.⁴ Siewierska extends the condition to cover ‘dependent person markers’, which include unstressed pronouns (as in English) as well as incorporated pronouns (2004: 67):

...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.

However if ‘topicality’ is to have any independent explanatory power, 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:

³ The abbreviations used in this paper are 3P = third-person plural, 3S = third-person singular, AG = agreement, ART = article, FUT = future, HAB = habitual, IRR = irrealis, LOC = locative, NC = noun class, NEG = negative, PFV = perfective, PLAC = pluractional, POSS = possessive, PRO = pronoun, REL = relativiser, RLS = realis, VENT = ventive.

⁴ For example, Corbett (2003: 189) states that ‘if the marker is an incorporated pronoun, it will be linked via anaphoric agreement to the ‘topic’ of the sentence’.
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.5

The highlighted unstressed pronouns take this form not because the discourse is ‘about’ the offending object, but simply because it has just been mentioned. Siewierska’s ‘highly cognitively accessible’ and ‘topical’ are not the same thing, a fairly trivial point but one often missed in the literature on agreement. While English may use a single coding technique (unstressed pronouns) for both topical referents and non-topical highly-accessible referents, this is not true in general, as the Cicipu data in section 4 will show.

Culy (2000) takes the opposite view to Siewierska, and suggests that cross-linguistically, incorporated pronouns may index sentence topics, discourse topics, or non-topical referents. Thus he regards topicality as an independent dimension along which agreement markers can be classified, rather than one derivative of the grammatical vs. anaphoric agreement distinction.

Culy’s typology is a little suspect since it relies on work by scholars using very different methodologies to his own. LFG researchers often depend on introspection/elicitation for their insights, whereas Culy worked with texts from the extinct language Takelma. The former typically follow Lambrecht’s (1994) theory of sentence topic, which is concerned with the pragmatic presuppositions of speakers at a snapshot during a discourse, and how these affect the structure of a single clause or sentence. Culy’s conception of topic, however, is more in line with discourse topic (e.g. Chafe 1994, Dooley 2007), which is concerned with how topics are established and maintained throughout a paragraph or even a whole text. Another difficulty is that he compares markers across different languages – this is not surprising, since according to Siewierska (2004: 67):

…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.

Cicipu does have two series of dependent person markers for the same function, and this makes it a good test case to determine whether dependent markers can in fact vary with respect to their discourse function. As we will see in section 4, the Cicipu data supports Culy’s basic contention that topicality should be regarded as an independent dimension for the classification of agreement markers. In the case

5 Taken from a blog entry at http://sonmislocuras.com/page/2/.
of Cicipu it is discourse topic rather than sentence topic which allows us to distinguish between the two paradigms of Cicipu agreement markers.

3. CICIPU GENDER AND PERSON PARADIGMS

Like all Benue-Congo languages, nouns in Cicipu can be divided into classes according to the prefixes found on the singular and plural. For example, kà-ràkùmí /à-ràkùmí ‘camel/camels’ belongs to the KA class in the singular and the A class in the plural. Similarly mà-wáa /à-wáa ‘dog/dogs’ belongs to the MA class in the singular and the N class in the plural. Noun class or GENDER agreement is found on a great many agreement targets inside and outside the noun phrase. Example (3) shows agreement on a demonstrative modifier, the relativiser, and verbs.

(3) (a) kà-llù ké-llè kà-nà kà-yâa-nà
   NC1-hunger AG1-that AG1-REL AG1-arrive<RLS-PFV
   ‘that hunger which arrived’

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

Where Cicipu differs from the more well-known Bantu languages is that for a number of these targets, gender agreement is in competition with a separate paradigm, that of PERSON agreement. Example (4a) shows AG1 gender agreement on the verb, while (4b) shows gender-neutral 3PS person agreement.

(4) (a) kò-dôohò (b) ù-dôohò
   AG1-disappear<RLS AG1-disappear<RLS
   ‘it disappeared’ ‘it disappeared’

Since gender is not marked for first and second persons, the alternation is confined to the third person. The alternation also exists for plural nouns. (5a) shows gender agreement while (5b) shows (gender-neutral) person agreement:

(5) (a) bo-dôoh <ii> ò (b) ò-dôoh <ii> ò
   AG2-disappear<PLAC><RLS AG3-disappear<PLAC><RLS
   ‘they disappeared’ ‘they disappeared’

Although this paper concentrates on subject agreement prefixes, the same alternation is also found on pronouns, the article, the interrogative quantifier bo-o ‘which’, demonstrative pronouns and modifiers, and the copula. Thus there is a series of competing paradigms across different agreement targets.

---

6 There is no object agreement in Cicipu.
It remains to set out the paradigms themselves. There are nine syntactic noun classes in Cicipu, the agreement prefixes of which are as follows:7

Table 6
Cicipu gender agreement prefixes

<table>
<thead>
<tr>
<th>Class</th>
<th>Prefix</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>ka-</td>
<td>ha-</td>
<td>yi-</td>
<td>ma-</td>
<td>mi-</td>
<td>ti-</td>
<td>wu-</td>
<td>Ø-/vi-</td>
<td>ku-</td>
</tr>
</tbody>
</table>

The person system is the common six-way distinction involving person and number:

Table 7
Cicipu person agreement prefixes

<table>
<thead>
<tr>
<th></th>
<th>Singular</th>
<th>Plural</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Ø-</td>
<td>ti-</td>
</tr>
<tr>
<td>2</td>
<td>Ø-/v-</td>
<td>i-</td>
</tr>
<tr>
<td>3</td>
<td>u-</td>
<td>a-</td>
</tr>
</tbody>
</table>

In principle, every Cicipu noun has the possibility of triggering either gender agreement or person agreement. In the following section I turn to the factors influencing which paradigm of prefixes will in fact occur.

4. FACTORS RELEVANT FOR THE GENDER/PERSON ALTERNATION

4.1. Other Kainji languages
The gender/person alternation found in Cicipu is also present in other Kainji languages, as well as the related Plotoïd branch of Niger-Congo. Hoffman drew attention to the alternation in Central Kambari (Kainji), calling it ‘a very peculiar feature’ (1963: 168). For him DEFINITENESS was the key to the alternation; indefinite subject NPs trigger gender agreement, whereas definite subjects trigger person agreement.

According to Crozier (1984: 215-222), although Hoffmann’s analysis works most of the time, definite subjects in Central Kambari 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. Crozier’s account is closer to what we are about to find for Cicipu, but it does not provide the whole story.

---

7 Cicipu has ten noun class prefixes, but two of these share the same agreements, and so we can say that there are ten morphological noun classes but only nine syntactic noun classes. Singular and plural noun classes pair up in different ways to make seven paired genders and nine single class genders. See McGill (2007) for details.
4.2. Cicipu agreement prerequisites

For cases of variation in agreement, Corbett (2006) uses the term CONDITIONS to refer to the factors that determine which of the available agreement patterns occurs. He is careful to distinguish between conditions and PREREQUISITES:

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. (Corbett 2006: 183)

The factors involved in Cicipu are complex and it is helpful to separate them out into prerequisites and conditions, with the latter only coming into play in environments where there is a choice between gender and person agreement.

The most obvious prerequisite concerns the subject NP. If there is an explicit subject NP (i.e. grammatical rather than anaphoric agreement), then usually there is no choice, and the features indexed on the verb are a straightforward copy of the features on the NP. In the grammatically correct examples (6a, 7a) the features on the subject and verb match – gender in (6a) and person in (7a). Where the features do not match, as in (6b) and (7b), the sentences are ungrammatical. This is the case even if the antecedent of èvì in (7b) is an NC1 noun such as kà-ràkùmì ‘camel’.

(6) (a) kà-ràkùmì kò-dôohò NC1-camel AG1-disappear\RLS ‘the camel disappeared’
    (b) *kà-ràkùmì ù-dôohò NC1-camel 3S-disappear\RLS

(7) (a) èvì ù-dôohò 3S.PRO 3S-disappear\RLS ‘it disappeared’
    (b) *èvì kò-dôohò 3S.PRO AG1-disappear\RLS

Subjects lacking a specification for gender (such as personal pronouns, personal demonstrative pronouns, and the names of people) may only co-occur with person subject agreement, not gender subject agreement. As with (7b), (8b) would be ungrammatical even if the antecedent of èmpè was an NC1 noun.

(8) (a) èmpè ù-dôohò 3S-this 3S-disappear\RLS ‘this one disappeared’
    (b) *èmpè kò-dôohò 3S-this AG1-disappear\RLS

There is one scenario in which person agreement is possible with a gender-marked subject NP – if the subject NP is NC8, and the verb stem is vowel-initial – but only if the referent is sufficiently animate or discourse-topical. Example (9) illustrates the choice for an inanimate referent; (9b) is only possible if the Koran is a discourse topic.
In summary, person agreement is possible if one of the following three prerequisites is met: (i) there is no subject NP, (ii) there is a subject NP but it is unspecified for gender, (iii) there is an NC8 subject NP and the verb stem is vowel-initial. Otherwise, only gender agreement can occur.

4.3. Cicipu agreement conditions

Agreement conditions apply when more than one possibility for agreement exists. In order to determine the relevant conditions for Cicipu I examined the distribution of the two different types of agreement over a multi-genre interlinear corpus of 12,000 clauses. This consists partly of texts collected through a project funded by the Hans Rausing Endangered Languages Documentation Programme, and partly of texts collected specifically to investigate this alternation. A qualitative analysis of subject agreement in these texts reveals that animacy and discourse topicality are two independent conditions influencing the alternation between gender and person agreement, in those environments where both can occur.

4.3.1. Inanimate referents

Referents which are both inanimate and non-topical rarely function as the subjects of clauses. On occasion, however, anaphoric chains of inanimate non-topical subject referents do occur, in which case they are without exception coded using gender agreement. Inanimate topical referents may trigger person agreement, although they do so rarely. Examples of inanimate referents triggering person agreement include beer, gin, tobacco, the Bible, the Koran, palm-wine, fish trapping, trees, a lake, a fence, and clothing. In each case the referent is a discourse topic.

4.3.2. Animals

Non-topical animal referents are without fail encoded using gender agreement. In contrast, 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. Thus animals clearly reflect the importance of topicality. The following two examples, from separate texts, form a "minimal pair". The first example (10) is from a text about monkeys, the second (11) 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. Observe the contrasting means by which the dog is referred to in each of the examples:

(10) ñ w-îndà mè-wà a mò-wísô / ëlì m-ëñwà kà-sàu

and 3s-see/RLS NC4-dog AG4-bark/RLS even AG4-see/RLS NC9-smell

kwéevì mëvùngènò mò/kânyì nûlài ñ /
kwéevì mëvùngènò mò/kânyì nûlài ñ /

AG9-3s.Poss AG4-rise/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’

 tô, n=ù-sùmà / ù-nàhà ù-gùtò-nà mà-kàmà /
OK and=3s-run/RLS 3s-leave/RLS NC7-return-VENT hurriedly

‘OK, when it runs / it won’t come back in a hurry’

(11) tô, lòòkàcí wú-nà mà-wà a ì më-dìnò mà-dìyà /

OK time 3s- NC4-dog and AG4-follow/RLS NC4-hare

ART

n=ù-hùnà më-dìyà ì=/kà-dìbà /
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 cè ³ / ù-sì-ràa cè /
3s-HAB-eat NEG there 3s-HAB-eat NEG

sée ù-dòmò ì /
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 /’

In (10) the non-topical dog is indexed by subject gender agreement (mA-), while the topical monkey is indexed with person agreement (in a variety of morphosyntactic positions). In (11) the topical dog is at first marked with gender agreement, but then it graduates to person agreement. This is an instance of a general property of discourse-topical referents, i.e. they undergo a coding progression from LEXICAL NP > GENDER AGREEMENT > PERSON AGREEMENT. The hare, on the other hand, is only marked with gender agreement. It is important to note that the propositions encoded in several of the individual clauses in (10) are about the dog, and so the dog qualifies as a sentence topic.

8 Bold references are to the dog, underlined references are to the monkey and the hare.
9 See Lambrecht (1994: 149) for analysis of a similar example.
(10) is not of *intrinsic interest* (Dooley 2007) to the speaker, instead its role is to tell us something about the monkey. It is only by considering the role of the participants in the wider discourse (i.e. discourse topicality) that the differences in morphosyntactic coding seen in (10) and (11) can be accounted for.

The third intonation unit in (10) (beginning *n=ù-súmá*) provides an interesting parallel to the second intonation unit in (11) (beginning *n=ù-húnà*). In both cases the previous clause encoded the dog as subject by means of anaphoric gender agreement. In (11) the subject remains the same going into the second intonation unit, seamlessly transitioning from gender agreement to person agreement. However in (10), the subject of the verb *ù-súmá* 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 (10) the monkey and not the dog is topical, the speaker can use the 3PS marker *u-* without fear of ambiguity. Since animal referents can only be encoded using person agreement if they are discourse-topical, *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. In (10) 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 in Cicipu are concerned, allows them to graduate from anaphoric gender agreement to anaphoric person agreement.

4.3.3. Humans

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. Briefly, the answer is that human referents in the corpus are usually marked with person agreement, but gender agreement is likely to persist longer (i.e. across intonation units) for non-topical referents.
5. CONCLUSION

We have seen that there are two paradigms of agreement morphology in Cicipu, one inflected for gender and the other for person, and in section 4 I set out the factors relevant to this alternation with regard to subject agreement. Whenever there is an explicit subject NP (i.e. grammatical agreement) the features on the verb must match the features on the subject NP. When there is no subject NP (i.e. anaphoric agreement) either gender or person agreement may occur.

Anaphoric chains can progress from LEXICAL NP > GENDER AGREEMENT > PERSON AGREEMENT. This happens more quickly if the referent is high in topicality and/or animacy, and the effect of topicality is particularly clear with animals.

Anaphoric agreement markers occur with both topics and non-topics in Cicipu. Therefore topicality should be considered an independently-varying dimension with respect to the classification of agreement markers, not one determined by the grammatical vs. anaphoric agreement distinction, and thus it is possible to have distinct dependent markers encoding different information statuses (contra Siewierska 2004).

Finally, neither sentence topic nor referential density are sufficient to account for the examples in section 4. Instead we need a theory of discourse topic that also takes into account the intrinsic interest which the referent holds for the speaker. Unlike generativists, most typologists hold that linguistics should be grounded in empirical data. Consequently typological claims about the nature of agreement and incorporated pronouns should take into account the possibility of grammatical structure being influenced by discourse topic as well as sentence topic.

REFERENCES


10 The exception to this being the co-occurrence of an NC8 controller and a vowel-initial verb, as mentioned in section 4.2.


