

## Interaction between Discourse Functions and Agreement in Setawana

KATHERINE DEMUTH and MARK JOHNSON

### 1. INTRODUCTION

In this paper we examine the interaction between discourse function and agreement in Setawana, a northern dialect of Setswana.<sup>1</sup> In particular, we explore the grammatical and anaphoric agreement properties of Setawana subject and object markers from the Lexical-Functional Grammar (LFG) perspective of Bresnan and Mchombo (1987) (henceforth B&M). The paper focusses on subject and object asymmetries in Setawana and their treatment within the framework developed by B&M. We discuss a number of analyses of the Setawana data, including the hypothesis that both the subject marker (SM) and object marker (OM) are incorporated pronominals, and that the subject marker (SM) is not ambiguously a grammatical agreement marker as it is in Chicheŵa.

In section two we discuss a variety of pro-drop, word order and wh-extraction phenomena in Setawana and locate the Setawana subject marker (SM) and object marker (OM) within the typology of agreement-marker/discourse function interactions. In the third section we present additional evidence for our analysis by examining Setawana cleft constructions, the role of question words in passives and clefts, and the behavior of adverbials, tone and present indicative tense forms. We conclude with a discussion of the problematic issues that this analysis of Setawana raises for B&M and for linguistic theory in general.

Our analysis and the work of B&M is conducted within the framework of Lexical Functional Grammar as developed in Bresnan 1982. Because this framework may be unfamiliar to the reader we summarize some of

\* This material is based upon work supported by the National Science Foundation under Grant No. BNS-8609642. An earlier version of this paper was read at the 17th Annual African Linguistics Conference, Bloomington, Indiana in April of 1986. We thank Resiame Moetse and Eagilwe Segosebe, speakers of the Setawana dialect of Setswana (northern Botswana - Lake Ngami and Maun region), for supplying the data for this paper. Comparative data from other Setswana dialects was supplied by Mpho Molomo (Sekgatla) and Olga Mangope (Sefurutshe). We thank also Joan Bresnan and the Stanford Bantu Working Group for comments and suggestions. The final interpretations, however, remain our own.

its distinctive components below, and refer the reader for further detail to the articles in Bresnan 1982, particularly Kaplan and Bresnan 1982.

Lexical Function Grammar (LFG) describes an utterance in terms of several distinct representations. The constituent structure (c-structure) represents the surface constituency relationships that hold between the words of an utterance, whereas the functional structure (f-structure) represents the function-argument structure of an utterance. The c-structure is a conventional tree structure with words as its terminals, and the f-structure is a recursive attribute-value structure, which resembles the hierarchical "frame" structures used in knowledge-representation systems. The c- and f-structures that are associated with the utterance *Mary chases John* are given in Figure 1.

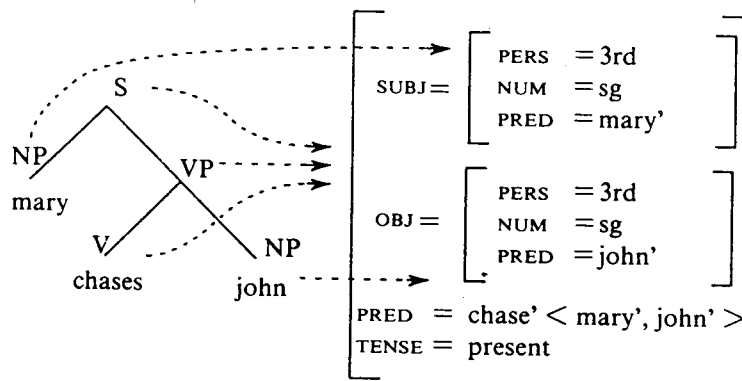


Figure 1: Sample c- and f-structures.

C-structures are specified using context-free phrase-structure rules and lexical entries, and the associated f-structures are specified by annotations that appear on these rules and lexical entries. The f-structures associated with a given c-structure are not determined via a derivational relationship, but by constraints simultaneously imposed by the syntactic, morphological and lexical structure of the language. If there is no f-structure which simultaneously satisfies all of the constraints, then there is no f-structure for the utterance and that utterance is ill-formed.

This can be made formally precise, as was done in Kaplan and Bresnan 1982 and Johnson 1988, but the relationship between c-structure and f-structure can also be understood at a less formal and more intuitive level as well. We view each c-structure node as associated with an element of the f-structure of the utterance as a whole; the arrows in Figure 1 show the association in this example. This association is formally a mathematical function; i.e. while each c-structure node is associated with exactly one f-structure element, an f-structure element may be associated with zero, one or more than one c-structure node. The annotations on the lexical



include topicalized elements, while constituents that typically carry the FOC discourse function include interrogative constructions. B&M hypothesize that incorporated pronominals are topic-oriented; that is, they can only anaphorically link to items filling the TOP function. Since question-words are assumed to fill the FOC function, an incorporated pronominal cannot link to a question word.

Critical to both TOP and FOC is that they must satisfy the Extended Coherence Condition; i.e. both TOP and FOC must be linked to predicate argument structure by either being functionally identified with or anaphorically linked to an argument. Functional identification takes place via the f-structure equation annotations in phrase-structure rules and lexical entries in the manner described above, while anaphoric linking only takes place between an anaphoric element such as a pronominal and a discourse-salient entity.

## 2. DISCOURSE PROPERTIES OF THE SETAWANA SM AND OM

In this section we argue that, in terms of the typology of agreement markers developed by B&M, the Setawana SM and OM are both pure anaphoric agreement markers, and that a lexical NP filling the OBJ function must be adjacent and following the main verb.

### 2.1. *Pro-drop Phenomena*

Setawana, like Chicheŵa, exhibits what is commonly called “pro-drop” phenomena for both subject and object argument NPs, as shown in (3). This suggests that the SM and OM are either optionally or obligatorily incorporated pronominal elements.

- (3) *ó-e-bidítse*  
SM-OM-lashed  
's/he lashed it'

### 2.2. *Word Order Variation*

B&M used the variation of word order freedom in Chicheŵa with respect to the presence or absence of the OM to support their claim that the Chicheŵa OM is an incorporated pronoun. Here we use the variation of word order freedom in a similar way to determine the properties of the SM and OM in Setawana.

(4) shows the grammaticality of the different word order permutations of a simple transitive clause without an object marker.

- (4) a. *Thabo ó-bidítse ntsá*  
 Thabo SM-lashed dog  
 'Thabo lashed the dog'  
 b. *ó-bidítse ntsá Thabo*  
 c. \**ntsa o-biditse Thabo*  
 d. \**o-biditse Thabo ntsa*  
 e. \**Thabo ntsa o-biditse*  
 f. \**ntsa Thabo o-biditse*

(5) shows the grammaticality of the different word order permutations of a simple transitive clause when an object marker is included.

- (5) a. *Thabo ó-e-bidítse ntsá*  
 Thabo SM-OM-lashed dog  
 'Thabo lashed it, the dog'  
 b. *ó-e-bidítse ntsá Thabo*  
 c. *ntsa ó-e-bidítse Thabo*  
 d. *ó-e-bidítse Thabo ntsá*  
 e. *Thabo ntsá ó-e-bidítse*  
 f. *ntsa Thabo ó-e-bidítse*

The pattern of grammaticality with respect to word order variation and object marking is summarized in (6).

(6) Order	Without om	With om
SVO	yes	yes
VOS	yes	yes
OVS	*	yes
VSO	*	yes
SOV	*	yes
OSV	*	yes

This is the same word-order distribution pattern found by B&M for Chicheŵa, it also appears in several other Bantu languages (e.g. Kiswahili - Wald 1979; Xhosa - Visser 1985; Makua - Stucky 1983, 1985; Kihaya - Byarushengo, Hyman & Tenenbaum 1976; Byarushengo & Tenenbaum 1976; Zulu - Wald 1979).

Following B&M, if we suppose that a lexical NP can only fill the OBJ function when it is adjacent and to the right of the V (i.e. the c-structure rule expanding VP is as given in (7)), we account for the ungrammaticality of (4c)-(4f).

- (7) VP → V NP  
OBJ

On the other hand, when the OM is present it fills the OBJ function and links anaphorically to the NP that fills the TOP function. If we assume that the TOP function is not associated with an adjacency restriction such as that associated with the OBJ function, we can account for the increase in word order freedom associated with the presence of the OM.

Thus the word order data provide evidence that the OM is either optionally or obligatorily an anaphoric agreement marker, and that a lexical NP that fills the OBJ function must be adjacent to and follow the main verb.

### 2.3. Question Word Constructions

B&M hypothesize that question words obligatorily fill the FOC function, and as such cannot serve as the antecedents for anaphoric agreement markers. They thus provide a means to test whether the agreement markers are purely anaphoric agreement markers.

(8a) shows that a question word may fill the OBJ function. The ungrammaticality of (8b) follows if we assume that the OM is purely anaphoric in nature, i.e. a pronominal (filling the OBJ function) which cannot anaphorically link to the question word that fills the FOC function for the reasons discussed above. (8c) and (8d) show that a question-word cannot be functionally identified with or anaphorically linked to the SUBJ function. In this regard Setawana differs from Chicheŵa; B&M report the Chicheŵa equivalent of (8c) to be grammatical.

- (8) a. *Thabo ó-bónye máng?*  
Thabo SM-saw who  
'Who did Thabo see?'  
b. \**Thabo o-m-monye máng?*  
Thabo SM-OM-saw who  
'Who did Thabo see?'  
c. \**mang o-bonye Thabo?*<sup>2</sup>  
who SM-saw Thabo  
'Who saw Thabo?'  
d. \**o-bonye Thabo máng?*  
SM-saw Thabo who  
'Who saw Thabo?'

The following examples show that this pattern (i.e. WH-words cannot appear linked to either a SM or an OM) also holds in main clause questions where the questioned element is located in a complement clause. (8h) shows

how a cleft construction can be utilized to form an acceptable subject question.

- (8) e. *John ó-dúméla góre Bill ó-súnné máng?*  
 John SM-believe that Bill SM-kissed who  
 'Who does John believe Bill kissed?'  
 f. \**John o-dumela gore Bill o-mo-sunne mang?*  
 John SM-believe that Bill SM-OM-kissed who  
 'Who does John believe Bill kissed?'  
 g. \**John o-dumela gore mang o-sunne Mary?*  
 John SM-believe that who SM-kissed Mary  
 'Who does John believe kissed Mary?'  
 h. *John ó-dúméla góre ké máng yó ó-súnné-ńg Mary?*  
 John SM-believe that be who RM SM-kissed-REL Mary  
 'Who does John believe kissed Mary?'

Examples (8i)-(8l) show that the same pattern also holds with embedded questions.

- (8) i. *John ó-nágana góre Bill ó-súnné máng?*  
 John SM-wonder that Bill SM-kissed who  
 'John wonders who Bill kissed?'  
 j. \**John o-nagana gore Bill o-mo-sunne mang?*  
 John SM-wonder that Bill SM-OM-kissed who  
 'John wonders who Bill kissed?'  
 k. \**John o-nagana gore mang o-sunne Mary?*  
 John SM-wonder that who SM-kissed Mary  
 'John wonders who kissed Mary?'  
 l. *John ó-nágana góre ké máng yó ó-súnné-ńg Mary?*  
 John SM-wonder that be who RM SM-kissed-REL Mary  
 'John wonders who kissed Mary?'

Just as for example (8b) involving the OM, the ungrammaticality of (8c) and (8d) would follow if we assumed that the SM is a purely anaphoric agreement marker. But if the SM is a purely anaphoric agreement marker then given the morphological requirement that every tensed verb have a SM, it follows that the SM is the subject of every tensed sentence. This appears to be a rather strange conclusion, since it implies that lexical NPs are never subjects in Setawana. We discuss below our attempts to find alternative evidence bearing on this hypothesis.