Typology of variation in Bantu Object Markers

Natalie DelBusso & Hazel Mitchley

April 2021

1 Introduction

There is significant variation in patterns of object marking in Bantu languages, and cross-linguistically more generally. When object markers (OMs) occur, their distribution depends on syntactic, semantic, and discourse factors. The typology of OM patterns has been subject to much theoretical research (Baker 2018; Riedel 2009; van der Wal 2015; Zeller 2014). This paper develops an analysis in Optimality Theory (OT; Prince and Smolensky 2004) of variation in OM patterns. Both single- and double-object constructions are analyzed, with variation in the form of the object: in-situ lexical object, pro, and a displaced topicalized object. To focus on the key properties of the analysis, other points of variation are omitted, such relative clauses or questions and variation based on semantic features (see Marten and Kula (2012) a.o. for additional points of variation).

After reviewing the attested empirical OM patterns (§2), the paper develops the system and the resulting typology (§3). We then analyze the structure of the typology, using Property Theory (Alber and Prince prep; Alber et al. 2016; DelBusso 2018) (§4). There is a set of crucial constraint conflicts that define all the languages, and the analysis identifies precisely which conflicts align with which extensional traits. The agreement constraints that favor having an OM are ranked relative to different configurational and structural constraints that favor lack of an OM in specific contexts. When languages share a property value, they share a specific trait. For example, there is a ranking defining all language allowing multiple OMs, the reverse of which holds in all non-multiple OM languages. Each of these groups is further divided along additional parameters. Different OM patterns result from different sets of property values.

2 Object Marking in Bantu

Both subject markers and object markers occur in Bantu languages, but OMs display a greater range of patterns, conditioned by various restrictions (see also Marten and Kula (2012), Zeller (2014), etc.). OMs are not limited to Bantu, and occur cross-linguistically, for example in Palauan (Austronesian; Woolford 2001) and Amharic (Semitic; Baker and Kramer 2018). Related phenomena include object clitics, common in Romance languages, and object shift, as in some Germanic languages (Baker and Kramer 2018).

In Bantu languages, OMs generally occur as affixes in the verbal complex and agree with the object(s) in features, including noun class. Not all Bantu languages have OMs; for example, Basaa lacks them (van der Wal 2015). In other languages, OM distribution is restricted to particular constructions or contexts. Languages that have OMs all allow (or even require) an OM when the object is non-overt (pro), as in the Swahili example below.[1]

(1) A-li-**wa**-on-a
1.SM-PST-OM2-see-FV
'He saw them.' (Swahili, Riedel 2009:4)

There is a substantial literature on OMs in Bantu languages (Baker 2018; Baker and Kramer 2018; Bresnan and Mchombo 1987; Bresnan and Moshi 1990; Marten and Kula 2007, 2012; Riedel 2009; van der Wal

[1] Abbreviations used in glosses: 1/2/3SG/PL: 1st/2nd/3rd person singular/plural; APPL: applicative; AUG: augmentive; CJ/DJ: conjoint/disjoint; DEM: demonstrative; FV: final vowel; OM: object marker; PERF: prefect aspect; PST/PRES/FUT: past/present/future tense; SM: subject marker; numbers following SM and OM are noun classes. OMs are bolded.

2015; Zeller 2014). Much of this work has addressed the status of OMs as agreement markers or incorporated pronouns, using the distinction to explain points of variation between languages. Other analyses argue for one or the other, such as Riedel (2009), who analyzes all OMs as agreement markers. This paper is similar in analyzing all OMs as agreement, with variation arising from conflict with other syntactic constraints.

The following reviews several salient points of variation in OM patterns, drawing examples from the following languages that represent the different types of OM patterns: Chaga (E62), Chichewa (N31), Haya (JE22), Lozi (K21), Sambaa (G23), Swahili (G42), and Zulu (S42). Data comes from studies of these languages, especially the parametric work of Marten and Kula (2007, 2012).[2]

The first point of variation is whether an OM can co-occur with an overt referential object in the VP. Such co-occurrence is called *doubling*, and is not possible in languages such as Zulu, but can occur in Swahili (though there are finer degrees of variation in obligatoriness/optionality depending on factors of animacy, specificity, etc., which are not the focus here).

(2) Ngi-(***yi**)-theng-a le moto.
1SG-OM9-buy-FV DEM9 9.car
'I'm buying this car.' (Zulu, Zeller 2012:222)

(3) Ni-li-**mw**-on-a Juma
1SG-PST-OM1-see-FV Juma
'I saw Juma.' (Swahili, Marten and Kula 2012:5)

However, doubling or co-occurrence restrictions are often limited to in-situ objects. When the object is displaced outside of the VP (as in topicalization), an OM often becomes possible even in non-doubling languages. As rightward object displacement is often less evident than leftward, there are several diagnostics to identify rightward object displacement outside of the VP: i) an atypical word order, such as when an adverbial intervenes between the verb and the direct object; ii) a "disjoint" affix on the verb;[3] and/or iii) a prosodic break between the verb and the displaced object, sometimes indicated by phrase-final intonation or tonal patterns. In the example from Zulu in (4), an adverb intervenes between the verb and the direct object, indicating displacement; in such constructions, an OM occurs.

(4) Si-**zi**-bon-e kaningi i-zi-tshudeni
PL-OM8-see-PST often AUG-8-student
'We saw the students often' (Zulu, Van der Spuy 1993:364)

A second point of variation is the occurrence of multiple OMs in double-object constructions. Multiple OMs (with pro objects) are possible in some languages (Chaga (5)), while others are restricted to a single OM (Swahili (6)).

(5) N-a-i-**ki-m**-lyi-i-a.
FOC-SM1-PRES-OM7-OM1-eat-APPL-FV
'He is eating it for her.' (Chaga, Bresnan and Moshi 1990:151)

(6) a. Ni-li-**m**-p-a.
SM1SG-PAST-OM1-give-FV
'I gave him (it).'

b. *Ni-li-**i-m**-p-a.
SM1SG-PAST-OM9-OM1-give-FV
Intd.: 'I gave him it.'

[2]This study does not make empirical claims about these languages. For example, Downing (2018) argues that Chichewa allows some doubling, though it has been previously analyzed as non-doubling (Bresnan and Mchombo 1987). The decision of Chichewa's correct classification could change which language it matches in the typology here but would not change the overall argument, which concerns OM patterns rather than particular languages.

[3]The conjoint/disjoint distinction is a verbal alternation in many Bantu languages that aligns with other syntactic properties, in particular whether or not anything vP internal follows the verb (van der Wal and Hyman 2016).

c. *Ni-li-**m-i**-p-a.
SM1SG-PAST-OM1-OM9-give-FV
Intd.: 'I gave him it.'

(Swahili, Marten and Kula 2012:12)

A third point of variation is which object can be OM-ed in a double-object construction. 'Symmetrical' languages like Zulu (7), allow an OM for either indirect or direct objects, while 'asymmetrical' languages like Chichewa (8) restrict OMs to the structurally highest object (the indirect object).

(7) a. U-Langa u-**m**-phek-el-a i-nyama (u-mama)
AUG-1a.Langa SM1-OM1-cook-APPL-FV AUG-9.meat (AUG-1a.mother)
'Langa is cooking meat for her (mother).'

b. U-Langa u-**yi**-phek-el-a u-mama (i-nyama)
AUG-1a.Langa SM1-OM9-cook-APPL-FV AUG-1A.mother (AUG-9.meat)
'Langa is cooking it for mother (the meat).'

(Zulu, Zeller 2012:227)

(8) a. A-lenje a-ku-**wá**-phík-ir-a zí-túmbúwa (anyani)
2-hunters SM2-PRES-OM2-cook-APPL-FV 8-pancakes 2.baboons
'The hunters are cooking (for) them (the baboons) some pancakes.'

b. *A-lenje a-ku-**zí**-phík-ir-a anyani (zí-túmbúwa)
2-hunters SM2-PRES-OM8-cook-APPL-FV 2.baboons 8-pancakes
'The hunters are cooking them (the pancakes) for the baboons.'

(Chichewa, Marten and Kula 2007:37, citing Mchombo and Firmino 1999)

Interestingly, this asymmetry is seen even in languages where multiple OMs are possible. In Sambaa double-object constructions, for example, the lower (direct) object can only be OM-ed if the higher (indirect) object also is. It is possible to have OMs for both objects or just for the indirect, but not for the direct object alone.

(9) a. Mbegha a-za-**(u)-m**-nka ng'wanae uzumbe
1Mbegha SM1-PERF.DJ-OM14-OM1-give 1child.POSS.3S 14kinghood
'Mbegha gave his child the kinghood'

b. *Mbegha a-za-**u**-nka ng'wanae uzumbe
1Mbegha SM1-PERF.DJ-OM14-give 1child.POSS3S 14kinghood
Int: 'Mbegha gave his child the kinghood

(Sambaa, Riedel 2009:3)

These three points of variation – doubling, multiplicity and symmetry – divide the seven languages as shown in the following table. One combination is missing from this table: non-doubling, multiple, asymmetric. Notably, none of the languages classified in Marten and Kula (2012) have this combination. Our analysis below derives this systematic gap.

(10) Main points of OM variation

Language	Doubling	Multiple	Symmetric
Chichewa	no	no	no
Zulu	no	no	yes
Chaga	no	yes	yes
Swahili	yes	no	no
Lozi	yes	no	yes
Sambaa	yes	yes	no
Haya	yes	yes	yes

The problem this data raises is how to account for the similarities in OM-ing while also explaining the variation. The following section takes the approach that all OMs result from satisfying agreement constraints, but that these can be subordinated to other constraints on syntactic structures.

3 Analysis

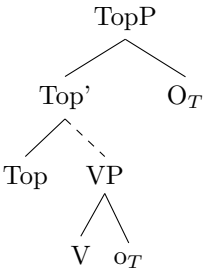
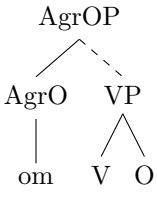
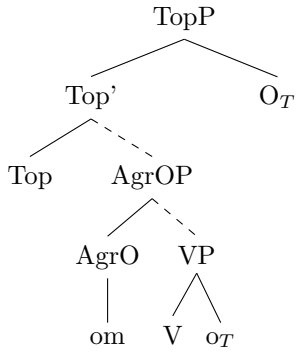
This different possible OM patterns are explained here as resulting from the interaction of three types of constraints: agreement, configurational, and structural. Candidates with OMs have an additional projection compared to those without, an AgrOP headed by the OM.[4] Such candidates satisfy agreement constraints, but incur violations of structural and co-occurrence constraints. In candidates with a topic object, the topic is displaced to a rightward specifier of a higher TopP. These candidates also add violations of structural constraints, but satisfy constraints against co-occurrence of O and OM within the agreement domain when there is also an AgrOP and OM.

3.1 Gen

Inputs consist of a verb (V) and one or two objects (in a single-object construction or double-object construction, respectively). In double-object constructions, indirect and direct objects are indicated by 'I' and 'D'. Objects vary in three possible forms: lexical object, pro, or topic (a discourse-marked element; indicated by a subscripted T in trees below and just T in candidates).

All output structures are left-headed, binary branching trees consisting minimally of a VP with an object as complement. Head and specifier direction are held constant across candidates (Bantu languages are generally SVO in neutral word orders). In double-object constructions, the direct object is the complement and the indirect is in the specifier of the VP. Output structures may have AgrOPs above the VP headed by an OM and/or TopPs with rightward specifiers to which a topic object is displaced, leaving a lower, unpronounced copy in-situ. For inputs with topic objects, all outputs have a TopP and displacement; for non-topic object, only non-displacement (no TopP) outputs are generated.[5] Gen does not produce empty projections: candidates without OMs or displacement lack AgrOPs or TopPs, respectively. Example output structures for inputs with single objects are shown in (11).[6]

(11) Output structures for single-object constructions

No OM or topic	No OM, topic	OM, no topic	OM, topic
			

3.2 Con

There are five constraints that assess agreement and syntactic structures. Agreement constraints are violated when the ϕ features of the specified argument of a verb (V) are not realized by a functional head (X) in the verbal extended projection (EP_V; following Grimshaw's (2005) definition of Extended Projection). Such constraints are satisfied by candidates with an OM heading AgrOP, as the OM encodes/realizes the features.

[4] This is similar to Woolford's (2001) OT analysis, though she assumes that object agreement occurs when the object moves to Spec AgrOP. Whether an object moves depends on their features and the rankings of constraints referencing those features.

[5] Object position is fixed in Gen (i.e., O in-situ and T ex-situ) to study the relationship of OMs to displacement. Putting O and T displacement in Con adds additional dimensions of variation (producing a larger typology).

[6] Some accounts assume more articulated structures for double-object constructions, for example, with the indirect object introduced by an ApplP. What is essential for the present analysis is that the indirect is structurally higher than the direct, as in the structures defined above (which are similar to those in (Riedel 2009:ch.1). The structures do not show verb movement that would be needed for V to be adjacent to the OMs. Additionally, projections between VP and AgrOP or AgrOP and TopP are omitted (indicated with a dotted line), as they are not relevant to the main focus of the analysis and are taken to be the same across all candidates.

Two such constraints are used in this analysis (12): the more general agr.O applies to any object; the more specific agr.HO requires agreement only for the structurally highest object. An object is 'highest' if there is none higher; this groups single objects (which are vacuously highest) with indirect objects as distinct from direct objects. This predicts that the treatment of objects is based on position and not e.g., theta roles, and that OM patterns with single objects will mirror those with indirect objects. The two agreement constraints are equivalent for single-object inputs, assessing the same number of violations. For double-object constructions, however, they differ: agr.HO is satisfied in candidates with an OM for the indirect object, but agr.O is only satisfied if there are OMs for both objects.

(12) Agreement constraints

a. agr.O: *O: O is an object of V and $\nexists$ X $\in$ EP_V s.t. ϕ (O) $\in$ X.
Prose: Assign a violation for each object O of V such that there is no head X in the extended projection of V that realizes the ϕ features of O.

b. agr.HO: *O: O is an object of V and $\nexists$ O' s.t. O' $>$ O and $\nexists$ X $\in$ EP_V s.t. ϕ (O) $\in$ X.
Prose: Assign a violation for each object O of V such that there is no other object O' higher than O and there is no head X in the extended projection of V that realizes the ϕ features of O.

There are two co-occurrence constraints that are violated by multiple occurrences of an element or feature. The constraint m.ff assigns violations to candidates with multiple overt realizations of the same ϕ features within the AgrOP (including the projections it dominates). Candidates with an OM and in-situ object violate this constraint because the object features (f) are realized both in the OM and in the object itself. The constraint is satisfied in candidates with: i) no OM; ii) pro (non-overt); or iii) displaced objects (a topic, which is not overt in the AgrOP).

The second constraint, m.XX, assigns violations for co-occurrence of the same syntactic head (X) in and EP. This is violated by structures with multiple OMs, as each is the head of an AgrOP.[7]

(13) Co-occurrence constraints

a. m.ff: *$[\text{f}\ldots\text{f}]_{AgrOP}$.
Prose: Assign a violation for each pair of elements realizing identical ϕ features in the AgrOP.

b. m.XX: *$[\text{X}\ldots\text{X}]_{EP}$.
Prose: Assign a violation for each pair of identical syntactic heads in an EP.

The final constraint refers to properties of syntactic structure (Grimshaw 2001). CpL is violated by complements that are not left-aligned within their maximal projections because either the head and/or a specifier precedes it. Grimshaw (2001) shows that the interaction of CpL with SpecL and HdL and obligatory element constraints ObHd and ObSp gives rise to economy of structure effects. In the present system they have similar effects, resulting in optimality of non-OM structures as CpL is violated more in candidates with an AgrOP.

(14) CpL: *(Cp, Y) $\in$ XP: $[\ldots\text{Y}\ldots\text{Cp}\ldots]_{XP}$.
Prose: Assign a violation for each syntactic element, head or specifier, that precedes the complement in the immediately dominating projection.

3.3 Candidates and tableaux

Given Gen, there are two possible output candidates for inputs with a single pro object, [Vp], shown in the violation tableau (VT) in (15). The first, (a), has no OM, violating both agreement constraints; the second, (b), has an OM, violating CpL because the complement of the AgrO head (the VP) is not aligned with the edge of AgrOP. The same two outputs—with and without an OM—are possible with single-object construction with a topic object and incur the same violations, as a topic always moves out of the VP.

[7] These constraints are more targeted variants of Grimshaw's (1997) *XX constraint: "Sequences of identical functional heads are ill-formed". Alternatively, this constraint could be formulated as an alignment constraint aligning each AgrOP with the left edge of a higher domain.

(15) VT: single pro and topic

Input	Output	agr.O	agr.HO	m.ff	m.XX	CpL
Vp	a. [Vp]	1	1			
	b. [om[Vp]]					1
VT	a. [[Vt]T]	1	1			
	b. [[om[Vt]]T]					1

However, with an overt, non-topic object, the candidate with an OM (b) also violates m.ff as there are two copies of the object features in the AgrOP.

(16) VT: single object

Input	Output	agr.O	agr.HO	m.ff	m.XX	CpL
VO	a. [VO]	1	1			
	b. [om[VO]]			1		1

With a double-object construction (17), there are four possible outputs (multiplying the OM/no OM options for each object): neither object has an OM (a); only one has an OM (b;with 'im' for 'indirect object marker') and (c; with 'dm' for 'direct object marker'); or both have OMs (d). The candidate with an OM for the direct object only (c) is harmonically bounded: it has all of the same violations as the candidate with an OM for the indirect object only, plus a violation of agr.HO, as the highest (indirect) object is not agreed with.

(17) VT: double-object construction, two lexical objects

Input	Output	agr.O	agr.HO	m.ff	m.XX	CpL
VID	a. [IVD]	2	1			
	b. [im[IVD]]	1		1		1
	c. [dm[IVD]	1	1	1		1
	d. [im[dm[IVD]]]			2	1	2

However, when the direct object is pro or a topic, it is possible to have an OM for the direct only (when the indirect is not a pro or topic). For this input (18), the direct object OM candidate (c) does not violate m.ff, and so is not harmonically bounded by candidate (b).

(18) VT: double-object construction, topic direct object

Input	Output	agr.O	agr.HO	m.ff	m.XX	CpL
VIT	a. [[IVt]T]	2	1			
	b. [[im[IVt]]T]	1		1		1
	c. [[dm[IVt]]T]	1	1			1
	d. [[im[dm[IVt]]]T]			1	1	2

3.4 Factorial Typology

The set of twelve inputs and five constraints generate a factorial typology of 8 languages that differ systematically in OM occurrence.[8] Despite the variation, several implicational generalizations hold across the system: if an OM occurs with an object in a language, then an OM occurs with pro. The distribution of OMs with pro tracks that of OMs with T. If an OM occurs with a direct object, then it occurs with a indirect object of the same type. Additionally, if a language allows only a single OM, it is for the indirect object if the indirect and direct objects are of the same type. However, if the direct is pro and indirect is a lexical object, it is possible to have an OM for the direct.

The typology is shown in (19) with a subset of the inputs that serves to distinguish all of the languages (because OMs with topics tracks OMs with pros—as shown in the 3rd and 4th columns—optima for other inputs with topics are predictable and are not included here). Cells are shaded as follows: blue for a single or indirect object OM in a double-object construction, purple for multiple OMs, and green for a direct object OM in a double-object construction.

[8]The factorial typology was calculated and analyzed using OTWorkplace (Prince et al. 2020).

(19) Factorial Typology

Lg	Vp	VO	VT	Vpp	VpD	VIp	VID
L1	[Vp]	[VO]	[[Vt]T]	[Ip[VDp]]	[p[VD]]	[I[Vp]]	[I[VD]]
L2	[om[Vp]]	[VO]	[[om[Vt]T]	[im[Ip[VDp]]]	[im[p[VD]]]	[I[Vp]]	[I[VD]]
L3	[om[Vp]]	[VO]	[[om[Vt]T]	[im[Ip[V Dp]]]	[im[p[V D]]]	[dm[I[Vp]]]	[I[VD]]
L4	[om[Vp]]	[VO]	[[om[Vt]T]	[im[dm[Ip[VDp]]]]	[im[p[VD]]]	[dm[I[Vp]]]	[I[VD]]
L5	[om[Vp]]	[om[VO]]	[[om[Vt]T]	[im[Ip[VDp]]]	[im[p[VD]]]	[im[I[Vp]]]	[im[I[VD]]]
L6	[om[Vp]]	[om[VO]]	[[om[Vt]T]	[im[Ip[VDp]]]	[im[p[VD]]]	[dm[I[Vp]]]	[im[I[VD]]]
L7	[om[Vp]]	[om[VO]]	[[om[Vt]T]	[im[dm[Ip[VDp]]]]	[im[p[VD]]]	[im[dm[I[Vp]]]]	[im[I[VD]]]
L8	[om[Vp]]	[om[VO]]	[[om[Vt]T]	[im[dm[Ip[VDp]]]]	[im[dm[p[VD]]]]	[im[dm[I[Vp]]]]	[im[dm[I[VD]]]]

The typology is repeated in (20) below, labeling the columns according to the parameters from §2. There is an additional distinction of OM/no OM, aligning with choice of optima for the input [Vp]. This separates L1 from all other languages in which an OM is possible in some context. The Doubling property aligns with the optima for input [VO]; Multiplicity aligns with that for [Vpp]. Symmetry is identified by the combination of two inputs, [VIp] and [VpD]; a language is symmetric if it treats the pro's in these inputs the same or differently.

(20) Factorial Typology, parameterized

Lg	Any OM	Doubling	Multiple	Symmetric
L1	no	no	no	(yes)
L2	yes	no	no	no
L3	yes	no	no	yes
L4	yes	no	yes	yes
L5	yes	yes	no	no
L6	yes	yes	no	yes
L7	yes	yes	yes	no
L8	yes	yes	yes	yes

This typology matches that in (10) (except for the inclusion of L1, which accounts for OM-less languages like Basaa), including the absence of a language with no doubling, multiple OMs, and asymmetry. The following section explains how the constraint interactions generate the possible and impossible combinations of traits.

4 Property Analysis

Each of the parameters defining the typology is linked to one or more specific constraint interactions, or *properties*. Each language represents a unique combination of parametric choices, or traits. The correlating grammars consist of a unique combinations of the rankings from the properties. The following section examines the properties of the OM typology and their associated traits, decomposing the system into a set of core rankings that define all of the grammars.

To do so, we use Property Analysis (PA) (Alber and Prince prep; Alber et al. 2016; DelBusso 2018). A PA consists of a set of properties encoding these core rankings as conflicts between 2 sets of constraints, the antagonists A and B, written as P: A <> B. The values are the rankings in either direction (α. A$\gg$B and β. B$\gg$A). Property values relate to specific traits/inputs, such as whether an OM occurs with an overt object (doubling). Properties have scopes: if all grammars have a value, the property is wide-scope; if only some do, it is narrow-scope, with the scope defined by the value(s) of other properties (Alber et al. 2016; Alber and Prince prep).

There are six properties in the PA of the OM system. While three of these align with specific parameters—any OM, doubling, and multiplicity—the remaining three all align with symmetry. The constraints involved in the choice of (a)symmetry depends on different the values of the properties associated with other traits.

4.1 P.OM: OM-ability

The basic choice of whether an OM is possible for any input is determined by the ranking of both agreement constraints relative to CpL, as stated in P.OM (21). The agreement constraints form a class and CpL

interacts with the whichever of the two is dominant, indicated by the operator *dom* (Alber and Prince prep; Alber et al. 2016; DelBusso 2018). For non-agreement (no OMs), CpL must dominate the dominant member of this class, thereby transitively dominating both. If either agreement constraint dominates CpL, an OM occurs at least with a single pro input.

(21) P.OM: {agr.O, agr.HO}.dom <> CpL

For single-object construction inputs, agr.O and agr.HO assign equivalent violations (as the sole object is also the highest). P.OM values align with whether an OM is possible (α) or not (β). The value ERCs[9] are shown in (22), using the input [Vp].

(22) P.OM values

Value	[Vp]→	agr.O	agr.HO	m.ff	m.XX	CpL
α	[om[Vp]]~[Vp]	W	W			L
β	[Vp]~[om[Vp]]	L	L			W

This property divides the typology in (20) into two sets, as shown in (23). Only L1 has P.OM.β; in all other languages, an OM is possible in some contexts.

(23) P.OM value table

L	P.OM	OM
L1	β	no
L2-L8	α	yes

4.2 P.D: Doubling

Among the set of languages where an OM is sometimes possible (P.OM.α), a key point of variation is whether doubling is possible (an OM with a lexical object). The ranking determining this optima also involves the agreement constraints, this time in conflict with m.ff, as stated in P.D.

(24) P.D: {agr.O, agr.HO}.dom <> m.ff

The value ERCs are shown in (25). As with P.OM, the α value, where at least one of the agreement constraints is dominant, correlates with the optimality of an OM; the β value, where both are dominated, correlates with the lack of OM (no doubling). However, for P.D, the conflicting constraint is m.ff rather than CpL. Note that CpL still prefers the no-OM candidate (b); however, its ranking relative to the set of agreement constraints is determined by P.OM so its preferences are not shown in the ERCs.

(25) P.D values

Value	[VO]→	agr.O	agr.HO	m.ff	m.XX	CpL
α	[om[VO]]~[VO]	W	W	L		
β	[VO]~[om[VO]]	L	L	W		

The values of this property divide only the languages with P.OM.α. For L1, with P.OM.β, the property is *moot*: the constraints in P.D are not crucially ranked in the grammar of this language. The *scope* of P.D is thus P.OM.α. The resulting split of the typology is shown in (26).

(26) P.OM + P.D value table

L	P.OM	P.D	OM	Doubling
L1	β		no	
L2-L4	α	β	yes	no
L5-L8	α	α	yes	yes

[9] An Elementary Ranking Condition (ERC; Prince 2002) is a ranking condition from the comparison of two candidates. It is a vector with a value for each constraint: W, if the constraint prefers the winner; L if it prefers the loser; and *e* if it does not distinguish the candidates. The winner is optimal if all Ls in the ERC are preceded by some W. In the following, ERCs are written with the constraint order agr.O-agr.HO-m.ff-m.XX-CpL.

These two properties together determine the treatment of all single-object constructions, defining three types of languages: those with no OMs whatsoever (like Basaa, L1); those with OMs for pro, but no doubling (like Zulu); and those with doubling (like Swahili). These three types are further divided by multiplicity and symmetry, determined by the remaining properties.

4.3 P.M: Multiplicity

The languages allowing an OM with a pro object (P.OMα) differ in whether they have OMs for *all* pro objects or only one. The general agreement constraint, agr.O, is only satisfied when there is an OM for each pro object. Candidates with multiple OMs, however, incur additional violations of both CpL and m.XX, which is violated when there are multiple AgrO heads (one for each OM). P.M encodes the ranking between these. Here, CpL and m.ff form a class with the .dom operator; it is the dominant member of this class that interacts with agr.O. The value ERCs are shown in (28) deciding whether the optimal output for the [Vpp] input has at most one or multiple OMs.

(27) P.M: agr.O <> {CpL, m.XX}.dom

(28) P.M: values[10]

Value	[Vpp]→	agr.O	agr.HO	m.ff	m.XX	CpL
α	[im[dm[ipVdp]]]~[im[ipVdp]]	W			L	L
β	[im[ipVdp]]~[im[dm[ipVdm]]]	L			W	W

In L1, where no OMs are possible, the value P.M.β is entailed by the value P.OM.β. If CpL dominates both agr.O and agr.HO (P.OM.β) then CpL or m.XX dominates agr.O (P.M.β). ERC-wise: `LLeeW` ⇒`LeeWW` by L-retraction and W-extension (Prince 2002). This is logical empirically as well: multiple OMs are possible only if the language allows any OMs. Among the other languages, P.M values determine multiplicity, adding another dimension of variation.

(29) P.OM + P.D + P.M value table

L	P.OM	P.D	P.M	OM	Doubling	Multiplicity
L1	β		β	no		no
L2-L3	α	β	β	yes	no	no
L4	α	β	α	yes	no	yes
L5-L6	α	α	β	yes	yes	no
L7-L8	α	α	α	yes	yes	yes

Multiplicity is independent of doubling. The values of the properties determining these traits combine freely, yielding four types. This matches the empirical facts: while multiplicity concerns the *number* of OMs, doubling concerns the *type* of object that can have an OM. It is possible to be non-doubling but have multiple OMs, as in Chaga, or conversely, doubling but not multiple, as in Swahili.

4.4 P.S*n*: Symmetry

The final point of variation is symmetry. This does not correlate with a single property, but rather with a set of properties, P.S1, P.S2, and P.S3, that have different scopes. Which property is relevant in a language depends on the values of P.D and P.M. In addition, they do not align with the optima for single inputs because symmetry is determined by (non)parallel treatment of indirect and direct objects (of the same type).

Consider first non-doubling languages (P.D.β). In these languages, only pro or topic objects have OMs. For inputs of the form [VpD], these languages will have an OM for the indirect object pro. Symmetric languages will also have an OM for a direct object pro in parallel inputs like [VIp]. An asymmetric language, however, forgoes OMs altogether for such an input. The ranking determining this choice is stated in P.S1. Symmetric languages have value α, satisfying agr.O in having agreement, and asymmetric have β, satisfying CpL. The value ERCs are shown in (31).

[10]The values are shown using the comparison between the 2-OM and 1-OM candidates, which is the relevant competition for all languages with OMs (L2-8, P.O.α). For L1 (P.Oβ), the relevant comparison is that between no OMs, [pVp] and 1 OM.

(30) P.S1: agr.O <> CpL

(31) P.S1: values

Value	[VIp]→	agr.O	agr.HO	m.ff	m.XX	CpL
α	[dm[IVp]]~[IVp]	W				L
β	[IVp]~[dm[IVp]]	L				W

The antagonists are familiar from P.OM and P.M, but in those properties one of them was in a class with another constraint. In languages allowing multiple OMs, symmetry is entailed: such a language has an OM for any pro. This is entailed by the P.M value. P.M.α (`WeeLL`), where agr.O dominates CpL and m.XX, entails P.S1.α (`WeeeL`), where agr.O dominates CpL. In non-multiplicity languages (P.M.β), either value of P.S1 is possible, separating L2 and L3 in the typology. L2, with P.S1.β, represents an asymmetric language like Chichewa, and L3, with α, represents a symmetric counterpart like Zulu.

(32) P.OM + P.D + P.M + P.S1 value table

L	P.OM	P.D	P.M	P.S1	OM	Doubling	Multiplicity	Symmetry
L1	β		β		no		no	
L2	α	β	β	β	yes	no	no	no
L3	α	β	β	α	yes	no	no	yes
L4	α	β	α	α	yes	no	yes	yes
L5-L6	α	α	β		yes	yes	no	
L7-L8	α	α	α		yes	yes	yes	

Among the doubling languages (P.D.α) the multiplicity distinction determines which constraints are involved in symmetry. All languages allow OMs with both pro and lexical objects. In non-multiple languages, however, the number of OMs is limited to one. An input of the form [VIp] raises a conflict in which object is marked: the indirect, satisfying agr.HO in marking the highest object but violating m.ff in having doubling; or the direct, violating agr.HO but satisfying m.ff because the direct object is pro. These are the antagonists of P.S2, stated in (33), with the value ERCs in (34).

(33) P.S2: m.ff <> agr.HO

(34) P.S2: values

Value	[VIp]→	agr.O	agr.HO	m.ff	m.XX	CpL
α	[dm[IVp]]~[im[IVp]]		L	W		
β	[im[IVp]]~[dm[IVp]]		W	L		

The values of P.S2 divide only those languages at the intersection of P.D.α and P.M.β (L5 and L6). A P.S2.β value defines an asymmetric language like Swahili, where the indirect object is always marked regardless of the form of the objects. The opposite value defines a symmetric language, like Lozi, where the form of the object (pro or overt object) determines the OM rather than its position.

(35) P.OM + P.D + P.M + P.S1 + P.S2 value table

L	P.OM	P.D	P.M	P.S1	P.S2	OM	Doubling	Multiplicity	Symmetry	
L1	β		β			no		no		
L2	α	β	β	β		yes	no	no	no	
L3	α	β	β	α		yes	no	no	yes	
L4	α	β	α	α		yes	no	yes	yes	
L5	α	α	β		β	yes	yes	no		no
L6	α	α	β		α	yes	yes	no		yes
L7-L8	α	α	α			yes	yes	yes		

Finally, among the doubling, multiple languages, symmetry manifests in whether all lexical objects are doubled, or only the highest (all pros are always doubled). This choice is encoded in a final P.S3. For an input of the form [VID], OMs occur for both objects when agr.O dominates m.ff, which is violated in outputs

with doubling.[11] The reverse ranking results in the optimality of the form with an OM for the indirect object only, lessening the violation of m.ff (37).

(36) P.S3: agr.O <> m.ff

(37) P.S3: values

Value	[VID]→	agr.O	agr.HO	m.ff	m.XX	CpL
α	[im[dm[IVD]]]~[im[IVD]]	W		L		
β	[im[IVD]]~[im[dm[IVD]]]	L		W		

Adding P.S3 yields the final value table (38). Each language is defined by a unique combination of property values, aligning with OM traits along the parameters considered here. P.S3 values split L7 and L8. L7 represents the asymmetric language, like Sambaa, while L8 is the symmetric, like Haya.

(38) All properties value table

L	P.OM	P.D	P.M	P.S1	P.S2	P.S3	OM	Doubling	Multiplicity	Symmetry		
L1	β		β				no		no			
L2	α	β	β	β			yes	no	no	no		
L3	α	β	β	α			yes	no	no	yes		
L4	α	β	α	α			yes	no	yes	yes		
L5	α	α	β		β		yes	yes	no		no	
L6	α	α	β		α		yes	yes	no		yes	
L7	α	α	α			β	yes	yes	yes			no
L8	α	α	α			α	yes	yes	yes			yes

4.5 Properties and languages

The typology languages are matched with the extensional languages discussed in §2 based on where they allow OMs for the majority of objects (which may be optional or required depending on the object). The present study abstracts away from further distinctions made on the basis of features such as specificity or animacy. For example, L5 represents a language like Swahili, in which doubling is possible (P.OMα), multiple OMs are not (P.Mβ), and objects are treated asymmetrically (P.S2α). The facts in Swahili are more complex in that doubling is required for animate objects but optional for inanimates (though with an interpretative difference in definiteness/specificity; (Marten and Kula 2012:(16))).[12]

While properties describe ranking conditions holding over multiple languages within the typology, it also defines the languages at the individual level, as each is a distinct combination of the choices represented by property values. To show this, the properties of two languages are discussed in detail below. L4 represents a Chaga-like language, while L7 is Sambaa-like. While both allow multiple OMs, they differ in doubling and symmetry.

The four property values defining L4 are given in (39a), with the associated trait in the final column. The resulting grammar is represented in the Hasse diagram in (39b).

(39) L4 (Chaga)

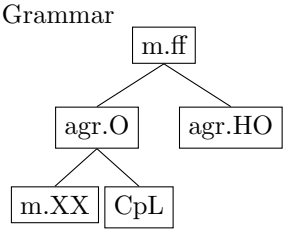
a. Property values

Value	agr.O	agr.HO	m.ff	m.XX	CpL	Extensional trait
P.OMα	W	W			L	OM: [om[Vp]]
P.Dβ	L	L	W			no doubling: [VO]
P.Mα	W			L	L	multiple: [im[dm[pVp]]]
P.S1α	W				L	asymmetric: [dm[IVp]]

[11] Recall that for doubling to be possible (P.D.α), either agr.O or agr.HO dominates m.ff; this still allows for m.ff to dominate agr.O (when agr.HO is the dominator).

[12] It would be possible to account for some of these factors through additional agreement constraints specific to animate and/or definite objects. Such additions would increase the typology, though the structures of the smaller one described here would be preserved.

b. Grammar



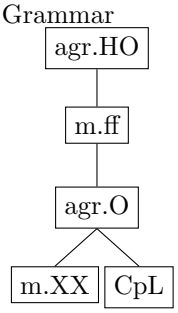
L7 differs from L4 in the value of P.D, the doubling-related property. Therefore, asymmetry results from the value of P.S3 rather than P.S1. The values are shown in (40a) and the Hasse diagram grammar in (40b).

(40) L7 (Sambaa)

a. Property values

Value	agr.O	agr.HO	m.ff	m.XX	CpL	Extensional trait
P.OMα	W	W			L	OM: [om[Vp]]
P.Dα	W	W	L			doubling: [om[VO]]
P.Mα	W			L	L	multiple: [im[dm[pVp]]]
P.S3β	L		W			asymmetric: [im[IVD]]

b. Grammar



Evident from comparing the two grammars is that they differ minimally: only in the relative ranking of agr.HO and m.ff. This comes from their differing values on P.D.

While the focus has been on Bantu languages, the analysis extends to other languages. English exemplifies an OM-less language; Nahuatl has an obligatory OM; Spanish clitic doubling in certain contexts parallels context-specific OMs.

5 Summary and Conclusions

The analysis developed in this paper explained patterns of OM variation through the interaction of agreement constraints with structural and configurational constraints that are violated by the structures with OMs. Rather than positing distinct types of OMs, such as agreement markers versus pronouns, all languages select from the same possible forms. All share the same agree-promoting constraints, but these are variously subordinated to different antagonists violated by aspects of the agreeing candidates. It is possible to have both agreement and non-agreement in the same language, for different inputs.

References

Alber, B., DelBusso, N., and Prince, A. (2016). From intensional properties to Universal Support. *Language: Phonological Analysis*, 92.2:e88–e116.

Alber, B. and Prince, A. (In prep.). *Typologies*. Ms, Free University of Bozen-Bolzano & Rutgers University.

Baker, M. (2018). On the status of object markers in Bantu languages. In O. Adesola, A. A. and Orie, O. O., editors, *Data-Rich Linguistics: Papers in honor of Yiwola Awoyale*, pages 2–40. Cambridge Scholars Publishing.

Baker, M. and Kramer, R. (2018). Doubled clitics are pronouns: Amharic objects (and beyond). *NLLT*, 36.4:1035–1088.

Bresnan, J. and Mchombo, S. (1987). Topic, pronoun, and agreement in Chichewa. *Language*, 63:741–782.

Bresnan, J. and Moshi, L. (1990). Object asymmetries in comparative Bantu syntax. *Linguistic Inquiry*, 21.2:147–185.

DelBusso, N. (2018). *Typological structure and properties of Property Theory*. PhD diss, Rutgers University.

Downing, L. (2018). Differential object marking in chichewa. In Seržant, I. and Witzlack-Makarevich, A., editors, *Diachrony of differential argument marking*, pages 35–59. Language Science Press, Berlin.

Grimshaw, J. (1997). The best clitic: Constraint conflict in morphosyntax. In Haegeman, L., editor, *Elements of grammar: A handbook of contemporary syntactic theory*, pages 169–196. Kluwer, Dordrecht.

Grimshaw, J. (2001). Economy of structure in OT. Technical report, Rutgers University.

Grimshaw, J. (2005). *Words and structure*. Chicago: CSLI Publications.

Marten, L. and Kula, N. C. (2007). Parameters of morpho-syntactic variation in Bantu. *Transactions of the Philological Society*, 105.3:253–338.

Marten, L. and Kula, N. C. (2012). Object marking and morphosyntactic variation in Bantu. *Southern African Linguistics and Applied Language Studies*, 30.2:237–253.

Mchombo, S. and Firmino, G. (1999). Double object constructions in Chichewa and Gitonga: A comparative analysis. *Linguistic Analysis*, 29.1-2:214–233.

Prince, A. (2002). Entailed Ranking Arguments. Technical report, Rutgers University.

Prince, A., Merchant, N., and Tesar, B. (2020). OTWorkplace. https://sites.google.com/site/otworkplace/.

Prince, A. and Smolensky, P. (1993/2004). *Optimality Theory: Constraint interaction in generative grammar*. Wiley Blackwell, Malden, MA.

Riedel, K. (2009). *The syntax of object marking in Sambaa*. PhD diss, University of Leiden.

Van der Spuy, A. (1993). Dislocated noun phrases in Nguni. *Lingua*, 90.4:335–355.

van der Wal, J. (2015). Bantu object clitics as defective goals. *Revue Roumaine de Linguistique*, LX.2-3:277–296.

van der Wal, J. and Hyman, L., editors (2016). *The Conjoint/Disjoint Alternation in Bantu*. De Gruyter Mouton, Berlin.

Woolford, E. (2001). Conditions on object agreement in Ruwund (Bantu). In Benedicto, E., editor, *The UMass Volume on Indigenous Languages*. GLSA, Amherst, MA.

Zeller, J. (2012). Object marking in isiZulu. *Southern African Linguistics and Applied Language Studies*, 30.2:219–235.

Zeller, J. (2014). Three types of object marking in Bantu. *Linguistische Berichte*, 239:347–367.