

On the Status of Object Markers in Bantu languages¹

Mark C. Baker
Rutgers University
January 2016

1. Introduction

A classic issue in the syntax of Bantu languages is the status of “object markers” (OMs)—a topic that has much discussed since at least Bresnan and Mchombo’s (1987) classic study of Chichewa. Chichewa, for example, allows the range of structures seen in (1) (among others).

- (1) a. Njuchi zi-na-lum-a alenje. (p. 744)
bees SM-PST-bite-FV hunters
‘The bees bit the hunters.’
- b. Njuchi zi-**na**-wa-lum-a.
bees SM-PST-OM-bite-FV
‘The bees bit them.’
- c. Njuchi zi-na-**wa**-lum-a alenje. (p. 744)
bees SM-PST-OM-bite-FV hunters
‘The bees bit (them,) the hunters.’

Much of this literature has concentrated on choosing between one of two analyses: either the OM is a pure agreement marker, analogous to subject markers/agreement in Bantu and other languages, or it is a cliticized pronoun that counts as the true object of the verb even though it appears on the verb on the surface. On the latter analysis, the overt NP in (1c) is not the true grammatical object of the clause, but rather some kind of dislocated adjunct, while the OM is the true object in (1c) as it is presumably in (1b).

Bresnan and Mchombo themselves use various diagnostics to argue in favor of the cliticized pronoun analysis (they call it an incorporated pronoun), and their conclusion seems to carry over to many other Bantu languages—for example, Lubukusu, Kinande, and Zulu. But there are a great many Bantu languages, and there is substantial diversity among them—including diversity in the behavior of OMs. Hence it is quite unlikely that “one size fits all” in this respect. In contrast, Riedel (2009) argues that in the superficially similar paradigm in Sambia shown in (2), the OM *-mu-* is best understood as being a realization of object agreement.²

¹ This paper is dedicated to Professor Yiwola Awoyale (with a twinge of regret that I have nothing new to say about serial verbs...) and with him the many great African linguists working on their languages, often despite many practical difficulties, whom it has been my pleasure to meet. May their numbers only increase!

I thank Ruth Kramer and Ken Safir for discussions of many of the issues touched on here. Errors are my own responsibility

² I have sought to make the glosses of Bantu examples taken from different sources and different languages as uniform as I could, in the interests of comparability. Abbreviations used in the glosses include the following: ACC, accusative; APPL, applicative; AUX, auxiliary; CAUS, causative; CJ, conjunct

- (2) a. N-za-ona ng'wana.
 1sS-PERF.DJ-see 1child
 'I saw the/a child.' (Riedel 2009:46)
- b. Ni-ta-za-na-**mu**-itanga.
 1sS-PRES.DJ-PERF.DJ-MD-OM1-call
 'I might possibly call her.' (p. 31)
- c. N-za-**mw**-ona ng'wana.
 1sS-PERF.DJ-OM1-see 1child
 'I saw the/a child.' (Riedel 2009:46)

This is also a common view for at least the obligatory OM that goes with human (class 1 and 2) objects in many Swahili dialects. And by all accounts there are many mixed or transitional cases too—perhaps more mixed cases than clear ones. With this in mind, Riedel (2009) claims that there is a continuity in behavior across languages, and that no sharp boundaries can be drawn. From this impression, taken together with her conviction that OMs are agreement in Sambia, she infers OMs must be agreement in more or less all of the Bantu languages.

Less radically, but in a similar spirit, Zeller (2012) claims that OMs in Zulu are in the process of changing from clitic pronouns to agreement, but they currently have a somewhat mixed transitional status. He thus conjectures that we should be looking for some kind of intermediate analysis, different from both the pure pronominal clitic analysis and the pure agreement analysis. After all, it is all very well to talk about historical changes in process, but the “transitional” Bantu languages must have some kind of coherent grammatical representation in the minds of the many people who speak them.

In this paper, I seek to provide a third analysis along the lines that Zeller calls for, one that mixes aspects of an agreement account and a pronoun account in a principled way. I also claim that, once we know more precisely what to look for, we may not find continuity among the Bantu languages in the way that Riedel claims, but rather evidence of some bona fide parametric clustering. The fact that this clustering is centered around three stable prototypes rather than just two helps to give the appearance of continuity, and this appearance is further enhanced by there being other variation in (what I take to be) the “accidents” of the construction—properties that do not directly concern the status of the OM as agreement or pronoun.

The new(ish) analysis that I present here is an analysis in terms of true clitic doubling. In particular, I appeal specifically to clitic doubling as it is found in another, quite different African language, namely the Ethiosemitic language Amharic, as described and analyzed in recent work

form; DAT, Dative; DEF, definite; DJ, disjunct form; FOC, focus; FV, final vowel (mood marker); GEN, genitive; M, masculine; MD, mood; NEG, negative; OM, object marker; PERF, perfect(ive); PF, perfective (Amharic); PL, plural ; POSS, possessor ; PRES, present; PST, past tense; PTPL, participle; Q, question particle; REFL, reflexive; RS, relative suffix; SM, subject marker; T, tense marker. Glosses for agreement markers in Bantu include 1sS for first person singular subject and 2sS for second singular subject, Glosses for agreement markers and OMs in Amharic follow a similar format, but see Baker and Kramer (2015) for further details. Apart from these, roman numerals refer to noun classes in the various Bantu languages.

by myself and Ruth Kramer (2015) (henceforth B&K). Amharic has paradigms like (3), clearly analogous to (1) and (2) from Bantu languages, except for predictable differences in word order and some (more idiosyncratic) differences in morpheme order.

- (3) a. Läm̄ma **wiffa-w-in** ayy-ä
 Lemma dog-DEF-ACC see.PF-3MS.S
 Lemma saw the dog.
- b. Läm̄ma ayy-äw
 Lemma see.PF(.3MS.S)-3MS.O
 Lemma saw it/him.
- c. Läm̄ma **wiffa-w-in** ayy-äw
 Lemma dog-DEF-ACC see.PF(.3MS.S)-3MS.O
 Lemma saw the dog.

(3a), like (1a) and (2a), has a direct object and no OM, although in Amharic the direct object comes before the verb, not after it, in keeping with Amharic's status as a head final language. In (3b), like in (1b) and (2b), the equivalent of a weak pronoun is expressed as a kind of morpheme attached to the verb in a different place in the string from where the full DP object would be: after the verb root in Amharic; before the verb root in the Bantu languages. In (3c), both the OM and the object co-occur, as they do in (1c) and (2c). So Amharic is clearly comparable to the Bantu languages in these respects, despite belonging to a different family, not being in direct contact with the Bantu languages, and having different ordering properties for words and morphemes.

What is particularly interesting is that Amharic has mixed properties compared to prototypical Bantu languages with pronominal OMs and prototypical Bantu languages with agreement OMs. First, the overt object nominal does not (necessarily) seem to be dislocated to the edge of the clause. It is neither at the right edge of the clause, nor at the left edge of the clause in (3c); rather, it follows the subject and precedes the verb. Nor is there a noticeable intonation break between it and the verb. Also possible are examples like (4), where the object follows an adverb—indeed, the sort of adverb that is usually low in the clause (Cinque 1999).

- (4) Läm̄ma ahunimm **wiffa-w-in** y-ay-äw-al.
 Lemma still dog-DEF-ACC 3MS.S-see-3MS.O-AUX(3MS.S)
 'Lemma still sees the dog.'

In this respect, the OM in Amharic seems to behave like agreement—like Sambia, not like Chichewa. However, the OM in Amharic is like a pronoun in that it is incompatible with less-than-fully referential objects. For example, the *wh*-phrase in-situ in (5) cannot be doubled by an OM.

- (5) Girma tinant männ-in ayy-ä(*-w)
 Girma.M yesterday who-ACC see.PF-3MS.S-3MS.O
 'Who did Girma see yesterday? (Kramer 2014:601)

In this respect, OMs in Amharic are like OMs in Chichewa (Bresnan and Mchombo 1987:760) and not like OMs in Smbaa

- (6) a. (Kodi) mu-ku-(??chi)-fun-a chiyani? (Chichewa)
 Q 2sS-PRES-OM7-want-FV what.7
 ‘What do you want?’
- c. U-wa-ene (wa)-ndayi? (Smbaa)
 2sS-OM2-see.PERF (PL-)who
 ‘Who (all) did you see?’ (Riedel 2009:158)

In short, then, the Amharic OM construction has properties that are truly in between those of the two Bantu prototypes. It is thus worth investigating the prospects of generalizing the account of clitic doubling in Amharic to Bantu and seeing what this does for our understanding of the diversity of OMs in Bantu. That is what I do here, in a somewhat preliminary fashion—preliminary in that the literature relevant to this topic is vast, and I cannot survey or systematize all of it here. Rather, I concentrate on a few languages and issues that seem to me to be immediately relevant, paying special attention to Riedel’s (2009) study, for reasons that should become apparent.

My discussion unfolds in the following stages. In section 2, I review the Baker-Kramer account of clitic doubling in Amharic, showing how it accounts for the mixture of properties in (3)-(5) by saying that the object nominal—technically a “determiner phrase” (DP)—moves to Spec vP (object shift) but then is reduced down to its D head, resulting in the OM. This D head is then interpreted as a pronoun at LF, which derives restrictions like the one in (5) from known conditions of grammar, such as the Weak Crossover Condition. Along the way I discuss the prospects for adapting the account to the overall grammatical design of the Bantu languages, and I review how B&K distinguish languages like French—and Chichewa—which have clitics but not true clitic doubling from languages like Amharic which do allow clitic doubling. In section 3, I apply the account to particular Bantu languages, considering especially the Bantu languages of Tanzania, which seem especially agreement-like, based on Riedel 2009. I show that OMs in Smbaa and one variety of Swahili are indeed agreement, as Riedel says, but another dialect of Swahili and Haya are systematically different, having clitic doubling in the Amharic sense. There are other differences, particularly in terms of optionality, obligatoriness and animacy restrictions, but these are not directly relevant to the agreement/pronoun distinction, I claim. Section 4 outlines what would remain to be done from this perspective, and concludes.

2. Clitic doubling in Amharic

2.1 Overview of the analysis

The paradox of a true clitic doubling construction is that both the full DP and the OM seem to be objects. The DP’s claim to be an object is that it seems to be in the true direct object position in examples like (3) and (4)—unlike the object in (1c) in Chichewa, which is outside the VP, as Bresnan and Mchombo show using evidence from word order and from phrasal phonology. The OM’s claim to be the object comes from the fact that it is interpreted as a pronoun; that is presumably why it is incompatible with nonreferential DPs like *wh*-phrases, as shown in (5). If it

were true agreement, there should be no incompatibility. For example, having ‘who’ in subject position in Amharic is perfectly compatible with there being third masculine singular subject agreement on the verb, as shown in (7).

- (7) Man mät’t’a? (Leslau 1990:68)
 who come.PF-3MS.S
 ‘Who came?’

B&K’s strategy for resolving this paradox is to get the fact that there seem to be two objects out of the idea that the object moves. This is a tempting theoretical possibility because movement intrinsically involves copying within Minimalist-style theories (see Chomsky 1993, 1995). The sort of derivation that we have in mind is outlined in (8) for a sentence like (3c).

- (8) a. [TP Lämmä [VP [VP [DP D [NP dog]] see]] v] T] → v agrees with DP
 b. [TP Lämmä [VP [VP [DP D [NP dog]] see]] v] T] → DP moves (copies into) Spec vP
 c. [TP Lämmä [VP [DP D [NP dog]] [VP [DP D [NP dog]] see] v] T] → DP reduces to D
 d. [TP Lämmä [VP [D] [VP [DP D [NP dog]] see] v] T] → D in Spec vP adjoins to v
 d. [TP Lämmä [VP [∅] [VP [DP D [NP dog]] see] v+D] T] (surface form)

I now discuss each stage in this derivation in more detail, together with the prospects of applying the same operations in a Bantu language.

2.2 The role of agreement in clitic doubling structures

The first step in (8) is to have v enter into an Agree relationship with the object DP that will end up doubled. Such agreement is taken to be a precursor to movement in many instances (Chomsky (2000) says that this holds in all cases, but that claim might be too strong). For Amharic, using Agree in this way captures the locality that is characteristic of clitic doubling structures. In particular, when there are two DPs inside the VP, only the higher one can be clitic doubled—the goal argument ‘Almaz’ but not the theme argument ‘book’ in (9), for example.

- (9) a. Girma **lä-Almaz** mäs’haf-u-n sät’t’-at (*sät’t’-ä-w)
 Girma.M DAT-Almaz.F book-DEF.M-ACC give(3MS.S) -3FS.O give-3MS.S-3MS.O
 ‘Girma gave the book to Almaz.’ (Kramer 2014:600)

This restriction holds because when v probes down into VP it necessarily finds the goal before theme and Agrees with it, given that the goal c-commands the theme (cf. Barss and Lasnik 1986, Larson 1988, Marantz 1993), and v c-commands them both.

Indo-European languages like Greek seem to be different from Amharic in this respect, in that one can often double the theme argument even in the presence of an undoubled goal, as shown in (10a). But the difference turns out to be more apparent than real. Anagnostopoulou

(2001) shows that in constructions in which the goal has the same case as the theme (accusative), only the goal can be clitic doubled, as in (10b). This is just like Amharic in (9).

- (10) a. (To) edosa tu Petru to vivlio (EA 2001:15)
 it.ACC gave.1sS the.GEN Peter.GEN the.ACC book.ACC
 ‘I gave (it) Peter the book.’
- b. (*Tin) didaksa ta pedhia tin grammatiki ton arxion
 ellinikon
 it.ACC taught.1sS the.ACC children.ACC the.ACC grammar the.GEN ancient
 Greek
 ‘I taught (it) the children the grammar of ancient Greek.’ (EA 2001:12, 17)

This supports the idea that Agree plays a role, since Agree crucially depends on the features of both the probe and the goal. When the two DPs have the same features, including case, the probe cannot distinguish them and can only find the higher one, as in (10b). When the two DPs have different features, *v* could probe for the feature dative, or it could probe for the feature accusative. When *v* probes for accusative, the goal doesn’t match, and Agree ignores it, as in (10a) in Greek. In Amharic, *v*’s features are different: it probes for any DP, regardless of case. As a result, it only finds the highest DP regardless of case, so dative intervenes for accusative in Amharic. Exactly what can be doubled then depends on both the features of the DP (accusative versus dative, (10a) versus (10b) in Greek) and on features of the Agreeing head *v* (sensitive to case in Greek, not sensitive to case in Amharic). This is what we expect if the mechanism of Agree, whose job is to matching the features of the two syntactic objects, is involved.

There is reason to think that Agree is relevant in Bantu too—even in languages where the OM has pronominal properties. This is the fact that, in many languages, the OM can only express/be related to the goal argument in ditransitives and applicatives, as in Swahili and Chichewa. This is true for Sambia, where OMs are arguably agreement, for Chichewa, where OMs are presumably cliticized pronouns, and for Swahili as shown in (11)—even dialects that I will claim to have true clitic doubling (see below) (see Riedel 2009 for further discussion).

- (11) Stella a-li-m-pa mtoto kitabu pale. (*a-li-ki-pa) (Riedel p.131)
 Stella SM1-PST-OM1-give 1child 7book 16there (SM1-PST-OM7-give)
 ‘Stella gave the child a book there.’

This is directly analogous to Amharic. And we expect this intervention pattern, given that Swahili, like other Bantu languages, does not have a distinction between dative case and accusative case on objects, as Greek does (indeed, they do not have any case distinctions at all). Therefore, there is utility to saying that Agree is involved in OM phenomena in Bantu languages too—even in those where the OM is a cliticized pronoun and not just a manifestation of agreement.³

³ Of course, there is also variation on this point in Bantu languages: some allow an OM to express either object of a double object construction, and some allow OMs to express both objects simultaneously. This is the famous symmetrical object parameter, distinguishing (e.g.) Kinyarwanda and Kichaga from Chichewa and Swahili (Kimenyi 1980, Bresnan and Moshi 1990, many others). I have nothing new to say here about the nature this parameter. one class of proposals is that locality is somehow parameterized in Bantu languages (e.g. Riedel 2009); an

2.3 Object shift-style movement

The next ingredient in the B&K analysis of clitic doubling is movement of the DP that *v* agreed with to Spec *v*P. This is equivalent to object shift in Germanic languages. This step has two main effects in our analysis. First, it relates clitic doubling to specificity: definite DPs can be doubled, but nonspecific indefinites cannot be, as shown for Amharic in (12). This is parallel to, and related to, the fact that definite DPs can move out of the VP in Dutch and German, but nonspecific indefinites cannot, as studied by many in the tradition of Diesing's (1992) Mapping Hypothesis.⁴

- (12) a. Läm̄ma wiff̄fa ayyä-(*w).
Lemma.M dog.M 3MS.S-see-AUX(3MS.S)
'Lemma saw a dog.'
- b. Läm̄ma **wiff̄fa-w-in** ayyä-w
Lemma dog-DEF-ACC see.PF-3MS.S
Lemma saw the dog.

Second, the fact that the landing site of this movement is *v* fits well with the fact that the OM ultimately shows up attached to the main verb in Amharic, below the finite auxiliary, as shown in the present tense-imperfective example in (13). This is roughly where we expect *v* to be: at the top of the thematic domain that also includes *V*, but below the various tense and aspect heads that may be present in a particular language.

- (13) s'ähafi-wa-n i-fällig-**at**-allä-hu
secretary-DEF.F-ACC 1S.S-want-3FS.O-AUX-1SS
'I will look for the secretary.'

Note that OMs in Amharic are different from clitic pronouns in IE languages like Romance and Greek in this way: in the IE languages clitics attach the highest tense-marked auxiliary, rather than to the lowest verb. This may mean that object shift targets a higher head than *v* in IE, or that clitics initially target *v* but then "climb" to a higher head in IE (cf. Roberts 2010).

alternative is that movement can happen internally to VP in some of the languages, changing the locality relationships (e.g. McGinnis 2001). Either way, the expectation here is that this important dimension variation should be independent of whether OMs are agreement, or undoubleable clitics, or doubleable clitics. That seems to be true. Among languages where the OM can only be the higher object, Chichewa has undoubleable clitics, Mainland Tanzanian Swahili has doubleable clitics, and Zanzibari Swahili has true agreement. Among languages where the OM can express the lower object (under some conditions), Kinyarwanda has undoubleable clitics, Haya has doubleable clitics, and Sambia has true agreement (although it is not a paradigmatic symmetrical object language). Of course a wider typology supporting this claim of logical independence would be most welcome.

⁴Note that there may be crosslinguistic variation in whether/which specific indefinites can undergo this object shift, which I do not analyze here. For example, even specific indefinites cannot be doubled by an OM in Amharic and Greek, but some specific indefinites can be doubled in Spanish and Romanian (see Anagnostopoulou 1999 for a brief survey). See note 17 on the possibility of similar variation among the Bantu languages.

These considerations that point to object shift targeting Spec vP in Amharic are also relevant to at least some Bantu languages, as has been previously recognized (see, for example, Woolford 2001 on object shift in Ruwund, and Baker 2008 for a similar proposal for Zulu). First, some Bantu languages do not like to use OMs with nonspecific indefinite objects, even if they are possible with specific or definite objects (although OMs in other Bantu languages are not sensitive to specificity; see below). (14) gives a pair from a dialect of Swahili that is analogous to (12) from Amharic.

- (14) a. Ni-li-mw-ona mwanawe Swahili (unspecified) p. 46
 1sS-PST-OM1-see 1child.POSS.3S
 ‘I saw his child.’
- b. Ni-li-ona mtoto. Mainland Tanzanian Swahili
 1sS-PST-see 1child.
 ‘I saw a child.’

It is also robustly true that in Bantu languages OMs attach near the main verb, where we expect *v* to be. They attach inside of the tense marker in all instances, and when there are auxiliary verbs, in so-called complex tenses, the OM is consistently on the main verb, not the auxiliary, as in Amharic. (15) is an example from Kinande.

- (15) N-a-by-a n-ga-ki-to-y-a.
 1sS-T-be-FV 1sS-PTPL-OM7-fall-CAUS-FV
 ‘I was making it fall.’

So B&K’s motivations for having a movement that targets vP as part of the account of clitic doubling in Amharic readily carry over to Bantu languages too.

2.4. Reduce

The next step in the derivation in (8) is the most distinctive one in the B&K account: the DP that has moved to Spec vP reduces to its D head. In other words, everything inside DP other than D itself is deleted, including the NP complement of D, any possessor that might be in Spec DP, adjuncts that might adjoin to DP, and so on. We further assume that a D head that has phi-features but no NP complement is essentially a (weak) pronoun, following Paul Postal’s idea that pronouns are essentially intransitive determiners. Part of what this means is that a weak pronoun can surface in (or near) Spec-vP, but an overt DP cannot. This is perhaps even clearer in Bantu languages, which have Spec-Head-Complement word order, so it is relatively easy to see whether an element is in specifier or complement position, and if it is a specifier, which head it might be the specifier of, as compared to a Spec-Complement-Head language like Amharic. For example, it is plausible to say that the OM in an example like (1b) or (14a) is in (or near) Spec vP, but a full DP certainly cannot appear in this position; nothing like (16) is possible.

- (16) *Ni-li mwanawe ona Swahili
 1sS-PST 1child.POSS.3S see
 ‘I saw his child.’

For B&K, the operation Reduce also plays an important role in deriving clitic doubling, because it bleeds deletion of the lower copy in the Object Shift chain, allowing an overt DP to appear in situ inside the VP along with the D (=OM) created by Reduce. This accounts for one side of the paradox about true clitic doubling structures: the fact that in some languages one can have an object pronoun and a full object both, without the full DP being dislocated. Generally speaking, according to standard proposals like that of Nunes (2004), the lower of two copies in a movement chain is deleted (unless another condition intervenes, perhaps). However, Reduce applying to the higher copy makes the two copies nonidentical, so that neither deletes, according to Kandybowicz’s (2007) amendment to Nunes’s principles. This is crucial to explaining why an example like (4) is possible in Amharic—and if we discover that true clitic doubling is attested in the Bantu family (sections 3.2 and 3.3), it will be important to explaining that too.

Third, by transforming a DP into a bare D head, Reduce essentially creates a new pronoun that would not otherwise have been present in structure in Amharic examples like (3b). LF then interprets the D in Spec vP as a pronoun, distinct from the DP object to which it is derivationally related, for purposes of conditions like the Weak Crossover Condition and the Binding theory. It is then the impossibility of properly relating the pronoun in Spec vP with the quantificational DP inside VP that induces the badness of examples like (5) in Amharic. Reduce thus is also the key to understanding the other side of the clitic doubling paradox, the fact that the OM seems to an object inasmuch as an object pronoun is present in the structure. I defer until section 2.6 the task of filling out how the assumption that the OM is a pronoun in (roughly) Spec vP contributes to an account of why examples like (4) are bad in Amharic, and discuss in section 3 whether the same forces can be seen at work in Bantu languages or not.

An important feature of B&K’s conception of Reduce is that it is parameterized, such that it can happen in some languages (and in some constructions in a given language) but not others. Indeed, we took this idea from Diercks and Sikuku’s (2013) analysis of OMs in the Bantu language Lubukusu. It is well-known that languages that have pronominal clitics can vary in whether they allow true clitic doubling or not. For example, French, Italian, Spanish, and Romanian all have pronominal clitics, but the first two do not allow clitic doubling whereas the last two do (I refer to the “standard” varieties of French and Italian, and to South American Spanish; nonstandard dialects also vary in this respect). More precisely, French and Italian do not allow an object pronoun to be doubled by a direct object in situ, although it is possible for a sentence to have an object pronoun on the verb and a “dislocated” DP associated with it, in the so-called clitic left dislocation and clitic right dislocation constructions (see, e.g., Cinque 1990). Lubukusu seems to be very much like French and Italian in this respect.

- (17) a. N-a-**ba**-bon-a.
 1sS-PST-**OM2**-see-FV
 ‘I saw them.’ (Diercks and Sikuku 2013:9)
- b. N-a-(***ba**)-bon-a **baa-somi**
 1sS-PST-**OM2**-see-FV 2-students
 ‘I saw the students.’ (Diercks and Sikuku 2013:9)
- c. babaana, mayi a-ba-siima
 2-2-children 1mother SM1-OM2-like

‘The children, the mother likes them.’ (p. 11).

- d. N-a-ba-bon-e (likolooba), baa-somi (p. 12, 13)
1sS-PST-OM2-saw-PST yesterday 2-students
‘I saw them yesterday, the students.’

Other Bantu languages with the sort of profile include Chichewa (Bresnan and Mchombo 1987), Kinande (Baker 2003), and Zulu (Zeller 2012), among others. Indeed, the majority of Bantu languages outside of roughly Tanzania may be like this.

With this sort of variation in mind, B&K proposed that *v* heads across languages are parametrized in terms of what sort of EPP features they have. These features specify what kind of phrase *v* tolerates as its specifier, and what kind of syntactic operations it sanctions in order to meet the requirement. We imagine at least the following range of possibilities:

- (18) a. [EPP: Max] Dutch, German, Icelandic object shift
b. [EPP: Min] MSc object shift, French, Lubukusu
c. [EPP: Reduce-to-Min] Amharic, Bulgarian, Greek, Rioplatense Spanish
d. [EPP: null] English, etc.

(18a) says that *v* can allow any kind of DP in its Spec position; this possibility is familiar from studies of object shift in a range of Germanic languages, including Dutch, German, and Icelandic. At the other extreme, (18d) says that *v* does not allow any DP in its Spec position at all, except perhaps a null DP, i.e. the trace of *wh*-phrase that has moved through Spec *v*P on its way to Spec TP to satisfy cyclicity/locality requirements related to phases (Chomsky 2001). This is the familiar case of English. (18b) is a new value for this feature: the idea is that *v*P allows a specifier, but only a small one, a single (functional) head, perhaps for prosodic reasons.⁵ Apart from clitic languages, the difference between (18a) and (18b) can be seen in Scandinavian languages, in that full DPs can undergo object shift to a higher position (above VP-adverbs) in Icelandic, but only weak pronouns can do so in Danish, Swedish, and Norwegian. The contrast between examples like (1b) with an OM as opposed to the ungrammatical (16) with a whole fronted DP in Standard Average Bantu can be seen as essentially the same.

Within this general conception, (18c) is crucially different from (18b) in that a *v* with this feature not only insists that its head needs to be small—only a D^0 —but can invoke the operation Reduce to achieve this goal. This then is what B&K take to be the root of the difference between (for example) French and Greek. The *v* head in French says that its Spec can only be a D^0 , and no operation is triggered to make it so. The implication is that a DP that moves to Spec *v*P in French must be a D^0 from the beginning, generated as such in the object thematic position. It follows, then, that no full DP can be generated in the object position: either there is no other phrase associated with the object theta role, or there is a DP generated in the clause periphery and linked to the pronominal object by coreference. This analysis looks appropriate for

⁵ The conjecture that there might ultimately be a prosodic motivation for the parameter settings in (18) is based on the following. If T is the highest head in the clause, Spec TP will be at the edge of the clause, where it is easy to pronounce as a separate prosodic phrase if necessary. Hence T is usually (but not always) specified as [EPP Max]. However, *v* is not the highest head in the clause, so Spec *v*P is not generally at the edge of the clause, but internal to it. This might create prosodic problems if Spec *v*P can be of arbitrary complexity and length. [EPP Min], [EPP Reduce-to-Min] and [EPP: null] could be seen as formal tools that languages use to “manage” this potential problem in pronounceability.

- (20) Girma tinant männ-in ayy-ä(*-w)
 Girma.M yesterday who-ACC see.PF-3MS.S-3MS.O
 ‘Who did Girma see yesterday?’ (Kramer 2014:601)

Our strategy was to explain this seeming anomaly by saying that the OM is in fact a pronoun, combining this with the well-known fact that pronouns can only be referentially dependent on *wh*-phrases when certain specific structural conditions are met. For example, the pronoun can be a variable bound by the *wh*-phrase in (21a) but not in (21b) in English; this is the well-known Weak Crossover effect.

- (21) a. Who -- loves his mother?
 b. ?*Who does his mother love -- ?

The contrast between (21b) and (21a) is normally captured by a condition that can be stated as in (22).⁷

- (22) A pronoun cannot be referentially dependent on a quantifier unless it is c-commanded by the quantifier (if it is in its A-position) or the trace of the quantifier (if it has A-bar moved).

I have stated (22) abstracting away from the difference between languages in which *wh*-movement occurs and languages in which *wh*-phrases are left in situ (like Amharic, and an option in many Bantu languages), since languages with *wh*-in-situ generally do not allow examples of the form “His mother loves who?” on a par with the badness of (21b) in English. (This disjunction can be avoided in the familiar way if all *wh*-phrases in fact move at LF.) According to this condition, (21a) is possible on the relevant interpretation because the trace of ‘who’ in the subject position does c-command the pronoun that depends on it inside the direct object. In contrast, (21b) is not possible on the relevant interpretation (it does not have the obvious interpretation of the passive version *Who is loved by his mother?*) because the trace of the *wh*-phrase in the object position does not c-command the pronoun that depends on it inside the subject phrase.

Now by hypothesis the structure of an example like (20) in Amharic is given in (23) (cf. (8d); here I abstract away from Cliticize, which B&K assume leaves a copy in Spec vP anyway).

- (23) [TP Girma [vP [DP D] [VP [DP D [NP who]] see] v] T]

This representation also violates the Weak Crossover condition given in (20), because the pronoun D in Spec vP is not c-commanded by the quantifier DP ‘who’ inside VP; on the contrary D c-commands ‘who’.⁸ The badness of (20) can thus be explained in the same way as the familiar contrast between (21a) and (21b).

⁷ See B&K for discussion of how more recent and technical versions of the Weak Crossover Condition, like those of Safir (2004) and Büring (2005) work to rule out (20). The differences between the various versions are not important here.

⁸ Thus (23) is technically a strong crossover violation rather than a weak crossover violation, according to the standard taxonomy. This distinction is not important here, however.

Part of the enduring interest of the Weak Crossover phenomenon is that it applies not only to *wh*-phrases, but also to noninterrogative quantified expressions. Thus, the possessive pronoun in (24a) can be understood as a variable bound by the negative quantifier in (24a) but not in (24b) in English. This is parallel to the contrast in (21), and also follows from (22).

- (24) a. Nobody loves his own mother.
 b. ?*His own mother loves nobody.

Similarly, the possessive pronoun in (25a) can be a variable bound by the singular distributive universal quantifier in (25a) but not in (25b)—a third instance of weak crossover in English.

- (25) a. Every boy loves his mother.
 b. ?*His mother loves every boy.

Given then that negative quantified expressions and universal distributive quantified expressions are like *wh*-phrases in counting as quantifiers for purposes of the condition in (22), we expect that expressions of this type should also resist being doubled by an OM in Amharic, since the representation would be essentially the same as (23). Indeed, this is true, as shown by (26).

- (26) a. Läṃma mann-in-imm al-ayy-ä-(*wi)-mm.⁹
 Lemma.M one-ACC-FOC NEG-see.PF-3MS.S-(*3MS.O)-FOC
 ‘Lemma saw nobody; Lemma didn’t see anybody.’
- b. Läṃma hullu-n-imm säw ayy-ä. (*ayy-äw)
 Lemma.M every-ACC-FOC person see.PF-3MS.S see.PF(3MS.S)-3MS.O
 ‘Lemma saw everyone.’

Other arguably quantified DPs that cannot be doubled by an OM in Amharic include generic DPs, DPs with a weak quantifier like ‘three’ or ‘many’, and certain kinds of specific indefinites (e.g. those that include forms of *and* ‘one’).

However, the boundaries of the class of quantifiers are known to somewhat subtle, in that superficially comparable elements that intuitively seem to be somewhat more referential do not trigger weak crossover in English (or trigger it to a lesser degree). For example, DPs containing the plural universal quantifiers ‘all’ do not necessarily induce weak crossover effects in English, as shown by the relative acceptability of (27a) (contrast (25b)). Similarly, a D-linked *wh*-phrases of the form ‘which N’ do not necessarily induce weak crossover as shown in (27b) (contrast (21b)—although judgments on this vary a certain amount; see Safir 2015).

- (27) a. Their incautious statements ended up ruining all my friends. (Cinque 1990:11, Safir 2015)
 b. ?*Which picture did the man who painted it refuse to sell? (Wasow 1979, Safir 2015)

⁹ (26a) may not be exactly analogous to (24b) in English, if *mann-in-imm* is better analyzed as a negative polarity item in the scope of negation, marked on the verb, rather than a quantifier that has negative force in its own right. However, the distinction is probably not crucial, since NPIs also show weak crossover asymmetries in English:

- (i) a. I don’t think that anyone will denounce his own father.
 b. ?*I don’t think that his own son will denounce anyone.

The crosslinguistic expectation, then, is that analogous DPs in Amharic should be able to be doubled by an OM on the verb. This is also true. (28a) has a plural OM doubling the plural expression ‘all the students’ (note that ‘every’ and ‘all’ are homonyms in Amharic); contrast (26b). Similarly, (28b) has an OM doubling ‘which student’—a minimal pair with (20).

- (28) a. Almaz hullu-n tāmari-wotʃf agāññ-ätʃf-atʃfäw
 Almaz all-ACC student-PL meet.PF-3FS.S-3PL.O
 ‘Almaz met all the students.’
- b. Almaz tinant yätīñnaw-in tamari ayy-ätʃf-iw?
 Almaz.F yesterday which.M-ACC student see.PF-3FS.S-3MS.O
 ‘Which student did Almaz see yesterday?’ (Kramer 2014:601)

The evidence, then, that the OM counts as a pronoun inducing weak crossover effects in Amharic is quite rich. A crucial fact here is that the class of DPs that cannot be doubled by an OM does not look like an arbitrary or idiosyncratic one: it is rather precisely the class of DPs that one expects to be subject to weak crossover on general grounds. This is very different from the idea that languages might differ in idiosyncratic ways in what can be clitic doubled. The same distinctions seem to be at work in the familiar clitic doubling languages of Europe, such as Rioplatinese Spanish, Greek, and Bulgarian, according to B&K and their sources. This should have diagnostic value in the Bantu languages as well.¹⁰

B&K also use the idea that OMs are intrinsically pronominal in Amharic to explain the fact that an OM cannot double a reflexive anaphor, as shown in (29).

- (29) Läm̄ma ras-u-n gäd̄däl-ä. (*gäd̄däl-äw)
 Lemma.M self-his-ACC kill-3MS.S kill(3MS.S)-3MS.O
 ‘Lemma killed himself.’

Here, however, the engine of explanation for this type of sentence is not the Weak Crossover condition, but rather the Binding theory. The structure of (29) with an OM on the verb would be (30).

- (30) [TP Lemma [vP [DP D] [VP [DP D [NP self]] kill] v] T]

Since D in Spec vP counts as a pronoun, it is subject to Condition B of the Binding theory (cf. Chomsky 1981). As such, it cannot be coreferential with a c-commanding DP in the same clause. But in (29)/(30) it is: namely, with the subject. (And if it were not, then either the anaphor ‘self’ would fail to be bound in the clause, or D and ‘self’ would be referring to different entities, a Theta criterion violation.) So (29)/(30) is ruled out by Binding theory—for essentially

¹⁰ A related prediction might be that OMs could double relative operators—at least in nonrestrictive relative clauses—although they cannot double interrogative operators. This would be related to the observation that WCO effects are found in questions but not always in otherwise similar relative clauses (Lasnik and Stowell 1999). That is true in Amharic, but questions arise about how the possibility of true resumptive pronouns in relative clauses should play into a full account, so B&K do not pursue the matter, and neither will I here. The role of OMs in relative clauses in Bantu is a much discussed aspect of variation across the languages (e.g. Riedel 2009, Henderson 2006, Diercks and Sikuku 2013), so this could be a fertile area for further investigation.

the same reason that an OM that matches the subject in phi-features but does not double a reflexive anaphor cannot have a reflexive interpretation, as shown in (31) (see also (3b)).

- (31) Lämma gäddäl-äw.
 Lemma.M kill(3MS.S)-3MS.O
 ‘Lemma killed him (him≠Lemma).’

Examples like (29) play an important role in the B&K project because they give independent evidence that the OM in Amharic is intrinsically pronominal, apart from the Weak Crossover condition. However, it will not be very applicable to discerning the status of OMs in Bantu languages, simply because canonical Bantu languages do not have overt reflexive DPs analogous to *herself* in English and *rasu* in Amharic (see the discussion of (40) below).

One final application of the B&K approach that is of potential interest is sentences like (32) in Amharic. This is perfectly grammatical, but ‘his’ inside the object cannot be understood as a variable bound by the quantified subject when the object is doubled by a matching OM.

- (32) hullu säw lidz-u-n yi-wädd-(#əw)-all
 Every person child M-his-ACC 3MS.S-love-(3MS.O)-AUX.3MS.S
 ‘Everyone loves his child.’ (Bad as ‘For all x, x loves x’s child’)

There is no agreement-theoretic reason why (32) should be ruled out: ‘his child’ surely has a normal complement of phi-features, and it even counts as a definite DP, which normally facilitates object agreement. But if the OM is taken to be a pronoun, then this effect can be explained in terms of weak crossover, by a generalization of the account of (20) and (26). A phrase like ‘his child’ is not structurally a quantifier, but it is semantically like a quantifier in that it ranges over different semantic values as its possessor ‘his’ ranges over different values by virtue of being bound by the quantifier. It is semantically plausible, then, to say that an expression that contains a variable bound by a quantifier itself counts as a quantifier for purposes of the condition in (22). (See B&K for further discussion, and independent evidence for this from English.) Given this, the subtle and interesting facts summarized in (32) follow immediately. This effect also does not have immediate diagnostic value for Bantu languages, since (as far as I know) such structures have not been much investigated. However, the Bantu languages have the grammatical resources to construct such sentences, so it should be a good place to look for new evidence bearing on the status of OMs in future work.

In this section, then, I have reviewed the idea that OMs in Amharic are the result of Reduce applying to a copy of the object DP in Spec vP, which is then interpreted as a pronoun at LF by familiar principles. This proposal gives an account of the characteristic mixed behavior of clitic doubling constructions: the full DP object is in situ, not subject to chain deletion, but restrictions on the object are derived from saying that the OM is a pronoun distinct from it. Moreover, we have seen that there is no intrinsic barrier to applying the same analysis to OMs in Bantu, given locality, where the OM appears, and it being attached to the verb. The question, then, is whether this analysis fits well the facts of any particular Bantu language(s) discussed to date. In other words, is there evidence that some Bantu languages have true clitic doubling as distinct from both agreement and clitic dislocation constructions.

3. Sorting out agreement and clitic doubling in Eastern Bantu

1sS-PERF-(OM1)-give 1student 7book
 ‘I gave the student a book.’

- b. *N-za-(m)-nka kitabu ng’wanafunzi (Riedel p., 66)
 1sS-PERF-(OM1)-give 7book 1student
 ‘I gave the student a book.’

Again, this contrasts with B&M’s Chichewa (although they may have overstated the badness of OM+V IO DO order in that language; see also Zeller (2012) on the possibility of both NPs being right dislocated, albeit with only one OM in Zulu). So Sambia is not a language in which OMs are possible only with null objects or dislocated objects; rather doubled objects are possible in situ, as they are in Amharic in (3) and (4).

Now consider the effect of varying the referentiality of the object on object marking in Sambia. Over all, it seems that there are essentially no referentiality limits on what can be doubled by an OM in Sambia. Riedel shows that OMs are possible with both definite and indefinite objects (animate or inanimate). For example, (35a, b) can have either interpretation.

- (35) a. N-za-(mw)-ona ng’wana. (p. 46)
 1sS-PERF.DJ-OM1-see 1child
 ‘I saw the/a child.’
- b. N-za-(chi)-ona kitezu.
 1sS-PERF.DJ-OM7-see 7basket
 ‘I saw the/a basket.’

Riedel mentions that neither definiteness or specificity is marked on NPs in Sambia, and her glosses do not reveal whether the indefinite ones are specific or nonspecific. But given that she sees no need to make a distinction in specificity (although she is clearly aware of the issue) and OMs are possible with both definites and nonspecific indefinites, it is reasonable to suppose that they are possible with specific indefinites as well. So for example, an example like (35a) presumably does not rule out a meaning of ‘a certain child’ (one that the speaker was consciously looking for) or a meaning ‘any old child’ (one that the speaker just happened to see, perhaps in a place where no child would have been expected).

Moving to quantificational data, Riedel shows that OMs are possible with negative DPs items in Sambia, as in (36).

- (36) Si-**chi**-on-iyē **kintu chochoshe.**
 NEG.1sS-OM7-see-PERF 7thing 7any
 ‘I didn’t see anything.’ (Riedel 2009:50)

She also shows that they are possible with a range of *wh*-in-situ expressions, including non D-linked ones (‘who’), D-linked ones (‘which child’), and ones that can have a range of meanings (‘how many people’).¹¹

¹¹ However, the OM is not possible doubling ‘what’ or ‘which book’ (p. 156). This is presumably not a difference in specificity, quantification, referentiality – anything relevant to a syntactic principle like Weak Crossover. Rather

- (37) a. U-(**wa**)-ene **wa-ndayi?**
 2sS-OM2-see.PERF (PL-)who
 ‘Who (all) did you see?’ (Riedel 2009:158)
- b. U-(wa)-alik-e wantu wangahi?
 2sS-OM2-invite-PERF.CJ 2people 2how.many
 ‘How many people did you invite?’

Riedel does not talk specifically about nonreferential quantifiers, but she gives the following example in the course of an argument to show that the goal object c-commands the theme object in Sambia.¹²

- (38) N-za-m-somea [kia mwandisi] kitabu chakwe.
 1sS-PERF.DJ-OM1-read.APPL every 1writer 7book 7his
 ‘I read each author his book.’

Note that this universal quantifier is singular, not plural, so it is presumably truly quantificational. Indeed, we know independently that it is the kind of element that is subject to WCO, because Riedel contrasts (39) with (40), which violates WCO.

- (39) N-za-m-somea mwandisi wakwe kia kitabu.
 1sS-PERF.DJ-OM1-read.APPL 1writer 1its every 7book
 No bound reading: I read its author each book.

Nevertheless, the universally quantified IO is doubled by an OM in (38). Therefore, the OM must not be a pronominal, subject to weak crossover in this language. Overall, there is no sign that OMs in Sambia are pronominal in the sense of inducing discernable weak crossover effects. Importantly, this holds uniformly for the whole class of known quantificational elements.

Ideally, we would like to confirm this by seeing if an OM can double a reflexive anaphor, as they can in (say) Burushaski (B&K), but not in Amharic (see (29)). However, as already mentioned in section 2.6, this test is not relevant in Sambia, or indeed most Bantu languages. First, they do not have (clear cases of) reflexive anaphoric DPs; second, they do have a special reflexive form of the OM. So the way to express a reflexive situation is as in (40).

- (40) Ni-**ki**-ghul-iye kitabu. (p. 87)
 1sS-REFL-buy-PERF.CJ 7book
 ‘I bought a book for myself.’

Riedel says that OMs are less and less common as the NP gets less animate in Sambia (Ch.3), and this is presumably a somewhat heightened version of this effect, where “rare” may cross over into “never”.

¹² Note that there is no need to clitic-double the goal object in a DOC in Sambia (see (33) p. 106). In general, there is little or no reason to say that OMs doubling goals in Sambia have any different status from OMs doubling themes, as has been claimed for Romance languages, where dative clitics might be agreement and accusative clitics are true pronominal clitics (see for, example, Sportiche 1996). Recall that there is no dative/accusative distinction among OMs in the Bantu languages being discussed.

Without taking a firm stand on what the reflexive marker is exactly in Bantu languages, I count this as “not applicable” with respect to the question of agreement versus clitic doubling. The issue of doubling reflexive anaphors in Bantu languages does not arise, because they do not have the right sort of overt DP anaphors to test this.

The final part of the proposed diagnostic cluster is whether an OM can double an object that contains a possessor interpreted as a bound variable, in an example like ‘Everyone loves his child’. This is unknown for Sambia—not surprisingly, given that this is a newly discovered diagnostic. The prediction is that this should be possible in Sambia.

I conclude, in agreement with Riedel, that OMs are agreement markers in Sambia, not pronominal clitics.¹³

This conclusion also presumably holds for the Mozambican language Makhuwa, where agreement is obligatory with all class 1/2 objects (van der Wal 2009: 85), but the data is somewhat fragmentary. At least doubling is not qualified for the definiteness or specificity of the object, and wh objects like ‘who’ can be doubled.

3.2 Variation in Swahili dialects

Now let us consider Swahili. It is often said that OM is obligatory with class 1 and class 2 (i.e., human) objects, but the larger picture is more complex than this. Here we find some evidence of relevant variation—not surprising, perhaps, given the wide area and range of circumstances in which Swahili is spoken.

Riedel is not as explicit about whether objects related to OMs are in the VP in Swahili, but the assumption is that they are. Swahili does not have the disjunct/conjunct distinction to use as a test. I don’t know if Swahili has the right sort of phrasal phonological processes to speak to the issue. But one thing that is clear is that one can have V-IO-DO order where the OM doubles the IO, even though it is evidently VP-internal, not at the edge of the clause.

(41) A-li-m-nunulia Juma kitabu. (Riedel p. 80)
 SM1-PST-OM1-buy.APPL 1Juma 7book
 ‘She bought Juma a book.’

There is no sign of Swahili varieties varying in this respect. So they are presumably either languages with true object agreement or clitic doubling, not languages where overt DPs in the presence of an OM must be dislocated.

¹³ Note that this diagnostic cluster gives a different result from other diagnostics proposed in the recent literature. The OMs in Sambia do not vary with the tense-aspect of the clause (Riedel gives only one form in Table 2.3 p. 21), so they would be clitics by Nevins’ (2011) criterion. And when there is no object in the clause, one does not get a default OM, like one does for subject agreement, so they are clitics by Preminger’s (2009) criterion. These morphological/PF-oriented tests give a different result from our syntactic tests. B&K conclude that these other diagnostics are not reliable for discerning whether an OM is an agreement marker or a pronoun in the syntax, because their connection to the known theory of pronouns is at best more tenuous. If they distinguish agreement from clitics at all, it is for a different sense of clitic than the normal one that Bantu linguistics and other studies of morphosyntax have been interested in.

Consider now nonreferential objects. Here Riedel reports some significant variation. In Kiunguja Swahili, spoken in parts of Zanzibar, OMs are possible (indeed required) with nonspecific human objects (pp. 49-50). In contrast, certain speakers from Mainland Tanzania do not tend to use object marking for non-specific human objects, as seen in (42b).

- (42) a. Ni-li-*(mw)-ona mtoto. Kiunguja Swahili
 1sS-PST-OM1-see 1child
 ‘I saw a child.’
- b. Ni-li-ona mtoto. Mainland Tanzanian Swahili
 1sS-PST-see 1child.
 ‘I saw a child.’

The difference is also seen with negative items: Kiunguja Swahili requires doubling with such objects when they are human, whereas MT Swahili avoids it.

- (43) a. Si-ku-*(mw)-ona mtu yeyote. Kiunguja Swahili
 NEG.1sS-NEG.PST-OM1-see 1person 1any
 ‘I didn’t see anyone.’
- b. Si-ku-(??mw)-ona mtu yeyote Mainland Swahili
 NEG.1sS-NEG.PST-OM1-see 1person 1any
 ‘I didn’t see anyone.’

In a different part of her work, Riedel discusses variation in *wh*-constructions in Swahili. She reports that some speakers of Swahili judge (44) to be perfectly acceptable, whereas others rule it out.¹⁴

- (44) %U-li-ki-pata nini? (p. 157)
 2sS-PST-OM7-get what

¹⁴ In contrast, all Swahili speakers apparently allow an OM doubling the human *wh*-word ‘who’, in examples like (i)—even Mainland Tanzanian speakers. Also all speakers allow an OM doubling the D-linked *wh*-phrase in (ii).

- (i) U-li-mw-ona nani? (p. 155)
 2sS-PST-OM1-see who
 ‘Who did you see?’
- (ii) U-li-ki-pata kitabu kipi? (p. 157)
 2sS-PAST-OM7-get 7book 7which
 ‘Which book did you get?’

The contrast between (ii) and (44) for some speakers is expected; it is the same as the contrast between (20) and (28b) in Amharic, and fits within B&K’s account in terms of WCO. I tentatively say that the goodness of (i) might be similar to that of (ii): it would be related to the fact that questions with ‘who’ have more presuppositions than questions with ‘what’, in particular, the presupposition that the answer is a human. As such, (i) is on its way to having a D-linked reading, not so different from ‘which person’. If so, the goodness of (i) may reduce to the goodness of (ii). (There is no difference between *who* and *what* with respect to WCO in English, however.)

‘What did you get?’

This would fit very well with the B&K analysis if the speakers who reject (44) are (more or less) the same as the speakers who avoid an OM (42) and (43)—i.e. Mainland Tanzania speakers. I hereby express the hope that this is true.

One would also like to know about universal quantifiers, predicting that they too could be doubled by OMs for some Swahili speakers, including those of the Kiunguja dialect, but not for others, e.g. those from Mainland Tanzania. Unfortunately, Riedel does not discuss this directly. She does give in passing one example of an OM doubling a universal quantifier, namely (45), originally from Marantz 1993 (credited to Vicki Carstens pc). She implies that she replicated this with an informant, but does not say from what dialect group.

(45) Ni-li-m-somea kila mwandishi kitabu chake. (p. 105)
1sS-PST-OM1-read.APPL every writer 7book 7his
‘I read for each author his book.’

Overall, then, some Swahili speakers clearly have object agreement, as in Sambia. Others, however—especially those from Mainland Tanzania—show signs of having a systematically different grammar, one with clitic doubling, as in Amharic. This is consistent with the idea that there is something intrinsically pronominal about the OM. There is more data to gather to see if this is as systematic as it should be, but what is known looks promising. If everything comes out right, then Mainland Tanzanian Swahili comes out as a clitic-doubling language—replicating the Amharic pattern (at least to a good first order) and contrasting minimally with languages in which OMs are true agreement markers in a cluster of ways.¹⁵

3. Clitic doubling in Haya

Another Bantu language with Amharic-style clitic doubling is Haya, based on Riedel’s analysis.

Previous literature said that OMs force right dislocation in Haya, but Riedel argues against this. The contested evidence came from tone patterns: with no object marking, only the last object has a falling tone, whereas with (two) OMs on the verb, both objects all have falling

¹⁵ Keach (1995) presents a somewhat different constellation of facts for the version of Swahili that she studied (Zanzibari)—but still one consistent with what is allowed by the B&K view. For human-animate (class 1/2) objects, the OM is required, even with question words (‘who’). Her (4a) shows an OM with a generic object as well (“Kenyan people like children”). However, for other noun classes, Keach says that the OM is bad with question word objects, and with idiomatic objects (“I (*OM)-hit iron” = I ironed). This suggests a view in which class 1/2 OM is pure agreement, but other OMs are pronominal. (Keach assumes, following Bresnan and Mchombo, that an overt inanimate NP that co-occurs with an OM is dislocated, but her word order data does not show this to be so. It could be that her Swahili is a clitic doubling language when it comes to inanimates.) If everyone is right about their facts, then, we have three kinds of Swahili:

- animate and inanimate OMs are pure agreement; animate ones are obligatory, inanimate ones are not (Riedel 1)
- animate and inanimate OMs are pronominal; animate ones want to be obligatory, but this can be trumped by nonreferentiality (Riedel 2).
- animate OMs are agreement; inanimate ones are pronominal (Keach)

What we do not find is mixtures where an OM is possible for some less than fully referential NPs but not others (pace note 12). That is what is crucial for the proposal.

tone. There is also a morphological difference in the Past 1 tense: without the OM, the conjunct form is used; with the OM the disjunct form is used (as in Zulu). However, Riedel observes that the conjunct/disjunct distinction in Haya is much less elaborated than in Sambia (or Zulu): it is realized only with one tense, and there it is marked by vowel length, which is also affected by the subject marker. The boundary tone evidence is stronger, but we know that there can be mismatches between syntactic phrases and phonological phrases.

Crucially, word order does *not* reveal a difference between when an OM is present and when it is not. DOs and IOs are freely ordered in Haya, with or without OMs, so no clear test is available in that domain. But adverbs like ‘day before yesterday’ are relevant. NPs can appear after the adverb only if there is a corresponding OM; that is an indication of right dislocation. Undoubled objects can only appear before the adverbs, because they have to be in situ inside the verb phrase. The crucial fact is that objects doubled by an OM are also perfectly acceptable before the adverb, just as undoubled ones are, as seen in (46). This shows that doubled objects can nevertheless still be inside the verb phrase.

- (46) Y-aa-(mu)-bona Kato kileki. (p. 71)
 SM1-PST1.DJ-OM1-see 1Kato today
 ‘He saw Kato today.’

Indeed, this is the preferred order for a doubled IO:

- (47) a. N-ka-ki-mu-gulira omwaana ijo ekitabo. (p.71)
 1sS-PST3-OM7-OM1-buy.APPL 1child day.before 7book
 ‘I bought it for him the day before yesterday, the book, the child.’
- b. ?N-ka-ki-mu-gulira ijo ekitabo omwaana. (p.71)
 1sS-PST3-OM7-OM1-buy.APPL day.before 7book 1child
 ‘I bought it for him the day before yesterday, the book, the child.’

In addition, Riedel’s informants accept and spontaneously produce examples in which a doubled NP comes before an undoubled NP. If the undoubled NP must be inside VP (as all agree), then the doubled NP must be too (although this is different from Tenebaum’s 1977 data).

- (48) Ba-ka-mu-cumbila Kakulu enkoko. (p. 72)
 SM2-PST3-OM1-cook.APPL 1Kakulu 9chicken
 ‘They cooked Kakulu the chicken.’

Indeed, a doubled NP (theme) can even precede a *wh*-in-situ object, which is intrinsically focal, and cannot itself be dislocated outside of VP (p. 73). Finally, Riedel contrasts postposed subjects with doubled objects: the former have to be after a time adverb, whereas the later does not have to be. So if objects are right dislocated, they must be right dislocated to a lower position than subjects—a somewhat odd state of affairs.

- (49) a. *Ba-ka-goba abakazi ijo. (OK only with marked pause before *ijo*)
 SM2-PST3-arrive 2woman day.before
 ‘The two women arrived the day before yesterday.’

- b. Ba-ka-goba ijo, abakazi.
 SM2-PST3-arrive day.before 2women
 ‘They arrived the day before yesterday, the two women.’

Overall, then, word order evidence shows that DPs doubled by an OM can be inside VP in Haya, and do not need to be dislocated (although there is a mismatch with phonological phrasing to be concerned about).¹⁶ Therefore, Haya has either agreement or clitic doubling.

To choose between these options, we should then consider the referentiality of the doubled NP. Riedel (2009: 186) reports that it was hard to distinguish clearly definite, specific, and nonspecific indefinite readings of NPs in Haya (recall that these distinctions are not marked on the DP itself in many Bantu languages). However, OMs are not obligatory with any class of NP in Haya (different from Sambia and Swahili), so there is certainly the possibility of omitting the OM when an indefinite interpretation is intended. Moreover, the fact that the doubled object in (50) cannot have a nonspecific indefinite interpretation, as an existential with narrow scope with respect to negation, suggests that doubling enforces definiteness (or at least specificity).

- (50) Ti-n-a-ki-bona ekintu.
 NEG-1sS-PST1-OM7-see 7thing
 Not: *‘I didn’t see a thing.’ Only: ‘I didn’t see the thing.’

Moreover, the OM is clearly incompatible with specialized negative expressions in Haya, like in (51), where the object lacks the augment vowel and can be modified by ‘any’.

- (51) Ti-n-a>(*ki)-bona kintu kyonakyona). (p. 186)
 NEG-1sS-PAST1>(*OM7)-see 7thing 7any
 ‘I didn’t see anything.’

It is also clear that no simple *wh*-word (neither ‘who’ nor ‘what’) can be clitic doubled in Haya:

- (52) a. W-a(*mu)-bona owa?
 2sS-PST1.CJ>(*OM1)-see who
 ‘Who did you see?’
 b. O-ka(*ki)-bona ki?
 2sS-PST3(*OM7)-see what
 ‘What did you see?’

Finally, Riedel gives examples with a universally quantified object, intended to be directly parallel to Swahili and Sambia examples shown above (replicating the result that the IO c-commands the DO for Haya too). But whereas her Swahili and Sambia examples have the IO doubled by an OM, the Haya example notably does not:

¹⁶ For what it is worth, Riedel also claims that clitic doubled NPs in Haya do not have the semantic or pragmatic properties of RD in English, where they are afterthoughts that give additional information.

- (53) N-ka-shomera buli mwandiki ekitabo kye. (p. 123)
 1sS-PST3-read.APPL every 1writer 7book 7POSS1
 ‘I read every writer his book.’

This is at least consistent with, and may actively suggest, the claim that OMs are not compatible with quantified NPs in Haya. Overall, then, we have a fairly complete data set for Haya,¹⁷ and it turns out to another language with clitic doubling, like Amharic.

Overall, then, we do not see chaotic or continuously varying patterns in this domain after all. Rather, we see relatively stable patterns of facts, centering around two clusters: an agreement cluster (Sambaa, Zanzibari Swahili) and a clitic doubling cluster (Mainland Tanzanian Swahili, Haya). And that is somewhat remarkable, given the superficial impression that Bantu languages vary enormously in their object marking properties.

3.4 Why OM behavior in Bantu may look continuous

This putative success in teasing out a consistent pattern comes primarily from two factors. First, I have opened up a third theoretical possibility, not widely used in the Bantu literature heretofore. That means that I am fitting the observed data to three points, not to two. This eliminates some otherwise problematic-looking intermediate cases.

Second, and perhaps equally important, I am completely ignoring one other dimension along which OMs in Bantu clearly vary considerably: the question of whether an OM doubling a certain class of DPs is obligatory or optional, and how this varies with the humanness or animacy of the DP. OMs on the verb are optional with all classes of NPs in Amharic (except maybe pronouns). OMs are optional with many classes of NPs in Bantu too. Where OMs are obligatory, that needs some other kind of account, not part of my core analysis of either agreement or clitic doubling. But that is arguably where the really fine-grained variation in the (Eastern) Bantu languages is. For example, Sambaa has obligatory OM with first and second person pronouns and proper names (also kin terms and unique titles), but nothing else. In contrast, many Swahili dialects have obligatory OMs with all humans and even animates. Reidel (2009: 53) considers six Bantu languages in some detail and shows that they have six different patterns with respect to what NPs are obligatorily doubled as opposed to optionally doubled. She presents the following table, summarizing her results.

(54) Categories which trigger obligatory object marking per language

Property	Sambaa	Swahili	Nyaturu	Ruwund	Mahkuwa	Haya
1/2 person	Yes	Yes	Yes	Yes	Yes	No
Proper names	Yes	Yes	Yes	Yes	Yes	No
Definite humans	No	Yes	Yes	Yes	No	No
Specific humans	No	Yes	Yes	No	No	No

¹⁷ The notable missing element is an object that contains a pronoun bound by a quantifier. The prediction would be that such an object cannot be doubled by an OM in Haya (or Mainland Tanzanian Swahili).

Humans	No	Yes	No	No	No	No
Animates	No	Yes	No	No	No	No
Class 1/2	No	(yes)	No	No	Yes	no

This is a bit of a mess, and here we do seem to see a continuity of patterns, with no clear dividing lines. It is true that there are associations between this issue of optional vs obligatory object marking and the theoretical distinction between agreement and clitic doubling: agreement is normally thought to be obligatory wherever possible, whereas clitic doubling might well be optional (as object shift sometimes is). However, that simple way of distinguishing agreement from clitic doubling does not correlate well with B&K’s diagnostic cluster involving referentiality restrictions. For example, Sambia and Kiunguja Swahili differ in whether an OM is required with common nouns that denote humans, but when the OM is present both behave like they have true agreement by the referentiality tests. Conversely, Kiunguja Swahili and Mainland Tanzanian Swahili do not differ in whether an OM is required with definite human objects, but the OM behaves like agreement in one and like a pronoun in the other by the referentiality test.

Suppose then that we simply deny that there is a direct connection between whether a marker is obligatory or optional and whether it is agreement, and concentrate on whether the OM obligatory-or-optional as opposed to forbidden.¹⁸ Then we potentially find an island of relative stability. And that is good, because it is this part of the paradigm that can be attributed to venerable syntactic conditions, notably the Weak crossover condition. I am not so naïve as to think that there will be no noise in the data at all, as a result of dialect mixture, language contact and change (see for example, note 12 on ‘who’ triggering OMs in a dialect of Swahili where that is anomalous). But the claim is that there is a third “strong attractor” underneath the apparent chaos here.¹⁹

4. Conclusion

In this paper, I have considered the prospects of generalizing Baker and Kramer’s (2015) analysis of “object markers” as pronominal clitics created within the syntactic derivation by an operation of Reduce from Amharic (and various Indo-European languages) to certain Bantu languages. First, we saw that there is nothing about the superficial syntactic design of Bantu languages that makes such a generalization look unlikely or problematic. Then we saw that,

¹⁸ This is not a trivial shift theoretically speaking: there are good reasons from canonical cases of agreement (e.g. subject agreement in IE languages) to think that agreement is obligatory wherever it is possible. One would either need to find a subtle structural difference between examples in which the verb agrees with the object in (say) Sambia and examples in which it does not, or change the theory of agreement somehow to allow it to be optional. The proposal, however, is that the latter option may well be the way forward.

¹⁹ Perhaps the most potentially problematic aspect of (54) is the contrast between Nyaturu and Ruwund. The fact that OMs are present with definite humans but not with all humans suggests that referentiality conditions are at play in these two languages. But then the fact that specific indefinites count as referential enough to have an OM in Nyaturu but not in Ruwund needs some account that does not follow from the B&K proposal. In fact, there is some variation among the clitic doubling languages of Europe (Spanish, Romanian, and Greek) in whether doubling is allowed with specific indefinites, and if so, which kinds. This variation does not seem to have to do with Weak Crossover; rather, B&K speculate that it has to do with the conditions on exactly what sorts of DPs undergo object shift, something that may vary a bit across languages.

despite what looks like rather chaotic variation across languages, some seem to fit the profile expected of languages in which OMs are true agreement (Sambaa, Zanzibari Swahili), whereas others fit quite well the profile expected of languages in which OMs are pronominal clitics in the Amharic sense (Haya, Mainland Tanzanian Swahili). I conclude, then, that at least a few Bantu languages do have clitic doubling as opposed to mere agreement, and that distinction is a discrete parametric factor that contributes to the observed diversity in this domain. Other Bantu languages, in which the presence of an OM forces the dislocation of an overt DP object to the right or left edge of the clause can be modeled by saying that *v* is parametrically specified as licensing the operation Reduce in some languages but not others—similar to the fact that some Romance languages allow both clitic dislocation and true clitic doubling, whereas other languages have only clitic dislocation. I have also made some suggestions about how further data could clarify the situation for particular languages. This includes the possibility of investigating a wider range of quantified DPs than has usually been done so far—e.g. universally quantified DPs as well as interrogative ones—and the possible relevance of DPs that are not quantifiers themselves but contain variables bound by quantifiers, in sentences like ‘Everyone loves his mother.’ Hopefully future investigations will confirm whether the link between which DPs give weak crossover effects and which DPs resist doubling by an OM is as strong in the various Bantu languages as it is in Amharic.

References

- Anagnostopoulou, Elena. 1999. Conditions on Clitic Doubling in Greek. In *Clitics in the Languages of Europe*, Henk van Riemsdijk (ed.), pp. 761-798. Berlin: Mouton de Gruyter
- Anagnostopoulou, Elena. 2001. Two classes of double object verbs: the role of zero-morphology. In van Oostendorp, Marc & Elena Anagnostopoulou (Eds.) *Progress in Grammar. Articles at the 20th Anniversary of the Comparison of Grammatical Models Group in Tilburg* 1-27. Roquade, Amsterdam/Utrecht/Delft. Meertens Institute Electronic Publications in Linguistics (MIEPiL).
- Baker, Mark. 2003. Agreement, dislocation, and partial configurationality. In *Formal approaches to function in grammar*, eds. Andrew Carnie, Heidi Harley and Mary Ann Willie, 107-134. Amsterdam: John Benjamins.
- Baker, Mark. 2008. *The syntax of agreement and concord*. Cambridge: Cambridge University Press.
- Baker, Mark and Ruth Kramer 2015. Doubling Clitics are Pronouns: Agree, Move, Reduce, and Interpret. Ms. Rutgers University and Georgetown University.
- Barss, Andrew, and Lasnik, Howard. 1986. A note on anaphora and double objects. *Linguistic Inquiry* 17:347-354.
- Bresnan, Joan, and Sam Mchombo. 1987. Topic, pronoun, and agreement in Chichewa. *Language* 63:741-782.
- Bresnan, Joan and Lioba Moshi. 1990. Object asymmetries in Bantu syntax. *Linguistic Inquiry* 1990. 147-185.
- Büring, Daniel. 2005. *Binding Theory*. Cambridge: Cambridge University Press.
- Chomsky, Noam. 1981. *Lectures on Government and Binding*. Dordrecht: Foris.
- Chomsky, Noam. 1986. *Barriers*. Cambridge: MIT Press.

- Chomsky, Noam. 1993. A minimalist program for linguistic theory. In *The view from Building 20*, eds. Kenneth Hale and Samuel Jay Keyser, 1-52. Cambridge, Mass.: MIT Press.
- Chomsky, Noam. 1995. *The Minimalist Program*. Cambridge, Mass.: MIT Press.
- Chomsky, Noam. 2000. Minimalist inquiries, the framework. In Roger Martin, David Michaels and Juan Uriagereka, eds. *Step by Step: Essays on Minimalist Syntax in Honor of Howard Lasnik*. Cambridge: MIT Press. 89-155.
- Chomsky, Noam. 2001. Derivation by phase. In Michael Kenstowicz, ed. *Ken Hale: A Life in Language*. Cambridge: MIT Press. 1-52.
- Cinque, Guglielmo. 1990. *Types of A'-Dependencies*. Cambridge, Mass: MIT Press
- Cinque, Guglielmo. 1999. *Adverbs and Functional Heads: A Crosslinguistic Perspective*. Oxford: OUP.
- Diercks, Michael and Justine Sikuku. 2013. Object clitics in a Bantu language: deriving pronominal incorporation in Lubukusu. Ms., Pomona College and Moi University.
- Diesing, Molly. 1992. *Indefinites..* Cambridge, MA: MIT Press.
- Henderson, Brent. *The syntax and typology of Bantu relative clauses*. Ph.D dissertation, University of Illinois Urbana-Champaign.
- Kandybowicz, Jason. 2007. On fusion and multiple copy spell-out. In Norbert Corver and Jairo Nunes, eds. *The Copy Theory of Movement on the PF Side*. Amsterdam: John Benjamins. 119-150.
- Kimenyi, Alexandre. 1980. *A relational grammar of Kinyarwanda*. Berkeley, Calif.: University of California Press.
- Kramer, Ruth. 2014. Clitic doubling or object agreement: the view from Amharic. *Natural Language and Linguistic Theory* 32. 593-634.
- Larson, Richard. 1988. On the double object construction. *Linguistic Inquiry* 19:335-392.
- Lasnik, Howard and Tim Stowell. 1991. Weakest crossover. *Linguistic Inquiry* 22. 687-720.
- Leslau, Wolf. 1995. *Reference Grammar of Amharic*. Wiesbaden: Harrassowitz.
- Marantz, Alec. 1993. Implications of asymmetries in double object constructions. In Sam A. Mchombo, ed. *Theoretical Aspects of Bantu Grammar*. Volume 1. Stanford, CA: CSLI. 113-151
- McGinnis, Martha. 2001. Variation in the phase structure of applicatives. *Linguistic Variation Yearbook* 1:105-146.
- Nevins, Andrew. 2011. Multiple agree with clitics: person complementarity vs omnivorous number. *Natural Language and Linguistic Theory* 29. 939-971.
- Nunes, Jairo. 2004. *Linearization of Chains and Sideward Movement*. Cambridge: MIT Press.
- Preminger, Omer. 2009. Breaking agreements: distinguishing agreement and clitic doubling by their failures. *Linguistic Inquiry* 40. 619-666.
- Riedel, Kristina. 2009. The syntax of object marking in Sambiaa: a comparative Bantu perspective. Doctoral dissertation, Leiden University.
- Roberts, Ian. 2010. *Agreement and Head Movement*. Cambridge, MA: MIT Press.
- Safir, Ken. 2004. *The Syntax of (In)dependence*. Cambridge, MA: MIT Press.
- Safir, Ken. 2015. Weak Crossover. To appear in the *Blackwell Companion to Syntax*.
- Sportiche, Dominique. 1996. Clitic constructions. In J. Rooryck and L. Zaring, eds. *Phrase Structure and the Lexicon*. Dordrecht: Kluwer. 213-276.

- Wal, Jenneke van der. 2009. *Word order and information structure in Makuwa-Enahara*, Ph.D. dissertation, University of Leiden.
- Wasow, Thomas. 1979. *Anaphora in Generative Grammar*. Ghent: E. Story-Scientia.
- Woolford, Ellen. 2001. Conditions on object agreement in Ruwund (Bantu). In *The UMass volume on Indigenous Languages*, ed. by E. Benedicto. Amherst, MA: GLSA.
- Zeller, Jochen. 2012. Object marking in isiZulu. *Southern African Linguistics and Applied Language Studies* 30. 219–235.