There is significant synchronic evidence of the close relationship between anaphoric and grammatical agreement even within the grammatical structures of a single language. It is possible to predict clear syntactic differences between a grammatical agreement marker and a morphologically incorporated pronoun. What is required is a theory of grammatical function that integrates the properties of argument functions, such as subject and object; and discourse functions, such as topic and focus. An analysis of this aspect of Chichewa, a Bantu language, indicates that important parameters of change and variation lie in surface form—the external phonologically interpreted morphology and phrase structures. The pronominal incorporation property is such a parameter; when it is combined with postulations about grammatical theory and discourse function, fundamental differences appear to be explained between syntactic structures organized by grammatical agreement with governed functions and anaphoric agreement with discourse functions. (Author/MSE)
On Topic, Pronoun, and Agreement in Chichewa

by

Joan Bresnan and Sam A. Mchombo

from

Proceedings of the Second

EASTERN STATES CONFERENCE ON LINGUISTICS

The State University of New York at Buffalo
Buffalo, New York
October 3-5, 1985

Edited by:
Soonja Choi
Dan Devitt
Wynn Janis
Terry McCoy
Sheng-sheng Zhang
On Topic, Pronoun, and Agreement in Chichewa

by Joan Bresnan and Sam A. Mchombo

Typologists have maintained that grammatical agreement systems evolve historically from the morphological incorporation of pronouns into verbs or nominal heads, and it has also been claimed that there is no clear dividing line between grammatical agreement, such as subject-verb agreement, and incorporated pronominal anaphora to a copic. Current theories of formal grammatical structure provide little insight into the nature of grammatical and anaphoric agreement, why they are so closely related, and what significant differences there are between them. As we will show in this study, there is substantial synchronic evidence of the close relation between grammatical and anaphoric agreement even within the grammatical structures of a single language. But, as we will also show, it is possible to predict clear syntactic differences between a grammatical agreement marker and a morphologically incorporated anaphoric pronoun. What is required is a theory of grammatical functions that integrates the properties of argument functions, such as subject and object, and discourse functions such as topic and focus. This study is a step toward developing such a theory within the overall framework of the lexical-functional theory of grammar.

1 The Object Marker as an Incorporated Pronoun

Chichewa, like other Bantu languages, shows both subject and object agreement in its verbal morphology. In finite verb forms the subject marker (SM) is obligatory, while the single object marker (OM) is optional.

---

1We are grateful for the support of the Center for the Study of Language and Information at Stanford University. This paper is extracted from Bresnan and Mchombo to appear.
3See Givón 1976.
4Chichewa is a Bantu language spoken in East Central Africa, particularly in Malawi and its bordering countries Mozambique, Zambia, and Zimbabwe, where it is also known as Chinyanja. In Guthrie's 1967-71 classification, Chichewa belongs to zone N in the single unit N31, comprising Chinyanja, Chimang' anja, and Chichewa.
5All of our examples from Chichewa are represented in Chichewa orthography, with the addition of tone markings. For an explanation of the orthography, see Bresnan and Mchombo to appear. High, rising, and downstepped tones are designated by ′, , and ‒, respectively; low tones are not marked. The broad phonetic tonal transcriptions
The SM and OM show person, number, and gender of the subject and object, respectively. The gender classes for third person verb agreement are illustrated with their conventional numbering in the following table. Note that the OM is exactly the same form as the SM in every class but 1 and 2.4

(3) Gender Classes for Verb Agreement

<table>
<thead>
<tr>
<th>class</th>
<th>example</th>
<th>gloss</th>
<th>SM</th>
<th>OM</th>
</tr>
</thead>
<tbody>
<tr>
<td>1, 1A</td>
<td>mlenje</td>
<td>hunter</td>
<td>a, u</td>
<td>mu</td>
</tr>
<tr>
<td>2</td>
<td>alenje</td>
<td>hunters</td>
<td>a</td>
<td>wa</td>
</tr>
<tr>
<td>3</td>
<td>mkángo</td>
<td>lion</td>
<td>u</td>
<td>u</td>
</tr>
<tr>
<td>4</td>
<td>mikángo</td>
<td>lions</td>
<td>i</td>
<td>i</td>
</tr>
<tr>
<td>5</td>
<td>phiri</td>
<td>mountain</td>
<td>li</td>
<td>li</td>
</tr>
<tr>
<td>6</td>
<td>mapiri</td>
<td>mountains</td>
<td>a</td>
<td>a</td>
</tr>
<tr>
<td>7</td>
<td>chipéwa</td>
<td>hat</td>
<td>chi</td>
<td>chi</td>
</tr>
<tr>
<td>8</td>
<td>zipéwa</td>
<td>hats</td>
<td>zi</td>
<td>zi</td>
</tr>
<tr>
<td>9</td>
<td>njúchi</td>
<td>bee</td>
<td>i</td>
<td>i</td>
</tr>
<tr>
<td>10</td>
<td>njúchi</td>
<td>bees</td>
<td>zi</td>
<td>zi</td>
</tr>
<tr>
<td>12</td>
<td>kamwána</td>
<td>small child</td>
<td>ka</td>
<td>ka</td>
</tr>
<tr>
<td>13</td>
<td>tiána</td>
<td>small children</td>
<td>ti</td>
<td>ti</td>
</tr>
<tr>
<td>14</td>
<td>ulalo</td>
<td>bridge</td>
<td>u</td>
<td>u</td>
</tr>
<tr>
<td>6</td>
<td>maulalo</td>
<td>bridges</td>
<td>a</td>
<td>a</td>
</tr>
<tr>
<td>15</td>
<td>kuímba</td>
<td>to sing, singing</td>
<td>ku</td>
<td>ku</td>
</tr>
<tr>
<td>16</td>
<td>pamsika</td>
<td>at the market</td>
<td>pa</td>
<td>pa</td>
</tr>
<tr>
<td>17</td>
<td>kumudzi</td>
<td>to the village</td>
<td>ku</td>
<td>ku</td>
</tr>
<tr>
<td>18</td>
<td>m'nyúmba</td>
<td>in the room</td>
<td>mu</td>
<td>mu</td>
</tr>
</tbody>
</table>

Word order in Chichewa interacts with verb morphology in an interesting way. In simple transitive sentences, when there is no object marker on the verb, the object immediately follows the verb, while the subject may be reordered, as shown in (4):

(4) njúchi z-i-ná-lúm-á alenje
    bees SM-past-bite-indic hunters
    'The bees bit the hunters.'

(5) njúchi z-i-ná-wá-lúm-á alenje
    bees SM-past-OM-bite-indic hunters
    'The bees bit them, the hunters.'
But when the OM is present, all of the above orders are possible:

These facts can be explained as follows.

i) Let us first assume that the obligatory SM is a third person subject agreement marker that optionally incorporates a subject pronoun. From the uniqueness and completeness conditions of lexical-functional theory it follows that the third person pronominal interpretation of SM will arise when and only when there is no subject NP in the phrase structure. If we omit the subject NPs from all of the grammatical examples in (4)-(5), a pronominal subject interpretation in fact occurs.

ii) Let us next assume that the optional OM is not an agreement marker at all, but an incorporated object pronoun. From the uniqueness condition it follows that an object NP can occur in the phrase structure only when OM is lacking. This implies that what we have labelled the object (O) in (5) is in fact something else.

iii) Let us further assume that all object NPs in Chichewa are generated in a fixed postverbal position in a VP constituent:

---

7 In terms of the formal representation of Kaplan and Bresnan 1982, the semantic ([PRED] = 'PRO') attribute is optional for the subject marker. See also Andrews 1984, Ishikawa 1985, Simpson 1985, and Wager 1985.

8 In other words, the semantic ([PRED] = 'PRO') attribute is obligatory for the object marker.
Thus, a postverbal object can appear in the VP only if there is no OM on the verb.

iv) Let us moreover assume that S consists of an optional subject NP, a VP, and an optional topic NP, all unordered with respect to each other. To express the fact that the S constituents are unordered, we separate them by commas in the S rule:

\[
S \rightarrow (NP) \quad , \quad VP \quad , \quad (NP)
\]

Thus, the S rule allows six different orders of the subject NP, the VP, and the topic NP.

v) Finally, let us assume that the grammaticized discourse functions—FOC(us) and TOP(ic)—universally must satisfy an extended coherence condition. This demands that they be linked to predicate argument structure, either by functionally or anaphorically binding an argument. The apparent cooccurrence of OM and an object NP is thus explained as the anaphoric binding of an object pronoun incorporated in the verb to a topic NP in S.

The differences between (4) and (5) follow from (i)-(v). In (4) we have a transitive verb but no OM. The verb’s subcategorization for object can be satisfied by the postverbal NP generated by the VP rule in (iii). This object has a fixed position in the VP. The subject NP generated by the S rule in (iv) can be reordered before or after the VP, but not inside it. If a topic NP were also generated by the S rule, the extended coherence condition (v) would require that it be linked to the predicate argument structure. This can be done in Chichewa by generating an incorporated anaphoric object in the verb (the OM), which the topic NP anaphorically binds. The OM prevents having an object NP in the VP by functional uniqueness. Thus, the free-floating NP linked to the OM in (5) is not really an object, but a topic, as

---

*The topic designates what is under discussion, whether previously mentioned or assumed in discourse (cf. Chafe 1977, Givón 1976, Wald 1979). We discuss it further below.


11The extension of the coherence condition to discourse functions is proposed by Faassi-Fehri 1984 and Abd-Rabbo 1984. Our version of this condition can be formulated more precisely as follows. The extended coherence condition requires that all functions in f-structure be bound. An argument function (that is, a subcategorizable function such as SUBJ, OBJ, OBL) is bound if it is the argument of a predicator (PRED). An adjunct is bound if it occurs in an f-structure containing a PRED. Finally, a topic or focus is bound whenever it is functionally identified with or anaphorically binds a bound function.
hinted in our translations. As such, it is freely orderable with respect to the subject and the VP. Thus, we replace (5) with the more accurate description (6):

(6) a. \[\text{TOP S [VP V ] bees SM-past-OM-bite-indic hunters}
\]
   \[\text{The bees bit them, the hunters.}\]

b. \[\text{TOP S: xin'áwál'úmá alenje njúchi}
\]
c. \[\text{TOP [VP V ] S: alenje zin'áwál'úmá njúchi}
\]
d. \[\text{[VP V ] S TOP: zin'áwál'úmá njúchi alenje}
\]
e. \[\text{S TOP [VP V ]: njúchi alenje zin'áwál'úmá}
\]
f. \[\text{TOP S [VP V ]: alenje njúchi zin'áwál'úmá}
\]

Although we have seen how the word order differences between (4) and (5) follow from (i)-(v), we have yet to explain the central similarity: why does the anaphoric linking of topics to the argument structure look like agreement in Chichewa? More generally, why should pronominal anaphora so closely resemble agreement in some languages? There are two questions to be answered here: first, why is there gender class agreement between the topic NP and the incorporated object pronoun? and second, why must the topic NP be anaphorically linked to an incorporated pronoun, which suspiciously resembles an agreement marker, rather than an independent pronoun in the object NP position (as in the English example *I love him dearly, my father*)?

The first question is how to account for the agreement in gender class between the OM and the topic NP. The answer is straightforward: person, number, and gender are precisely the pronominal categories which universally show agreement in anaphoric relations. As we have already remarked, typologists have long maintained that grammatical agreement systems evolve historically from incorporated deictic and anaphoric pronominal systems, explaining the fact that the categories of grammatical agreement are pronominal in nature. Chichewa clearly shows gender class agreement in both discourse anaphora and deixis. Consider first (7) and (8).

(7) Fisi anapili chipéwa ku San Francisco dzulo.
   hyena bought hat(7) in S.F. yesterday

(Finn 1978 and Lehman 1982, 1984. While the categories of grammatical and anaphoric agreement coincide in Chichewa, in some languages they partially clash. In Swahili, for example, animacy is a category of verb agreement that overrides the gender class categories of nominal concord (Bokamba 1981). Lyons 1968 points out that the categories of verb agreement in Swahili correspond to those of pronominal agreement rather than nominal concord.

See n. 2.)
Madzulo anapfta ku S.Jose kuméné á-ná-ká-chi-gulitsá kwá
evening he-went to San Jose where he-pst-go-it(sea)-sell to
mlondá wá á méya.
guard of hon. mayor
'The hyena bought a hat in San Francisco yesterday. In the evening
he went to San Jose where he went to sell it to the mayor’s guard.'

(8) Fisi anagulá chipéwa ku San Francisco dzulo.
hyena bought hat(7) in S.F. yesterday
'Madzulo anapfta ku San Jose kuméné á-ná-ká-wá-gulitsá kwá
evening he-went to S.J. where he-pst-go-it(2)-sell to
mlondá wá á méya.
guard of hon. mayor
'The hyena bought a hat in San Francisco yesterday. In the evening
he went to San Jose where he went to sell it to the mayor's guard.'

The incorporated pronoun in (7) must agree in gender class with the antecedent chipéwa 'hat' in a previous sentence in the discourse; (8) shows that the class 2 OM -wa-, which disagrees with the class 7 antecedent, cannot be used to establish the anaphoric relation. Observe that this anaphoric relation crosses sentence boundaries in a discourse, and hence could not possibly be analyzed as agreement in the sense of a local syntactic relation between a grammatical agreement marker on the verb and an argument of the verb.

Consider now (9) and (10):

(9) (pointing to a lion lying on the ground)
Uwu.
this (class 3)

(10) (pointing to a lion lying on the ground)
*Ichi.
this (class 7)

The word for lion, mkángo, belongs to gender class 3. Deixis to a lion requires the class 3 deictic pronoun form; the class 7 form shown in (10) would be used for deixis to a hat or some other thing whose corresponding noun belongs to class 7. Again, these phenomena could not possibly be analyzed as syntactic agreement. Thus, the choice of agreement features of person, number, and gender in the anaphoric use of pronouns is independently motivated, and need not—indeed, should not—be accounted for by a sentence-internal syntactic agreement mechanism.

The second question is how to explain the use of incorporated pronouns to anaphorically link the topic NPs to the predicate argument structure.
It would seem more natural (to the English speaker, at least) to use an independent pronoun in the object NP position (as in the English example I love him dearly, my father), establishing a clearly anaphoric relation which no one would take for verb-object agreement. But naturalness apart, we need to explain the hypothesized anaphoric function of OM when a topic NP is present. The explanation lies in a fundamental typological difference between languages like Chichewa and languages like English. In Chichewa, independent object pronouns are used only to introduce new topics or for contrast. This is shown in (11)-(14).

(11) Fsi anadyá chfmanga. Á-tá-chi-dya, anapftá ku San Francisco.
    hyena ate corn(7) he-serial-it(7)-eat he-went to S.F.
    'The hyena ate the corn. Having eaten it, he went to S.F.'

(12) Fsi anadyá chfmanga. Á-tá-dyá icho, anapftá ku San Francisco.
    hyena ate corn(7) he-serial-eat (7)it he-went to S.F.
    'The hyena ate the corn. Having eaten it (something other than corn), he went to S.F.'

(13) Fsi anadyá mkango. Á-tá-dyá icho, anapftá ku San Francisco.
    hyena ate lion(3) he-serial-eat (7)it he-went to S.F.
    'The hyena ate the lion. Having eaten it, he went to S.F.'

(14) Fsi anadyá mkango. Á-tá-dyá fwo, anapftá ku San Francisco.
    hyena ate lion(3) he-serial-eat it(3) he-went to S.F.
    'The hyena ate the lion. Having eaten it (something other than the lion), he went to S.F.'

While the discourses in (11) and (13) are natural, those in (12) and (14) are bizarre. The independent pronouns are interpreted as referring to topics not mentioned in the previous sentence, even though they agree with the objects of the previous sentences in person, number, and gender class. Note from the translations of (11) and (13) that this is not at all a property of the English pronominal system.

Now within a sentence, the floating topic must be anaphorically bound to an argument in order to satisfy the extended coherence condition. Because of their contrastive discourse function, the independent pronoun objects of Chichewa cannot be used topic-anaphorically to satisfy this condition. As a result, sentences like the following are ungrammatical, although similar examples with independent pronouns in English are fine.

14These independent pronouns are morphologically distinct from the series of demonstrative pronouns in Chichewa.
It appears that noncontrastive anaphora to the topic, a communicative function that is borne by independent syntactic pronouns in languages like English, is carried by the incorporated object pronouns of languages like Chichewa.\(^\text{16}\)

We have now come to the conclusion that the OM, apparently an object agreement marker, is actually an incorporated object pronoun which may be anaphorically linked to a floating topic NP in the sentence. Our evidence has come from the interactions of word order with verbal agreement morphology.\(^\text{16}\) We have also drawn on research on typology and discourse to answer the question of why pronominal anaphora to the topic should so closely resemble agreement. First, discourse anaphoric relations and even deixis universally show agreement in the referentially classificatory categories of person, number, and gender class, and these are also the categories of grammatical agreement between a verb and its arguments, reflecting the historical derivation of many agreement systems from pronominal systems. Second, the independent object pronouns of Chichewa have a contrastive discourse use that makes them incompatible with anaphora to the topic, either within sentences or in discourses. Hence, the incorporated pronouns are the only pronominal objects that can serve to link the topic NP to the predicate argument structure.

These conclusions raise the theoretical question of how to distinguish agreement from incorporated pronominal anaphora in principle. For example, we analyzed the SM marker differently from the OM marker, in assuming that the former only optionally functions as a pronoun. But we could have analyzed the SM simply as an incorporated pronoun, like the OM. What is the principled basis for choosing between these alternatives? The answer

\(^{16}\)See Bresnan and Mchombo to appear for evidence from phrase-final retraction of high tone that the NP agreeing with the OM lies outside of the verb phrase. In their detailed study of tone in Kilaya, Byaruchengo, Hyman, and Tenenbaum 1976 discovered a similar phenomenon, from which they concluded that the object markers are incorporated pronouns related to the higher NP as in left or right dislocation. See also Byaruchengo and Tenenbaum 1976.

\(^{16}\)In Bresnan and Mchombo forthcoming we adduce further evidence from the morphology, syntax, and semantics of the incorporated reflexive pronoun, and in Bresnan and Mchombo to appear, we give corroborating evidence from the interactions of tone with phrase structure.
lies in the theory of argument functions and discourse functions.

2 Grammatical versus Anaphoric Agreement

2.1 Locality

Our theory tells us that grammatical agreement relations with non-controlled arguments can be distinguished from anaphoric agreement relations by locality: only the anaphoric agreement relations can be nonlocal to the agreeing predicate. The following reasoning supports this conclusion. First, only the argument functions, SUBJ, OBJ, etc., can be directly governed by predicates. In order to satisfy the completeness and coherence conditions, such argument functions must be either expressed syntactically within the phrasal structures headed by the predicates, expressed morphologically on the head itself, or else remain unexpressed (that is, anaphorically or functionally controlled by nonlocal structures). Hence, the government relation between a verb and its noncontrolled arguments must be structurally local to the verb. But verbs can agree grammatically only with their governable arguments. Therefore, grammatical agreement between a verb and any of its noncontrolled arguments must be structurally local to the verb.

In contrast, an incorporated pronoun is a referential argument itself governed by the verb. By functional uniqueness, an external referential NP cannot also serve as that argument. Hence, such an external NP cannot be related to that argument position of the verb by government, but only by anaphora with the agreeing incorporated pronoun. But anaphoric relations between (nonreflexive) pronouns and their antecedents are in general nonlocal to sentence structure, since their primary functions belong to discourse.

Because only the anaphoric agreement relations can be nonlocal to the agreeing predicate, we would expect that the relation between the OM in Chichewa and the floating NP with which it agrees can be nonlocal, if this is indeed anaphoric agreement. This prediction is correct:

(16) chigawéngá ichi asilikálé gányu a-na-úz-á
   terrorist(7) this soldiers of temporary.work SM-rmpst-tell-indic
   mtogolé wáthu kutsí a-ángáth-é ku-chi-gwfr-a
   leader our that not-SM-be.able-subjn inf-OM(7)-catch-indic
   'This terrorist, the mercenaries told our leader they cannot catch him.'

In (16) the class 7 noun chigawéngá 'terrorist' is a floating topic NP three levels of verbal embedding above the class 7 OM chi, that agrees with it. If we remove that OM, the sentence becomes ungrammatical:

17See Bresnan 1982b for the theory of governable functions assumed here.
The ungrammaticality follows from the extended coherence condition, requiring the topic NP to be bound to a lexical predicate argument structure, and from a fact of Chichewa grammar, that topicalizations are constructed by anaphoric binding only, and not by functional identification. Not only can the floating topic NP be nonlocal from the OM that it is linked to, but the nonlocal topic shows the same ordering possibilities within its higher clause that we found when it occurred in a monoclausal sentence with the OM. Thus, the topic NP in (16) can also appear sentence-finally, as in (18)a, and after the highest subject, as in (18)b, but not after the main verb inside the VP, as in (18)c, exactly as our analysis in Section 1 (i)-(v) predicts:

In contrast, English allows both constructions, as we see from the grammaticality of both translations in (16) and (17). Bantu languages vary in this respect. For example, Northern Sotho has both a preposed topic NP with anaphoric binding to the object prefix, and a preposed focus NP with no object prefix (Louwrens 1982), Kihung'aa, spoken in southwestern Congo (Kishasa), has a preposed focus NP construction with no object prefix (Takiala 1973), and Dzamba, spoken in the Equator province of Zaïre, has a preposed topic NP construction with no object prefix as well as a left-dislocated construction with OM (Bokamba 1981).

When two or more topics occur in the same sentence, there appear to be some constraints on their anaphoric relations. Although we have not yet investigated these in Chichewa, nesting constraints on multiple anaphoric binding of topics have been found in Arabic (Aoun n.d., Abd-Rabbo 1984, Fasih-Fehri 1984).
The floating topic NP can also be generated in an intermediate sentential clause between the main clause and the embedded complement verb bearing the OM:

(19) asilikálf á gányu a-ná-úz-á mtsogoleri wáthu kutf soldiers of temporary.work SM-past-tell-indic leader our that chigawéngá ichi a-ángáth-é ku-chí-gwfr-a terrorist(7) this not-SM-be.able-subjn inf-OM(7)-catch-indic 'The mercenaries told our leader that this terrorist, they cannot catch him.'

But the topic NP cannot appear between the second verb down and its infinitival complement:

(20) ?asilikálf á gányu a-ná-úz-á mtsogoleri wáthu kutf soldiers of temporary.work SM-past-tell-indic leader our that a-ángáth-é chigawéngá ichi ku-chí-gwfr-a not-SM-be.able-subjn terrorist(7) this inf-OM(7)-catch-indic 'The mercenaries told our leader that they cannot, this terrorist, catch him.'

If we assume that the infinitive is a direct VP complement to a-ángáth-é 'not be able', not immediately dominated by an S node, our analysis predicts this result. The reason is that topic NPs are generated under S and not under VP.

The following examples show that OM agreement is not only nonlocal, but has typical properties of pronominal relations, violating constraints on extraction (Ross 1967) and operator binding (Higginbotham 1980):

(21) a. chigawéngá ichi ndi-ku-fúné ku-dázwa ngati terrorist(7) this l-pres-want inf-know whether asilikálf amëně é-kù-bá nkhuńů záthu a-ŋa-fúné soldiers who SM-pres-steal chickens our SM-may-want ku-chí-gwfr-ts-a ntchfsto inf-OM(7)-grab-cause-indic work 'This terrorist, I want to know whether the soldiers who are stealing our chickens may want to make use of him.'

b. chigawéngá ichi alenje a-a-tf-tsimkizira kutf terrorist this hunters SM-perf-us-assure that maganizo wótf asilikálf ñwa a-ángáthé ku-chí-gwfr-a belief that soldiers these not-SM-be.able inf-OM(7)-catch
s-á-ku-wá-pátá má nthá
not-SM-pres-OM-give fear
'This terrorist, the hunters have assured us that the belief that
the soldiers cannot catch him does not give them any worries.'

c. amáíyi á mwná uyu á-ma-mu-ránuza
mother of child this SM-habit-OM-mistreat
'The mother of this child mistreats him.'

In contrast to the OM, which is not an agreement marker, but an incor-
porated pronoun, the SM on our analysis is a third person agreement marker
and only optionally pronominal. This implies that all simple Subject Verb
sentences are functionally ambiguous: the apparent subject NP could either
be a true subject with which the verb shows grammatical agreement, as in
(22), or it could be a topic NP related by anaphoric agreement to the subject
pronominal in the verb, as in (23).

(22)
In the former case, the subject NP must be local to the verb, but in the latter case, the floating topic NP may be nonlocal to the verb. Hence, we expect to find nonlocal subject agreement as well as nonlocal object agreement in Chichewa, and we do. In (24) the topic mkángó uwu 'this lion' appears three levels of embedding above its verb, in sentence-initial position in (24a), following the highest subject in (24)b, and in sentence-final position in (24)c.

(24) a. mkángó uwu, alenje a-ku-gánša kuti ụ-ma-fúná
lion(3) this hunters SM-pres-think that SM(3)-habit-want
ku-gúmáá nyumbá yá mfúná
inf-pull.down house of chief
'This lion, the hunters think that it wants to pull down the chief's house.'

b. alenje mkángó uwu a-ku-gánša kuti ụ-ma-fúná
hunters lion(3) this SM-pres-think that SM(3)-habit-want
ku-gúmáá nyumbá yá mfúná
inf-pull.down house of chief

c. alenje a-ku-‘ánša kuti ụ-ma-fúná
hunters SM-pres-think that SM(3)-habit-want
ku-gúmáá nyumbá yá mfúná mkángó uwu
inf-pull.down house of chief lion(3) this

In sum, we see that the SM can be used like the OM for nonlocal anaphora to the topic. However, on our analysis the SM is ambiguous: in addition to being an incorporated pronominal, it can also be used as a true grammatical agreement marker, unlike the OM. Hence, we should expect asymmetries to arise between the patterns of subject agreement and object agreement. We take these up next.
2.2 Subject versus Topic

When the SM is used as a grammatical agreement marker, it agrees with a nominal that has the SUBJ function; when the SM is used for anaphoric binding, its antecedent within the sentence has the TOP function. Thus the theory of functions should provide a basis for predicting and explaining certain syntactic differences between grammatical and anaphoric agreement.

Grammatical functions in our theory can be partitioned into argument functions, such as SUBJ, OBJ, OBL (lique), and nonargument functions, such as TOP, FOC, and ADJUNCT. Argument functions are directly mapped onto semantic or thematic roles in lexical predicate argument structures. They serve to designate the participants in events in a cross-lexically invariant way (Simpson 1983). In contrast, nonargument functions, by the extended coherence condition, must be linked to other grammatical functions (or, in the case of adjuncts, must cooccur with a PRED attribute); hence nonargument functions are only indirectly associated with predicate argument structure. They serve to structure the information content of an utterance so as to facilitate communication between the speaker and the hearer. Argument functions must be unique in their clauses, while nonargument functions may admit of multiple instances.21

We will adopt three postulates about the role of the TOP and FOC functions in the grammars of natural language. First, in relative clauses the relative pronoun or relativized constituent universally bears the TOP function.22 For example, in (25) which is the topic of the clause which you don't want:

(25) The car which you don't want is a Renault.

The extended coherence condition requires that, like the floating topic NP, the relative topic be linked to the lexical predicate argument structure either

21Multiple instances are expressed in the formal language of lexical-functional grammars by the membership connective e (Kaplan and Bresnan 1982). Note that some multiple topics come from stacked S structures of the form S = NP, S, where NP has the TOP function, and S is an ADJUNCT (Fassi-Fehri 1984). These structures maintain the uniqueness of topics. Languages in which multiple grammatical focuses occur in clusters of preposed interrogative phrases are discussed by Wachowitsch 1974 and Ackerman 1981. See Bresnan 1982a on the nonuniqueness of adjuncts.

22A similar proposal is due to Kuno 1976, who uses the concept of 'theme'. Our terminology is consistent with that of Chafe 1976 and Givon 1976. Poulos 1981 adopts this analysis for Zulu relative clauses.
by functional identification or by anaphoric binding. The former mode of linkage is subject to well-known extraction constraints. Second, in interrogative clauses the interrogative pronoun or questioned constituent universally bears the FOC function. For example, in (26) what is the focus of the clause what you want:

(26) I know what you want.

\[
\text{I know [what you want ___]}
\]

\[
\text{FOCUS OBJ}
\]

Third, the same constituent cannot be both focus and topic of the same level of (functional) clause structure. Thus, in cleft constructions the same phrase is interpreted as both a focus and a topic, but at different levels of embedding. For example, in (27) my car is the focus of the main clause, and the relativized object is the topic of the embedded complement clause that you don't want:

(27) It's my car that you don't want.

\[
\text{[It is my car [that you don't want ___]]}
\]

\[
\text{FOCUS TOPIC OBJ}
\]

These three postulates ultimately derive from the theory of the role and interpretation of these functions in discourse. For example, because the topic designates what is under discussion, whether previously mentioned or assumed in discourse, it is presupposed. The interrogative focus designates

---

22 See Saiki 1985 for an exposition of new work on long-distance functional identification in lexical-functional grammar and an extremely interesting application to relativization in Japanese. She also shows that relativization is subject to different constraints from topicalization (thematization) in coordinate constructions, in apparent conflict with Kuno's 1976 hypothesis. However, her evidence is consistent with the hypothesis that the relativized element is the topic (or theme, in Kuno's terminology), if we assume that in Japanese the mode of linkage of topics differs in topicalizations and relative clauses, as in fact it does in Chichewa. While Chichewa employs only anaphoric binding of the floating topic, it employs both anaphoric binding and functional identification of relative topics.

24 For arguments that support this postulate, see Dik 1976. See also Myers 1971 for evidence of a right focus position in Kikuyu for both emphatic, or contrastive, phrases and interrogative phrases.
what is not presupposed as known, and is contrasted with presupposed material. Hence, allowing the same constituent to be both topic and focus of the same clause leads to inconsistent presuppositions. Until we have more explicit theories of the interpretation of these functions in discourse, however, we will adopt the strategy of simply postulating properties of the grammaticized discourse functions in order to derive explicit predictions. We can then explain the contrasts between (28)a,b and between (29)a,b:25

(28) a. (Mary asked) what it was that Fred cooked.
   b. ??(Mary ate) what it was that Fred cooked.

(29) a. (I asked) who it was that Marilyn suspected.
   b. ??(I met) the person who it was that Marilyn suspected.

Examples (28)a and (29)a contain interrogative clauses, while examples (28)b and (29)b contain a so-called headless and a headed relative clause, respectively. These examples show that although it is perfectly natural to question the clefted noun phrase in a cleft construction, it is much less so to relativize it. To see in detail why this is so, consider (30), which schematically displays the analysis of (29)a.

(30)

(i) the cleft construction:

[it was who [that Marilyn suspected ___]]

FOCUS TOPIC OBJ

(ii) the interrogative clause:

[who it was ___ [that Marilyn suspected ___]]

FOCUS FOCUS TOPIC OBJ

Here the cleft NP in (i) is questioned in (ii). Since the cleft NP and the questioned phrase both have FOC functions, there is no violation of our postulates. But now consider (31), which schematically illustrates the analysis of (29)b.

25 Takizawa 1973 makes this point explicitly.
26 The observation of contrasts of this kind in English is due to Baker 1970. See also Chiba 1973.
(31) (i) the cleft construction:

\[
\text{[it was who [ that Marilyn suspected ]]} \quad \text{FOCUS} \quad \text{TOPIC} \quad \text{OBJ}
\]

(ii) the relative clause:

\[
\text{the person [ who it was [ that Marilyn suspected ]]} \quad \text{TOPIC} \quad \text{FOCUS} \quad \text{TOPIC} \quad \text{OBJ}
\]

Here the cleft NP in (i) is relativized in (ii). Since the cleft NP has the FOC function and the relative pronoun, the TOP function, and since these conflicting discourse functions occur in the same level of functional clause structure, the result is ill-formed.

This theory leads us to the following five predictions about Chichewa.

**Prediction 1.** In Chichewa, questions are formed with the question word in place in a within-clause position. In simple (noncleft) interrogative clauses, there should therefore be an agreement asymmetry between subjects and objects: it should be possible to question the subject with SM but not the object with OM. The reason is that the OM is an incorporated object pronoun, so an object question word in the same clause must be interpreted as a floating topic NP anaphorically linked to the OM. But then the question word will be both FOC and TOP of the same clause, violating our third hypothesis about discourse functions. In contrast, the SM is an agreement marker for grammatical subjects, and only optionally used as a referential pronoun. Thus the interrogative constituent can simply be the subject of the verb, without also being interpreted as TOP. These predictions are correct.27

(32) (Kodzi) mu-ku-fún-a chiyáni?
Q you-pres-want-indic what
‘What do you want?’

\[
\text{[ kodzi [ mu-ku-fún-a chiyáni ] ]} \quad \text{Q} \quad \text{SUBJ} \quad \text{FOCUS} \quad \text{OBJ}
\]

---

27A similar asymmetry appears in KiHaya. Bennett 1977 observes in passing that a question word cannot be used with the object pronoun prefixes, but it does appear with the subject prefix. This is particularly interesting in the light of the tonal evidence for the pronominal status of the object prefixes in KiHaya (n. 15).
(33) ?(Kodi) mu-ku-chi-fún-á chiyáni?
Q you-pres-OM(?) want-indic what(?)
'What do you want ("it")?'

[ kodí [ mu-ku-chi-fúná ] chiyáni ]
Q SUBJ OBJ FOCUS TOPIC

function clash 

anaphoric binding

(34) (Kodi) chiyáni ch-i-náoneka?
Q what(?) SM(?) past-happen-indic
'What happened?'

[ kodí chiyáni ch-i-náoneka ]
Q SUBJ SM

grammatical agreement

In (34) the SM is interpreted as a grammatical agreement marker. The interpretation of the SM as an incorporated subject pronominal anaphorically linked to a topic NP is ruled out by the same function clash that appears in (33)—namely, a single constituent is both a focus and a topic of the same clause:

(35)

[ kodí chiyáni ch-i-náoneka ]
Q FOCUS SUBJ TOPIC

function clash --> 

anaphoric binding

The functional ambiguity of subject-verb agreement saves the example by providing the structure shown in (34). This is striking evidence for the difference between grammatical agreement, shown by the SM, and incorporated pronominal anaphora, shown by the OM. It also shows that the SUBJ function is grammatically distinguishable from the TOP function.²⁸

Prediction II. Recalling the nonlocality property of anaphoric agreement discussed above, we can derive a further prediction from our theory: in

²⁸One might wonder whether the restriction on OM with interrogatives reflects a more general restriction against using OM with indefinite or nonspecific objects. See Bresnan and Mchombo to appear for evidence against (33).
contrast to local subjects, the nonlocal subjects described above should not allow questioning in place. For example, in contrast with (36)a, (36)b should be ill-formed:

\[(36)\]

\[a. \quad (\text{Kodl}) \text{mu-ku-fun-å kutf chfyåni chi-onék-e?}\]
\[Q \text{you-pres-want-indic that SM-happen-subjn}\]
\[\text{`You want what to happen?'}\]

\[b. \quad ??(\text{Kodl}) \text{chfyåni mu-ku-fdn-i kutf chi-onek-e?}\]
\[Q \text{what you-pres-want-indic that SM-happen-subjn}\]
\[\text{`What do you want to happen?'}\]

In (36)a, chfyåni `what' is a subject questioned in place, and is grammatical for the same reason that (31) is. In (36)b, however, it is a floating topic, anaphorically bound to the pronominal SM on the embedded verb. The TOP function is incompatible with an interrogative FOC function, so the example is ill-formed.\(^{29}\) Thus, only subjects locally governed by the verb can be questioned in non-cleft constructions.

**Prediction III**. In the examples above, the question is formed with the question word chfyåni in place. There is an alternative construction for questions in Chichewa, in which the question word is clefted and the content of the question is expressed within a relative clause. The relative clause may contain an OM to which the relative pronoun is anaphorically bound. Because clefting splits the FOC and TOP functions into two different clauses, our theory predicts that the subject-object asymmetry should disappear in these constructions, with both SM and OM possible within the embedded clause.\(^{30}\) This prediction is correct.

---

\(^{29}\)The presence of the complementizer adjacent to the subject gap is irrelevant to the ill-formedness of this example, as one can see from the grammaticality of example (51).

\(^{30}\)As noted above (n. 23), Chichewa employs two relativization strategies—anaphoric binding and functional identification. Only with the former will the asymmetry disappear. Since the OM is an incorporated pronoun, the principle of functional uniqueness would preclude functional identification of the OM with an object relative pronoun in a cleft (or relative) construction. It follows that in a language which has an incorporated pronoun OM and which employs only the functional identification strategy for relativization (at least within the domain permitted by island constraints—cf. Clements 1985), no OM will appear in a cleft interrogative construction questioning the object; Kiheng'an appears to be such a language (Takizala 1973).
(37) Kodif ndi chiyani ch'-m'ene mu-ku-chi-funa
Q copula what(?) 7-rel you-pres-OM(7)-want-indic
‘What is it that you want?’

[ kodif ndi chiyani [ ch'-m'ene mu-ku-chi-funa ]]
Q copula FOCUS [ TOPIC OBJ ]

(38) Kodif ndi chiyani ch'-m'ene ch'-n'a-oneka
Q copula what(?) 7-rel SM(7)-past-happen-indic
‘What is it that happened?’

[ kodif ndi chiyani [ ch'-m'ene ch'-n'a-oneka ]]
Q copula FOCUS [ TOPIC SUBJ ]

(39) (Kodif) ndi chiyani ch'-m'ene mu-ku-funa kuti chi-onek-e?
Q copula what(?) 7-rel you-pres-want-indic that SM(7)-happen-subjn
‘What do you want to happen?’

[ kodif ndi chiyani [ ch'-m'ene mu-ku-funa [ kuti chi-onek-e ]]]
Q copula FOCUS [ TOPIC SUBJ ]

Prediction IV. It is a further consequence of our theory that the subject-object agreement asymmetry found in simple questions should not appear in relative clauses.\(^{32}\) While the question word is a focus, and hence could not also be a topic in the same level of clause structure, the relative pronoun is a topic, and hence consistent with anaphoric binding of both OM and SM. This prediction is correct.

(40) munthu a-mene ndi-nu-mu-yendera
person(1) 1-rel 1-past-OM(1)-visit
‘the person that I visited’

\(^{31}\)Relative verbs in Chichewa show an initial high tone. See Mtenje in preparation for discussion.

\(^{32}\)See note 30.
Thus (40) and (41) contrast with the examples with both SM and OM given under the heading of Prediction I above.

*Prediction V.* While both definite and indefinite noun phrases can be used to represent information previously mentioned in the discourse, and so can be linked anaphorically to the OM or SM as topics, idiomatic objects and cognate objects are usually not used in this way, perhaps because they merely elaborate on the meaning of the verb. These NPs are therefore difficult to topicalize. In (42)a, *bóndo* 'knee' is an idiomatic object of the verb *-nong’oneza* 'whisper to', yielding the meaning 'to feel remorse'.33 The presence of OM makes the result bad, as in (42)b. Yet the object can undergo passivization, as in (42)c, showing that the SM, unlike the OM, serves as a grammatical agreement marker.

(42) a. *chifukwá* chá mwáno wáke Mavuto tšópáno
because of rudeness his Mavuto now
*a-ku-nong’oneza-á bóndo*
SM-pres-whisper-to-indic knee
'Because of his rudeness, Mavuto is now whispering to his knee (that is, feeling remorse).'

b. ??*chifukwá* chá mwáno wáke Mavuto tšópáno
because of rudeness his Mavuto now
*a-ku-Il-nong’oneza-á bóndo*
SM-pres-OM(5)-whisper-to-indic knee(6)
'Because of his rudeness Mavuto is now whispering to it, his knee.'

---

33 This meaning is evoked by the image of a person sitting doubled up hugging his knees with his head bowed, whispering.
Similarly, in (43a) the verb -lotu 'dream' has the cognate object maloto 'dreams'. Again the presence of the OM makes the result bad, as in (43)b. And again, passivization of the cognate object is possible, as in (43)c, showing that the SM, unlike the OM, functions as a grammatical agreement marker.

(43)  

a. mlenje a-na-lot-ə maloto ówópsya usiku  
hunter SM-rmpst-dream-indic dreams frightening night  
'The hunter dreamed frightening dreams last night.'

b. ??mlenje a-na-wá-lot-ə maloto ówópsya usiku  
hunter SM-rmpst-OM-dream-indic dreams frightening night  
'The hunter dreamed them last night, frightening dreams.'

c. maloto ówópsya a-na-lot-ɛdwa ndf mlenje usiku  
dreams frightening SM-rmpst-dream-past-indic by hunter night  
'Frightening dreams were dreamed by the hunter last night.'

The cognate object can be topicalized in certain circumstances, as in (44), where the recurrence of the same dream is referred to.

(44)  

maloto awa mlenje a-na-wá-lot-ə kasanu  
dreams these hunter SM-rmpst-OM-dream-indic five-times  
'These dreams, the hunter dreamed them five times.'

Likewise, the verb -vina 'dance', usually intransitive, does take as an object the name of a dance, as in (45)a. This object resists topicalization and consequently the OM, as (45)b shows. But it does passivize, allowing the SM, as (45)c shows.

(45)  

a. mfúmu i-ná-vin-ə chiwoda  
chief SM-past-dance-indic chiwoda  
'The chief danced the Chiwoda dance.'

b. ??mfúmu i-ná-chi-vin-a chiwoda  
chief SM-past-OM(7)-dance-indic chiwoda(7)  
'The chief danced it, the Chiwoda dance.'

c. chiwoda chi-ná-vfn-tdw-ə ndf mfúmu  
chiwoda(7) SM(7)-past-dance-pass-indic by chief  
'The Chiwoda dance was danced by the chief.'
Thus, if we assume that SM is an agreement marker as well as an incorporated pronoun while OM is only an incorporated pronoun, our theory of argument functions and discourse functions predicts a number of subject-object symmetries and asymmetries in agreement patterns that are actually found to occur.

In Bresnan and Mchombo to appear, we argue from the systematic patterning of the anaphoric system within and across sentences that the TOP function does indeed derive its properties from discourse topics. Of course, further research into the role and interpretation of topic and focus in discourse structures is needed to extend our theory.

3 Typology

In addition to the locality property and the five predictions that we have just confirmed, our theory also suggests a basis for certain properties that appear to distinguish incorporated anaphora from grammatical agreement typologically. We have seen that Chichewa has two series of anaphoric pronouns, the OMs, used for anaphora to a topic, and the independent object pronouns, used to introduce new topics or for a 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 in discourse and those used for "contrast, emphasis, or focus". The former have less phonetic content than the latter. For example, in English the contrast arises between unstressed and stressed independent pronouns; in Latin, between the bound pronominal use of the verbal subject inflections and independent pronouns; and in Japanese, between zero pronominals and independent pronouns. Since incorporated pronominal arguments generally have less phonetic content than independent pronouns, this observation suggests that the fundamental typological property that distinguishes the uses of independent pronouns in Chichewa from those in English is simply that Chichewa employs the morphological incorporation of referential pronominal arguments into the lexical categories that govern them. Let us call this typological property the pronominal incorporation property.

It is an immediate consequence of the principle of functional uniqueness that languages having the pronominal incorporation property must show 'Pro-Drop' (Perlmutter 1971), that is, ellipsis of nominal arguments with consequent pronominal interpretation. For incorporated pronominal arguments are incompatible with the corresponding syntactic NP arguments by functional uniqueness, and so can be employed only when the latter can be omitted. We have seen that Chichewa has both subject and object pro-drop,

Kameyama notes that her two anaphoric pronominal functions are implicit in Givón's 1983 proposed universal scale of referring expressions from a typological perspective.
in the sense that the SM optionally, and the OM obligatorily, has pronominal function.

From the principle of functional uniqueness it also follows that in languages with the pronominal incorporation property, a verb or other head cannot govern the case of any referential nominals with which its incorporated pronouns agree. For if the incorporated pronoun is a referential argument itself governed by the verb, then an external referential NP cannot also serve as that argument, by functional uniqueness. Hence, such an external NP cannot be related to that argument position of the verb by government, but only by anaphora with the agreeing incorporated pronoun. But the categories of agreement in these anaphoric relations are universally the referentially classificatory properties—person, number, and gender, but not grammatical case. For example, in She knows I admire her and She enjoys herself, accusative her and herself show person, number, and gender agreement with their nominative antecedents she, but differ in grammatical case. Fassi-Fehri 1984 shows that in Arabic, invariant case is associated with topic NPs that anaphorically bind the incorporated pronominal arguments of verbs, while case government is associated with argument NPs that grammatically agree with the verb. His work strikingly confirms the prediction that verbally governed case on the full nominal is inconsistent with the anaphoric linking of the nominal to an incorporated pronoun.

Chichewa, of course, lacks grammatical case-marking of dependent nominals. The independent pronouns, for example, are invariant in form whether occurring as subjects, objects, or prepositional objects. We are therefore unable to test exactly the prediction that verbal case government is inconsistent with pronominal incorporation in Chichewa. However, Chichewa is typologically consistent with our prediction. Nichols 1985 proposes a typological opposition between head-marking languages, in which the relation of arguments to predicates is registered on the predicate, and dependent-marking languages, in which it is registered on the arguments (Nichols 1985, Van Valin 1985). Chichewa clearly exemplifies the head-marking type. The reason that head-marking languages do not show case government of dependent nominal arguments may simply be that head-marking is a reflection of the pronominal incorporation property.

Chichewa, of course, lacks grammatical case-marking of dependent nominals. The independent pronouns, for example, are invariant in form whether occurring as subjects, objects, or prepositional objects. We are therefore unable to test exactly the prediction that verbal case government is inconsistent with pronominal incorporation in Chichewa. However, Chichewa is typologically consistent with our prediction. Nichols 1985 proposes a typological opposition between head-marking languages, in which the relation of arguments to predicates is registered on the predicate, and dependent-marking languages, in which it is registered on the arguments (Nichols 1985, Van Valin 1985). Chichewa clearly exemplifies the head-marking type. The reason that head-marking languages do not show case government of dependent nominal arguments may simply be that head-marking is a reflection of the pronominal incorporation property.a

---

See Lehmann 1982, 1984. In Warlpiri nominal adjuncts show case concord with the arguments they modify, but these nominal adjuncts are nonreferential and are used to attribute properties to the arguments they concord with (Simpson 1983: 252 et passim; Jelinek 1984). Case and not person agreement is a general property of nominal adjunct agreement (Lehmann 1982, 1984).

Nichols 1985 notes the existence of 'double-marking' languages, which have both head-marking and dependent-marking morphology. Our theory is consistent with double-marking as a historical development, but strongly constrains the synchronic analysis of
Evidence that the pronominal incorporation property may indeed be typologically significant in predicting the discourse role of independent pronouns, the presence of pro-drop, and the absence of grammatical case-marking on nominals is provided by Coleman's work on Kunparlang, a non-Pama-Nyungan language spoken in Northern Australia. Coleman 1985a observes:

No morphological marking occurs on Subject or Object nominals; moreover, word-order is largely unconstrained. Verbs govern their Subjects and Objects by means of cross-reference prefixes. Historically, these prefixes are incorporated pronouns; synchronically they also function as anaphoric pronouns in discourse, where heavy ellipsis of nominal arguments of predicates occurs. (Analytic pronouns may not be used anaphorically in discourse except to introduce new topics and to provide deixic contrast between arguments.)

Elsewhere, in a discussion of topic, pronominalization, and grammatical agreement, Coleman 1985b notes that the subject prefix (SA) in Kunparlang is obligatory, while the object prefix (OA) is optional, used under complex conditions of discourse salience including the following condition:

... when both Subject and Object arguments are people, the presence of the Object agreement affix indicates that the Object argument is sentential Topic:

/nga-pun-pum/
i1sgSubj-3sgObj-hit Pst Real
'I hit him'

/ngirra nga-pun-pum/
(1) that 1sgSubj-3sgObj-hit Pst Real
'That (male one), I hit him'

She then makes this remarkable observation:

In Kunparlang, it is possible to question either Subject or Object argument. When the Object argument is questioned, the Topic-marking OA affix may not occur; this is predicted by the definitions given above for the pragmatic notions of Topic and such languages; for example, verbally governed grammatical case-marking on a given nominal argument is completely inconsistent with the anaphorically linked topic analysis of that nominal. Fassi-Fehri's 1984 work on Arabic provides a striking illustration in support of this consequence.
Focus. When the Subject argument is questioned, however, the Subject affix always occurs; this suggests that the SA affix does not necessarily function to mark the subject as a Topic. For example:

Questioning the Subject argument:

/na-gaypi ka-ngun-pum/
1-who 3:Subj-2sgObj-hit Pst Real
‘Who hit you?’

Questioning the Object argument:

/na-gaypi ki-pum/
1-who 2sgSubj-hit Pst Real
‘Who did you hit?’

/na-gaypi ki-pun-pum/
1-who 2sgSubj-3sgObj-hit Pst Real
‘Who did you hit him?’

These remarkable parallels between Chichewa and Kunparlang suggest that the morphological incorporation of pronouns into predicators represents a fundamental typological property from which a theory of grammatical structure and discourse functions can derive a variety of deeper characteristics.

Another clue that the pronominal incorporation property may indeed be typologically significant is provided by an observation of Chafe (1976: pp. 37–8) about Iroquoian:

In some languages, where the role of given nouns is captured primarily through agreement in the verb, independent pronouns appear to be used mainly to express a focus of contrast. In Seneca, for example, and in the Iroquois languages generally, a first person referent is normally expressed only through a verbal prefix. There is, however, a separate Seneca pronoun i’ ‘I’ which appears typically in sentences like:

1 ononó’tá’ kyéthwas
I potatoes I-plant
I plant potatoes.

The context might be, “Other people may plant other things, but ...” This is clearly a contrastive function. The independent
pronouns for other persons and genders are typically used in the
same way."

The same clustering of properties has been observed to occur in other lan-
guages that have the pronominal incorporation property, such as Cree, an
Algonquian language (Dahlstrom in preparation), and Lakota, a Siouan
language (Van Valin 1985).

Finally, evidence from Aghem, a Grasslands Bantu language spoken in
Cameroon, may also be significant for our typological hypothesis. From Hy-
man ed. 1979, we see that Aghem is clearly an isolating language, lacking
the pronominal incorporation property. Its pronouns are independent, be-
ing conjoinable and separable from the verb by direct objects. There are
no pronominal prefixes on the verb. Texts show that both subject and ob-
ject (independent) pronouns allow anaphora to topic, and that pro-drop is
rarely if ever used. Moreover, there are different morphological forms for
subject and nonsubject pronouns, suggesting case-like differentiation. Thus,
although Aghem is a Bantu language, it is typologically different in all three
of the properties implied by the pronominal incorporation property: the
contrastive use of independent pronouns, the presence of pro-drop, and the
absence of verbally governed case-marking.

In sum, our theory implies that pronominal incorporation can be distin-
guished from grammatical agreement typologically by a cluster of at least
three properties: the contrastive discourse role of the independent pronouns,
the presence of pro-drop, and the lack of verbally governed grammatical case
marking on the nominal that is anaphorically linked to the incorporated pro-
noun. These are all typological properties of Chichewa and Kunparang they
all appear to be lacking in a nonincorporative Grasslands Bantu language,
Aghem. Where both case-marking and pronominal incorporation are found,
as in Arabic, our theory correctly predicts a complementarity in their dis-
tribution, as found by Fassi-Fehri 1984.

4 Sources of Variation

Our analysis of Chichewa in Section 1 (iv) assumes that the subject and
topic NPs appear at the same level of structure in the S, i.e. with exactly the
same ordering possibilities. An alternative hypothesis is that the structural
position of the subject is fixed in Chichewa as [NP VP], and the post-VP
subject is really a postposed (right-dislocated) topic anaphorically linked to
the subject agreement marker (which is optionally pronominal, as we have
seen). The latter analysis would predict that the subject in VP-final posi-
tion cannot be questioned in place, for in that position the apparent subject

Such an analysis has been suggested for Kiifaya by Byaruhanga and Tenenbaum 1976.
is actually a postposed topic, and hence incompatible with the question
word's FOC function. But in Chichewa, the question word can follow the
VP.

(46) (Kodi) chi-ná-ónék-a chifyáni?
    Q SM(7)-past-happen-indic what(7)
    'What happened?'

This confirms that the subject NP in Chichewa is unordered with respect to
the VP.

Another alternative analysis is that both the subject and topic NPs are
postposable, but the topic lies outside of the subject structure at a higher
level of S (or $S$).

\[
\begin{align*}
S & \rightarrow \text{NP} & S \\
& (\uparrow \text{TOPIC}) = \downarrow & S \\
S & \rightarrow \text{NP} & \text{VP} \\
& (\uparrow \text{SUBJ}) = \downarrow & \text{VP}
\end{align*}
\]

Because of the independence of structure and function in our theory, gram-
matical functions need not be represented by distinctive phrase structural
configurations in this way: the choice between the flat-structure topic analy-
sis in Section 1 (iv) and the hierarchical analysis in (47) is thus an empirical
issue. If the topic NP is generated either initially or finally, at a higher
S level than the subject NP, then the subject must always be adjacent to
the VP. On this analysis, therefore, the [V TOP SUBJ] order could only
be generated by analyzing the final S as an apparent subject. The apparent
subject would actually be another topic NP generated at the topmost level
of S structure and anaphorically linked to the SM, which optionally functions
ronominally, as we have seen:

(48)

\[
[S [S [S \text{SM} -\text{OM} \text{-V}] \text{TOP}] \text{TOP}]
\]

Since interrogative words cannot be topics, this hypothesis predicts that in
non-cleft questions, questioning the subject should be possible only when
the subject is adjacent to the VP. Our analysis, in contrast, predicts that in
such cases questioning the subject should be possible even when the subject is separated from the VP by a topic NP. Observe now that the question word subject can be separated from the verb phrase by a topic NP linked anaphorically to the OM.

\[(49)\]

\[\text{a. (Kodi) chi-ku-f-fdn-â mîchfrâ yâ mbeâa chiyâni [V TOP S]}
\]

Q SM-pres-OM-want-indic tails of mice what

"What wants them, mouse tails?"

\[\text{b. (Kodi) chiyâni mîchfrâ yâ mbeâa chi-ku-l-fân-â [S TOP vi]}
\]

Q what tails of mice SM-pres-OM-want-indic

"What wants them, mouse tails?"

This confirms that the subject NP is at the same level as the topic NP, as in our analysis given in Section 1 (iv).

We see, then, that there is good evidence for the analysis we gave in Section 1 (iv), in which the SUBJ and TOP NPs occur unordered at the same level of sentence structure. Although the SUBJ function is grammatically distinguishable from the TOP function in Chichewa, as we have seen, the subject NP is indistinguishable from the topic NP in its phrase structure properties.

Phrase structures on our theory, like word structures, vary across languages. These are the grammatical structures that give external expression to the abstract functional structure; they are the phonologically interpreted structures. Therefore, we expect other languages to fix properties of the phrasal structure that encodes the TOP function in different ways. Indeed, many Bantu languages differ from Chichewa in fixing the topic in sentence-initial position. In such languages the OM will be in obvious complementary distribution with the object NP in the verb phrase. The postposable topic construction that we find in Chichewa (and also in KiHaya (n. 16)) masks the pronominal status of the OM by giving the appearance of an agreement marker cooccurring with an object NP.

In our theory, the difference between an incorporated object pronoun and a grammatical object agreement marker is merely the presence or absence of the referential property, which is represented by the semantic PRED feature. There are exactly three possibilities that can arise: the PRED feature is obligatory, it is optional, or it is absent. In Chichewa, the OM illustrates

\[\text{In our theory, the phrase-structure properties are dominance, precedence, and structural category, as determined by word order, word structure, the post-lexical phonological interpretation of phrasing, and the like. Case government, agreement, and anaphoric binding are determined at structure.}

\[\text{Examples include Dzamba (Bokamba 1976), Kikuyu (Bergvall to appear), and Kichaga (Lloha Moahi, personal communication). According to Wald 1979, this construction is the most widespread and represents the older Bantu pattern.} \]
the first possibility and the SM, the second possibility. We must look
to non-pro-drop inflections, such as English subject agreement, for the third
possibility. This theory predicts the existence of true grammatical object
agreement parallel to true grammatical subject agreement.

In fact, some Bantu languages are now undergoing grammaticization of
the pronominal OM into an object agreement marker, parallel to the hypo-
thesized earlier evolution of the SM (Givón 1976, Wald 1979). In our theory,
what must happen in this process is simply the loss by the pronominal OM
of its PRED feature. Once the PRED feature is lost, functional uniqueness
will no longer prevent the cooccurrence of the OM with an object NP within
the verb phrase. The uniqueness condition will require only that all of the
remaining pronominal features—number, gender class, and person—be con-
sistent with the features of the NP object.40

Such a development has happened in Makua (Stucky 1981, 1983), and
appears to be underway in Kiswahili (Wald 1979). In the Imithupi dialect
of Makua studied by Stucky, the OM is obligatory with the human classes:41

(50) a. Aráárima á-hó-líh-a mwaáná
    Aráarima SM-T/A-OM-feed-T/A child
    ‘Aráarima fed a child.’
b. *Aráárima á-hó-líh-a mwaáná
    Aráarima SM-T/A-feed-T/A child

Given the context provided by example (50)a, one can ask the question
shown in (51)a. Example (51)b shows that the OM is obligatory with the
interrogative object as well.

(51) a. Aráárima a-n-líh-fre mpánf
    Aráarima SM-OM-feed-T/A who
    ‘Who did Aráarima feed?’
b. *Aráárima a-líh-fre mpánf
    Aráarima SM-feed-T/A who

In standard Kishwahili, according to Bokamba 1981, the occurrence of the
object prefix is optional when the object is inanimate, but obligatory when
it is animate:

(52)
a. Maryamu a-li-wa-onyesha watoto kisu
   Maryamu SM-past-OM-show children knife
   'Maryamu showed the children a/the knife.'

   Maryamu SM-past-OM-show children knife

In another context, Bokamba shows that the animate object can be questioned in place in Kiswahili, and cooccurs with the OM:

(53) a. Bakari a-na-wa-som-e-a watoto hadithi maktaba-ni?
    Bakari 3M-pres-OM-appl-indic children stories library-loc
    'Bakari is reading stories to/for the children in/at the library.'

b. Bakari a-na-wa-som-e-a nani hadithi maktaba-ni?
    Bakari SM-pres-OM-appl-indic who stories library-loc
    'To/for whom is Bakari reading stories in/at the library?'

If these are true cases of grammatical object agreement, then our theory predicts a range of correlated phenomena, which future research must test.

Finally, in some Bantu languages, an interrogative pattern occurs which seems at variance with that of Chichewa, in that the subject cannot be questioned in place. In Dzamba, for example, it is possible to question all VP constituents in place, but subjects cannot be questioned in the initial subject position (Bokamba 1981). To question a subject, it is necessary to use a differ-at construction altogether, based on a headed or headless relative clause:

(54) a. 6-Nebo a-imol-aki 6-Biko e-kondo loo mé
    'Nebo told Biko a story/tale today.'

b. 6-Nebo a-imol-aki nzanyi e-kondo loo mé?
    'Nebo told who a story today?'

c. *Nzanyi 6-wimol-aki 6-Biko e-kondo loo nyi?
    'Who told Biko a story/tale today?'

d. '6-Moto 6-wimolaki 6-Biko e-kondo loo nzanyi?
    'The person who told Biko a story/tale today is who?'

What could be the explanation for this pattern on our theory? Observe that precisely this result would follow if in such languages the sentence-initial position for the SUBJ function also had the TOP function. Now in Dzamba, unlike Chichewa, there are nominal preprefixes. Bokamba 1981 has shown that these prefixes are used to definite noun phrases. They are obligatory on subjects and they are obligatory in topicalizations.42

42It would be simplistic, however, to identify the preprefixes solely as topic markers. For discussion of some of the semantic complexities of their use in Chichewa, see Givón 1969.
In this way our theory of argument and discourse functions may illuminate a range of variation. If we looked only at the structural aspects of agreement, the real generalizations would never emerge, because at that level the facts conflict: object agreement occurs with interrogatives in Makua and Swahili, but not in Chichewa and Dzamba; subjects can be questioned in place in Chichewa but not in Dzamba; and so forth. At this level, the facts are chaotic. But once we see that each language encodes the same functions in slightly differing ways, the results appear totally predictable.

At the same time, our study indicates that important parameters of change and variation lie in surface form—the external, phonologically interpreted morphology and phrase structures. The pronominal incorporation property is such a parameter. Together with our postulates about grammatical theory and discourse function, it appears to explain fundamental differences between syntactic structures organized by grammatical agreement with governed functions on the one hand, and, on the other, those organized by anaphoric agreement with discourse functions.
References


Bokamba, Eyamba G. (1975) "Observations on the Immediate Domi-


Bresnan, Joan (1982a) "Polyadicity", in J. Bresnan, ed.

Bresnan, Joan (1982b) "Control and Complementation", Linguistic Inquiry 13; also in J. Bresnan, ed.


Bresnan, Joan and Sam A. Mchombo (forthcoming) "Agreement and Pronominal Incorporation in Chichewa"


Chiba, Shuji (1973) "On the Movement of Post-copular NP's in English", *Studies in English Linguistics* 2, pp. 1-17.


Coleman, Carolyn (1985b) "Topic, Pronominalization, and Grammatical Agreement", unpublished paper, Department of Linguistics, Stanford University.


Givón, Talmy (1976) "Topic, Pronoun, and Grammatical Agreement", in Charles N. Li, ed.


Greenberg, Joseph (1978) "How Does a Language Acquire Gender Markers?", in J. H. Greenberg et al., eds.


Kuno, Susumu (1976) "Subject, Theme, and the Speaker's Empathy-A Reexamination of Relativization Phenomena", in Charles N. Li, ed.


Li, Charles N., ed. (1977) *Mechanisms of Syntactic Change*, University of Texas Press, Austin.

Li, Charles N. and Sandra A. Thompson (1976) "Subject and Topic: A New Typology of Language", in Charles N. Li, ed.


York, 123-148.


