AGREEMENT AND COORDINATION IN XITSONGA, SESOTHO AND ISIXHOSA: AN OPTIMALITY THEORETIC PERSPECTIVE

A thesis submitted in fulfilment of the requirements for the degree of

MASTERS IN LINGUISTICS

at

RHODES UNIVERSITY

by

HAZEL MITCHLEY

July 2015

Abstract

This thesis provides a unified Optimality Theoretic analysis of subject-verb agreement with coordinated preverbal subjects in three Southern Bantu languages: Xitsonga (S53), Sesotho (S33), and isiXhosa (S41). This analysis is then used to formulate a typology of agreement resolution strategies and the contexts which trigger them.

Although some accounts in the Bantu literature suggest that agreement with coordinate structures is avoided by speakers (e.g. Schadeberg 1992, Voeltz 1971) especially when conjuncts are from different noun classes, I show that there is ample evidence to the contrary, and that the subject marker used is dependent on several factors, including (i) the [±HUMAN] specification on the conjuncts, (ii) whether the conjuncts are singular or plural, (iii) whether or not the conjuncts both carry the same noun class feature, and (iv) the order of the conjuncts.

This thesis shows that there are various agreement resolution strategies which can be used: 1) agreement with the [+HUMAN] feature on the conjuncts, 2) agreement with the [-HUMAN] feature on the conjuncts, 3) agreement with the noun class feature on both conjuncts, 4) agreement with the noun class feature on the conjunct closest to the verb, and 5) agreement with the noun class feature on the conjunct furthest from the verb. Not all of these strategies are used by all languages, nor are these strategies interchangeable in the languages which do use them – instead, multiple factors conspire to trigger the use of a specific agreement strategy within a specific agreement featural context.

I show that these effects can be captured using Optimality Theory (Prince and Smolensky 2004). The analysis makes use of seven constraints: RES#, MAX[+H], MAX[-H], DEP[-H], MAXNC, DEPNC, and AGREECLOSEST. The hierarchical ranking of these constraints not only accounts for the confinement of particular strategies to specific agreement featural contexts within a language, but also accounts for the cross-linguistic differences in the use of these strategies. I end off by examining the typological implications which follow from the OT analysis provided in this thesis.

Acknowledgments

Writing a thesis is never easy, and although people might tell you this before you start, you never fully understand until you find yourself in a dark tunnel, desperately looking for the light that signals the end of the journey. Bearing this in mind, I owe a huge debt of gratitude to the people who helped me, in any way, make it to the end of this thesis.

There are in particular two people who I cannot thank enough: My supervisors, Dr Mark de Vos and Dr Will Bennett. Mark went above and beyond the duties of a supervisor – in addition to guiding me in my research, he also dedicated a significant amount of time discussing the various analyses, dissecting the pros and cons of wording constraints in one way or another, considering how the ordering of chapters affects continuity, and counselling me through the inevitable writer's block. He also showed me how to use LaTeX, which made formatting this thesis a dream and eliminated hours spent trying to work out why a tab keeps moving, or trying to fix all the unavoidable cross-referencing errors. Without Mark this thesis would be very different. I am also incredibly grateful to Will, who only arrived at Rhodes after the start of this thesis, but was nonetheless willing to come onboard as a co-supervisor. I have learnt so much from Will, and I know that this thesis would be significantly poorer if it were not for his input.

I owe a debt of gratitude to the all the staff and post-graduate students in the Rhodes linguistics department: They became a second family to me, and were always there through the ups and downs of both thesis-writing and life in general. Of the staff, I must particularly thank Dr Ron Simango, who was always willing to talk about aspects of my data which I found puzzling, and draw parallels with data from other Bantu languages. Dr Sally Hunt, in her role as the post-graduate course-coordinator also deserves a special mention for, amongst many other things, her constant encouragement, and her much-needed reminders that writing a thesis is like eating an elephant. Finally, I must express my gratitude for Nomvula Sitole, our department administrator, who patiently put up with my (at times frequent) questions about isiXhosa.

I would like to express my gratitude to all the people who participated in this study: Thina Hlatswayo, Thandi Bombi, Trecia Makhubele, Nhlanhla Hlatshwayo, Sinesipho Mbusi, Zondelela Njaba, Lerato Lelia, Lesa Zwane, Tshwarelo Sikhweni and Kay Mosiane. I am extremely grateful for their invaluable help.

I've also benefited from the questions and comments raised by the audiences of LSSA/SAALA/SAALT conferences of 2012 and 2013, as well as the SAMWOP workshop of 2013.

The financial assistance of the National Research Foundation (NRF) towards this research is hereby acknowledged. Opinions expressed and conclusions arrived at are those of the author and are not necessarily to be attributed to the NRF.

Finally, to the friends and family who have supported me through everything, I offer my sincerest thanks. Your support and encouragement means more than words can express. I would especially like to thank my mother, Bobby, father, Mike, and sister, Jacyn, for everything they have done: They kept me sane, and helped me retain my sense of humour when I needed it most. Unfortunately Dad did not get to see the end of this thesis. This thesis is dedicated to him.

List of glosses

1ST	First person
3RD	Third person
DEF	Definite article
CAUS	Causative
FEM	Feminine
FV	Final vowel
LOC	Locative
MASC	Masculine
NEUT	Neuter
PASS	Passive
PAST	Past tense
PERF	Perfect
PL	Plural
POSS	Possessive
PRES	Present tense
RECIP	Reciprocal
SG	Singular
SM	Subject marker
STAT	Stative

Contents

Chapter 1

Introduction

In the thesis I examine the issue of subject agreement in Bantu languages, and explore the agreement resolution strategies which speakers use when a coordination phrase (&P) occupies the subject position (Spec-TP), resulting in multiple potential controllers of subject agreement. In particular, I explore the different factors which may trigger the use of different agreement resolution strategies. I then go on to show that these strategies are not, in fact, ontological primitives, but rather the surface manifestations of the interactions between underlying constraints. The analysis of these interactions is formalized within an Optimality Theoretic framework (Prince and Smolensky 2004).

The thesis first focuses on the agreement resolution strategies employed in three Southern Bantu languages, and then goes on to explore a typology of agreement resolution strategies and the contexts in which they can occur. The languages chosen for the study each come from a separate subgroup of the S-zone group of the revised Guthrie classification system (Guthrie 1948; Maho 2009): Sesotho, S33, from the S30 group (the Sotho-Tswana family); isiXhosa, S41, from the S40 group (the Nguni family); and Xitsonga, S53, from the S50 group (the Twsa-Rhonga family). Two varieties of Xitsonga emerged during the course of the study; as both varieties are descriptively interesting, both are discussed and analysed in this thesis.

A recurring theme in some studies of agreement with conjoined noun phrases in Bantu languages is that it is often avoided where possible (Schadeburg (1992:22) says this of Kiswahili, for example), or that it is entirely ineffable – Voeltz (1971:44), for example, says that "in Xhosa it is not possible to resolve agreement conflict of conjoined nouns of

different genders". According to Voeltz, in isiXhosa the only way to grammatically express a sentence which involves coordination of conjuncts from two different noun classes is to use the comitative construction: One of the conjuncts is extraposed to the end of the sentences, and the non-extraposed conjunct controls the agreement. Thus, for example, if one wanted to say '*The doctor and the dancer are going home*', one would have to phrase it as either (1a) or (1b):

(1) a. *I-gqira* *li-ya-goduka* *nom-danisi*
5-doctor SM5-PRES-go.home with.1-dancer
"The doctor is going home with the dancer."

b. *Um-danisi* *u-ya-goduka* *ne-gqira*
1-dancer SM1-PRES-go.home with.5-doctor
"The dancer is going home with the doctor." (IsiXhosa; own data)

But what if the dancer isn't *that* kind of dancer, and the two are actually going home separately? As it turns out, it is actually possible to agree with coordinated DPs in isiXhosa. There are various strategies which speakers use to resolve agreement, and these strategies are employed in different contexts.

For this thesis I collected several hundred sentences involving coordinated DPs in subject position, and then closely examined the data to determine what resolution strategies are possible in what context. I then account for both the strategies and their confinement to particular contexts with reference to universal markedness and faithfulness constraints in an Optimality Theoretic (OT) framework.

§1.1 describes in more detail the problem which is at the heart of this thesis, and discusses the scope of inquiry, while §1.2 lays out the assumptions which form the base of the Optimality Theoretic analyses of the data. Finally, §1.3 outlines the structure of this thesis, and briefly delineates the content covered in each of the following chapters.

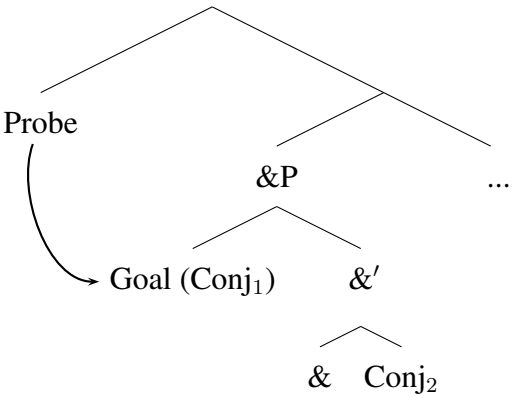
1.1 The problem

When a coordination phrase (&P) occupies subject position (Spec-TP), it is conventionally accepted that there are at least two potential controllers of subject verb agreement[1]: The

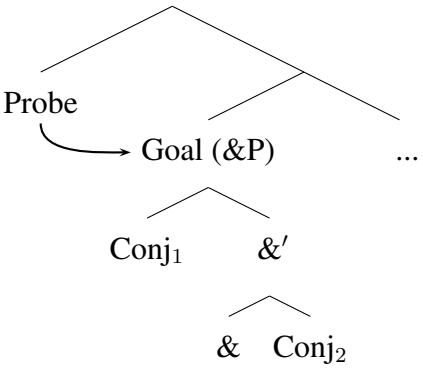
[1]Although this thesis does not have a Minimalist focus, the Minimalist view of agreement is assumed. These assumptions, and the terminology associated with Minimalist syntax, are explained in greater detail in § 2.1.2 on page 20.

agreeing element (the probe) could either agree with the conjunct which is hierarchically closest (2a), or with the whole &P (2b) (van Koppen 2005).

(2) a. Agreement with one conjunct



b. Agreement with whole &P



(Adapted from Riedel 2009:192)

If the probe agrees fully with one of the conjuncts, this is considered to be syntactic agreement: AGREE has simply checked uϕ-features of the verb against the iϕ-features on the conjunct in question. By contrast, if the probe agrees with neither conjunct, it is said to be agreeing with the &P complex; this is typically through some mechanism of semantic agreement, as the &P does not possess inherent ϕ-features, and therefore must first acquire features though some sort of semantic resolution, which resolves the conflict caused by the two conjuncts bearing incompatible ϕ-feature bundles.

What such analyses do not account for is that syntactic and semantic agreement can manifest in different surface resolution mechanisms. For example, syntactic agreement can emerge as 1) partial agreement – i.e. agreement with one of the conjuncts (as in (2a) above), but also as 2) morphological agreement (agreement with both conjuncts, if the conjuncts have identical features), or 3) agreement with both conjuncts if the conjuncts have different features, but the agreement morphs for both conjuncts have identical phonological forms –

this is called phonological resolution. Additionally, semantic agreement can surface either in the form of a default agreement marker, the choice of which is entirely unmotivated (for example, the use of masculine agreement as the default in French), or as a default marker whose use is motivated by the presence of certain semantic features on the conjuncts.

Previous studies of agreement with &Ps in Bantu (Givón 1970; Roberts and Wolentis 1974; Corbett and Mtenje 1987; Bosch 1988; Moloto 1992; Marten 2000, 2005; Riedel 2009; Simango 2012) have found both syntactic and semantic agreement to be attested strategies. Semantic agreement is typically achieved through the use of a default agreement marker, the choice of which, in Bantu, is governed by the animacy, or, more accurately, the 'humanness' of the conjuncts. Syntactic agreement, on the other hand, manifests differently in different languages. Some studies report that morphological agreement is the only way in which syntactic agreement can arise, although a language which allows morphological agreement may also allow phonological agreement if there are homophonous agreement markers in the language (Voeltz 1971). Partial agreement appears to be less popular, although it is still used in some languages (such as Kiswahili (Marten 2000)). However, while previous studies have identified such strategies, what is missing is 1) a comprehensive description of which environments favour which strategies, and 2) an analysis of *why* such environments favour those strategies, and why this differs cross-linguistically.

As chapters 4-6 show, the value of the [NUMBER], [NOUN CLASS], and [±HUMAN] features all have the potential to affect the choice of the agreement strategy, as does the issue of whether the conjuncts are **NC-balanced** (i.e. both conjuncts carry the same noun class feature) or **NC-unbalanced** (i.e. the conjuncts carry different noun class features).

The table in (3) below, for example, shows how speakers of isiXhosa, Sesotho and two varieties of Xitsonga resolve agreement in four different scenarios: 1) where the conjuncts are plural, NC-balanced, and both bear a [+HUMAN] feature; 2) where the conjuncts are plural, NC-balanced, and both bear a [-HUMAN] feature; 3) where the conjuncts are plural, *NC-unbalanced*, and both bear a [+HUMAN] feature, and finally, 4) where the conjuncts are *singular*, NC-unbalanced, and both bear a [+HUMAN] feature. The table ignores surface resolution strategies, and focuses exclusively on whether the language uses semantic or syntactic agreement in these contexts.

(3) Differences in choice of agreement strategies with the same 4 sets of feature bundles

	Feature bundles on each conjunct			
	[NC6, PL, +H] & [NC6, PL, +H]	[NC6, PL, -H] & [NC6, PL, -H]	[NC6, PL, +H] & [NC8, PL, +H]	[NC5, SG, +H] & [NC7, SG, +H]
Xitsonga1	Semantic	Semantic	Semantic	Semantic
Xitsonga2	Syntactic	Semantic	Semantic	Semantic
Sesotho	Syntactic	Syntactic	Semantic	Semantic
isiXhosa	Syntactic	Syntactic	Syntactic	Semantic

The table above shows how just changing the [±HUMAN] feature on the conjuncts can affect the choice of syntactic/semantic agreement in one variety of Xitsonga, whereas this does not affect the choice in Sesotho and isiXhosa.[2] Additionally, by comparing the third column with the first column, one can see that Sesotho is sensitive to whether the conjuncts are NC-balanced or NC-unbalanced, and that this will affect the choice between syntactic and semantic agreement; by contrast, isiXhosa is not affected by this. The choice of agreement strategy in isiXhosa is, however, affected by the number feature on the conjuncts: By comparing the third and forth columns, one can see that when all else is held constant[3] but the number feature changes, the agreement strategy changes too.

The table in (3) represents only four points of data from a much larger set of data collected for this thesis, and this is merely provided here to illustrate the fact that the coordination of conjuncts with the same sets of feature bundles can trigger different resolution strategies across different languages. However, it is not satisfying to merely list which strategies are used in which contexts; providing stipulations such as "in language L, in context C, agreement resolution strategy S is used" offers no theoretical explanation of why certain contexts allow for the use of one strategy, but not another. It is more interesting to try and understand what it is about those contexts which causes them to trigger the use of particular agreement strategies, and why this varies across different languages. This thesis sets out to do precisely this, within the framework of Optimality Theory.

[2] At least not in terms of syntactic or semantic agreement strategies; as will be shown in chapters 5 and 6, in cases where semantic agreement is used, the value of the [±HUMAN] feature will affect the choice of subject marker.

[3] As noun class and number features are inherently linked, it is not possible to keep the noun class constant while changing the number feature; nonetheless, as NC5 is the singular class which corresponds to NC6, and likewise for NC7 with NC8, the *gender* of the conjuncts has been held constant. Please see page 18 for a discussion on the difference between noun class and gender.

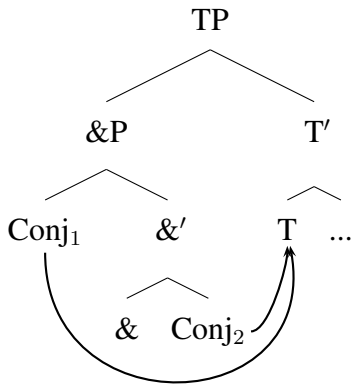
1.2 A brief explanation of the assumptions underlying the OT analyses in this thesis

The way in which I apply Optimality Theory in this thesis is somewhat different from the way it is applied in phonology and other morphosyntactic studies.[4] This section therefore explains the assumptions regarding the input, the candidate set and the constraints used in the following chapters.

1.2.1 The input and output

The most important difference between this application of OT and most other OT studies is that it does not involve a simple one-to one mapping from input to output. Instead, it involves a two-to-one mapping: There are two feature bundles in SpecTP – one on each conjunct – and yet only one feature bundle appears on the verb in T (4).

(4)



The fact that only one feature bundle can appear on the verb means that one's grammar must have a mechanism in place to decide which features to copy across, and which to ignore. Thus the *two* feature bundles provided by the conjuncts act as the input, while the subject marker which appears on the verb, bearing only *one* feature bundle, is the output of the mapping operation.

It is up for debate as to precisely where in the derivation the two sets of feature bundles are mapped onto one bundle. One could argue that it is at the level of the &P: That the two feature bundles percolate up to the &P, resolution happens here, and then AGREE checks the probe against this resolved feature bundle. This however posits a look-ahead mechanism,

[4] §2.3 and §2.4 provide a detailed explanation of the Optimality Theory framework, as well as how it has been applied in other morphosyntactic studues.

which is not traditionally favoured in Minimalist syntax. Alternatively, one could argue that AGREE happens first, where both features bundles are accessed simultaneously by the probe, and are both copied to T, and then that it is at this point that resolution occurs, mapping the two different feature bundles to one bundle which can be matched with a form in the lexicon. While this approach appears to be more elegant, it seems to discount the hierarchical structure of the &P, and the unequal status held by the two DPs (see Zhang (2009)). However, the site of resolution does not affect my analysis, as this thesis is more interested in exploring the mechanisms of resolution, rather than the locus of resolution. What *is* important is the stipulation that the two feature bundles in the input are *ordered*, and that the mapping operation is sensitive to this order.

It is also important to clarify exactly which features make up the feature bundles in the input and output. The bundles each consist of valued versions of the following features:

a. [NUMBER],

b. [NOUN CLASS], and

c. [±HUMAN].

These are somewhat similar to the traditional ϕ-features: [NUMBER], certainly, is one of the basic ϕ-features, and [NOUN CLASS] is very similar to [GENDER].[5] The last of these features, [±HUMAN], is not traditionally considered as a ϕ-feature, however, I argue that as some languages seem to actively encode animacy in their syntax,[6] information about animacy must be available to the syntax in the form of features.

To summarize then: The input consists of *two* features bundles, one bundle from each conjunct, and each feature bundle consists of [NUMBER], [NOUN CLASS], and [±HUMAN] features. The output, on the other hand, consists of *a form* (the subject marker) which carries only *one* feature bundle.

[5]I do not consider Noun Class and Gender to be exactly the same - please see the discussion on page 18 in chapter 2. I accede to the pairing of a singular noun class with its corresponding plural noun class being considered as a gender (e.g.: gender 1/2, which encompasses noun classes 1 and 2, and gender 3/4, which encompasses noun classes 3 and 4, etc). However, the Noun Class distinction is important to this analysis, as the notion of gender seems to play almost no role in agreement resolution - for example, two conjuncts from NC3 will take default agreement, instead of, as one might expect, NC4 agreement - and thus the analyses in this thesis focus more on the [NOUN CLASS] feature than the [GENDER] feature.

[6]In Dutch several substantive qualifiers require the suffix '-n' when they modify human referents (de Swart et al. 2008); In Malayalam, accusative case can only be marked on animate objects (*Ibid*), which parallels several Bantu languages (such as Sambaa, Kiswahili, Ruwund, and Kirimi) which largely limit object marking to animate objects (Riedel 2009; Woolford 1999).

1.2.2 GEN and the candidates

As the output is a morph (i.e. a specific form associated with specific information), this logically means that the candidates must all be morphs too. This naturally has implications for the way in which GEN is conceived. Traditionally, GEN is considered to be infinite, in that the candidate set which it produces is not restricted to forms available in a particular language. While this works for phonological applications of OT, as they involve combinations of segments, this cannot be the case for morphosyntactic analyses: If the output is a morph, then the candidate set can only consist of forms which are recognised as morphs in the language in question. Thus GEN cannot indiscriminately produce different combinations of features and randomly associate them with a form; instead, GEN draws from the lexicon, and only produces candidates which exist as morphs in the language.[7] Specifically, in the analyses in this thesis the candidate set consists of all the subject markers in that language. The implication of this is that the candidate sets cannot be identical across different languages, as not all languages use the same form to encode the same information.

As noted, the morphs which constitute the candidate set each come with a predefined set a features which they encode. Unlike the feature bundles in the input, which are determined by the features on the conjuncts, the features on the subject markers are stipulated as part of the language. While this seems intuitive for the [NUMBER] and [NOUN CLASS] features, it is less so for the [±HUMAN] feature. However, I argue that each subject marker encodes one of three options: 1) human ([+H]), or 2) nonhuman ([-H]), or 3) under-specified regarding the [±HUMAN] feature ([ØH]). The assumptions about the value of the [±HUMAN] feature are derived from the data: If human-denoting conjuncts systematically trigger the use of one subject marker, and nonhuman-denoting conjuncts systematically trigger the use of a different subject marker, it is argued that these two subject markers respectively bear a [+HUMAN] feature and a [-HUMAN] feature. If a subject marker is not systematically triggered as default agreement with [+HUMAN] or [-HUMAN] conjuncts, then that subject marker is assumed to have an under-specified human feature.

Importantly, the value of the [±HUMAN] feature on the subject marker is not determined by the value of the [±HUMAN] feature of the conjuncts, nor does it restrict the class of entities which fall into the corresponding noun class. This means that even though SM8, for ex-

[7] This mirrors other candidate sets in OT analyses of morphosyntactic data - see for example Grimshaw (1997a); Bresnan (2001).

ample, might bear a [-HUMAN] feature, it does not follow that NC8 is devoid of [+HUMAN] nouns, or that SM8 cannot be used to agree with a [+HUMAN] noun.

Another important issue to consider is the issue of homophony. Some languages have two or more homophonous subject markers (SMs). In some cases the homophonous SMs are distributionally distinct, while in other cases they are not. Consider, for example, the following two pairs of subject markers from Sesotho: SM8 & SM10 (*di*), and SM4 & SM9 (*e*). Chapter 5 shows that although Sesotho typically uses 'morphological agreement' if there is a plural subject marker which will agree with the NC-feature of both conjuncts, there is an exception to this: If morphological agreement would result in the use of a subject marker which clashes with the [±HUMAN] feature on the conjuncts, default agreement is resorted to instead. This happens in two contexts: 1) with plural NC-balanced [+HUMAN] conjuncts from NC8 (as illustrated in 5a), and 2) with plural NC-balanced [+HUMAN] conjuncts from NC10 (as illustrated in 5b) – in these cases, SM2 is used instead of SM8 and SM10, respectively.

(5) a. *Di-ra* *le* *di-tsibi* ***ba (/∗di)*** *bua*
8-enemy and 8-expert **SM2 (/∗SM8)** talk
'The enemies and the experts are talking.'

b. *Di-nese* *le* *di-ngaka* ***ba (/∗di)*** *bua*
10-nurse and 10-doctor **SM2 (/∗SM10)** talk
'The nurses and the doctors are talking.'

Instead of assuming that Sesotho uses two different plural subject markers as a default SM for [-HUMAN], and that they both just happen to be realised as *di*, it is much simpler to stipulate that there only one [-HUMAN] default, *di*, and that it carries both SM8 and SM10 features.[8] By contrast, SM4 and SM9 behave differently. Firstly, these subject markers have different number features (PL and SG, respectively), and secondly, when conjuncts from NC4 and NC9 are coordinated, the SM used is not *e*, even though on a phonological level it could be said to agree with the individual conjuncts; its failure to agree with both simultaneously is taken as evidence of there being two separate subject markers with the same form, rather than one subject marker with both NC4 and NC9 features.[9] Thus subject

[8]Please note that this does not mean that NC 8 and NC10 are actually the same noun class: The noun prefixes in these classes are distinct – the prefix for NC8 is *di-*, while the prefix for NC10 is typically *diN-*.

[9]It should also be noted the SM1 and SM3 are also homophonous in Sesotho, however, as *both* these subject markers are singular, the issue of whether or not they should count as one candidate is moot, as singular subject markers were never chosen for agreement with &Ps.

markers which are homophonous but distributionally distinct are considered to be separate candidates, whilst subject markers which are homophonous and *not* distributionally distinct are considered to be one and the same.

1.2.3 The constraints

This thesis makes use of seven constraints, one of which is a markedness constraint. This constraint, RES#, addresses the computational work done by an &P, as it requires that an &P triggers plural agreement, regardless of the number features on the individual conjuncts.

RES# : Coordinated DPs in subject position trigger plural agreement on the predicate.

The remaining six constraints are all faithfulness constraints which fall mainly into two families: 1) the DEP-type, and 2) the MAX-type constraints. In each case these constraints address faithfulness to the [NOUN CLASS] and [±HUMAN] features.

1.2.3.1 The DEP family

In phonology DEP is traditionally used to mean 'don't epenthesize', and I have extended this meaning to a prohibition against features in the output which don't appear in the input. The DEP[X] constraints therefore use the following templatic definition:

DEP[X] : If feature [X] is in the output, then there must be some corresponding feature [X] in the input.

Thus a violation mark is assigned to a candidate if the candidate has feature [αX], but there is no corresponding [αX] feature on *either conjunct*. The similarity to the traditional 'don't epenthesize' function of DEP is in the fact that DEP[X] penalizes the presence of something in the output for which there is no correspondence in the input.[10]

There are two DEP constraints which are used in this thesis: DEPNC, and DEP[-H]. The effect of DEPNC is to penalize any candidate which does not agree with the NC-feature of at least one of the conjuncts, while DEP[-H] penalizes candidates which carry a [-HUMAN] feature when there is no [-HUMAN] conjunct in the input.

[10]This family of constraints thus fulfill the same function as the FILL FEATURE constraints proposed by Grimshaw (1997a).

1.2.3.2 The MAX family

There are three MAX constraints used in this thesis: MAXNC, MAX[+H], and MAX[-H]. In phonology MAX traditionally means 'don't delete', and I have extended this to input-to-output feature faithfulness. The MAX[X] family of constraints demands that if feature [αX] is in the input, there must be a corresponding feature [αX] on the subject marker in the output.[11] The MAX[X] constraints have the following templatic definition:

MAX[X] : If there is feature [αX] in the input, there must be a corresponding feature [αX] in the output.

The matter is somewhat complicated, however, by the fact that two feature bundles are being mapped to one feature bundle, which almost inevitably results in one or more features from the input being left by the wayside, as *typically* subject markers do not carry multiple values of the same type of feature. This two-to-one mapping problem is reflected in the way in which MAX[X] evaluates candidates: It penalizes a candidate if the value of feature [X] on the subject marker does not match the value of feature [X] on *both* conjuncts.

For example, MAXNC demands that *all* the NC-features in the input must be reflected on the subject marker. The evaluation process is binary in this respect: If the NC-feature on the SM corresponds to both NC-features in the input, then there is no violation; if the NC-feature on the SM does not correspond to *both* NC-features in the input, then a violation mark is assigned. One violation is assigned regardless of whether the subject marker's NC-feature corresponds to one or none of the NC-features in the input. This means that if the two conjuncts bear NC4 and NC10 features, candidates SM4, SM10, and SM6 will all receive *one* violation mark each, even though SM4 and SM10 each agree with the NC-features of one of the conjuncts, while SM6 agrees with the NC-feature of neither conjunct.

MAX[+H] and MAX[-H] follow the same templatic definition as given above, yet there is a small but significant difference between these two constraints and MAXNC. The latter constraint does not specify a value of the feature it demands faithfulness to, meaning that all instances of any NC feature must correspond to the value of of the NC feature on the SM. By contrast, MAX[+H] and MAX[-H] specify the value of the [$\pm$HUMAN] feature, and thus

[11] This family of constraints thus have a similar function to the PARSE FEATURE constraints proposed by Grimshaw (1997a).

don't require that *all* [±HUMAN] features in input must be reflected in the output; instead, they demand faithfulness to a specific value of that feature.

An illustration of the differences between MAX[X] and DEP[X]

There are thus two significant differences between the DEP[X] constraints and the MAX[X] constraints. Firstly, DEP[X] requires that feature X on the output must faithfully correspond to a feature on *at least one* of the conjuncts, while Max[X] requires that *all* instances of feature X in the input must agree with feature X on the subject maker; thus DEP[X] seems to work with an existential quantifier, while MAX[X] seems to work with a universal quantifier.

The second, and more fundamental, difference is that DEP[X] penalizes the presence of features in the output that aren't in the input, while MAX[X] penalizes outputs which don't express features that are present in the input. The tableaux in (6) illustrate this difference.

(6) a.

2[+H] & 6[+H]	MAX[-H]	DEP[-H]
a. ☞ SM2 [NC2, +H, PL]		
b. ☞ SM6 [NC6, ØH, PL]		
c. SM8 [NC8, -H, PL]		*!

b.

10[-H] & 6[-H]	MAX[-H]	DEP[-H]
a. SM2 [NC2, +H, PL]	*!	
b. SM6 [NC6, ØH, PL]	*!	
c. ☞ SM8 [NC8, -H, PL]		

DEP[-H] penalizes a subject marker if it has a [-HUMAN] feature when the input has no such feature; in (6a), the input has no [-HUMAN] features, and yet candidate (c) carries a [-HUMAN] feature. It therefore violates DEP[-H]. By contrast, none of the other candidates have a [-HUMAN] feature, and therefore they are not penalized by this constraint. On the other hand, in (6b), there are [-HUMAN] features present in the input, and so candidate (c) does not violate DEP[-H] here. Again, as the other candidates don't have a [-HUMAN] feature, DEP[-H] doesn't evaluate them against the input, and so they vacuously satisfy the constraint. *Thus Dep[-H] can only ever penalize a candidate which has a* [-HUMAN] *feature, and only if the input does not contain a* [-HUMAN] *feature.*

In contrast, MAX[-H] demands that that if there is a [-HUMAN] feature in the input, then there must be a [-HUMAN] feature in the output. As there is no [-HUMAN] feature in the input for (6a), all of the candidates vacuously satisfy the constraint in this case, regardless of their specification of the [±HUMAN] feature. On the other hand, [-HUMAN] features are present in the input in (6b), and so candidates which do not bear a [-HUMAN] feature are penalized. *Thus Max[-H] only ever penalizes candidates which* ***do not*** *have a [*-HUMAN*] feature, and only if the input* does *contain a [*-HUMAN*] feature.*

1.2.3.3 AGREECLOSEST

While the MAX[X] and DEP[X] constraints account for five of the six faithfulness constraints, the remaining faithfulness constraint could be argued to be part of either family, just with locality conditions added. AGREECLOSEST (abbreviated as CCA in the tableaux) requires that the NC-feature on the subject marker must agree with the NC-feature on the conjunct linearly closest to the verb. The reason that this constraint could be argued to be part of either family is that it is in a stringency relation with both MAXNC *and* DEPNC, albeit in different roles. McCarthy (2008:32) gives the following formal definition for stringency: "Constraint CONST1 is more stringent than constraint CONST2 if every violation of CONST2 is also a violation of CONST1, but there are some violations of CONST1 that are not violations of CONST2."

(7) a.

6[+H] & 6[+H]	Max[NC]	CCA	DepNC
a. SM2 [NC2, +H, PL]	∗!	∗!	∗!
b. ☞ SM6 [NC6, ØH, PL]			
c. SM8 [NC8, -H, PL]	∗!	∗!	∗!

b.

6[+H] & 8[+H]	Max[NC]	CCA	DepNC
a. SM2 [NC2, +H, PL]	∗	∗!	∗!
b. SM6 [NC6, ØH, PL]	∗	∗!	
c. ☞ SM8 [NC8, -H, PL]	∗		

The tableaux above demonstrate the violations assigned by AGREECLOSEST are always a subset of the violations assigned by MAXNC, but also a superset of the violations assigned by DEPNC. This means that MAXNC is more stringent than AGREECLOSEST, and that AGREECLOSEST is more stringent than DEPNC.

Although AGREECLOSEST could therefore be argued to be related to both MAXNC and DEPNC, I argue that seems more closely related to DEPNC, as both these constraints act as mechanisms which allow partial agreement to surface: DEPNC favours any subject marker whose NC-feature agrees with the NC-feature of at least on of the conjuncts, regardless of whether it be the closest conjunct or the furthest conjunct; AGREECLOSEST has a similar function, but with the added stipulation that the agreement happens with the closest conjunct. By contrast, as will be shown in the following chapters, MAXNC acts as a mechanism which allows for 'morphological agreement' (agreement with both conjuncts). Thus, based on the types of resolution strategies these three constraints allow to surface, I argue that AGREECLOSEST is more closely related to DEPNC.

1.3 Overview of the thesis

This thesis is structured as follows. Chapter 2 reviews some literature and theory which is relevant to the topic of the dissertation, particularly regarding 1) agreement, and 2) the Optimality Theory model. The first section is designed as a brief refresher on agreement, what it is, and how it works, while the next section discusses previous work on agreement with &Ps, and looks at different ways in which speakers resolve agreement in such instances. Next follows a brief explanation of Optimality Theory, written with the hope of making the analyses in this thesis more accessible to those not entirely familiar with the framework. I then look at different ways in which OT has been applied to morphosyntactic data, and compare these applications of the framework with the one presented in this thesis. The final section looks at what gaps exist in the literature, and presents research questions which follow from these gaps. Answering these research questions forms the crux of this thesis.

Chapter 3 discusses the methodology used in this study, outlining the data elicitation process used, and the various controls put in place to try and ensure that the constructions being elicited matched what I thought I was eliciting. Additionally, a brief description of the demographics of the consultants used for the study is given.

The next three chapters then go on to describe and analyse the data from each of the three languages in this study. Chapter 4 covers Xitsonga. Two varieties of Xitsonga emerged, based on the resolution strategies employed. One systematically made use of only default agreement even when syntactic/morphological agreement was expected; this generalization

is shown to be the effect of MAX[+H] and MAX[-H] dominating all the constraints requiring faithfulness to noun class features. The second variety is somewhat more complex. Although default agreement is still typically favoured in this variety, there is one context where morphological agreement is preferred: When the conjuncts are plural, NC-balanced, and *both* bear a [+HUMAN] feature. Unexpectedly, the same does not hold true for conjuncts which are plural, NC-balanced and [-HUMAN]. This discrepancy is caused by MAX[+H] and MAX[-H] being ranked differently relative to MAXNC.

Sesotho (chapter 5) does not restrict morphological agreement to plural, NC-balanced, [+HUMAN] conjuncts, but also allows it with plural, NC-balanced, [-HUMAN] conjuncts. Unexpectedly, however, there is a small subset of data from this group which systematically uses default agreement instead. I show that this set of exceptions are the result of DEP[-H] being highly ranked in Sesotho.

Chapter 6 looks at how isiXhosa deals with agreement with conjoined subjects. While default agreement is also widely used in this language, isiXhosa, unlike the other two languages, also allows for closest agreement in certain contexts. This shows the need for a constraint which references the ordering of the conjuncts (the AGREECLOSEST constraint). Surprisingly, in very specific circumstances, agreement with the furthest conjunct is not only allowed, but is in fact *preferred*. This is shown to be the effect of DEPNC interacting with other constraints.

Chapter 7 then discusses a factorial typology of agreement with coordinated subjects. The typology was generated in OTWorkplace (Prince et al. 2014). Although there are 5040 possible linear rankings of the 7 constraints, there are only 159 distinct combinations of agreement strategies which can surface in various contexts. This chapter looks at a set of 47 of these rankings and their outputs (the agreement effects of which are described in full in appendix A), and examines the implications of the ways in which the constraints interact with each other.

Finally, chapter 8 concludes by revisiting the research questions, and showing how they were answered, before suggesting avenues for future research.

Chapter 2

Theoretical Overview

This chapter covers two broad topics which are central to this thesis, namely agreement (§2.1-§2.2) and Optimality Theory (§2.3-§2.4). Each of these two topics is split into two sections, the first outlining the absolute basics, and the second taking a more in depth look at the parts of the topic most relevant to what follows. Finally, §2.5 sets out the research questions which the thesis aims to answer.

2.1 A brief overview of agreement

This section covers the basics of agreement. §2.1.1 explains *what* agreement is, and distinguishes between grammatical gender and nouns, while §2.1.2 briefly discusses *how* agreement works in Minimalist syntax.

2.1.1 What agreement is

The notion of *grammatical agreement* has perhaps been best defined by Steel (1978:610), as the "systematic covariance between a semantic or formal property of one element and a formal property of another". To unpack the definition with reference to a specific example, let us consider the Italian phrase below.[1]

(1) *i* *bei* *ragazzi*
the.PL.MASC beautiful.PL.MASC boy.PL.MASC
'the beautiful boys'

[1]The Italian examples in this chapter were provided by myself.

Here the noun *ragazzi* bears two features (or 'properties', as Steel terms it) in addition to the content morpheme *boy*, namely [PLURAL] and [MASCULINE]; these two features also appear on the determiner and the adjective. What Steel means when she refers to 'systematic covariance', is that whenever determiners and/or adjectives modify a noun which carries the features [PLURAL] and [MASCULINE], then the determiners and/or adjectives will also carry these features. By contrast, if the noun which they modify bears the features [SINGULAR] and [FEMININE], then the modifiers would bear these features, as (2) shows:

(2) *la* *bella* *ragazza*
the.SG.FEM beautiful.SG.FEM girl.SG.FEM
'the beautiful girl'

It is clear that, in examples (1) and (2), the noun determines the features that appear on the determiner and adjective, and thus the determiner and the adjective are said to *agree* with the the noun. Following Corbett (1983b, 2006, *inter alia*), the element which determines the features which appear on other elements is called the *controller*, and the element which agrees with the controller is called the *target*. As is the case with the two Italian phrases above, there may often be more than one target per controller. More complex, however, is the situation where there is more than one controller per target; this issue is discussed in more detail in section 2.2, and is central to this thesis.

2.1.1.1 Grammatical gender

The gender features [MASCULINE] and [FEMININE], mentioned above, are not entirely semantically determined. In many languages, including Italian, all nouns in the language are classified as belonging to one or other 'gender'. Thus even inanimate objects such as *book* or *scarf* may be classified as being 'masculine' or 'feminine', as the the two examples below show.

(3) *il* *bel* *libro*
the.SG.MASC beautiful.SG.MASC book.SG.MASC
'the beautiful book'

(4) *la* *bella* *sciarpa*
the.SG.FEM beautiful.SG.FEM scarf.SG.FEM
'the beautiful scarf'

Thus in Indo-European languages gender is not determined by the biological sex of the entity in question, although for human entities there is often some correspondence between biological sex and grammatical gender.[2] Such languages may be said to assign gender on a *partially* semantic basis: That is, nouns of the same biological sex are usually grouped together, and the remaining nouns are divided out among the genders with no discernable form of systematic semantic grouping.

Some languages assign nouns to genders on a more systematic semantic basis. For example, Tamil, a Dravidian language, has three genders, and nouns are assigned a gender on a strict semantic basis: The masculine gender is reserved for gods and male humans, the feminine gender is reserved for goddesses and female humans, and all the remaining nouns (called the 'residue') are assigned to the neuter gender (Corbett 1991:9).

Another example of a language with a slightly more complicated semantic gender system is Dyirbal, an endangered Australian language. Dyirbal has four genders which Dixon (1972) refers to as Genders I, II, III, and IV. He identifies the semantic core of each of the four genders as follows: Gender I consists primarily of male humans, and non-human animates, Gender II of female humans, water, fire, and fighting, Gender III of non-flesh foods, and the residue falls into Gender IV. Although there are many exceptions, most of them can be accounted for by three principles; unfortunately, owing to space limitations, these principles cannot be discussed here.

As demonstrated in Dyirbal, gender features do not always have to be *masculine, feminine* or *neuter*, but can instead refer to groups which are merely designated by numbers. This is particularly useful when a language has more than three grammatical genders, as is the case in Bantu languages, which all have more than 5 genders, or 10 noun classes.

2.1.1.2 Noun classes

At this point it is necessary to distinguish between grammatical gender and noun classes. Although some (e.g. Corbett (1985)) indicate that the two terms essentially refer to the same phenomenon, I distinguish between the two in this dissertation. The difference comes down to the issue of singular/plural. In languages such as Italian, Tamil, and Dyirbal, nouns

[2]Many exceptions do exist, however: For example, in German, *Mädchen (girl)* is neuter, whilst in French *sentinelle (sentry)* is feminine, even in contexts where the sentry in question is unquestionably male, such as the phrase *the bearded sentry* (Badecker 2007:1543).

typically remain in the same gender, regardless of whether they are singular or plural. By contrast, in Bantu languages a noun class may include only singular items, or only plural items, but the singular and plural forms of the same noun will never be found in the same noun class. For example, in isiXhosa the noun *umfundi* (*'student')* belongs to noun class 1, but *only* when the it is in its singular form; when the noun is in its plural form (*abafundi*), it is assigned to noun class 2. Thus together noun classes 1 and 2 form the equivalent of a grammatical gender. This distinction between gender and noun class is an important one for this thesis.

Givón (1971) posits that Proto-Bantu had a predominantly semantic system for gender assignment. He postulates that originally nouns were categorised according to the schema laid out in (5).

(5) Semantic assignment in Proto-Bantu

class 1/2	humans	class 12/13	small objects
class 3/4	plants	class 14	masses
class 5/6	fruits	class 6	liquids
class 7/8	inanimates	class 15/6	paired body parts
class 9/10	animals	class 15	infinitive
class 11/10	elongated objects		nominalizations

(Givón 1971:33)

Whether Givón's hypothesis is true is still a matter for debate; the fact is that today the semantic categories are spread across the noun classes, with only class 1/2 being reserved almost exclusively for human entities (although humans may also be found in other noun classes).

The noun class system is an integral part of Bantu grammars, as noun class agreement shows up on many other elements in the sentence, including demonstrative determiners, pronouns, adjectives, and verbs, as well as on relative, attributive and possessive morphemes. The verb has the potential to carry two noun class agreement morphemes: One which agrees with the subject, and one which agrees with the object; these morphemes are called the subject marker (SM) and object marker (OM), respectively.[3] The alternation of prefixes on the

[3]Interestingly, some Bantu languages allow the verb to carry additional object markers: The extra object marker(s) can be used to agree with other DPs in the sentence, which could be anything from a benefactor to

numeral and verb in (6) present a simple demonstration of the agreement system.

(6) a. *Um-fundi* ***om****-nye* ***u****-ya-thetha*
1-student **1**-one SM1-PRES-speak
'One student is speaking.'

b. *I-polisa* ***eli****-nye* ***li****-ya-thetha*
5-policeman **5**-one **SM5**-PRES-speak
'One policeman is speaking.' (isiXhosa)

In (6a), the subject *student* belongs to noun class 1, and thus the numeral and the verb both carry a noun class 1 agreement marker. By contrast, in (6b), the noun belongs to class 5, and thus the numeral and the verb carry morphemes with noun class 5 agreement features instead.

Although agreement is ubiquitous in the Bantu languages, this thesis limits itself to the study of subject-verb agreement.

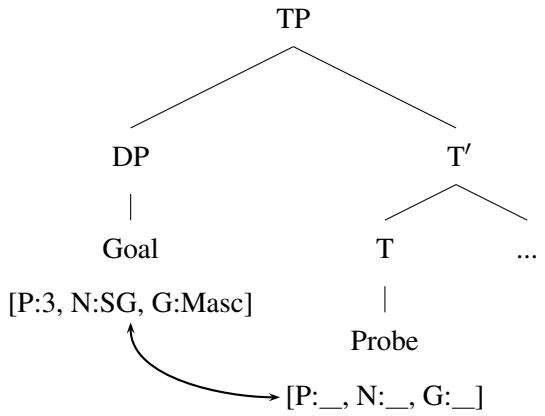
2.1.2 How agreement works

Although this thesis does not have a Minimalist focus, a Minimalist model of agreement (Chomsky 1995, 2000) is assumed. In this model, agreement is the result of a feature checking process between a *probe* and a *goal* (these terms correspond, respectively, to the Corbettian terms *target* and *controller*, mentioned on page 17). The probe is said to have uninterpretable ϕ-features which it checks against the interpretable ϕ-features on the goal.

In simpler terms, what this means is that there is an element (the goal) which has a set of valued features on it (such as [PERSON: 3], [NUMBER: SINGULAR] and [GENDER: MASCULINE]), and an element (the probe) which has a set of unvalued features (such as [PERSON:_], [NUMBER:_] and [GENDER:_]). The presence of these unvalued features triggers agreement: The probe searches within its local C-command domain, looking for valued features which match its unvalued ones, and then 'copies' these valued features.

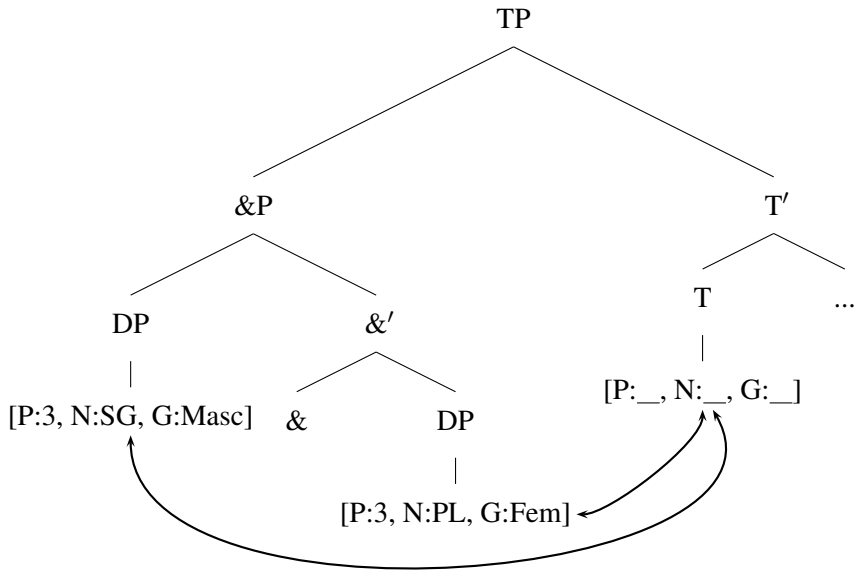
locative phrase. In Chaga, for example, the sentence *The chief sent him there with it* would result in three OMs on the verb: One for *him*, one for *there*, and one for *it* (Moshi 1998:142-3).

(7)



In instances of coordination, however, a problem occurs: there are multiple sets of interpretable ϕ-features against which the probe could check its uninterpretable ϕ-features. An example of this is given in (8).

(8)



As coordination is fairly common in language, grammars must have a way of determining what agreement features to use in such instances. The next section covers different strategies of resolving this agreement problem.[4]

2.2 Strategies for agreeing with coordinated DPs

When two DPs with different ϕ-features are coordinated in Spec-TP, at least three different types of clashes can occur: 1) A clash in number features, 2) A clash in person features, or 3)

[4]Please note that as this thesis focuses on verbal agreement with coordinated subjects, issues such as complementizer agreement and object agreement are not discussed.

A clash in gender/noun class features. The first of these clashes is semantically resolvable: The coordination of nouns with conflicting number features will typically result in plural agreement features on the probe – in fact, this is even the case when both nouns have the [SINGULAR] feature, as semantically coordination always entails more than one entity.

In the second case, where the conjoined nouns carry different person features, there are two resolution rules which tend to be followed cross-linguistically:

(9)
1. If one of the conjuncts carries a first person feature, first person agreement is used.
2. If one of the conjuncts carries a second person feature, second person agreement is used.

These rules are ordered, such that if the two rules conflict with each other, the first one takes precedence over the second one, as the following Czech example shows.

(10) *já* *a* *ty* *zůstaneme* *doma*
I and you will.stay.**1ST**.PL at.home
'You and I will stay at home.' (Trávníček 1949:433, cited in Corbett 1983b:176)

This example shows that when a DP with a first person feature is coordinated with a DP with a second person feature, first person agreement is used, because the rule requiring first person agreement is ordered before the one requiring second person agreement. The consistency with which the resolution rules in (9) are applied cross-linguistically leads Corbett (1983b) to argue that the rules are semantically motivated.

By contrast, unlike number and person features, grammatical gender/noun class is not semantically grounded, and thus the coordination of nouns with conflicting gender/noun class features is more complex, and the strategies used to solve this problem vary significantly cross-linguistically. It is thus the strategies that speakers use to resolve agreement with these conflicting features which will form the focus of this dissertation.

There are four main strategies which speakers tend to use to resolve clashes in gender or noun class agreement: 1) **avoidance**, 2) **default agreement**, 3) **partial agreement**, and 4) **phonological resolution**. However, before discussing these strategies, it is worth noting that in cases of coordination in which there is *no* clash of gender or noun class features, one would expect what Marten (2000) calls **morphological agreement**.

2.2.1 Morphological agreement

Morphological agreement refers to when two conjuncts of the same gender[5] use the plural agreement marker for that gender. For example, in (11), two singular conjuncts with NC1 features are coordinated, and the resulting agreement on the verb is SM2, as NC2 is the plural class for NC1.

(11) *Mw-alimu na mw-anafunzi w-ake wa-li-kuja*
1-teacher and 1-student 1-POSS **SM2**-PAST-come
'The teacher and his student came.' (Kiswahili; Marten 2000:pg5 of MS)

Similarly, in the following Chichewa example, two plural conjuncts from NC13 are coordinated, and SM13 is used on the verb.[6]

(12) *Ti-a-kazi ndi ti-a-na ti-ku-gona*
13-2-woman and 13-2-child **SM13**-PRES-lie
'The little women and the infants are lying down.'
(Chichewa; Corbett and Mtenje 1987:31)

However, morphological agreement is less common than one might assume, especially when conjuncts are singular. For example, when the nouns from example (12) are singular, morphological agreement is no longer allowed, as can be seen in (13).

(13) **Ka-m-kazi ndi ka-mwa-na ti-ku-gona*
12-1-woman and 12-1-child **SM13**-PRES-lie
Int: 'The little woman and the infant are lying down.'
(Chichewa; Corbett and Mtenje 1987:31)

Thus even when conjuncts are NC-balanced (i.e. they come from the same noun class), resulting in no clash in noun class features, morphological agreement is not always allowed.

2.2.2 Avoidance

One option when faced with the issue of coordination resulting in conflicting gender or noun class features is to simply *avoid* using coordination. In such instances the comitative

[5] i.e. from the same singular/plural noun class pairing, in the case of Bantu languages.

[6] Gender 12/13 in Chichewa is used to make nouns diminutive. Note, however, that the NC12 and NC13 prefixes attach not to roots, but to whole *words*. Thus a word retains its original noun class prefix in addition to the diminutive prefix, but the original noun class no longer controls the agreement (Simango, personal communication). The glosses in examples (12) and (13) have been changed from the original to reflect the presence of the NC2 and NC1 prefixes, respectively.

construction can be used. Thus, instead of saying *John and Mary are walking*, one would say *John is walking with Mary*.

This strategy is reported to be common in Bantu languages. Schadeburg (1992:22) describes avoidance as the "most common strategy" for dealing with conflicting noun class features in Kiswahili. Voeltz (1971) takes it further, arguing that it is the *only* strategy available in isiXhosa, unless phonological resolution is possible.[7] By contrast, other languages such as Chichewa (Corbett and Mtenje 1987) and Luganda (Givón 1970) only use avoidance as the preferred strategy under very specific circumstances, such as when human denoting nouns are coordinated with nouns which denote non-human entities.[8] If coordination is forced in such instances, the resulting sentence is ungrammatical, regardless of which agreement marker is used, as illustrated in (14a); instead, the comitative *must* be used, as illustrated in (14b).[9] It is worth noting that the conjunctive morpheme in (14a) and the comitative morpheme in (14b) are homophonous. It would be an understatement to say that this is very common in Bantu languages.

(14) a. **Omu-sajja ne em-bwa-ye **ba/bi**-a-lab-w-a omu-kazi*
1-man and 9-dog-POSS **SM2/SM8**-PAST-see-PASS-FV 1-woman
'The man and his dog were seen by the woman.' (Luganda; Givón 1970:253-254)

b. *Omu-sajja **y**-a-lab-w-a omu-kazi ne em-bwa-ye*
1-man **SM1**-PAST-see-PASS-FV 1-woman with 9-dog-POSS
'The man was seen with his dog by the woman.' (Luganda; Givón 1970:254)

Importantly, the comitative construction is only allowed in such instances if the comitative meaning is intended: If the man and the dog were not seen *together* in (14b), then the sentence would not felicitous – this would be the case even if the possessive marker on *embwa* were absent.

2.2.3 Default agreement

Another option available, if one truly wishes to use coordination, is to resolve the conflict by using some sort of default gender agreement marker, which may, in fact agree with neither

[7] See §2.2.5 below.

[8] §2.2.3 further discusses the role of humanness in agreement in Bantu languages.

[9] Glosses added by me, with the aid of Crabtree (1923). FV stands for 'final vowel': In Bantu, all verbs end in a vowel; this vowel may change according to the tense and mood of the verb, but it is very rarely, if ever, the sole indication of a different tense or mood. Furthermore, the same final vowel may appear in multiple tenses or moods; therefore, rather than attempting to assign a meaning to this segment, it is traditionally just glossed as FV.

of the conjuncts. This is the case in Serbo-Croat, as the following example shows.

(15) *Znanje* *i* *intuicija* *su* *kod* *njega*
knowledge.**NEUT.SG** and intuition.**FEM.SG** have in him
sarađivali *i* *dopunjavali* *se*
worked-together.**MASC.PL** and supplemented.**MASC.PL** each-other
'Knowledge and intuition have worked together and supplemented each other in him.' (Andrić, *Travnička Hronika*, cited in Corbett 1983a:187)

In this example the conjuncts are neuter and feminine, respectively, and yet neither neuter nor feminine agreement is reflected on the verb. This is because in Serbo-Croat the default gender is masculine: When nouns of different genders are coordinated, masculine agreement will be used, even if neither conjunct is masculine.

In some instances the choice of which gender to use as the default is merely an arbitrary convention, as is illustrated by the fact that the default is masculine in Italian and Serbo-Croat (Corbett 1983a), neuter in Icelandic and Faroese (Steinmetz 2006), and feminine in Maasai and Mohawk (Newell 2005). In other instances, however, the choice of default may, in fact, be based on notional/semantic criteria.

It has been shown that in many Bantu languages the animacy status of the noun may often affect the agreement on the verb, especially when there is a conflict in the noun class features on the conjuncts (Corbett and Mtenje 1987, Chichewa; Marten 2000, Kiswahili; Moloto 1992, Northern Sotho; Simango 2012, ciNsenga). As shown in (5) on page 19, in Bantu languages noun class 1/2 is typically reserved purely for human-denoting nouns, with very few exceptions.[10] Perhaps for this reason, the agreement markers for NC1/2 have come to be seen as default agreement markers for nouns denoting humans, even if they come from a different noun class. This is exemplified in (16), where the nouns are from classes 9 and 7, but the subject marker is from class 2, because the nouns denote humans.

(16) ***M**bala* *imodzi* *ndi* ***chi**tsilu* *chimodzi* ***a**-ku-meny-ana*
9.thief 9.one and **7**.fool 7.one **SM2**-PRES-hit-RECIP
'One thief and one fool are fighting.' (Chichewa; Corbett and Mtenje 1987:32)

For non-human conjuncts there is a different default that is used. Cross-linguistically within the Bantu languages, this default for non-humans usually seems to originate from

[10]Note, this does not mean that *all* human-denoting nouns are found in NC1/2, it merely means that most of the nouns in NC1/2 denote humans. Human-denoting nouns can be found in most non-locative/non-infinitive noun classes.

noun classes 8 or 10. This is most likely because historically gender 7/8 was reserved for inanimate entities, and 9/10 was reserved for animals (see (5) on page 19). Thus in (17), SM8 is used even though neither of the conjuncts belongs to NC8.

(17) *Mbale na ka-temo* ***v**-a-sowˆ-a*
9.plate and **12**-axe **SM8**-PAST-miss-FV
'The plate and the axe are missing.' (ciNsenga; Simango 2012:178)

For *some* Bantu languages, such as Sesotho (de Vos and Mitchley 2012), isiZulu (Bosch 1988), and Chichewa (Corbett and Mtenje 1987) the choice of subject markers from classes 8/10 may be driven by homophony: In these languages, the SM for class 8 is phonologically identical to the SM for class 10. It may thus simply be the case that the relevant SMs have been heard agreeing with non-human DPs more often than any other SM, and have accordingly come to be perceived as default subject markers for non-human DPs. This reasoning, however, does not apply in languages such as ciNsenga, Kiswahili and Luganda, where the SMs are not phonologically identical, and the default SM is specifically that of class 8 (Simango 2012; Marten 2000; Givón 1970), although class 10 has also been reported as the inanimate default in some dialects of Kiswahili (Schadeburg 1992).

Regardless of which subject marker is used as the default, the facts remain that 1) the use of a default agreement marker is a very productive strategy used to resolve agreement conflicts in Bantu languages, and 2) the choice of default changes according to the [±HUMAN] feature on the conjuncts. Bearing the second point in mind, it becomes apparent as to why some Bantu languages are resistant to the coordination of human and nonhuman conjuncts: In addition to a clash between the conjuncts' gender features, there is now also a clash between their [±HUMAN] features. Not all Bantu languages insist on the comitative construction in such instances, however; some allow for the [+HUMAN] default SM to be used, whilst others prefer the [-HUMAN] default.

It should be noted that while many studies seem to indicate a [±HUMAN] division in Bantu (de Vos and Mitchley 2012; Simango 2012; Moloto 1992; Corbett and Mtenje 1987; Givón 1970, *inter alia*), there are some languages which seem to draw the line elsewhere: In Sambaa, for example, when two animal-denoting nouns are coordinated, one uses the same subject marker as the one used with [+HUMAN] conjuncts (Riedel 2009), although the default used for inanimate nouns is also allowed. Additionally, Bosch (1988) reports that

isiZulu appears to make a tripartite distinction between humans, animals, and inanimates, and that this distinction is reflected through the use of a different default for each category: SM2 for humans, SM10 for animals,[11] and the 'indefinite concord *ku-*' for inanimates (as in example 18). Buell (2012) identifies *ku-* as SM17.[12]

(18) *Izin-kuni* *nama-lahle* ***ku****-phel-ile.*
10-firewood and.**6**-coal **SM17**-finish-PERF
'The firewood and coal are finished.' (isiZulu; Nyembezi 1970:4, quoted in Buell 2012)

One final note on the use of default agreement in Bantu: It is often common to use default agreement even in contexts where one would expect to find morphological agreement, as (19) shows.

(19) a. **I-puleti* *ne-bhodwe* ***a****-ngcolile*
5-plate and.**5**-pot **SM6**-dirty
Int.: 'The plate and the pot are dirty.'

b. *I-puleti* *ne-bhodwe* ***ku****-ngcolile*
5-plate and.**5**-pot **SM17**-dirty
'The plate and the pot are dirty.' (isiZulu; Bosch 1988:73)

In the example above, two conjuncts from NC5 have been coordinated, but the plural agreement marker from that gender, SM6, is considered ungrammatical, as illustrated in (19a). By contrast, default agreement is preferred, as illustrated in (19b). Likewise, in example (13) on page 23, Chichewa was also shown to not allow morphological agreement with *singular* NC-balanced examples; in such instances, even though there is no clash in noun class features, Chichewa will also use a default subject marker, as can be seen in (20), where SM2 is used to agree with the two [+HUMAN] nouns, rather than using the plural subject marker from gender 12/13.

(20) *Ka-m-kazi* *ndi* *ka-mwa-na* *a-ku-gona*
12-1-woman and 12-1-child **SM2**-PRES-lie
'The little woman and the infant are lying down.'
(Chichewa; Corbett and Mtenje 1987:31)

[11] Bosch lists the default for animal conjuncts as SM10, but SM8 and SM10 are homophonous in isiZulu, which means that, in fact, the default could be either.

[12] Please see Buell (2012) for his motivation for this subject marker as a default, rather than it being used as an expletive in such instances.

Corbett and Mtenje (1987) argue that in such cases, the gender resolution is triggered by number resolution: As both conjuncts are singular, the agreeing SM must undergo number resolution in order to gain a PLURAL feature, and as number and noun class are intrinsically linked in Bantu, this means that the noun class feature will also change. While this does not account for why the plural SM of the relevant gender is not used, it does go to show that semantic agreement, through the used of a default determined by the [±HUMAN] features of the conjuncts, is a common agreement resolution strategy in Bantu.

2.2.4 Partial agreement

In the previous section it was shown that agreement can be licensed through the use of a default agreement marker, which can either be arbitrarily chosen (such as in Serbo-Croat), or semantically motivated, as is the case in Bantu. In contrast to semantic agreement, a degree of *syntactic* agreement can also be used to license agreement. This can be done through the use of *partial agreement*. Partial agreement occurs when *all* the agreement features (such as person, number, *and* gender) on the probe reflect the iϕ-features of only *one* of the conjuncts, whilst not taking into account any of the iϕ-features of the other conjunct (Badecker 2007).

An example of partial agreement in Czech is given in (21), where the verb *seděla* faithfully reflects all the agreement features of only one of the conjuncts. This results in a [SINGULAR] feature appearing on the verb, even though semantically the action is performed by two entities.

(21) *Na* *rohožce* *seděla* *kočka* *a* *pes*
on mat was.sitting.**FEM.SG** cat.**FEM.SG** and dog.MASC.SG
'The cat and the dog were sitting on the mat.'
(Czech, Short 1993, quoted in Badecker 2007:1542)

The conjunct which is agreed with is not arbitrarily chosen; instead, partial agreement is sensitive to the order of the conjuncts, as the Slovenian examples below show: In (22a), the verb agrees with the neuter plural conjunct, but when the order of the conjuncts is switched as in (22b), the verb agrees with the feminine plural conjunct instead. Thus it is typically the conjunct linearly closest to the verb that controls the agreement – this is called closest conjunct agreement (CCA).

(22) a. [*Krave* *in* *teleta*] *so* *odšla* *na* *pašo.*
[cow.FEM.PL and calf.**NEUT.PL**] aux went.**NEUT.PL** on grazing
'Cows and calves went grazing.'

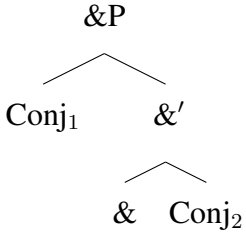
b. [*Teleta* *in* *krave*] *so* *odšle* *na* *pašo.*
[calf.NEUT.PL and cow.**FEM.PL**] aux went.**FEM.PL** on grazing
'Calves and cows went grazing.' (Slovenian; Marušič and Nevins 2009:2)

An important issue to consider with regard to closest conjunct agreement is the order of the probe relative to the goal. When the goal precedes the probe (i.e. the &P appears in a preverbal position, also called 'source target' order), the CCA takes the form of second conjunct agreement, as in example (22) above. By contrast, when the probe precedes the goal (i.e. the &P appears in a post-verbal position, also called 'target source' order), CCA takes the form of first conjunct agreement, as in the Arabic example in (23) below.

(23) *ža* *ʕomar* *w* *karim*
came.3.MASC.SG Omar and Karim
'Omar and Karim came.' (Morrocan Arabic; Aoun et al. 1999:669)

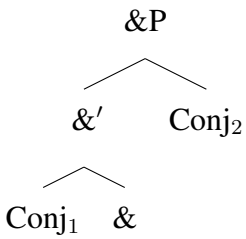
Although both first and second conjunct agreement are attested in various languages, it has been shown that cross-linguistically partial agreement is more likely to be licensed when there is a 'target-source' order, making first conjunct agreement a more common form of CCA than second conjunct agreement (Badecker 2007, Corbett 1991, Johannessen 1996). In fact, it has been argued that partial agreement is not a matter of *linear* proximity, but rather a matter of *hierarchical* proximity (see, for example, van Koppen (2005)), and that this can account for why first conjunct agreement with post-verbal subjects is less marked than second conjunct agreement with preverbal subjects. According to structural analyses of coordinated phrases which argue that conjunctions are the heads of &Ps (e.g. Johannessen 1996, 1998; Zoerner 1995), the second conjunct is more deeply embedded than the first, as (24) shows, thus making it hierarchically further away from the verb.

(24)



However, Johannessen (1996) proposes that head-final languages merge Spec-&P to the right (as in (25) below), rather than to the left, as above. This means that in these languages second conjunct agreement is possible, as the second conjunct is now hierarchically closer to the verb.

(25)



This proposal, however, does not account for the fact that some languages, such as Kiswahili, allow for both first conjunct agreement in post-verbal coordination, as well as second conjunct agreement in preverbal coordination (Marten 2000). In (26), for example, we see a case of the coordinated subject preceding the verb, and the verb carrying a subject marker which agrees with the second conjunct.

(26) *M-guu wa meza na ki-ti* ***ki****-mevunjika*
3-leg of table and **7**-chair SM**7**-be broken
'The leg of the table and the chair are broken.'

(Kiswahili; Bokamba 1985:45, quoted in Marten 2000:pg8 of MS)

By contrast, in (27), the coordinated subject appears post-verbally, and the verb now agrees with the first conjunct.[13]

(27) ***A****-li-kuja* *Haroub na Naila*
SM**1**-past-come **1**.Haroub and 1.Naila
'Haroub and Naila came.' (Kiswahili; Marten 2000:pg11 of MS)

It is worth noting that although Kiswahili licenses CCA both preverbally and post-verbally, it places semantic restrictions on CCA with preverbal conjuncts: The conjuncts may *not* be [+HUMAN]. If the preverbal conjuncts carry a [+HUMAN] feature, then 'full agreement' (i.e. agreement with the entire &P, either through morphological agreement or

[13] Although, theoretically, it could be agreeing with the second conjunct, as both conjuncts belong to NC1. Another explanation is that this is actually a case of clausal coordination, with ATB extraction of the verb, as per the analysis for first conjunct agreement in Arabic presented by Aoun et al. (1999).

the use of a default, see van Koppen (2005)) is obligatory, as illustrated in (28a). Partial agreement in such instances is considered ungrammatical, as (28b) shows.[14]

(28) a. *Haroub na Naila* ***wa****-li-kuja*
1.Haroub and **1**.Naila **SM2**-PAST-come
'Haroub and Naila came.'

b. **Haroub na Naila* ***a****-li-kuja*
1.Haroub and **1**.Naila **SM1**-PAST-come
Int.: 'Haroub and Naila came.' (Kiswahili; Marten 2000:pg9 of MS)

Thus both the position of the &P relative to the probe as well as semantic factors such as animacy may affect the licensing of closest conjunct agreement.

2.2.5 Phonological resolution

The last strategy which will be discussed is phonological resolution. This is more restricted than the other three strategies discussed above, as it is not available as a strategy in all languages. Even in the languages in which it is available, it is only applicable in very restricted circumstances, as will be seen below.

Phonological resolution refers to agreement being made possible by the fact that the agreement markers which agree with each of the conjuncts happen to share a phonological form. Pullum and Zwicky (1986:753-754) demonstrate this with the English examples in (29). In (29a), coordination is disallowed, as none of the three forms of the copula agrees with both conjuncts. By contrast, in (29b), *are* agrees with both the 3rd person plural pronoun, as well as the 2nd person pronoun.

(29) a. Either they or I *are/*am/*is going to have to go.

b. Either they or you are going to have to go.

Thus phonological resolution is only available as a strategy if a language has two or more agreement markers which are homophonous. For example, although the coordination

[14]Interestingly, animacy restrictions on partial agreement have also been reported in Bulgarian, which has determiner-noun agreement: When two singular inanimates are coordinated, the determiner can agree with the closest conjunct, even though this results in singular agreement; by contrast, when two singular animates are coordinated, CCA is not allowed, and plural agreement is required (Hristov 2011). Gender is not marked on plural agreement markers in Bulgarian, so it is not possible to determine whether the same restrictions apply with plural conjuncts.

of human and non-human conjuncts is typically prohibited in Chichewa, as (30) shows, if the subject markers which agree with the conjuncts in isolation are homophonous, agreement is licensed through phonological resolution, as illustrated in (31).

(30) **Mu-ntu ndi ng'ombe **a/zi**-ku-yenda*
1-person and **10**.cow **SM1/2/SM10**-PRES-walk
Int.: 'The person and the cows are walking.' (Chichewa; Corbett and Mtenje 1987:34)

(31) *A-na ndi ma-lanje **a**-ku-sowa*
2-child and **6**-orange **SM2/6**-PRES-missing
'The children and the oranges are missing.' (Chichewa; Corbett and Mtenje 1987:33)

In Chichewa, SM2 and SM6 are both realised as *a-*, thus in (31) the predicate *akusowa* agrees with both conjuncts on a phonological level.

Voeltz (1971) reports that phonological resolution is also used in isiXhosa. Moreover, not only does he say that it is available as an agreement resolution strategy, he also claims that it is the *only* agreement resolution strategy available, outside of morphological agreement with plural conjuncts from the same noun class. This effectively limits NC-unbalanced coordination in isiXhosa to the coordination of conjuncts from NC8 with conjuncts from NC10, as these are the only two plural noun classes with homophonous subject markers (both are realised as *zi-*). All other cases of NC-unbalanced coordination are ineffable, according to Voeltz.

When to use which strategy?

This section has shown that in addition to morphological agreement, there are several different strategies that are used to resolve the issue of coordination and agreement, namely 1) avoidance of coordination, through the use of the comitative construction; 2) use of a default, which, in Bantu, is determined by the animacy of the conjuncts; 3) partial agreement, typically through closest conjunct agreement; or 4) phonological resolution, which allows the coordination of conjuncts with homophonous agreement markers. This dissertation aims to bring all these strategies together coherently, by 1) looking at what agreement feature contexts favour which strategies in three Bantu languages, 2) determining the factors at play, and the relative importance of these factors in the different languages, and 3) generating a typology of agreement strategies. This will be done within an Optimality Theoretic framework.

2.3 Optimality Theory

This section serves as an introduction for those with little to no knowledge of Optimality Theory. Optimality Theory (OT) is a linguistic model originally proposed by Prince and Smolensky (1993/2004) as an alternative to rule-governed generative phonological frameworks, such as Chomsky and Halle's (1968) rewrite rules. Whilst rewrite rules identified particular environments in which particular phonological processes applied, they generally failed to identify or explain the underlying phonological "conspiracy" (Kisseberth 1970); that is to say, a specific constraint on surface forms may manifest itself differently in different contexts, but rewrite rules only identify the manifestation and the context, without identifying the constraint itself. By contrast, OT provides a systematic means of mapping the underlying phonological representations (the input) to surface realizations (the output), by taking into account the phonological system as a whole, not just specific phonological processes (§2.3.1 expounds on this). Although originally designed as a phonological framework, OT has since been extended into other fields of linguistics, such as syntax and morphosyntax (Grimshaw 1997a,b; Bresnan 2000, 2001; Trommer 2001; Xu 2007; Aronoff and Xu 2010, *inter alia*). However, as Optimality Theory started out as a purely phonological model, this section will use phonological examples to illustrate how OT works. §2.4 goes on to discuss applications of OT to the domain of morphosyntax.

2.3.1 An overview of Optimality Theory

The principal premise of Optimality Theory is that there is a set of universal violable constraints which languages apply in order to resolve internal conflicts in their grammars, and that different languages will order these constraints differently, according to their relative importance in each language.

In addition to the set of ranked constraints, there are two additional key components to OT: the generator, and the evaluator. The generator (henceforth GEN) is responsible for generating a list of possible surface realizations of the underlying representation (the 'input'). This list is called the *candidate* set, and it is potentially infinite (McCarthy 2008). The evaluator (henceforth EVAL) then applies the ranked constraints to the set of candidates provided, and evaluates which is the optimal candidate, based on which one incurs fewer violations, especially amongst the higher ranked constraints. The optimal candidate is then chosen

as the final output – that is, the final surface realisation of the underlying representation. McCarthy (2008:19) uses the flowchart shown in (32) to represent OT diagrammatically.

(32) /input/ $\rightarrow$ GEN $\rightarrow$ $\{cand_1, cand_2...\}$ $\rightarrow$ EVAL $\rightarrow$ [output]

This process is typically represented in a tableau of the sort in (33): The input is written in the top left corner, and GEN then produces various candidates, which are represented in the first column. The constraints are put in the top row, with the leftmost constraint being the most highly ranked, and the rightmost one being the lowest ranked.

(33)

Input	Constraint 1	Constraint 2	Constraint 3
a. ☞ Candidate 1		$*$	
b. Candidate 2		$*$	$*!$
c. Candidate 3	$*!$		

The candidates are evaluated against the constraints; if a candidate violates a constraint, this is indicated with an asterisk. This violation is considered to be terminal (indicated by '$*!$') if other available candidates fare better with this constraint; once a candidate has incurred a terminal violation, that candidate is no longer considered to be available as an option for the output, regardless of whether it satisfies all the other constraints – this is indicated by shading the remaining blocks in the row grey (as was done for candidate 3 in (33)). If two available candidates incur equal violations under a particular constraint, and there is no available candidate which fares better (see, for example, candidates 1 and 2 with respect to constraint 2 in the tableau above), then both candidates are still available to be evaluated by the next ranked constraint. When all candidates bar one have incurred terminal violations, the remaining candidate is said to be the optimal candidate; the index ☞ indicates which candidate is optimal in the tableau.

2.3.1.1 A simple example of an OT analysis

Baković (2006) shows how OT can be used to explain the morphophonemic alternation of the past tense morpheme in English. What follows is a considerably simplified discussion of Baković 's analysis. The first step to doing an OT analysis is to list the various surface

realisations, as well as the contexts in which they appear. These descriptive generalizations for the morphophonemic alternation of the English past tense morpheme are given below.

(34) (1) [t] is used when the verb stem ends in a voiceless phoneme, except for when that phoneme is [t]

(2) [d] is used when the verb stem ends in a voiced phoneme, except for when that phoneme is [d]

(3) [əd] is used when the verb stem ends in a [t] or [d]

For simplicity's sake, let us assume that /d/ is the underlying form of the past tense morpheme, as it is used in more phonetic environments than [t] and [əd] (/d/ is used after resonants and voiced obstruents, as opposed to only after voiceless obstruents in the case of [t], or after [t] and [d] in the case of [əd]).

The next step in the OT analysis is to identify the underlying phonological conspiracies which give rise to (34). For example, the exceptions listed for the first two generalizations – [t] may not occur after [t], and [d] may not occur after [d] – speak to an unwillingness to allow geminate consonants. This then can be the first constraint, one which prohibits geminate consonants; this shall be called the NO-GEMINATE (NO-GEM henceforth) constraint. The fact that sequences such as [tt] and [dd] do not occur at the ends of words in the past tense in English is a clear indication that this constraint is almost always obeyed, and therefore must be ranked quite highly.

The next constraint needs to capture the phonological conspiracy which is responsible for the content in generalizations 1 and 2 above, namely that [t], a voiceless phoneme, is used after other voiceless phonemes, and that [d], a voiced phoneme, is used after other voiced phonemes; this constraint clearly needs to capture the fact that the underlying /d/ must assimilate to match the voicing of the stem-final phoneme; Baković (2006) calls this the SEQ(voi) constraint. As with the previous constraint, the fact the one does not find sequences such as [bt] and [pd] do not occur at the ends of words in the past tense in English indicates that SEQ(voi) must be ranked quite highly.

An additional constraint is needed to capture the fact that the epenthesis used in generalization 3 is something of a last resort, and that, in general, unmotivated epenthesis is prohibited – without such a constraint, one might question why English-speakers don't use [əd] in

all circumstances, and avoid the trouble of phonologically conditioned allomorphs. Such a constraint should specify that the existence of each segment in the output should should be dependent in the existence of a corresponding segment in the input, and that epenthesis is typically avoided. This constraint is referred to as the DEPENDENCY constraint (hereinafter DEP). However, it is important to note that epenthesis clearly *does* occur in certain contexts, and thus DEP must be ranked lower than the previous two constraints, as the violation of this constraint is significantly less marked than the violation of the previous two constraints.

Before demonstrating how these constraints interact in such a way as to yield the generalizations in (34), the constraints are listed and summarised in (35) for ease of reference. Note that these constraints aren't context specific, or even language specific; this makes then preferable to the rewrite rules on which the generalizations in (34) are modeled.

(35) (1) NO-GEM: Tautosyllabic geminates are prohibited
(2) SEQ(voi): Tautosyllabic obstruent clusters have the same [± voice] specification
(3) DEP: Output segments must have input correspondents

Tableau (36) shows how the constraints in (35) interact to yield [bækt] as the optimal candidate for the past tense of the verb *to back*.

(36) Past tense of *to back*

/bæk + d/	NO-GEM	SEQ(voi)	DEP
a. ☞ [bækt]			
b. [bækd]		∗!	
c. [bækəd]			∗!

In (36), candidate (b) violates the SEQ(voi) constraint, because /k/ and /d/ differ in their [±voice] specifications. This violation leads to (b) being disregarded as a potential output. Only candidates (a) and (c) are thus left as potential surface realisations at this stage. But candidate (c) violates the next constraint, DEP, as it has a segment (schwa) in the output which has no corresponding segment in the input. This violation effectively removes (c) from the viable candidate set, and leaves (a) as the optimal candidate.

If one replaces the the final voiceless segment of the verb stem with a voiced segment, the pattern of violations changes, as can be seen in tableau (37) where the output needed is the past tense of *to bag*.

(37) Past tense of *to bag*

/bæg + d/		NO-GEM	SEQ(voi)	DEP
a.	[bægt]		∗!	
b. ☞	[bægd]			
c.	[bægəd]			∗!

In this instance, it is *candidate (a)* which violates the SEQ(voi) constraint, leaving candidates (b) and (c) as the only remaining potential outputs. But as before, (c) is eliminated by the DEP constraint, and this leaves (b) as the optimal candidate.

In (36) and (37), the NO-GEM constraint seems to play no role, and the epenthetic candidate (c) always seems doomed to be eliminated, because it will *always* violate DEP. In tableau (38), however, it can be seen that even though a candidate violates one of the constraints, it may still be the optimal candidate, *provided* all the other candidates have been eliminated by more highly ranked constraints.

(38) Past tense of *to bead*

/biːd + d/		NO-GEM	SEQ(voi)	DEP
a.	[biːdt]		∗!	
b.	[biːdd]	∗!		
c. ☞	[biː dəd]			∗

In this instance, candidate (b), which most closely mirrors the underlying representation, is eliminated because it violates the NO-GEM constraint, as it results in a tautosyllabic [dd] sequence. As in tableau (37), candidate (a) violates the SEQ(voi) constraint, and is thus eliminated. This leaves candidate (c) as the only remaining candidate, and thus (c)'s violation of DEP is not a terminal violation, and (c) is rendered the optimal candidate.

An important part of an Optimality Theoretic analysis is that the order in which the constraints are ranked has been justified. For instance, if DEP were ranked above NO-GEM

and SEQ(voi), then the analysis would incorrectly predict that [biːdəd] would be the least optimal candidate, as tableau (39) shows.

(39) Past tense of *to bead*

/biːd + d/	DEP	NO-GEM	SEQ(voi)
a. [*biːdt]			∗
b. [*biːdd]		∗	
c. [biːdəd]	∗!		

This shows that DEP *must* be ranked lower than the other to constraints. By contrast, the NO-GEM and SEQ(voi) constraints can be ranked in any order in relation to each other – whether NO-GEM is ranked first or second will never make a difference to the analysis, as tableaux (40a,b) exemplify. This because a candidate which violates the No-Geminate constraint will never violate SEQ(voi), as geminates, by their very nature, will always share the same [±voice] specification.

(40) (a) Past tense of *to bag*

/bæg + d/	SEQ(voi)	NO-GEM	DEP
a. [bægt]	∗!		
b. ☞ [bægd]			
c. [bægəd]			∗!

(b) Past tense of *to bead*

/biːd + d/	SEQ(voi)	NO-GEM	DEP
a. [biːdt]	∗!		
b. [biːdd]		∗!	
c. ☞ [biː dəd]			∗

Thus the two constraints work independently of each other, and, as their order is reversible, they are equally ranked. The broken line between the two constraints is used to signify this.

2.3.2 The nature of OT constraints

It would be appropriate at this point to consider the nature of the constraints. Prince and Smolensky (1993/2004) classify all constraints as belonging to one of two categories, either a) *markedness constraints*, or b) *faithfulness constraints*. Markedness constraints are those which constrain the output form, based on how structurally well-formed it is; for example, the NO-GEM and SEQ(voi) constraints above are both markedness constraints, as it is considered 'marked' to have word-final geminates, or tautosyllabic adjacent obstruents which differ in their [±voice] specification. Thus this type of constraint ignores the input form, and it is only concerned with whether the output form meets certain structural requirements.

By contrast, faithfulness constraints are aimed at ensuring that the output form is as similar as possible to the input form; thus this type of constraint pays attention to both input *and* output forms. The DEP constraint discussed above is an example of a faithfulness constraint, as it demands that every segment in the output must appear in the input, thereby taking into account both input and output forms. Another example of a common faithfulness constraint is the MAXIMALITY (henceforth MAX) constraint. MAX demands that every segment in the input must have a corresponding segment in the output. Whilst this may sound similar to the DEP constraint, it in fact has a very different effect. When applied on a phonological level, DEP penalizes insertion, whilst MAX penalizes deletion.

Without some form of faithfulness constraints the output form would not be obliged to share any structural or phonological similarities with the input form, and thus faithfulness constraints are necessary to eliminate candidates which in no way resemble the underlying form. Without some form of markedness constraint, on the other hand, there would be not requirement for surface forms to differ at all from the underlying forms, and this would give rise to incredibly marked forms, which may be difficult to articulate (such as */bi:dd/). Thus faithfulness constraints interact with the markedness constraints to balance the tension of not producing marked forms, while keeping the output form as close as possible to the input. An OT analysis will thus typically makes use of both types of constraints.

2.3.3 Using OT to explain different grammars

Although constraints are strictly ranked for specific languages, the ordering of the constraints may differ cross-linguistically. This difference in ordering gives rise to different

typologies. For example, consider the phenomenon of final devoicing in Germanic. Whilst most Germanic languages, including Dutch, German and Afrikaans, devoice word-final obstruents, English does not. Kager (1999) explains this difference by showing that although all these languages have the same constraints in this regard, English ranks the constraints differently to the other Germanic languages. There are two constraints at play here, as listed and explained in (41).

(41) 1. *VOICED-CODA: Obstruents must not be voiced in coda position.
2. IDENT-IO(voice): The specification for the feature [voice] of an input segment must be preserved in its output correspondent. (Kager 1999:14)

In Dutch, the markedness constraint *VOICED-CODA dominates (i.e. is ranked above) the faithfulness constraint IDENT-IO(voice), as tableau (42) demonstrates.

(42) Surface realization of /bɛd/ in Dutch

/bɛd/	*VOICED-CODA	IDENT-IO(voice)
a. ☞ [bɛt]		*
b. [bɛd]	*!	

By contrast, English ranks the faithfulness constraint above the markedness constraint, thus giving rise to voiced codas, as can be seen in tableau (43).

(43) Surface realization of /bɛd/ in English

/bɛd/	IDENT-IO(voice)	*VOICED-CODA
a. [bɛt]	*!	
b. ☞ [bɛd]		*

Although this example is quite simple, and only two typologies are possible, matters become considerable more complex when more than two constraints are used; for example, when three different constraints (A, B, and C) interact, this may yield as many as six different grammars, according to the ranking of the constraints: ABC, ACB, BAC, BCA,

CAB, and CBA. The more constraints involved, the number of possible grammars increases exponentially: 4 constraints give rise to 24 possible grammars, 5 give rise to 120, etc. This is known as a factorial typology, and it is a useful tool for exploring not only cross-linguistic differences, but also similarities, which can then give rise to a better understanding of what is and is not possible in language.

2.4 Optimality Theory and morphosyntax

Although originally conceived of as a phonological framework, Optimality Theory's ability to account for linguistic choices from competing forms, as well as typological differences between languages, has made it applicable to other linguistic domains, such as morphology, syntax and semantics. The application of this framework to these domains, however, is far from standardized, and several different approaches have been proposed. This section briefly covers three approaches which have been used in applying OT to morphosyntactic data: 1) The application of 'classical' OT, which most closely mirrors the phonological framework (§2.4.1); 2) Distributed Optimality, an approach combining aspects of Distributed Morphology with OT (§2.4.2); and 3) Realization OT, which incorporates an inferential-realization approach to inflectional morphology with OT (§2.4.3). The approach used in this thesis, as laid out in § 1.2 on page 6, is most like the 'classical' approach, with some minor differences, as discussed in §2.4.4.

2.4.1 A 'classical' OT approach to morphosyntax

The assumptions implicit to the approach discussed in this section most closely emulate those of traditional OT: The input typically consists of grammatical information, which may/may not be attached to a lexeme; the candidates are potential realizations of this grammatical information; and the constraints are universal (i.e. applicable to other languages), and they evaluate how faithful the candidates are in expressing the grammatical information in the input, and how marked they are in terms of form or placement on a markedness hierarchy.

A significant departure from traditional OT, however, lies in the assumptions about GEN. Traditionally GEN is considered to be unrestricted, and the list of possible candidates which

it produces is, theoretically, infinite. By contrast, in the approach described below GEN is constrained: GEN does not generate a list of every conceivable combination of the features in the input, but instead the candidates are morphemes drawn from the lexicon.

An example of such an approach is Grimshaw's (1997a) analysis of object pronoun clitics in Romance languages. The paper examines instances in which clitics have combined with each other to produce an unexpected form which does not faithfully reflect all of the grammatical features in the input. The core of the analysis rests on the two families of faithfulness constraints (PARSE FEATURE and FILL FEATURE), and the one markedness constraint (*XX), defined in (44).

(44) **PARSE FEATURE**: All features in the input must appear in the output

FILL FEATURE: Only features in the input can appear in the output

***XX**: Sequences of identical functional heads are ill-formed

(Grimshaw 1997a:170)

The PARSE constraints mirror the function of the MAX constraints in phonology, as they prohibit the 'deletion' or non-expression of a feature which appears in the input. The FILL constraints, on the other hand, mirror the function of DEP, as they penalize output candidates which carry features that are not found in the input. On the other hand, the *XX constraint, which here penalizes combinations of clitics which bear identical features, clearly parallels a number of constraints found in OT phonology, particularly the OCP type constraints,[15] including the NO GEMINATES constraint mentioned in (35) on page 36.

Grimshaw stipulates that the constraints in (44) work in tandem with two universal markedness hierarchies, listed in (45). These two hierarchies are typically left out of the tableaux, and only become relevant if the ranking of the constraints produces two optimal candidates – at this point the hierarchies will apply, and favour the candidate whose person and/or case features are less marked.

(45) Universal Markedness Hierarchies

Person: *2 > *1 > *3 (Grimshaw 1997a:170)

[15]The OCP family of constraints developed from the Leben's (1973) Obligatory Contour Principle, which was originally restricted to a prohibition against sequences of identical tones within a morpheme, but was later extended to precluding sequences of identical features in other domains of phonology – e.g. McCarthy's (1986) use of OCP to explain antigemination effects.

Case: *DAT > *ACC

The features relevant to the analysis are those found on the object clitics, namely case 'C', person 'P', reflexivity 'R', number 'N', and gender 'G'. Grimshaw (1997a:171) assumes the following features for Italian clitics (brackets around a feature indicate its absence on the clitic in question):

(46) Features on Italian Clitics

si	(R)(P)(N)(G)(C)	self
mi	(R) 1 sg (G)(C)	me, to me, (self)
ti	(R) 2 sg (G)(C)	you, to you, (self)
ci	(R) 1 pl (G)(C)	us, to us, (self)
vi	(R) 2 pl (G)(C)	you, to you, (self)
lo/la	-R (P) sg m/f acc	him, her, it
li/le	-R (P) pl m/f acc	them
gli/le	-R (P) sg m/f dat	to him, her, it

For each feature there exists a PARSE constraint and a FILL constraint, e.g. PARSE R and FILL R. For the purpose of this brief summary, however, the constraints concerning gender and case will be omitted, as they play no major role in the following tableaux.

Grimshaw shows that the ranking below best accounts for the patterns of clitic usage in Italian:

*XX » PARSE R » FILL R » PARSE P » FILL P » PARSE N » FILL N

Not only does this ranking account for the use of clitics which only partially express the features in the input, it also accounts for the use of the impersonal *si*, which carries no features whatsoever. Tableau (47) shows how *si* is rendered the optimal candidate for an input such as [3 pl] (Grimshaw assumes a plural feature on the input for the impersonal, as Chinque (1988) points out that participles carry plural agreement when used with impersonal subjects).

(47)

Input: [3 pl]			*XX	Parse R	Fill R	Parse P	Fill P	Parse N	Fill N
a. ☞	si	(R)(P)(N)(G)(C)				*		*	
b.	mi	(R) 1 sg (G)(C)				*	*!	*	*
c.	ti	(R) 2 sg (G)(C)				*	*!	*	*
d.	ci	(R) 1 pl (G)(C)				*	*!		
e.	vi	(R) 2 pl (G)(C)				*	*!		
f.	lo/la	-R (P) sg m/f acc			*!	*		*	*
g.	li/le	-R (P) pl m/f acc			*!	*			
h.	gli/le	-R (P) sg m/f dat			*!	*		*	*

(Tableau adapted from Grimshaw (1997a:176))

As there is no reflexivity feature in the input, none of the candidates are required to explicitly parse R; thus none of the candidates violate the PARSE R constraint, regardless of whether they can an R feature or not. By contrast, FILL R penalizes candidates (f)-(h), as they carry an R feature when there is none in the input; these are terminal violations for these candidates. The PARSE P constraint then penalizes any candidate which does not parse the third person feature in the input – which happens to be every candidate, as the clitic set only contains clitics which carry either a first person feature, or a second person feature, or no person feature at all. Only those candidates which carry no person feature (i.e. candidates (a) and (f)-(h)) satisfy the next constraint, FILL P; as candidates(f)-(h) have already been eliminated by FILL R, this renders candidate (a) – the impersonal clitic – as the most optimal candidate.

However, in instances where one would expect a sequence of two impersonal clitics before the verb root, '*si si*' is not used, as can be seen in (48):

(48) ***Ci si*** *lava.* Impersonal plus reflexive (**si si*)

'One washes oneself.' (Grimshaw 1997a:181)

While one can account for the fact that *si si* is not used by pointing out that it violates the restriction against sequences of identical functional heads, the use of the the first person plural clitic (see (46)) is counterintuitive. Nonetheless, Grimshaw's analysis can account for this choice.[16]

[16]For the sake of space, the gender and case features have been omitted from tableau (49); as mentioned earlier, they are not relevant to the data discussed here.

(49)

	Input: [3 pl] [R 3 pl]	*XX	Parse R	Fill R	Parse P	Fill P	Parse N	Fill N
a.	si+si (R)(P)(N) +(R)(P)(N)	*!	*		**		**	
b.	mi+si (R) 1 sg + (R)(P)(N)		*		**	*	**!	*
c.	ti+si (R) 2 sg +(R)(P)(N)		*		**	*	**!	*
d. ☞	ci+si (R) 1 pl + (R)(P)(N)		*		**	*	*	
e. ☛	vi+si (R) 2 pl + (R)(P)(N)		*		**	*	*	
f.	lo/la+si -R (P) sg + (R)(P)(N)		*	*!	**		**	*
g.	li/le+si -R (P) pl + (R)(P)(N)		*	*!	**		*	
h.	gli/le+si -R (P) sg + (R)(P)(N)		*	*!	**		**	*

(Tableau adapted from Grimshaw (1997a:181))

The tableau above shows that *si si* is eliminated as a possible candidate right from the start, as it violates the prohibition against two adjacent realizations of identical feature bundles. As before, candidates (f)-(g) are eliminated by FILL R, as they carry a [-R] feature which is not in the input. This leaves only four possible candidates: The first person singular, second person singular, first person plural, and second person plural clitics, each combined with the impersonal clitic. Each of these candidates receive at least one violation mark under PARSE N, as the impersonal clitic fails to parse the plural feature in the second half of the input. Candidates (b) and (c), however, each receive an additional violation mark, as they also fail to parse the plural feature in the first half of the input; this is a terminal violation for these two candidates.

We are now left with two candidates: *ci si* and *vi si*. As FILL N does not distinguish between the two,[17] now is the time to apply the universal hierarchy of markedness constraints, given in (45) on page 42. The feature which these two candidates differ on is Person, and thus it is the *2 » *1 » *3 hierarchy which is relevant here. The constraint penalizing second person features dominates the one penalizing first person features. Thus *vi si* is more marked than *ci si*, rendering the latter the optimal candidate.

This application of OT to morphosyntax closely mirrors classical OT, as it relies, for the most part, on the same sorts of constraints, and does not introduce significantly new types of mechanisms compared to those used in phonological analyses.

[17] Nor would PARSE C/G and FILL C/G, as both candidates are unspecified for Case and Gender.

2.4.2 Distributed Optimality

A different application of OT to the domain of morphology is Distributed Optimality (DO), as proposed by Trommer (2001). This model incorporates some of the explicit theoretical assumptions of Distributed Morphology (Halle and Marantz 1993), including the idea that certain morphological operations happen post-syntactically. Trommer argues that certain key morphological processes happen only once the syntactic tree has been derived, which means that morphological constraints cannot interact with syntactic constraints – this is a marked difference to other classical OT-morphosyntax analyses, such as Grimshaw (2001) and Bresnan (2001).

Trommer (2002:84) describes Distributed Optimality as a "constraint based, modular adaption" of Distributed Morphology. One of the core tenets of Trommer's model is that morphosyntax comprises three modules: 1) Syntax, 2) Chain Interpretation, and 3) Head Interpretation. The first of these modules creates syntactic chains from lexical items; the second module then maps these chains to heads, which are in turn mapped to vocabulary items. This last stage, mapping heads to vocabulary items, is achieved through the interaction of ranked constraints, rather than through the application of ordered rules, as per traditional Distributed Morphology.

To illustrate the difference between Distributed Optimality and Distributed Morphology, Trommer (2002) shows how the two approaches deal with the blocking of multiple person agreement markers on the verb in Turkana. In this language, verbs can carry person markers agreeing with either the subject or the object – but only one person marker can appear on the verb at any one time. If one of the arguments carries a third person feature, and the other argument does not, the verb will agree with the non-third person argument, regardless of whether it is the subject or the object. Thus in (50) below, the verb shows first person agreement in both instances, with the only difference being the presence of the inverse marker *k-* in (50b), which shows that "the object is higher in animacy than the subject" (Trommer 2002).

(50) a. *à-mɪn-à*
D-1-love-ASP
'I love her.' (Dimmendaal 1983:69, cited in Trommer 2002:85)

b. *k-à-mɪn-à*
D-1-love-ASP

'She loves me.' (Dimmendaal 1983:123, cited in Trommer 2002:85)

Halle and Marantz (1993) account for this phenomenon by arguing that this is the result of AgrS and AgrO fusing into a single head position; thus while the feature structures occupying these two nodes may remain intact, the head can be mapped to only one vocabulary item. For a sentence which consists of a third person subject and a first person object (as in (50b)), there will then be a conflict in the grammar regarding which person feature to realise: [+3] or [+1]. Distributed Morphology deals with this conflict by arranging the morphs in a list, such as (51), where morphs which have first priority in realization are top of the list, whilst morphs with lower realization priority are ordered lower in the list.

(51) 1. \à-\ ↔ [+1]

2. \e-\ ↔ [+3] (Trommer 2002:86)

For Turkana, the rule requiring the realization of the first person feature precedes the rule requiring the realization of the third person feature, which means that when the fused AgrS/AgrO head has both [+1] and [+3] features, the first person agreement marker will be realised on the verb, rather than the third person agreement marker.

Distributed Optimality, on the other hand, replaces the morphological rules and lists of prioritised morphs, such as in (51), with constraints, and treats the possible feature structures as candidates which are evaluated against the feature structure in the input. To account for the Turkana data, Trommer (2002:86-87) draws from two families of constraints which are ubiquitous in Distributed Optimality: BLOCK and PARSE.

The BLOCK class of constraints penalize candidates which carry multiple affixes which realise the same types of features (e.g. person or number); note that these constraints do not block the feature in question from being realised in any form, they merely block multiple realisations of the feature. Thus BLOCK [P] penalises candidates which carry more than one person feature.

The PARSE constraints are similar to those discussed in §2.4.1 above, except that contextual conditions are added to the parsing requirements. For example, PARSE $[P]^{[-3]/[+3]}$ stipulates that a non-third person [P] feature should be realised if it is adjacent to a third person [P] feature. This constraint interacts with BLOCK [P] to account for the fact that there is only ever one [P] feature on the verb in Turkana, and that first or second person

arguments will always control the agreement when the other argument carries a third person feature (52) (Trommer 2002:87).

(52)

Input: [+Nom +3][+Acc +1]	BLOCK [P]	PARSE [P]$^{[-3]/[+3]}$
a. ☞ [+1]		
b. [+3]		∗!
c. [+1][+3]	∗!	

Trommer (2001) argues that Distributed Optimality is superior to Distributed Morphology in that it replaces language-specific rules with universal constraints, which could, in theory, be applied to any language – just like the constraints in the 'classical' approach, described in §2.4.1. It differs from the 'classical' approach, however, with regards to the notion of candidates: In Distributed Optimality the candidates are features, rather than morphs. This means that, theoretically, any features and/or any feature combinations are feasible candidates, regardless of whether or not there are morphs in the lexicon which can realize them. The matching up of features with morphs only happens *after* the optimal candidate has been chosen.

2.4.3 Realization OT

A markedly different approach is that of Realization OT (Xu 2007, 2011; Aronoff and Xu 2010), which subscribes to an inferential-realizational model of morphology. One of the key tenets of inferential morphology is that inflection markers do not have morpheme status, and therefore are not considered to be part of the lexicon; instead, the relationship between a root and its inflected forms is expressed through rules (Stump 2001).[18] This is combined with a realizational model of morphology, which assumes that morphosyntactic information exists on a root before inflection markers are added – the inflection markers merely realize the morphosyntactic information which is present.

Following Russell (1995), Yip (1998), *inter alia*, Realization OT assumes that morphosyntactic information in the input is phonological unrealized, and only becomes realised

[18]This is in contrast to lexical models of morphology, which recognize inflectional markers as morphemes which are stored in the lexicon. The approaches described in §2.4.1 and §2.4.2 both assume a lexical realizational model of morphology.

through language specific constraints, which stipulate what phonological material will realize a given feature. These constraints, called realization constraints, take the form of *{morphosyntactic feature}: morphophonological form*, where the ':' means 'realized by'. For example, the realization constraint for the English past marker would be *{past}: -d*. If several features are realized by the same phonological form, then those features are listed together in the constraint, e.g. *{3, sg, pres.}: -s*. These realizational constraints interact with markedness constraints, such as *FEATURE SPLIT, which prohibits multiple phonological forms realizing the the same morphosyntactic or semantic feature (Xu 2011:468).

Xu (2011) shows how Realization OT can account for a set of Arabic data which shows instances of both blocking (where the realization of a morph is blocked by the presence of another morph with the same feature) as well as extended morphological exponence (which is precisely the opposite of blocking: Multiple affixes on the root realize the same morphosyntactic feature). The data in question comes from the jussive conjugation paradigm in Classical Arabic, which is shown below.

(53) Classical Arabic jussive conjugation (Xu 2011:470, adapted from Noyer 1997:4-5)

Person/gender	Singular	Dual	Plural
1	*ʔ-aktub*	*n-aktub*	*n-aktub*
2, masc	*t-aktub*	*t-aktub-aa*	*t-aktub-uu*
2, fem	***t-aktub-ii***	*t-aktub-aa*	*t-aktub-na*
3, masc	*y-aktub*	*y-aktub-aa*	*y-aktub-uu*
3, fem	*t-aktub*	*t-aktub-aa*	***y-aktub-na***

The first form of interest is the form with [2, fem, sg] features, *taktubii*, as it exhibits extended morphological exponence: *t-* realizes [2], while *-ii* realizes [2, fem, sg]; thus [2] is realized twice. On the other hand, blocking occurs in the form with the features [3, fem, plural], which is *yaktubna*, instead of *taktubna*, as one might expect. The [fem] feature on *-na* (which realizes [fem, pl]) blocks the realization of the [fem] feature on *t-* (which realizes [3, fem]); thus *y-* is used instead, as it realizes only *y-*.

Xu argues that the Realization Constraints {2, fem, sg}: *-ii*, and {2}: *t-* must dominate *FEATURE SPLIT (*FS), therefore allowing for extended morphological exponence (54). On the other hand, *FS dominates {fem, pl}: *-na*, {3, fem}: *t-*, and {3}: *y-*, which accounts for the blocking of the double realization of the [fem] feature (55).

(54)

Input: aktub 2, fem, sg	{2, fem, sg}: *-ii*	{2}: *t-*	*FS	{fem, pl}: *-na*	{3, fem}: *t-*	{3}: *y-*
a. ☞ t-aktub-ii			∗!			
b. aktub-ii		∗!				
c. t-aktub	∗!					

(55)

Input: aktub 3, fem, pl	{2, fem, sg}: *-ii*	{2}: *t-*	*FS	{fem, pl}: *-na*	{3, fem}: *t-*	{3}: *y-*
a. ☞ y-aktub-na					∗!	
b. t-aktub-na			∗!			∗

In (54), (a) is the optimal candidate, even though it expresses [2] twice, thus violating *FS, as candidates which are inflected with only *t-* or *-ii* violate the Realization Constraints which dominate *FS. By contrast, in (55) *FS dominates all the constraints which concern the features in the input, which results in each feature only being realized once. Additionally, {fem, pl}: *na* dominates {3, fem}: *t-*, which results in *-na* being chosen over *t-* to realize [fem]. However, the feature [3] still needs to be realized, and thus *y-*, which realizes *only* [3], does not incur a violation.

Thus Xu (2011) argues that Realization OT is best able to account for the phenomena of both blocking and extended morphological exponence, without having to resort to stipulations about primary and secondary exponents, and separate morphological rules for each (cf. Noyer (1997)).

2.4.4 The approach used in this thesis

One of the disadvantages of Realization OT is the use of language-specific constraints. The beauty of Optimality Theory, in my opinion, lies in the fact that the constraints are universal – or, if not *entirely* universal, at least relevant to other languages – and thus different rankings yield different grammars. Thus Realization OT is not conducive to cross-linguistic studies, such as the one presented in this dissertation.

The 'classical' OT approach and Distributed Optimality are similar in several regards, although there are significant differences. One of these differences is that Trommer (2002) stipulates that morphological constraints and syntactic constraints do not interact – and while

I agree that morphology does not influence syntax, I *do* believe that syntax affects morphology, e.g. in the case of closest conjunct agreement, where linear, and arguably hierarchical, structure affects the choice of agreement morpheme. Additionally, the assumptions about GEN appear to be significantly different: In Distributed Optimality GEN is not constrained, and generates a list of candidates consisting of features which may or may not match up with morphs in the lexicon. By contrast, in the 'classical' approach, GEN appears to draw the candidates from the lexicon, and thus the candidates are morphemes.

The approach used in this thesis is thus closer to the 'classical' OT approach described in §2.4.1 than it is to either Distributed Optimality or Realization OT, as the candidates are morphemes from the lexicon (specifically, the candidates are subject markers). There are, however, a few differences, as outlined in §1.2. The main difference is that OT is traditionally used for mapping inputs to outputs on a one-to-one basis, but I am using it here to map on a two-to-one basis, as there are two feature bundles in the input (one on each conjunct), but there can only be one feature bundle in the output (on the subject marker). The details of this two-to-one mapping, and the implications thereof, were discussed in § 1.2.1 on page 6.

2.5 Research questions

While various studies on agreement with &Ps in Bantu languages have been carried out before, and have highlighted common agreement resolution strategies, there are still significant gaps in the literature. For the most part, the literature available offers only a perfunctory indication of when some of the various strategies can be used. This thesis aims to explore this in more detail, looking for patterns in which features, or combination of features, on the conjuncts tend to trigger which agreement resolution strategies. Furthermore, and perhaps more interestingly, this study looks at *why* those featural contexts trigger the use of those particular strategies, using Optimality Theory to tease apart the covert processes that interact with each other to give rise to the use of the various strategies. This thesis therefore sets out to answer the following questions:

1. What strategies are used to resolve agreement with coordinated preverbal subjects in Xitsonga, Sesotho and isiXhosa?
2. a) Are these strategies restricted to certain agreement feature contexts?

b) If so, which agreement feature contexts trigger the use of which strategies?

3. a) Are these strategies ontological primitives, or merely the surface manifestation of a conspiracy between underlying constraints?

   b) Can the restriction of certain agreement strategies to specific agreement feature contexts be accounted for by the hierarchical ranking of such constraints within an Optimality Theoretic framework?

4. What insight can be gleaned from generating and examining a factorial typology of agreement strategies, based on the constraints proposed in (3)?

Chapters 4, 5, and 6 focus on answering questions (1)-(3b), while question (4) is explored in chapter 7.

Chapter 3

Methodology

The data used in this thesis was collected through a series of semi-structured interviews with L1 speakers of Sesotho, isiXhosa, and Xitsonga. The primary means of elicitation was a series of questionnaires containing English prompts using coordinated noun phrases in subject position, which the language consults were then asked to translate into their L1.

This chapter outlines and motivates for the methodology used for the data collection, focussing on the language consultants who participated in the study (§3.1) and the data collection process itself (§3.2).

3.1 The participants

As the object of study is I-language (the abstract representation of linguistic knowledge) as opposed to E-language (the linguistic habits of a community) (Chomsky 1986), only three or four speakers of each language were interviewed. The distinction between I-language and E-language is older than the labels. In 1916 Saussure made the distinction between *langue* and *parole*, which was then echoed by Chomsky (1964) when he characterized the difference between Competence and Performance. *Langue*, Competence, and I-language are all associated with the speaker's *abstract* knowledge of a language and its rules, while *Parole*, Performance, and E-language are all associated with the *use* of language, and the manner in which the rules are executed in order to achieve a functional end.

Chomsky (1965:4) argued that the object of linguistic study should be abstract system, saying that linguistic theory in fact tends to be mentalistic, "since it is concerned with dis-

covering a mental reality underlying actual behaviour". As every competent speaker of a language has access to this abstract system, in theory one speaker should be sufficient as a source for one's linguistic data. However, in order to mitigate any idiolectical effects, three to four speakers of each language were consulted, and the questionnaires were tailored to each participant (see §3.2).

The participants in this study were a group of volunteers who responded to a request sent out to the undergraduate students in the linguistics department. This produced, on average, two speakers per language, so the remainder were accessed by using Milroy's (1987) friend-of-a-friend snowballing strategy: The volunteers were asked if any of their friends might be interested in participating in the study. The participants were not paid, but after each interview they were given a small token of appreciation, such as a slab of chocolate. Although many indicated that it was not necessary, and that they were just happy that their language was receiving attention, I felt that this token made them more amenable to returning for further interviews.

The consultants were all students between the ages of 19 and 21. For isiXhosa all were from the Eastern Cape (one from Mthatha, one from King William's Town, and two from Butterworth), whilst the three of the Sesotho consultants were from Johannesburg (Gauteng), and one from Welkom (Free State). The Xitsonga speakers were more spread out in terms of hometowns: One was from Durban (KwaZulu Natal), another from Johannesburg (Gauteng), and a third from Bushbuckridge (Mpumalanga). The geographic spread of the Xitsonga speakers had implications for the data collected, as chapter 4 will show.

Whilst I was not looking for speakers of the standard varieties,[1] not all people who volunteered were used for the study. Speakers who knew less than 50% of the words on the word list (see §3.2.1) were not asked back for additional interviews. Other participants' data was removed from the study on the basis of subtle phonological cues that suggested that they might have been speaking Northern Sotho (S32) rather than Sesotho (S33).

[1] Webb (2008:8) argues that although the nine official Bantu languages of South Africa have been standardized, the standardized varieties "have not been generally accepted in all high-function contexts, in particular in schools, [and] learners (and probably even teachers) do not know these varieties effectively". Thus the standard varieties are not widely used.

3.2 The data collection process

The interviews were administered one-on-one in a variety of neutral and informal contexts, such as coffee shops or tea rooms. These everyday contexts were chosen to counteract any impression that the language consultants might have that they were being tested. The participants were also specifically told that they were not being tested, and that the textbooks for their languages do not have any rules pertaining to the structures being studied.

I conducted a number of interviews with each consultant. In each case the first interview fulfilled three main functions: 1) to familiarise myself with the consultants and get their biographical details (such as age, hometown, other languages spoken at home, etc.), 2) to explain to the consultants what the study would entail, and 3) to check the words lists (explained in §3.2.1) against their lexicons so that the questionnaires (§3.2.2) could be tailored to each consultant. All subsequent interviews were focussed on eliciting sentences with conjuncts from specific noun classes being coordinated in preverbal position.

3.2.1 The word lists

As the issue of noun classes is central to this research, the first step in setting up the questionnaires was to compile a list of nouns that fall into the various noun classes, and then check the word lists with each consultant. These word lists were used to tailor the questionnaires to each consultant, ensuring that 1) the questionnaire used words which they were familiar with, and 2) I was using appropriate words for the noun class combinations which I was trying to elicit.

The word lists consisted of a range of words from the various noun classes; for each noun class I tried to include words spanning the three main degrees of animacy: human, animal, and inanimate (see (1) for an example). These words were found by combing through various dictionaries, grammars and textbooks for each of the languages, namely Mokoena (1998), Kriel (1950), and Reynierse (1991) for Sesotho; Dowling (1998), Nabe et al. (1976), and Reynierse (1991) for isiXhosa; and Baumbach (1987), Chatelain and Junod (1909) and Cuenod (1967) for Xitsonga.

Using words with differing degrees of animacy was important, as Bantu languages tend to use different default agreement markers for human denoting nouns and for nouns denoting inanimate objects (see §2.2.3). Various languages differ in whether nouns denoting animals

trigger the same default as human nouns (e.g. Sambaa (Riedel 2009)) or the default used with inanimates (e.g. Luganda (Givón 1970)).

(1) Example of isiXhosa word list for NC9/10

Inanimate		Animal		Human	
scarf	iqiya	ape	inkawu	albino	inkawu
dress	ilokwe	mouse	impuku	girl	intombi
book	incwadi	baboon	imfene	hero	intshatsheli
pipe	inqawa	cow	inkomo	blind man	imfama
tea	iti	elephant	indlovu	chief	inkosi
coffee	ikofu	dog	inja	man	indoda[2]

One issue that arose was trying to find an even spread of nouns across the various noun classes – in many cases the dictionaries alphabetise on the basis of the root, not the prefix. Thus while some noun classes were easy to find words for, as certain semantic domains are common in them (e.g. NC1 has several occupation names in it, such as *cook, student* and *artist*), other noun classes (such as NC3 and NC11) where harder to find *suitable* words for as they do not contain a handy semantic domain which is used in everyday conversation to draw upon. 'Suitable' here means nouns which are not abstract, do not refer to specialised subject areas that my city-dwelling consultants are unlikely to be familiar with (such as specific types of trees or fish), and which are easy to coordinate in subject position in sentences which do not sound pragmatically strange. The issue of certain types of nouns being pragmatically strange in coordination led me to omit from this study the noun classes which are reserved almost exclusively for abstract nouns (class 14), infinitives (class 15) and locatives (classes 16-18).

Trying to find human, animal and inanimate nouns for each of the noun classes posed some trouble for certain noun classes: NC1 tends to be reserved almost exclusively for nouns denoting humans, therefore making it *almost* impossible to include NC1 nouns denoting animal and inanimate objects in this study (isiXhosa has a few exceptions to this trend). Conversely, NC3 appears to be devoid of any nouns denoting humans, and light on nouns denoting animals.

Another problem that arose at times was ascertaining the noun class of a given word, as the noun classes were often not stated in the dictionaries. In such cases I would use the prefix on the noun as an indicator of its class. This, however, was not always a reliable solution: Some nouns have a null prefix, e.g. *hempe* (*shirt*) in Sesotho; in other cases the

[2]Although *indoda* belongs to NC9 in the singular, it belongs to NC6 in the plural.

prefixes for different noun classes happen to be homophonous, such as classes 1a, 1 and 3 in isiXhosa, which all bear the prefix *um-*. The issue of homophony is further relevant to cases where an allomorph exists for one noun class which is homophonous with the primary prefix of a different noun class – in isiXhosa, for example, the NC5 prefix *ili-* is under certain circumstances reduced to *i-*, making it homophonous with the NC9 prefix. It is therefore hard to tell just from looking at it whether an isiXhosa noun such as *isele* (*frog*) belongs to NC5 or NC9.

To counter this issue, I then went through the word lists with each of the consultants for that particular language, firstly giving the consultants the English word, and asking them what word they would use in their own language – thus testing whether or not that word was in their vocabulary – and then asking them to give the plural of the word; this showed what noun class the word is assigned to in their particular lexicons, as the pair of prefixes used on a particular word in singular and plural contexts tends to be what defines it as belonging to a certain class. For example, *isele* takes the NC6 prefix *ama-* in the plural, instead of the NC10 prefix *ii-*, thus indicating that *isele* is from NC5, rather than NC9.

Another way of disambiguating was to ask the consultants to translate a simple sentence with the noun in subject position, and then determine the noun class on the basis of the subject marker. These two methods were at times used in conjunction with each other, in order to reveal *before* the questionnaire was created whether a noun behaved strangely in the plural, by not taking the expected plural noun class; for example, we know that the isiXhosa word *indoda (man)* belongs to NC9 as it takes SM9, *i-*, but in the plural becomes NC6 *(amadoda)*.

The SG/PL method also revealed whether or not the consultant considered a noun to be non-count, and therefore incapable of taking a plural – for example, one consultant was adamant that *family* could not be pluralised in isiXhosa, and another felt the same way about *sweet potato* in Sesotho. By knowing this before setting up the questionnaires, I could avoid using these words altogether.

3.2.2 The questionnaires

By checking if a particular word was in the participant's lexicon I was able to tailor the questionnaires to each of the consultants – this meant that each consultant for a particular

language answered a slightly different version of the questionnaire. This helped to reduce noise in the data, and ensure the validity of the questionnaire for each participant, as the same construction was tested in each case, just with different words from time to time.

The questionnaires consisted of English prompts which had &Ps in the subject position which the consultants where asked to translate. The &Ps were carefully constructed to include specific combinations of conjuncts from specific noun classes. The consultants were asked to translate four variations of each sentence, based on the following paradigm:

(2) a. $DP1_{[SG]}$ & $DP2_{[SG]}$ e.g. *The fool and the lawyer are arguing.*
b. $DP1_{[PL]}$ & $DP2_{[PL]}$ e.g. *The fools and the lawyers are arguing.*
c. $DP2_{[SG]}$ & $DP1_{[SG]}$ e.g. *The lawyer and the fool are arguing.*
d. $DP2_{[PL]}$ & $DP1_{[PL]}$ e.g. *The lawyers and the fools are arguing.*

The first two combinations functioned to determine if different resolution strategies are used when the conjuncts are both singular (a) or both plural (b). Combinations (c) and (d) then tested whether the order of the conjuncts has any effect on the subject marker used. The data collected was then analysed, and the hypotheses were drawn. Thereafter additional questionnaires were constructed in order to collect the pertinent data to test these hypotheses. This cycle of model generation and testing was repeated until I deemed that I had an adequate corpus of data on which to base my analyses for the three languages under study.

During this process of model generation and testing, the analysis for isiXhosa (see chapter 6) made some interesting predictions about a set of data not in the corpus, as the paradigm in (2) failed to elicit it. I therefore created a new paradigm (3) and conducted additional interviews to elicit this data, not only for isiXhosa, but also for Sesotho and Xitsonga.

(3) a. $DP1_{[SG]}$ & $DP2_{[PL]}$ e.g. *The fool and the lawyers are arguing.*
b. $DP1_{[PL]}$ & $DP2_{[SG]}$ e.g. *The fools and the lawyer are arguing.*
c. $DP2_{[SG]}$ & $DP1_{[PL]}$ e.g. *The lawyer and the fools are arguing.*
d. $DP2_{[PL]}$ & $DP1_{[SG]}$ e.g. *The lawyers and the fool are arguing.*

This paradigm tested the effects of coordinating DPs with different number specifications, and whether the number specification on the DP closest to the verb might affect the choice of subject marker. Although (a) and (b) would be adequate to determine this, (c) and

(d) where included in the paradigm to 1) add further data for this issue, and 2) ensure that the results were not simply caused by the interaction of noun classes which behave strangely (as §5.2.5 and §6.2.3 will show, certain noun classes did seem to deviate from the trends in Sesotho and isiXhosa).

The length of the questionnaires differed for the various languages. Data collection for isiXhosa was more comprehensive than for the other two languages, as I attempted to collect an example of every different possible combination of two noun classes, spanning the three degrees of animacy. This resulted in over 500 sentences in the isiXhosa questionnaires. By the time I started collecting data for the other two languages, I had ascertained that many of the combinations were redundant, and so I scaled back the length of the Sesotho and Xitsonga questionnaires to a mere 176 and 160 prompts, respectively. I will show in chapter 7 that in fact a mere 22 carefully chosen prompts can be used to determine the complete range of agreement resolution strategies used within a language; however, I believe that a degree of redundancy helps to ensure the accuracy of one's data, and is thus, to a certain extent, desirable.

3.2.3 Avoiding the comitative

One significant issue which had to be held in consideration throughout the study was the fact that *with* is usually homophonous with *and* in Bantu. It was thus important to distinguish between comitative and conjunctive constructions. Several steps were taken to try and ensure that the conjunctive construction was being elicited instead of the comitative. The first of these was to ask the consultants to keep the translations as close as possible to the original, while still sounding natural and grammatical; they were also told that if a sentence couldn't be translated in a grammatical way without having to alter the sentence, they simply had to tell me, and I would accept it – this was to allow for the possibility that coordination was not possible in certain circumstances. I would also watch out for signals which might indicate that the comitative was being used. One such signal was the extraposition of one of the DPs, as in (4b) (contrast with (4a), where the DP has not been extraposed.)

(4) a. *Aba-fundi naba.zobi ba-ya e-doloph-ini*
2-student and.2-artist 2-go LOC-town-LOC
'The students and the artists are walking to town.'

b. *Aba-fundi ba-ya e-doloph-ini naba-zobi*
2-student 2-go LOC-town-LOC with.2-artist
'The students are walking to town with the artists.'

If the consultant produced a translation with an extraposed DP I would attempt to disambiguate the translation by asking questions either about what the sentence means (e.g. "*Are the students and the artists necessarily going to town* ***together****?*"), or about whether it would be appropriate to use that sentence in certain scenarios (e.g. "*Imagine that the students are going to town at ten o' clock, and the artists are going to town at eleven o' clock. You ask what everybody is doing today, and somebody replies that the students and the artists are going to town. Would it be appropriate to say* 'Abafundi baya edolophini nabazobi' *in that context? And what about* 'Abafundi nabazobi baya edolophini', *is that okay to say in that context?*"). Generally, the sentences with extraposed DPs necessarily entailed that the entities denoted by the two DPs be acting together, whilst the sentences without extraposition allowed for optional independent action by the two entities. Thus sentences with the extraposed DP were taken to be comitative, rather than coordinative.

3.3 Summary

The data used in this thesis was collected during a series of comprehensive individualised interviews with three to four speakers of each language. This individualization process is important in compensating for the small sample size. In large scale data collection methods, noisy data is balanced out by the large sample size. The advantage of the individualized interviews, however, lies not only in the reduction of noise in the data, but also in the fact that 1) one has more freedom in adjusting the questionnaire on the spot, should the consultant indicate some pragmatic 'weirdness' in a sentence and 2) one can question the consultant on issues such as *why* something sounds strange, or on alternative ways of phrasing the proposition being elicited. This helps to ensure the validity of the data collected from each consultant.

Chapter 4

Coordination in Xitsonga

This chapter examines how agreement with a coordinated subject works in Xitsonga. During the course of data collection, it became apparent that one of the consultants spoke a different variety from the other two. Interestingly, the two varieties differ not only in terms of the lexicon, but also in terms of the strategies used to resolve agreement with coordinated subjects. The first variety, dealt with in §4.2, uses a fairly straightforward resolution mechanism, always resorting to one of two default agreement markers, even when morphological agreement would be expected. This is shown to be the result of the constraints which require faithfulness to the [±HUMAN] features on the conjuncts dominating the constraint requiring faithfulness to the noun class features. The second variety (§4.3) utilizes a slightly more complex set of resolution rules, allowing morphological agreement with plural [+HUMAN] conjuncts which come from the same noun class, but not with plural [-HUMAN] conjuncts from the same noun class. This generalization is captured by reranking the constraints introduced in §4.2.

4.1 Background information about Xitsonga

Part of the Tswa-Ronga group, Xitsonga (S53)[1] is the mother tongue of 4.5% of the South African population (Statistics South Africa 2012:24). Within South Africa, it is generally associated with the provinces of Limpopo and Mpumalanga, but the 2011 census showed that there is also a population of almost 797 000 L1 speakers in Gauteng (cf. Mpumalanga,

[1]Originally classified as S52 by Guthrie (1948), it was later reclassified as S53, according to the New Updated Guthrie List (Maho 2009).

with just under 417 000 speakers) (Statistics South Africa 2012:23). Lewis (2009) estimates that in 2006 there were approximately 3 669 000 speakers in South Africa, Mozambique, Swaziland, and Zimbabwe.[2]

Xitsonga has 16 noun classes, namely 1-11, and 14-18. This study only focuses on classes 1-10, the prefixes and subject markers for which are shown in Table (1). The classes which are reserved almost exclusively for abstract nouns (NC14), infinitives (NC15), and locatives (NC16-18) have been omitted from this study, as the coordination of these types of DPs appears to be pragmatically strange. NC11 was initially included, but later removed due to the consultants being unfamiliar with words in this class.[3]

(1) The Subject Markers in Xitsonga

Singular NC	Basic SM	Plural NC	Basic SM
1 & 1a	u	2 & 2a	va
3	wu	4	yi
5	ri	6	ma
7	xi	8	swi
9	yi	10	ti
11	ri		

In this thesis I argue that each subject marker carries a combination of features regarding number, noun class, and 'humanness'. Table (2) shows what features are assumed for each SM. The assumptions about the [±HUMAN] features are derived from the data in §4.2.2 and §4.3.2.

As SM2 appears to be the default SM for [+HUMAN] conjuncts, it is assumed that SM2 must carry a corresponding [+HUMAN] feature. Likewise, as SM8 appears to be the default for [-HUMAN] conjuncts, it is assumed that SM8 must carry a corresponding [-HUMAN] feature. For the sake of consistency, the singular subject markers which correspond to the two defaults are assumed to carry the same [±HUMAN] feature as the plural. All the remaining subject markers are assumed to be under-specified with regard to the [±HUMAN] feature. It is important to note that the features assumed in this table apply only to the subject markers,

[2]Note, however, that this was based on a number of about 1 940 000 speakers in South Africa at the time, whereas the 2011 census showed an increase of more than 337 000 speakers in the intervening five years.

[3]It should also be noted that what van Wyk (1957) calls 'the augmentative noun class', and likens to NC21 in other Bantu languages (and which Baumbach (1987) classifies as NC5a) has been omitted, as my Xitsonga consultants preferred to use periphrastic means to indicate augmentation.

and not to all the nouns within those particular noun class. Thus it is possible to have a human denoting noun in NC8, even though SM8 carries a [-HUMAN] feature. Additionally, nouns are assumed to be either [+HUMAN] or [-HUMAN], but not [ØHUMAN], even though they might belong to a noun class whose SM is under-specified for humanness.

(2) Features associated with the various subject markers in Xitsonga

Singular SMs	Features	Plural SMs	Features
SM1 (u)	[SG, NC1, +H]	SM2 (va)	[PL, NC2, +H]
SM3 (wu)	[SG, NC3, ØH]	SM4 (yi)	[PL, NC4, ØH]
SM5 (ri)	[SG, NC5, ØH]	SM6 (ma)	[PL, NC6, ØH]
SM7 (xi)	[SG, NC7, -H]	SM8(swi)	[PL, NC8, -H]
SM9 (yi)	[SG, NC9, ØH]	SM10 (ti)	[PL, NC10, ØH]
SM11 (ri)	[SG, NC11, ØH]		

Finally, it should be noted Xitsonga has something called a 'compound subject marker'. This is a combination of the basic subject marker with a morpheme which appears to mark either present tense or the progressive aspect. In variety 1 this morpheme is realised as *-o-*, while in variety 2 it is realised as *-a-*; the plural compound subject markers for variety 1 are thus realised as *vo, yo, mo, swo* and *to*, while in variety 2 they are realised as *va, ya, ma, swa*, and *ta*. I have glossed the compound subject marker as SM.PRES throughout this chapter.

4.2 Variety 1

The variety of Xitsonga discussed in this section[4] is descriptively interesting, as it utilizes only one agreement resolution strategy. This strategy is default agreement, where the choice of default is governed by the [±HUMAN] features on the conjuncts, as §4.2.2 shows. This strategy is used even in contexts where one might expect to find morphological agreement (i.e. when the conjuncts do not carry conflicting NC-features). This preference for semantic

[4]Please note that this is not intended as a description of 'standard' Xitsonga. This variety is described first, as it employs a relatively straightforward resolution strategy, which is useful for establishing basic concepts that are central to this dissertation. There are two noticeable departures from the standard variety: Firstly, the conjunction *and* takes the form of *ni* instead of *na*; secondly, the vowel in the compound subject marker assimilates to *-o-* instead of *-a-*.

agreement over syntactic agreement is shown in §4.2.3 to be the result of MAX[+H] and MAX[-H] dominating MAXNC.

4.2.1 Consultant's background

The consultant who speaks variety 1 lives in Durban in KwaZulu Natal. Whilst this is not an area traditionally associated with Xitsonga – it is the mother tongue of less than 0.1% of the province (Statistics South Africa 2012:25) – there is enough of a Xitsonga community there that she none-the-less has opportunity to speak the language outside of the family home. It is, however, worth noting that her primary interactions in Xitsonga occur with her parents and extended family, all of whom are from Maputo in Mozambique. It is also important to note that she never studied Xitsonga formally. Thus, despite being raised in Durban, it is likely that her variety is more influenced by the Xitsonga spoken in Maputo, rather than that spoken in Durban. It is further worth noting that she also speaks Portuguese at home, as well as isiZulu and English socially.

4.2.2 Descriptive generalizations

Like many other Bantu languages (see § 2.2.3 on page 24), this variety of Xitsonga employs default agreement as an agreement resolution strategy. The default form used is determined by the [±HUMAN] feature on the conjuncts. This section will show, however, that *unlike* the Bantu languages discussed in §2.2, this variety of Xitsonga *only* uses default agreement, whereas the other languages utilize multiple agreement strategies, such as morphological agreement. The agreement resolution patterns for this variety of Xitsonga can be captured by the following three generalizations.

(3)
1. When both conjuncts are human, SM2 is used as the default.
2. When both conjuncts are nonhuman, SM8 is used as the default.
3. When conjuncts differ in their [±HUMAN] specification, SM2 is used as the default.

Before discussing the data, however, it is worth first pointing out the obvious: In instances of coordination, the subject marker appears to always originate from one of the

plural classes. Intuitively, this is hardly surprising, as coordination inherently involves semantic plurality (see Corbett 1983b:177).

4.2.2.1 [+HUMAN] balanced DPs

This section shows that when both conjuncts carry a [+HUMAN] feature, SM2 is used on the verb, regardless of what other features the conjuncts carry. Note that while only a few examples are given below, 48 sentences involving the coordination of two [+HUMAN] conjuncts were elicited, and in every instance the consultant for this variety used SM2 on the predicate.

In example (4a) below, a conjunct with the features [SG, NC7, +HUMAN] is coordinated with a conjunct with [SG, NC9, +HUMAN] features – thus there is a clash in NC features. The subject marker which appears on the verb is SM2, whose noun class feature (NC2) does not agree with the NC feature on either conjunct. SM2 is also used to agree with [+HUMAN] conjuncts even when the conjuncts are NC-balanced – that is to say, when they both carry the same NC-feature – as can be seen in (4b), where two conjuncts each carrying [SG, NC5, +HUMAN] features are coordinated.

(4) a. *A* *xi-kelema* *ni* *buchara* ***vo*** *vulavula*
DEF[5] 7-scoundrel and 9.butcher **SM2.PRES** talk
'The scoundrel and the butcher are talking.'

b. *A* *hahla* *ni* *jaha* ***vo*** *tira*
DEF 5.twin and 5.young man **SM2.PRES** work
'The twin and the young man are working.'

In (4b) one might expect that SM5 should be allowed, as this would agree with the NC-feature on both of the conjuncts. However, the [SG] feature on SM5 would clash with the semantic plurality of the &P.[6] It is also worth noting that although NC6 is considered the plural noun class for NC5 nouns, SM6 is *not* used on the verb.

The use of SM2 to agree with with [+HUMAN] nouns extends into the contexts where the conjuncts are plural. In (5a) the conjuncts come from NC8 and NC10, and again SM2

[5]What appears to be an augment has been glossed as DEF and transcribed separately from the noun, following Duarte (2011). It should be noted that on presenting this data at SAMWOP2, 2013, a Xitsonga speaker indicated to me that her speech community does not use this definiteness marker, whilst the speakers of the dialect discussed in 4.3 indicated that it was optional for them. The speaker from whom this data was elicited insisted on the use of the augment in sentence initial position, but did not require it at all times in sentence medial position.

[6]This would be like saying '*The twin and the young man* ***is*** *talking'* in English.

is used on the verb. In (5b) the conjuncts match each other feature for feature – they both carry the features [PL, NC6, +HUMAN] – and thus SM6 could *plausibly* be used on the verb, because it would agree with both conjuncts, and not clash with the seemingly inherent [PL] feature of the &P. But, again, SM6 is *not* used; SM2 is used instead. Thus the consistent use of SM2 to agree with [+HUMAN] conjuncts is reasonable grounds for assuming that SM2 is a default agreement marker for [+HUMAN] conjuncts. This subject marker most likely chosen as the default for humans owing to the fact that NC2 seems to be reserved almost exclusively for [+HUMAN] nouns.

(5) a. *A swi-kelema ni ti-buchara **vo** vulavula*
DEF 8-scoundrel and 10-butcher **SM2.PRES** talk
'The scoundrels and the butchers are talking.'

b. *A ma-hahla ni ma-jaha **vo** tira*
DEF 6-twin and 6-young man **SM2.PRES** work
'The twins and the young men are working.'

Thus the data above shows that when both conjuncts are human SM2 is used regardless of whether the conjuncts are singular (as in (4)) or plural (as in (5)), NC-unbalanced (i.e. the conjuncts come from different noun classes, as in the (a)-examples), or NC-balanced (as in the (b)-examples).

Default agreement is also used when the conjuncts differ in terms of number features, as can be seen in the following examples.

(6) a. *A swi-kelema ni xi-gevenga **va** yiva a mi-ntshumu*
DEF 8-scoundrel and 7-robber **SM2** steal DEF 4-thing
'The scoundrels and the robber are stealing the things.'

b. *A wa-nuna ni tin-hlanyi **vo** hambana*
DEF 1-man and 10-lunatic **SM2.PRES** argue
'The man and the lunatics are arguing.'

The default is used regardless of whether the singular conjunct appears before the plural conjunct, as in (6a), or vice versa, as in (6b). The issue of whether the conjuncts are gender-balanced[7] does not have an effect on the choice of resolution mechanism either, as can be seen from (6b).

[7]Gender-balanced conjuncts are conjuncts which come from the same singular-plural paired noun classes – eg. NC1 takes its plural in NC2, therefore they can be considered to be one gender – see § 2.1.1.2 on page 18.

4.2.2.2 [-HUMAN] balanced DPs

As with [+HUMAN] conjuncts, when both conjuncts are [-HUMAN], a default subject marker is used, regardless of whether the conjuncts are singular (as in (7)) or plural (as in (8)), NC-balanced (as in the (a) examples) or NC-unbalanced (as in the (b) examples). In Xitsonga the default SM for coordinated [-HUMAN] DPs is SM8, *swi* (or *swo* in its compound form).

(7) a. *A mfenhe ni mbyana **swo** lwa*
DEF 9.baboon and 9.dog **SM8.PRES** fight
'The baboon and the dog are fighting.'

b. *A nkhantjyu ni buruku **swi** li ka ntiya*
DEF 3.dress and 5.pants **SM8** LOC LOC line
'The dress and pants [one pair] are on the line.'

In (7a) two singular [-HUMAN] from NC9 are coordinated, and SM8 appears on the verb; similarly, in (7b) two singular NC-unbalanced [-HUMAN] conjuncts (from NC3 and NC5, respectively) also trigger SM8 agreement. The same agreement form is used when the conjuncts are plural, as the examples in (8) show.

(8) a. *A ti-mfenhe ni ti-mbyana **swo** lwa*
DEF 10-baboon and 10-dog **SM8.PRES** fight
'The baboons and the dogs are fighting.'

b. *A mi-nkhantjyu ni ma-buruku **swi** li ka ntiya*
DEF 4-dress and 6-pants **SM8** LOC LOC line
'The dresses and pants [multiple pairs] are on the line.'

In comparing the (a) examples with the (b) examples, it becomes apparent that Xitsonga variety 1 makes no grammatical distinction between animals and inanimates, as the same default agreement form is used for both. It is for this reason that I use the feature [-HUMAN], rather than [+ANIMAL] and [+INANIMATE]. The coordination of inanimates with animals is thus treated identically, with *swi* used as the default, as illustrated in (9).

(9) *A mi-mpfundla ni mi-nhlata **swi** li handle*
DEF 4-hare and 4-sweet potato **SM8** LOC out
'The hares and the sweet potatoes are outside.'

As before, SM8 is even used in the context of plural NC-balanced conjuncts. This is the context which should be least problematic for morphological agreement, as the noun class

features match each other, and the number feature on the conjuncts does not clash with the semantic PLURAL feature on &P. Intuitively, one might expect that morphological resolution agreement should happen by copying the features of the conjuncts onto onto the verb, but as (8a) and (9) show, this does not happen.

The consistent use of SM8 as a default agreement marker with [-HUMAN] conjuncts is grounds for arguing that this subject marker carries a [-HUMAN] feature.

4.2.2.3 Mixed [±HUMAN] conjuncts

When the two conjuncts differ in terms of their [±HUMAN] status – i.e. a [+HUMAN] DP is coordinated with a [-HUMAN] DP – the consultant for this variety paused and indicated that she faced a slight dilemma: The use of the [+HUMAN] default seemed to bestow human status on non-human conjuncts, whilst using the [-HUMAN] default seemed to to dehumanize the the human conjuncts. She found that the former option was significantly better than the latter, thus used the [+HUMAN] default, as illustrated in (10).

(10) a. *A n'-anga ni mbyana* ***vo*** ***(/∗swo)*** *famba*
DEF 9-doctor and 9.dog **SM2.PRES** **(/∗SM8.PRES)** walk
'The doctor and the dog are walking.'

b. *A mi-nhlata ni tin-hlanyi* ***va*** ***(/∗swi)*** *li ko zinya*
DEF 4-sweet.potato and 10-lunatic **SM2** **(/∗SM8)** LOC LOC room
'The sweet potatoes and the lunatics are in the room.'

Again, there appears to be no grammatical distinction between inanimates and animals, as it is possible to coordinate both types of [-HUMAN] DPs with [+HUMAN] DPs. Furthermore, the ordering of conjuncts appears to make no difference to the choice of default, as SM2 is used regardless of whether the [+HUMAN] appears first, as in 10a), or last, as in (10b). One can therefore conclude that faithfulness to a [+HUMAN] feature is prioritized over faithfulness to a [-HUMAN] feature.

4.2.2.4 The generalizations

The data presented in this section can be captured with the following three generalizations:

(11) 1. When both conjuncts are human, SM2 is used as the default.
2. When both conjuncts are nonhuman, SM8 is used as the default.

3. When conjuncts differ in their [±HUMAN] specification, SM2 is used as the default.

Additionally, one must remember that number agreement is always resolved to plural agreement – this fact, however, is implicit in the three generalizations above, as SM2 and SM8 both carry a [PLURAL] feature.

4.2.3 An OT analysis

In this section I show that the generalizations in (11) arise as the result of the interaction of four constraints, namely MAX[+HUMAN], which requires the subject marker to parse any [+HUMAN] feature on the conjuncts, MAX[-HUMAN], which requires the subject marker to parse any [-HUMAN] feature on the conjuncts, RESOLVENUMBER, which penalizes singular agreement with coordinated DPs , and finally MAXNOUNCLASS, which requires the subject marker to reflect the value of all NC-features in the &P. These constraints are defined in (12).[8]

(12) Constraint definitions

1. MAX[+HUMAN] (MAX[+H]): The semantic [+HUMAN] feature on each conjunct must be expressed morphologically by the SM.
2. MAX[-HUMAN] (MAX[-H]): The semantic [-HUMAN] feature on each conjunct must be expressed morphologically by the SM.
3. RESOLVENUMBER (RES#): Coordinated DPs in subject position trigger plural agreement on the predicate.
4. MAXNOUNCLASS (MAXNC): The noun class feature on *each conjunct* should correspond to the noun class feature on the SM.

Importantly, MAX[+H] only requires the output to carry a [+HUMAN] feature if, *and only if*, there is at least one [+HUMAN] feature in the input. If there is no [+HUMAN] feature in

[8]The MAX[FEATURE] constraints used here are very similar to the PARSE[FEATURE] constraints used by Grimshaw (1997a), as they require features which are present in the input to be parsed in the output; in the analyses in this thesis, the input consists of the two feature bundles on the conjuncts, and the output is the SM (and the feature bundle it carries) which appears on the predicate. Please see § 1.2.1 on page 6 for a more detailed discussion about the assumptions regarding the input-output issue.

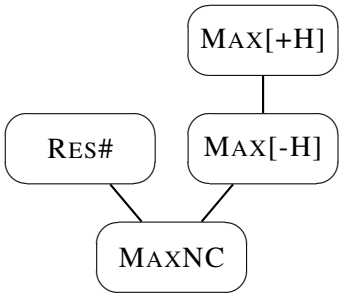
the input, MAX[+H] is vacuously satisfied by all of the candidate subject markers. Likewise, MAX[-H] only requires the output to carry a [-HUMAN] feature if, *and only if*, there is at least one [-HUMAN] feature in the input. Again, if there is no [-HUMAN] feature in the input, MAX[-H] is vacuously satisfied by all of the candidate subject markers.

By contrast, as all nouns in Xitsonga have a NC-feature, it is not possible for the input to not contain an NC-feature; this means that, unlike MAX[+H] and MAX[-H], the MAXNC constraint cannot be vacuously satisfied. Furthermore, MAXNC evaluates candidates in a slightly different way from the other two MAX constraints. MAXNC requires a candidate to parse all NC-features in the input, regardless of their values, and this evaluation process is binary: A candidate satisfies this constraint only if it agrees with the value of the NC-feature of *both* conjuncts; if a candidate parses the NC-feature of one conjunct but not the other, it violates the constraint just as badly as a candidate which does not parse the NC-feature of either conjunct. On the other hand, MAX[+H] and MAX[-H] specify the value of the [±HUMAN] feature which must be parsed – thus if the conjuncts carry different values for this feature, neither constraint requires the candidates to simultaneously parse both values.

There are three central facts about agreement with subject &Ps in this variety of Xitsonga that a formal analysis needs to account for: 1) &Ps always trigger plural agreement, 2) default agreement is always used, even when morphological agreement should be possible, and 3) faithfulness to a [+HUMAN] feature is prioritized over faithfulness to a [-HUMAN] feature.

This section shows that that the first fact is accounted for by the high ranking of RES#, while the preference for semantic agreement over morphological agreement arises as a result of MAX[+H] and MAX[-H] dominating MAXNC. Additionally, it is shown that MAX[+H]'s dominance over MAX[-H] gives rise to a preference for the use of the [+HUMAN] default SM to agree with mixed [±HUMAN] conjuncts. Thus, in this variety of Xitsonga, the constraints are ranked as follows.

(13) Ranking of constraints in Xitsonga variety 1



The ranking diagram in (13) shows MAX[+H] dominating MAX[-H], which in turn dominates MAXNC. RES# also dominates MAXNC, but no ranking relationship between RES# and MAX[+H] and MAX[-H] can be established.[9]

As §4.2.2.2 showed, when both conjuncts carry a [-HUMAN] feature, SM8 is used, because it also carries a [-HUMAN] feature; this is the case even when there exists a subject marker which agrees with the NC-feature on both conjuncts. Thus MAX[-H] must dominate MAXNC, as the latter constraint favours candidates which agree with the NC-feature on both conjuncts, whereas the the former favours candidates which carry a [-HUMAN] feature, as can be seen in tableau (14).[10]

(14) *A mi-mpfundla ni mi-hlwa **swo** dya*
DEF 4-hare and 4-termite **SM8.PRES** eat
'The hares and the termites are eating.'

4[-H] & 4[-H]	RES#	MAX[+H][11]	MAX[-H]	MAXNC
a. va (2)			*!	*
b. yi (4)			*!	
c. ☞ swi (8)				*
d. xi (7)	*!			*

In the tableau above, only SM4 (candidate (b)) reflects the NC4 feature found on both conjuncts, therefore only this candidate does not violate MAXNC. However, this candidate,

[9] As RES# is undominated, the ranking in (13) will be linearized with RES# in the first stratum of constraints in the tableaux below.

[10] In the tableaux in this chapter, the candidate subject markers are listed in their non-compound form – i.e. the form prior to the assimilation with the PRESENT morpheme.

[11] In tableau (14), MAX[+H] is vacuously satisfied by all the candidates, as there is no [+HUMAN] feature in the input.

along with candidate (a), violates MAX[-H], as it fails to parse the [-HUMAN] feature in the input. As MAX[-H] dominates MAXNC, candidate (b) is not chosen as the optimal candidate, despite its morphological agreement with each of the conjuncts. Candidates (c) and (d) *do* parse the [-HUMAN] feature in the input,[12] and thus they satisfy MAX[-H]. Of these two candidates, (c) has a plural feature, and so it does not violate RES#; candidate (d), however, carries a SINGULAR feature, and is therefore penalized by RES#. This leaves candidate (c), *swi*, as the optimal candidate.

As noted previously, no dominance relationship can be established between RES# and MAX[-H]: Candidate (c) would be chosen as the optimal candidate, regardless of whether RES# dominated MAX[-H] or MAX[-H] dominated RES#, as the following (minimal) comparison shows.

(15) a. RES# » MAX[-H]

4[-H] & 4[-H]	RES#	MAX[-H]
c. ☞ swi (8)		
d. xi (7)	∗!	

b. MAX[-H] » RES#

4[-H] & 4[-H]	MAX[-H]	RES#
c. ☞ swi (8)		
d. xi (7)		∗!

Thus, in this example, candidate (d) is actually harmonically bounded by candidate (c) – i.e. under no ranking of the constraints would candidate (d) emerge as the optimal candidate, as every constraint which it satisfies is also satisfied by candidate (c), but (c) also satisfies a constraint which (d) does not. Similarly, candidate (a) is harmonically bounded by candidates (b) and (c).

The tableau in (14), above, shows that MAX[-H] must dominate MAXNC in order for default [-HUMAN] agreement (as opposed to morphological agreement) to emerge as the preferred agreement resolution strategy with NC-balanced [-HUMAN] conjuncts. The tableau in (16), below, shows that something similar hold true for MAX[+H]: In order for default [+HUMAN] agreement (as opposed to morphological agreement) to emerge as the preferred agreement resolution strategy with NC-balanced [+HUMAN] conjuncts, MAX[+H] must also dominate MAXNC.

[12] Refer to table (2) on page 63 for an explanation of the features assumed for each subject marker.

(16) *A tin-hlanyi ni tin'-anga **vo** hambana*
DEF 10-lunatic and 10-doctor **SM2.PRES** argue
'The lunatics and the doctors are arguing.'

10[+H] & 10[+H]	RES#	MAX[+H]	MAX[-H][13]	MAXNC
a. ☞ va (2)				*
b. swi (8)		*!		*
c. ti (10)		*!		
d. u (1)	*!			*

In (16) two [+HUMAN] conjuncts from NC10 are coordinated. Although candidate (c) has the relevant NC-feature to simultaneously agree with the NC-feature on both conjuncts, thereby satisfying MAXNC, it fails to parse the [+HUMAN] feature in the input, and so violates MAX[+H]. The availability of candidates which satisfy MAX[+H] (specifically candidates (a) and (d), which *do* parse the [+HUMAN] feature) means that the violation assigned to candidate (c) by MAX[+H] renders NC10 suboptimal, because MAX[+H] dominates MAXNC. Of the two candidates which satisfy MAX[+H], candidate (d) carries a [SINGULAR] feature, which earns it a violation mark under RES#; candidate (a), however, has a [PLURAL] feature, and so satisfies RES#. Candidate (a) thus harmonically bounds candidate (d), and, despite the fact that SM2 does not parse the NC10 feature in the input, candidate (a) is chosen as the optimal candidate because it parses the [+HUMAN] feature, and also carries a [PLURAL] feature.

Note that the SM1 will not always be harmonically bounded by SM2: When NC-balanced [+HUMAN] conjuncts from NC1 are coordinated, SM1 will satisfy satisfy a constraint which SM2 does not, namely MAXNC. Thus RES# must dominate MAXNC, so that the plural [+HUMAN] agreement marker is chosen over the singular [+HUMAN] agreement marker which also happens to agree with the NC-feature on both conjuncts (17a), rather than the reverse scenario (17b).

(17) *A wa-nuna ni wa-nsati **vo** vulavula*
DEF 1-man and 1-woman **SM2.PRES** talk
'The man and the woman are talking.'

[13] In tableau (16), MAX[-H] is vacuously satisfied by all of the candidates, as the input does not have any [-HUMAN] feature which requires parsing.

a. RES# » MAXNC

1[+H] & 1[+H]	RES#	MAX[+H]	MAX[-H]	MAXNC
a. ☞ va (2)				*
b. u (1)	*!			

b. MAXNC » RES#

1[+H] & 1[+H]	MAX[+H]	MAX[-H]	MAXNC	RES#
a. va (2)			*!	
b. ☛ u (1)				*

Thus far I have shown that RES#, MAX[+H] and MAX[-H] must all dominate MAXNC. The tableau in (18) now shows that in order for the mixed [±HUMAN] conjuncts to trigger the [+HUMAN] default SM rather than the [-HUMAN] default SM, MAX[+H] must dominate MAX[-H].

(18) *A* *ma-hahla* *ni* *mi-mpfundla* ***vo*** *tlanga*
DEF 6.twin and 4-hares **SM2.PRES** play
'The twins and the hares are playing.'

6[+H] & 4[-H]	RES#	MAX[+H]	MAX[–H]	MAXNC
a. ☞ va (2)			*	*
b. yi (4)		*!	*	*
c. ma (6)		*!	*	*
d. swi (8)		*!		*

All of the candidates in this example have a [PLURAL] feature – they all, therefore, satisfy RES#. Candidate (a) parses the [+HUMAN] feature on the first conjunct, whilst the remaining candidates do not; (b) - (d) are therefore penalized under MAX[+H]. SM2, however, does not parse the [-HUMAN] feature on the second conjunct, and therefore violates MAX[-H]; SM8, by contrast, *does* parse the [-HUMAN] feature, and so does not violate MAX[-H]. Thus, in instances where the input consists of both a [+HUMAN] feature and a [-HUMAN] feature, MAX[+H] and MAX[-H] will favour different candidates. In order for SM2 to be rendered the optimal candidate, MAX[+H] *must* dominate MAX[-H].

Note that in (18), candidates (b) and (c) are harmonically bounded by candidates (a) and (d), despite the fact that (b) and (c) each parse one of the NC-features in the input; MAXNC evaluates candidates in a binary fashion: Either the candidate parses *all* the NC-features in the input, or it does not. Therefore, even though candidate (b) and (c) each parse one of the NC-features, these candidates both violate MAXNC, as neither candidate parses both of the NC-features.

In conclusion, the high ranking of RES# explains why Xitsonga Variety 1 always uses plural agreement with &Ps, while the dominance of MAX[+H] and MAX[-H] over MAXNC accounts for the preference for default agreement even in contests where morphological agreement is possible. Finally, the use of the [+HUMAN] default SM to agree with an &P which consists of both [+HUMAN] and [-HUMAN] conjuncts is accounted for by MAX[+H] dominating MAX[-H].

4.3 Variety 2

The second variety of Xitsonga has a slightly more complex set of resolution rules. In addition to default agreement, morphological agreement is also used. What is particularly interesting is that this variety treats [+HUMAN] and [-HUMAN] conjuncts differently, in a way which extends beyond merely the use of different defaults. After briefly describing the consultants' backgrounds (§4.3.1), I outline the generalizations in §4.3.2. I then show how the four constraints discussed above can also account for the agreement patterns observed in this variety (§4.3.3).

4.3.1 Consultants' backgrounds

The data for the second variety was collected from two students: One a resident of Bushbuckridge, Mpumalanga, and the other a resident of Johannesburg, Gauteng. The consultant from Bushbuckridge speaks Xitsonga both inside and outside the home. By contrast, the consultant from Johannesburg speaks Xitsonga exclusively at home, and thus her variety may be said to resemble that of her parents', rather than the variety(s) spoken elsewhere in Gauteng. Coincidentally, both her parents originate from Bushbuckridge; This variety may therefore be considered to be the Bushbuckridge variety. Neither consultant had studied

Xitsonga formally.

4.3.2 Descriptive generalizations

Initially this variety seems similar to the variety discussed in §4.2, as default agreement is the predominant strategy used. There is however, a significant difference: Plural NC-balanced [+HUMAN] conjuncts trigger morphological agreement. Interestingly the use of morphological agreement does not extend to conjuncts which [-HUMAN] conjuncts with otherwise identical features.

The generalizations for this variety are listed in (19).

(19)
1. Singular conjuncts (regardless of whether they are NC-balanced or NC-unbalanced) take default agreement: SM2 for Humans, and SM8 for non-humans.
2. NC-unbalanced plural conjuncts take default agreement.
3. NC-balanced [+HUMAN] plural conjuncts use morphological agreement.
4. NC-balanced [-HUMAN] plural conjuncts use default agreement.

For mixed [±HUMAN] conjuncts the case is less straight forward: One consultant found the coordination of human conjuncts with non-human conjuncts to be ungrammatical, while the other consultant found such coordination to be entirely acceptable. For the latter consultant, these mixed conjuncts behave like [+HUMAN] conjuncts: Morphological agreement is used when the conjuncts are plural and NC-balanced, and the [+HUMAN] default subject marker (SM2) is used in all other contexts.

4.3.2.1 NC-Unbalanced DPs

Like the variety discussed above, NC-unbalanced singular conjuncts take default agreement, regardless of whether they are [+HUMAN], in which case the default is SM2, *va*, as in (20) or [-HUMAN], in which case the default is SM8, *swi*, as in (21).

(20) a. *A hahla na xi-gevenga **va** ya a-sontw-eni*
DEF 5.twin and 7-robber **SM2** go LOC-church-LOC
'The twin and the robber are going to church.'

b. *Kokwana* *na* *n'-anga* ***va*** *vulavula*
1a.grandfather and 9-doctor **SM2.PRES** talk
'Grandfather and the doctor are talking.'

(21) a. *A* *mpfundla* *na* *khondlo* ***swa*** *tjutjuma*
DEF 3.hare and 5.rat **SM8.PRES** run
'The hare and the rat are running.'

b. *A* *xi-tulu* *na* *tafula* ***swi*** *le* *yindlw-ini*
DEF 7-chair and 5.table **SM8** LOC house-LOC
'The chair and the table are in the house.'

In (20a), two singular [+HUMAN] conjuncts, from NC5 and NC7, respectively, are coordinated; SM2 is used to agree with these conjuncts, even though neither conjunct carries a NC2 feature. The same is true of (20b): Neither of the [+HUMAN] conjuncts carries a NC2 feature, but, again, SM2 appears on the verb. By contrast, in (21a) and (21b) when singular [-HUMAN] conjuncts from different noun classes are coordinated, SM8 appears on the predicate instead of SM2. Thus SM2 is the default for [+HUMAN] conjuncts, while SM8 is the default for [-HUMAN] conjuncts. As with the variety discussed in 4.2, there appears to be no grammatical distinction between animals, as in (21a), and inanimate objects, as in (21b), as the same default is used for both groups.

The same default agreement markers are also used when the NC-unbalanced conjuncts are plural, again regardless of whether the conjuncts are [+HUMAN], as in example (22), or [-HUMAN], as in (23).

(22) a. *A* *ma-hahla* *na* *swi-gevenga* ***va*** *ya* *a-sontw-eni*
DEF 6-twins and 8-robber **SM2** go LOC-church-LOC
'The twins and the robbers are going to church.'

b. *Va-kokwana* *na* *tin'-anga* ***va*** *vulavula*
2a.grandfather and 10-doctor **SM2.PRES** talk
'The grandfathers and the doctors are talking.'

(23) a. *A* *mi-mpfundla* *na* *ma-khondlo* ***swa*** *tjutjuma*
DEF 4.hare and 6-rat **SM8.PRES** run
'The hares and the rats are running.'

b. *A* *swi-tulu* *na* *ma-tafula* ***swi*** *le* *yindlw-eni*
DEF 8-chair and 6-table **SM8** LOC house-LOC
'The chairs and the tables are in the house.'

Default agreement is further used when conjuncts differ in their SG/PL specification, as can be seen from (24).[14]

(24) a. *A hahla na ma-jaha **va** tira*
DEF 5.twin and 6-young man **SM2.PRES** work
'The twin and the young men are working.'

b. *A ma-jaha na hahla **va** tira*
DEF 6-young man and 5.twin **SM2.PRES** work
'The young men and the twin are working.'

The order of the conjuncts does not affect the choice of agreement resolution strategy: Default agreement is used regardless of whether the singular conjunct appears first, with the plural conjunct second (24a), or vice versa, with the plural conjunct appearing first, and the singular conjunct second (24b).

4.3.2.2 NC-Balanced DPs

The use of default agreement continues to make an appearance when NC-balanced singular DPs are coordinated, regardless of the [±HUMAN] status of the conjuncts, as (25) and (26) show.

(25) a. *A n-hlanya na n'-anga **va** tira*
DEF 9-lunatic and 9-doctor **SM2.PRES** work
'The lunatic and the doctor are working.'

b. *A hahla na jaha **va** tira*
DEF 5.twin and 5.young man **SM2.PRES** work
'The twin and the young man are working.'

(26) a. *A m-fenhe na m-byana **swa** lwa*
DEF 9-baboon and 9-dog **SM8.PRES** fight
'The baboon and the dog are fighting.'

b. *A hembe na buruku **swi** le mbedw-eni*
DEF 5.shirt and 5.pants **SM8** LOC bed-LOC
'The shirt and [one pair of] pants are on the bed.'

[14]Note that even though the conjuncts in (24) may be considered to be gender-balanced, they still count as NC-unbalanced, because NC5 and NC6 are different noun classes.

Where this variety of Xitsonga differs from the previous one is in its treatment of coordinated NC-balanced plurals. Example (27) shows that when the plural NC-balanced conjuncts both carry [+HUMAN] features, this variety uses morphological agreement, instead of the default agreement preferred by the variety discussed in §4.2.

(27) a. *A tin-hlanya na tin'-anga **ta** tira*
DEF 10-lunatic and 10-doctor **SM10.PRES** work
'The lunatics and the doctors are working.'

b. *A ma-hahla na ma-jaha **ma** tira*
DEF 6-twin and 6-young man **SM6.PRES** work
'The twins and the young men are working.'

In (27a), SM10 reflects the noun class of each of the conjuncts. Likewise, in (27b), both conjuncts come from NC6, and thus SM6 is used. This is in contrast with variety 1, which uses SM2 with the same conjuncts (c.f. example (16) on page 73 and example (5) on page 66, respectively).

Thus far, [-HUMAN] conjuncts seem to always have mirrored the strategies used by the [+HUMAN] conjuncts. As the below examples show, however, that is not the case in this instance. While plural NC-balanced conjuncts with [+HUMAN] features use morphological agreement, plural NC-balanced conjuncts with [-HUMAN] features use default agreement instead.

(28) a. *A tim-fenhe na tim-byana **swa** lwa*
DEF 10-baboon and 10-dog **SM8.PRES** fight
'The baboons and the dogs are fighting.'

b. *A ma-hembe na ma-buruku **swi** le mbedw-eni*
DEF 6-shirt and 6-pants **SM8** LOC bed-LOC
'The shirt and [multiple pairs of] pants are on the bed.'

Although (28a) initially seems to mirror (27a) in that both sentences conjoin two NC10 nouns in subject position, (28a) uses SM8 on the verb instead of SM10. Similarly, (28b) mirrors (27b) as the &P also consists of two NC6 nouns, but, again, SM8 appears on the predicate, instead of SM6.

The fact that the sentences in (27) and (28) differ in the [±HUMAN] status of the conjuncts plays a significant role in the choice of which resolution strategy is employed: Plural NC-balanced [+HUMAN] conjuncts use morphological agreement, whereas plural NC-balanced [-HUMAN] conjuncts use default agreement. Thus the [±HUMAN] features on the

conjuncts not only affect the choice of default agreement marker, but can also affect the choice of agreement resolution strategy. This strange asymmetry between the [+HUMAN] and [-HUMAN] conjuncts is something new which the OT analysis in 4.3.3 is going to have to account for.

4.3.2.3 Mixed [±HUMAN] conjuncts

There is no clear generalization for the coordination of human and non-human conjuncts: The two speakers of this variety of Xitsonga had different grammaticality judgements for such sentences. The consultant from Johannesburg was adamant that [+HUMAN] conjuncts may not be coordinated with [-HUMAN] conjuncts. For her, there is no way to resolve the clash of agreement features in a sentence such as (29).

(29) *A tin'-anga na tim-byana* ∗***va/∗swa/∗ta*** *famba*
DEF 10-doctor and 10-dog **SM2/8/10.PRES** walk
Int.: 'The doctors and the dogs are walking.'

She insisted that SM2 could not be used, as that would imply that the dogs had human qualities, and that likewise SM8 could not be used as that would make the doctors seem less than human. And even though both conjuncts are from NC10, and thus SM10 would not result in a clash of noun class features with either of the conjuncts, the consultant said the sentence still sounded strange when using this subject marker. For her there was no satisfactory way to coordinate the two conjuncts, and instead the comitative construction would be used.

By contrast, the consultant from Bushbuckridge did not find mixed [±HUMAN] conjuncts at all problematic. She resolved them in a similar fashion to the coordination of two [+HUMAN] conjuncts. When the conjuncts are both singular, SM2 is used as the default (30).

(30) a. *Jaha na mpfundla* ***va*** *le ngadin-i*
5.young man and 9.hare[15] **SM2** LOC garden-LOC
'The young man and the hare are in the garden.'

b. *A buruku na hahla* ***va*** *le kitshin-i*
DEF. 5.pants and 5.twin **SM2** LOC kitchen-LOC
'The pants [one pair] and the twin are in the kitchen'.

Like the consultant for the first variety, this consultant treated the coordination of human and animal conjuncts (see (30a)) the same as the coordination of human and inanimate conjuncts (see (30b)). Comparing (30a) with (30b) also shows that the relative ordering of the human and nonhuman conjuncts does not have any bearing on the choice of the subject marker: SM2 is used regardless of whether the [+HUMAN] conjunct appears before or after the [-HUMAN] conjunct.

In plural contexts mixed conjuncts continue to exhibit identical behaviour to coordination involving two human conjuncts. When the conjuncts are NC-unbalanced the [+HUMAN] default is used, as (31a) shows: Here the conjuncts come from NC6 and NC10, but SM2 is used on the verb.

(31) a. *Ma-jaha* *na* *ti-mpfundla* ***va*** *le* *ngadin-i*
6-young man and 10-hare **SM2** LOC garden-LOC
'The young men and the hares are in the garden.'

b. *A* *ma-buruku* *na* *ma-hahla* ***ma*** *le* *kitshin-i*
DEF. 6-pants and 6-twin **SM6** LOC kitchen-LOC
'The pants [multiple pairs] and the twins are in the kitchen'.

By contrast, when the conjuncts are NC-balanced, morphological resolution is used. This can be seen in (31b), where both conjuncts are from NC6, and SM6 is used on the verb.

4.3.2.4 The Generalizations

In summary, the following generalizations can be made about this variety of Xitsonga:

(32) 1. Singular conjuncts (regardless of whether they are NC-balanced or NC-unbalanced) take default agreement: SM2 for Humans, and SM8 for non-humans.
2. NC-unbalanced plural conjuncts take default agreement.
3. NC-balanced [+HUMAN] plural conjuncts use morphological agreement.
4. NC-balanced [-HUMAN] plural conjuncts use default agreement.

For mixed [±HUMAN] conjuncts there is not one clear cut generalization, as the consultants for this variety had differing grammaticality judgements. For one speaker, mixed

[15]Note that although the other Xitsonga consultants treated *mpfundla* as a NC3 noun, it became apparent during the word-list check that it was an NC9 noun for this consultant; subsequently it takes its plural in NC10, as can be seen in (31), rather than in NC4.

[±HUMAN] conjuncts triggered the same resolution strategies as [+HUMAN] conjuncts, whereas the other speaker found such sentences to be ungrammatical. For this reason I will treat the coordination of mixed [±HUMAN] conjuncts as outside of the scope of the analysis in §4.3.3.

4.3.3 An OT analysis

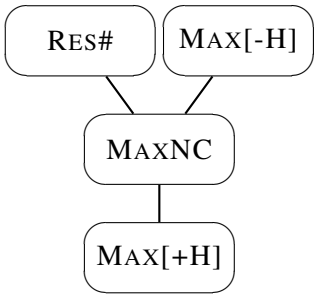
This section shows that the constraints used in §4.2.3 to account for the generalizations for Xitsonga variety 1 can also account for the generalizations for this second variety of Xitsonga. These constraints and their definitions are repeated here, for ease of reference.

(33) Constraint definitions

1. MAX[+HUMAN] (MAX[+H]): The semantic [+HUMAN] feature on each conjunct must be expressed morphologically by the SM.
2. MAX[-HUMAN] (MAX[-H]): The semantic [-HUMAN] feature on each conjunct must be expressed morphologically by the SM.
3. RESOLVENUMBER (RES#): Coordinated DPs in subject position trigger plural agreement on the predicate.
4. MAXNOUNCLASS (MAXNC): The noun class feature on *each conjunct* should correspond to the noun class feature on the SM.

As two [-HUMAN] conjuncts always take default agreement, and never morphological agreement, MAX[-H] must dominate MAXNC. By contrast, [+HUMAN] conjuncts sometimes trigger morphological agreement, which means MAXNC must dominate MAX[+H]. However, morphological agreement is only allowed when the [+HUMAN] conjuncts are both plural, and not when they are singular. This is the result of RES# dominating MAXNC. These ranking relations are represented schematically in (34), and discussed in more detail in the rest of this section.

(34) Ranking for Xitsonga Variety 2



MAXNC favours subject markers which parse all the NC-features in the input, while Max[+H] favours subject markers which parse the [+HUMAN] features in the input. Thus, in order for morphological agreement to surface instead of default agreement when [+HUMAN] conjuncts are plural and NC-balanced, MAXNC must crucially dominate MAX[+H], as can be seen in (35).

(35) *A swi-gevenga na swi-kelema **swa** tira*
DEF 8-robber and 8-scoundrel **SM8.PRES** work
'The robbers and the scoundrels are working.'

8[+H] & 8[+H]	RES#	MAX[–H]	MAXNC	MAX[+H]
a. va (2)			∗!	
b. ma (6)			∗!	∗
c. ☞ swi (8)				∗
d. u (1)	∗!		∗	
e. xi (7)	∗!		∗	∗

In (35), the conjuncts are two plural NC8 [+HUMAN] nouns. Of the potential subject markers, candidates (d) and (e) each bear a [SINGULAR] feature, and so violate RES#.[16] None of the candidates violate MAX[-H], as there is no [-HUMAN] feature in the input. Of the remaining candidates, MAXNC penalizes (a) and (b), as neither SM agrees with the NC8 feature on both conjuncts; candidate (c), however, satisfies this constraint, as it does carry a NC8 feature. As (c) is the only candidate to not have incurred any violations by

[16]In this example, all the other subject markers for the remaining singular noun classes (SM3, SM5 and SM9) incur the same violation pattern as candidate (e) in tableau (35), and so they have been left out of the tableau for the sake of space.

this point, it is selected as the optimal candidate. Note, however, that candidate (c) violates MAX[+H], as SM8 fails to parse the [+HUMAN] features in the input; by contrast, candidate (a) *does* parse these features, and so satisfies this constraint. Hence MAXNC must dominate MAX[+H], because if the relative ranking of these two constraints were reversed, MAX[+H] would favour SM2 (candidate (a)) over SM8.

RES#'s dominance over MAXNC explains why morphological agreement only happens when the NC-balanced conjuncts are *plural*. When the NC-balanced conjuncts are singular, as in (36), the only SM which would agree with both conjuncts is singular, and will therefore be penalized by RES#, as happens with candidate (d) in this example.

(36) *A* *xi-gevenga* *na* *xi-kelema* ***va*** *tira*
DEF 7-robber and 7-scoundrel **SM2.PRES** work
'The robber and the scoundrel are working.'

7[+H] & 7[+H]	RES#	MAX[-H]	MAXNC	MAX[+H]
a. ☞ va (2)			*	
b. ma (6)			*	*!
c. swi (8)			*	*!
d. xi (7)	*!			*

Again, MAX[-H] remains unviolated as there are no nonhuman features in the input. Candidate (d) is the only SM which satisfies MAXNC, but because it violates RES# it is no longer available for consideration. The remaining candidates all violate MAXNC, and, as a result, it is up to MAX[+H] to determine the optimal candidate. Only SM2 satisfies the faithfulness conditions of this constraint; candidate (a) thus is chosen as final output.

In (36) there are three constraints which distinguish between candidate (a) – the candidate which faithfully parses the [+HUMAN] features in the input – and candidate (d), the candidate which faithfully parses the NC-features in the input. These constraints are RES#, MAX[+H], and MAXNC. The latter favours candidate (d), while the other two favour candidate (a). Thus, in order for candidate (a) to be chosen as the optimal candidate, either RES# or MAX[+H] must dominate MAXNC. However, (35) demonstrated that MAXNC in fact dominates MAX[+H] – thus RES# must dominate MAXNC in order for SM2 to be used instead of SM7 in (36).

Not all instances of coordination involving plural [+HUMAN] conjuncts result in morphological agreement. When these conjuncts are NC-unbalanced, as in (37) below, default agreement is used.

(37) *A* *ma-hahla* *na* *tin'-anga* ***va*** *vulavula*
DEF 6-twin and 10-doctor **SM2.PRES** talk
'The twins and the doctors are talking.'

6[+H] & 10[+H]	RES#	MAX[-H]	MAXNC	MAX[+H]
a. ☞ va (2)			*	
b. ma (6)			*	*!
c. swi (8)			*	*!
d. ti (10)			*	*!
e. u (1)	*!		*	

In (37), the input consists of two different NC-features. None of the candidate subject markers parses BOTH NC-features, and so every candidate violates MAXNC. MAX[+H] then evaluates whether the candidates parse the [+HUMAN] features in the input. Only SM2 and SM1 bear a [+HUMAN] feature,[17] so only candidates (a) and (e) satisfy MAX[+H]. However, SM1 incurs a violation under the more highly ranked RES#, which leaves SM2 as the optimal candidate.

The tableau for choosing the optimal candidate to agree with singular NC-unbalanced [+HUMAN] conjuncts would look very similar to tableau (37). All singular subject markers would violate RES#, and thus be suboptimal. All of the remaining plural subject markers would vacuously satisfy MAX[-H], and then all violate MAXNC. Of the plural subject markers, only SM2 carries a [+HUMAN] feature, and so only SM2 would satisfy MAX[+H]; SM2 will therefore also be the optimal subject marker for agreeing with singular NC-unbalanced [+HUMAN] conjuncts.

MAX[-H] dominating MAXNC accounts for the fact that morphological agreement is not used with plural NC-balanced [-HUMAN] conjuncts. MAX[-H] favours subject markers which parse the [-HUMAN] feature, whilst MAXNC favours subject markers which parse

[17] Please refer back to table (2) on page 63 for the list of which features are associated with each subject marker in Xitsonga.

the value of the NC-feature. By ranking MAX[-H] above MAXNC, a grammar prioritizes the expression of the [-HUMAN] feature over the expression of the NC-feature.

In (38), two [-HUMAN] NC10 nouns are coordinated. Only SM10 parses both the NC-feature on each conjunct, and so only candidate (c) satisfies MAXNC. However, this candidate incurs a fatal violation under the more highly ranked MAX[-H], as it fails to parse the [-HUMAN] features on the conjuncts, and so is a suboptimal candidate. In Xitsonga, only SM7 and SM8 bear a [-HUMAN] feature (see (2) on page 63), and so only these two subject markers satisfy MAX[-H]. However, SM7 has a [SINGULAR] feature, and therefore violates RES#, and so is also considered suboptimal. This leaves SM8 as the optimal candidate.

(38) *A* *tim-fenhe* *na* *tim-byana* ***swa*** *lwa*
DEF 10-baboon and 10-dog **SM8.PRES** fight
'The baboons and the dogs are fighting.'

10[-H] & 10[-H]	RES#	MAX[-H]	MAXNC	MAX[+H]
a. va (2)		∗!	∗	
b. ☞ swi (8)			∗	
c. ti (10)		∗!		
d. xi (7)	∗!		∗	
e. yi (9)	∗!	∗	∗	

No ranking relation between RES# and MAX[-H] can be determined. As in variety 1, SM8 will always be chosen over SM7 to agree with two non-human conjuncts, regardless of the relative order of these two constraints, as the minimal comparison in (39) shows.

(39) a. RES# » MAX[-H]

10[-H] & 10[-H]	RES#	MAX[-H]
b. ☞ swi (8)		
d. xi (7)	∗!	

b. MAX[-H] » RES#

10[-H] & 10[-H]	MAX[-H]	RES#
b. ☞ swi (8)		
d. xi (7)		∗!

Note that the high ranking of MAX[-H] in this variety of Xitsonga means that SM8 will always be used to agree with two [-HUMAN] conjuncts, regardless of whether the conjuncts are NC-balanced (as in (38) above), NC-unbalanced, singular, or plural. In each case,

MAX[-H] would penalize every subject marker except for SM7 and SM8, as these are the only two which can parse the [-HUMAN] feature. Of these two, SM7 violates RES#, and so is suboptimal. Thus SM8 will always be the optimal subject marker when both conjuncts carry [-HUMAN] features.

Unresolved Issues

There is only one set of data which this analysis fails to account for: Coordination involving mixed [±HUMAN] conjuncts. For the consultant who found the coordination of [+HUMAN] nouns with [-HUMAN] nouns unproblematic, these &Ps used the same agreement resolution mechanisms as &Ps consisting of two [+HUMAN] nouns. As it stands, however, the ranking predicts that SM8, *swi*, should always be the optimal candidate, as it will satisfy the faithfulness conditions of MAX[-H] before SM2, *va*, has the opportunity to satisfy the faithfulness conditions of MAX[+H] in contexts where the [+HUMAN] default SM is required, as (40) illustrates.

(40) *Jaha* *na* *mpfundla* ***va*** *le* *ngadin-i*
5.young man and 9.hare **SM2** LOC garden-LOC
'The young man and the hare are in the garden.'

5[+H] & 9[-H]	RES#	MAX[-H]	MAXNC	MAX[+H]
a. ☞ va (2)		*!	*	
b. ma (6)		*!	*	*
c. ☛ swi (8)			*	*
d. ti (10)		*!	*	*
e. yi (9)	*!	*	*	*

In (40), MAX[+H] favours SM2, and MAX[-H] favours SM8. In order for SM2 to be chosen over SM8, Max[+H] must dominate MAX[-H]. However, as MAX[-H] dominates MAXNC, which in turn dominates MAX[+H], MAX[-H] transitively dominates MAX[+H]. Thus there is a ranking contradiction, and MAX[-H]'s transitive dominance over MAX[+H] results in the erroneous prediction that SM8 should be the optimal candidate (indicated by the ☛).

When the mixed [±HUMAN] conjuncts are plural and NC-balanced, they use morphological agreement, rather than default agreement. However, once again, the ranking erroneously predicts that the [-HUMAN] default SM should used, as (41) demonstrates.

(41) *A ma-buruku na ma-hahla **ma** le kitshin-i*
DEF. 6-pants and 6-twin **SM6** LOC kitchen-LOC
'The pants [multiple pairs] and the twins are in the kitchen'.

6[-H] & 6[+H]	RES#	MAX[–H]	MAXNC	MAX[+H]
a. va (2)		*!	*	
b. ☞ ma (6)		*!		*
c. ☛ swi (8)			*	*
d. ti (10)		*!	*	*
e. yi (9)	*!	*	*	*

Thus this analysis cannot account for the agreement patterns observed when conjuncts differ in their [±HUMAN] specification. However, as already noted, there is no one clear cut generalization about agreement with these conjuncts, as the other consultant for this variety deemed such coordination to be ungrammatical.

4.4 Concluding remarks

This chapter has shown how the same four constraints can account for agreement resolution strategies in two varieties of Xitsonga. In both varieties RES# is undominated, which results in a plural subject marker always being used to agree with the subject &P. The fist variety only uses default agreement, and never morphological agreement; this was shown to be the result of MAX[+H] and MAX[-H] (the two constraints which favour default agreement) dominating MAXNC (the constraint which favours morphological agreement). The second variety of Xitsonga uses morphological agreement with [+HUMAN] plural NC-balanced conjuncts, but not with [-HUMAN] plural NC-balanced conjuncts. This is the result of Max[+H] and Max[-H] being ranked differently relative to MAXNC.

The next chapter explores the agreement resolution strategies used in Sesotho, and then goes on to show how the constraint set introduced in this chapter can be supplemented so as capture the nuances of the Sesotho data.

Chapter 5

Coordination in Sesotho

This chapter examines the agreement resolution patterns found in Sesotho. There are some similarities between the patterns in Sesotho and those in the second variety of Xitsonga discussed in the previous chapter. However, unlike Xitsonga variety 2, Sesotho does not prohibit the use of morphological agreement with [-HUMAN] conjuncts. An additional difference is that Sesotho does not allow morphological agreement to be used when doing so would result in a clash between the [±HUMAN] feature on the subject marker, and the [±HUMAN] features on the conjuncts.

This chapter explores how the constraints proposed in the previous chapter account for the Sesotho generalizations, and proposes an additional constraint, DEP[-H], to account for a particularly interesting and unexpected subset of the data. I first contextualize this research in terms of background information on Sesotho and the consultants used (§5.1), before laying out the descriptive generalizations (§5.2) and discussing an Optimality theoretic analysis for the data (§5.3).

5.1 Background information

Part of the Sotho-Tswana group, Sesotho (S33) is spoken as an L1 by approximately 3.9 million speakers in South Africa (Statistics South Africa 2012), and a further estimated 1.8 million speakers in Lesotho, Botswana, and Swaziland (Lewis 2009). Within South Africa, it is primarily spoken in the Free State, with 1.7 million speakers, and Gauteng, with 1.4 million speakers (Statistics South Africa 2012).

Three of the four consultants used for the Sesotho study live in Johannesburg, Gauteng, and one lives in Welkom, Free State. All four are at least second generation residents of these areas, and none of them studied Sesotho at school. All of them were first year students at the time of data collection.

Sesotho has 15 noun classes, NC1-10, and NC14-18. As with Xitsonga, NC14-18 consist primarily of abstract nouns, infinitives, and locatives, and have thus been omitted from this study, as coordination involving these often yields pragmatically strange sentences; instead, classes 1-10 are focussed on. Table (1) shows the subject markers which are associated with each of these classes.

(1) The Subject Markers in Sesotho

Singular NC	Subject Marker	Plural NC	Subject Marker
1 & 1a	o	2 & 2a	ba
3	o	4	e
5	le	6	a
7	se	8	di
9	e	10	di

As with Xitsonga, I assume that each subject marker carries a combination of features regarding number, noun class, and 'humanness'. Table (2) shows the features which I have assumed for each subject marker. These features are important for the analysis which follows.

(2) Features associated with the various subject markers

Singular SMs	Features	Plural SMs	Features
SM1 (o)	[SG, NC1, +H]	SM2 (ba)	[PL, NC2, +H]
SM3 (o)	[SG, NC3, ØH]	SM4 (e)	[PL, NC4, ØH]
SM5 (le)	[SG, NC5, ØH]	SM6 (a)	[PL, NC6, ØH]
SM7 (se)	[SG, NC7, -H]	SM8/10 (di)[1]	[PL, NC8/10, -H]
SM9 (e)	[SG, NC9, -H]		

[1]It is impossible to tell whether the default is definitively SM8 or SM10 as they are homophonous: Both are realised as *di*. In this chapter I assume that SM8 and SM10 are in fact the same subject marker in Sesotho, and I thus use the label SM8/10. This assumption does not create any problems for the analysis, as far as I am aware.

The assumptions about the [±HUMAN] features of the subject markers are derived from the data which is laid out in §5.2, below. As SM2 appears to be the default SM for [+HUMAN] conjuncts, it is assumed that SM2 must carry a corresponding [+HUMAN] feature. Likewise, as SM8/10 (note that the two are homophonous) appears to be the default for [-HUMAN] conjuncts, it is assumed that SM8/10 must carry a corresponding [-HUMAN] feature. For the sake of consistency, the singular subject markers which correspond to the two defaults are assumed to carry the same [±HUMAN] feature as the plural, although the analysis would not be affected by assuming them to be under-specified for [±HUMAN]. All the remaining subject markers are assumed to be under-specified in this regard. It is important to note that the features assumed in this table apply only to the subject markers, and not to all the nouns within those particular noun class: Thus even though SM8/10 has a [-HUMAN] feature, it is still possible to have nouns with a [+HUMAN] feature in NC8 and NC10.

5.2 Descriptive generalizations for Sesotho data

Like the second variety of Xitsonga, Sesotho uses default agreement when the conjuncts have a mismatch in number or noun class features (i.e. are number-unbalanced, or NC-unbalanced), or when the [SINGULAR] feature on the conjuncts clashes with the semantic plurality of the &P. Where Sesotho differs from the second variety of Xitsonga is that it does not reserve morphological agreement solely for [+HUMAN] plural NC-balanced conjuncts – morphological agreement is also used when the plural NC-balanced conjuncts carry the [-HUMAN] feature. However, where morphological agreement would yield a subject marker which clashes with the [±HUMAN] feature of the conjuncts, morphological agreement is prohibited, and default agreement is used instead. These trends can be captured by the following generalizations.

(3)
1. When the conjuncts are **NC-unbalanced**, regardless of whether they are singular or plural, default agreement is used (SM2 for [+HUMAN] conjuncts, and SM8/10 for [-HUMAN] conjuncts).
2. When both conjuncts are **singular** and **NC-balanced**, default agreement is used.

3a. When both conjuncts are **plural** and **NC-balanced**, morphological agreement is used.

3b. Exception: When both conjuncts are **plural, [+HUMAN]**, and from **NC8 or NC10**, then default agreement is used (i.e. SM2).

The coordination of mixed [±HUMAN] conjuncts is ungrammatical, regardless of the subject marker used.

5.2.1 Singular conjuncts

Sesotho uses default agreement when both of the conjuncts are singular. As with Xitsonga, SM2 (*ba*) is the default for [+HUMAN] conjuncts. This can be seen from the fact that in coordination involving NC-unbalanced conjuncts, SM2 is used, even if none of the conjuncts have a conjunct from NC2, as illustrated in (4).

(4) a. *Mo-rena le le-kgoba* ***ba*** *ja*
1-king and 5-slave **SM2** eat
'The king and the slave are eating.'

b. *N-gaka le se-tsibi* ***ba*** *bua*
9-doctor and 7-expert **SM2** talk
'The doctor and the expert are talking.'

In (4a) the conjuncts come from NC1 and NC5, but SM2 is used even though it does not agree with either noun class. Similarly, in (4b) SM2 is used despite the fact that it does not agree with the NC9 feature on the first conjunct, or the NC7 feature on the second conjunct.

The fact that SM2 is used in contexts where [+HUMAN] conjuncts bear conflicting NC features – and where none of those NC features actually agrees with the NC feature of SM2 – is fairly convincing evidence on which to base the assumption that SM2 must have a [+HUMAN] feature. Further evidence comes from the fact that in singular NC-balanced contexts, SM2 is again used when both conjuncts denote human entities, as (5) shows.

(5) a. *Le-shodu le le-polesa* ***ba*** *lwana*
5-thief and 5-policeman **SM2** fight
'The thief and the policeman are fighting.'

b. *Se-tsibi le se-ra* ***ba*** *bua*
7-expert and 7-enemy **SM2** talk
'The expert and the enemy are talking.'

In (5a), above, the conjuncts both come from NC5, but instead of using SM5, which would agree with the NC5 feature on both conjuncts, SM2 is used, agreeing with the [+HUMAN] feature on the conjuncts. The same applies to (5b) in which both conjuncts are from NC7, but SM2 has been used. Additionally, although NC6 is the plural noun class for NC5, SM6 is not used in (5a); nor is SM8 used in (5b), even though NC8 is the plural noun class NC7.

Just as SM2 is used as a default for [+HUMAN] conjuncts, nonhuman conjuncts have their own default, namely SM8/10. Again, this can be determined from the fact that SM8/10 is used in instances of coordination involving NC-unbalanced [-HUMAN] conjuncts, as illustrated in (6).

(6) a. *N-tja le le-tata **di** tshimo-ng*
9-dog and 5-duck **SM8/10** garden-LOC
'The dog and the duck are in the garden.'

b. *Mo-ropa le se-lepe **di** nyametsa*
3-drum and 7-axe **SM8/10** disappear.CAUS
'The drum and the axe are missing.'

In (6a), the conjuncts come from NC9 and NC5, and yet the subject marker used is not associated with either of these noun classes; instead SM8/10 is used. Similarly, in (6b), the SM8/10 is used even though it does not agree with NC3 or NC7. Comparing (6a) with (6b) also shows that Sesotho, like Xitsonga, uses the same default for the animals as it does for inanimates.

The assumption that SM8/10 carries an explicit [-HUMAN] feature is further supported by the fact that it is used in contexts where the singular [-HUMAN] conjuncts are NC-balanced, as illustrated in (7).

(7) a. *Mo-nwana le mo-lala **di** bohloko*
3-finger and 3-neck **SM8/10** sore
'The finger and the neck are sore.'

b. *Le-tata le le-phela **di** ya tsamaya*
5-duck and 5-cockroach **SM8/10** PRES walk
'The duck and the cockroach are walking.'

In (7a), both conjuncts are from NC3, and thus SM3 would agree with both DPs; but SM3 is *not* used; nor is SM4, even though NC3 nouns take their plural form in NC4 –

instead SM8/10, *di*, is used. Likewise in (7b), both conjuncts are from NC5, but SM5, which would agree with both conjuncts, is not used. NC5 nouns take their plural for in NC6, but SM6 is not used either; once again, SM8/10 is used.

The only feature which the eight conjuncts in examples (6) and (7) have in common with each other, and do *not* have in common with the eight conjuncts in (4) and (5), is the [-HUMAN] feature. This is evidence for arguing that SM8/10 is used to agree with the [-HUMAN] feature on the conjuncts, and that, in order for it to be able to do so, the subject marker in question must carry a [-HUMAN] feature too.

Note that example (7) parallels (5), in that in each case there exists a subject marker which would agree with both conjuncts, but which is not used. However, if one assumes an inherent plural feature on &P (to match the semantic plurality yielded by coordination), then it is in fact not surprising that the agreeing subject marker is not used in these contexts: As the conjuncts all originate from singular noun classes, the subject markers in question would carry a singular feature, and this would clash with the plural feature on the &P.

5.2.2 Plural conjuncts

Although NC-balanced conjuncts behave the same way as NC-unbalanced conjuncts when both conjuncts are singular, they behave differently when both conjuncts are plural.

5.2.2.1 NC-unbalanced plurals

In plural contexts, NC-unbalanced conjuncts continue to trigger default agreement, as can be seen from examples (8) and (9).

(8) a. *Ba-rena le ma-kgoba* ***ba*** *ja*
2-king and 6-slave **SM2** eat
'The kings and the slaves are eating.'

b. *Din-gaka le di-tsibi* ***ba*** *bua*
10-doctor and 8-expert **SM2** talk
'The doctors and the experts are talking.'

(9) a. *Din-tja le ma-tata* ***di*** *tshimo-ng*
10-dog and 6-duck **SM8/10** garden-LOC
'The dogs and the ducks are in the garden.'

b. *Mo-ropa le se-lepe* ***di*** *nyametsa*
4-drum and 8-axe **SM8/10** disappear.CAUS
'The drums and the axes are missing.'

In each case in (8) and (9), the conjuncts carry NC features which clash with each other, making it impossible to supply one subject marker which would agree with both conjuncts. Instead, a subject marker which agrees with the [±HUMAN] feature of each conjunct is used. As already established in §5.2.1, this is SM2, *ba*, when both conjuncts are human, as in (8), and SM8/10, *di*, when both conjuncts are non-human, as in (9).

5.2.2.2 NC-balanced plurals

Plural NC-balanced conjuncts behave differently. Like with singular NC-balanced conjuncts, there exists a subject marker which agrees with both conjuncts, as both conjuncts carry the same NC features; unlike the singular contexts, however, the number feature on the subject marker *will* agree with the assumed plural feature on the the &P. This means that there is no clash which prevents this subject marker from appearing on the predicate. Thus morphological agreement is used with plural NC-balanced conjuncts, as illustrated in (10) & (11).

(10) a. *Me-nwana le me-lala* ***e*** *bohloko*
4-finger and 4-neck **SM4** sore
'The fingers and the necks are sore.'

b. *Ma-tata le ma-phela* ***ma*** *ya tsamaya*
6-duck and 6-cockroach **SM6** PRES walk
'The ducks and the cockroaches are walking.'

In example (10a) both conjuncts come from NC4. SM4 carries all the same morphological features as both conjuncts, and also agrees with the plural nature of the &P; in such contexts, therefore, agreement can be resolved morphologically, without any reference to the semantic [-HUMAN] feature on the conjuncts. The same applies in (10b), where both conjuncts come from NC6, and SM6 is used because it agrees with all the morphological features of both conjuncts, as well as the plural nature of the &P.

Likewise, in (11a), SM6 agrees with all the morphological features of both conjuncts (namely [PLURAL, NC6]), and does not clash with the assumed plural feature of &P; additionally, SM6 is under-specified with regard to the [±HUMAN] feature, and so does not

explicitly clash with the semantic [+HUMAN] feature on conjuncts. Similarly, in (11b) SM2 agrees with both conjuncts, as they both carry a NC2 feature. While this *could* also be interpreted as default agreement (as both conjuncts are [+HUMAN], and SM2 is the default agreement marker for [+HUMAN] conjuncts), the use of morphological resolution in other instances of plural NC-balanced coordination, including when both conjuncts are [-HUMAN] (e.g. (10), above) and when both are [+HUMAN], as in (11a), indicates that the agreement in (11b) is in all likelihood the result of morphological agreement, rather than default agreement.

(11) a. *Ma-shodu le Ma-polesa **a** lwana*
6-thief and 6-policeman **SM6** fight
'The thieves and the policemen are fighting.'

b. *Ba-nna le ba-sadi **ba** sebetsa*
2-man and 2-woman **SM2** work
'The men and the women are working.'

There are, however, two instances where default agreement is used in contexts of NC-balanced coordination, and that is when both conjuncts are [+HUMAN] and come from NC8, as in (12a), or NC10, as in (12b).

(12) a. *Di-tsibi le di-ra **ba (/*di)** bua*
8-expert and 8-enemy **SM2 (/*SM8/10)** talk
'The experts and the enemies are talking.'

b. *Di-ngaka le di-nese **ba (/*di)** bua*
10-doctor and 10-nurse **SM2 (/*SM8/10)** talk
'The doctors and the nurses are talking.'

If one were to follow the generalization that morphological agreement is used when conjuncts are plural and NC-balanced, then one would expect that when the conjuncts come from classes 8 or 10, the subject marker would be *di*, as per the agreement paradigm in table (1) on page 90. However, the study participants rejected this choice in these instances, indicating that SM2, *ba*, was significantly more preferred. This may be because *di*, the subject marker for classes 8 and 10, is the default SM for [-HUMAN], and this clashes with the [+HUMAN] features on the conjuncts. Thus when conjuncts were from classes 8 and 10, the participants reverted to semantic agreement instead of morphological agreement.

5.2.3 Number-unbalanced conjuncts

It is worth noting that number-unbalanced conjuncts are by their very nature NC-unbalanced conjuncts – this is because when one conjunct is singular and the other is plural, the very difference in number features means that the conjuncts will be in different noun classes. As such, coordination involving number-unbalanced conjuncts will thus behave exactly like coordination involving NC-unbalanced conjuncts: Default agreement will be used, and the form of the default will depend on the [±HUMAN] specification of the conjuncts.

(13) a. *Le-polesa le di-tsibi* ***ba*** *bua*
5-policeman and 8-expert **SM2** talk
'The policeman and the experts are talking.'

b. *Ma-polesa le se-tsibi* ***ba*** *bua*
6-policeman and 7-expert **SM2** talk
'The policemen and the expert are talking.'

(14) a. *Le-tata le din-tja* ***di*** *tshimo-ng*
5-duck and 10-dog **SM8/10** garden-LOC
'The duck and the dogs are in the garden.'

b. *Ma-tata le n-tja* ***di*** *tshimo-ng*
6-duck and 9-dog **SM8/10** garden-LOC
'The ducks and the dog are in the garden.'

Examples (13) and (14) both show instances where singular DPs have been coordinated with plural DPs. In (13), both conjuncts are [+HUMAN], and SM2 has been used accordingly. By contrast, when both conjuncts are [-HUMAN], SM8/10 is used, as illustrated in (14). By contrasting the (a) examples with the (b) examples, it can be seen that the order of the conjuncts has no effect on agreement resolution: Regardless of whether the singular conjunct precedes the plural conjunct (as in the (a) examples), or the plural conjunct precedes the singular conjunct (as in the (b) examples), default agreement is always used when the conjuncts are number-unbalanced.

5.2.4 Mixed Conjuncts

The Sesotho speakers I consulted for this study were very resistant when it came to coordinating conjuncts which differed from each other in their [±HUMAN] specification (i.e.

where one conjunct is [+HUMAN] and the other is [-HUMAN]). They argued that if one were to use *ba* (SM2) to agree with the human conjunct, then it would imply that the non-human conjunct has human status; conversely, to use *di* (SM8/10) to agree with the non-human conjunct would dehumanise the [+HUMAN] conjunct. Thus neither of the default agreement markers was considered appropriate (15a). Instead, the consultants insisted that coordination involving mixed [±HUMAN] conjuncts was just not allowed, and that a proposition such as *'The man and the dog are walking to town'* would have to be expressed using the comitative construction (15b).

(15) a. ∗*Mo-nna le n-tja* ***ba/di*** *ya teropo-ng*
1-man and 9-dog **SM2/SM10** go town-LOC
Int.: 'The man and the dog are going to town.'

b. *Mo-nna* ***o*** *ya teropo-ng le n-tja*
1-man **SM1** go town-LOC with 9-dog
'The man is going to town with the dog.'

Interestingly, one consultant pointed out that in some cases if the non-human conjunct was modified with a possessive pronoun – e.g. *'the man and* ***his*** *dog'* – then coordination was allowed. In such instances, the [+HUMAN] default, SM2, would be used (16).

(16) *Mo-nna le n-tja ya-hae* ***ba*** *ya teropo-ng*
1-man and 9-dog 9-3RD.POSS **SM2** go town-LOC
'The man and his dog are going to town.'

It is possible that in this instance the possessive pronoun '*his*' may be contributing a [+HUMAN] feature to the rest of the DP, thus counteracting the effects of the clash of the [±HUMAN] features. Unfortunately no definitive claims can be made about this however, as the use of a possessive pronoun modifying the non-human conjunct does not always ameliorate such a clash: My consultant informed me that a sentence such as *'the woman and her cat are sleeping on the couch'* would be just as ineffable in Sesotho as *'the woman and the cat are sleeping on the couch'*.

This resistance to mixed animacy coordination was not appeased when both conjuncts were plural and NC-balanced, even though this is precisely the context in which morphological resolution can occur with little reference to the [±HUMAN] specification of the conjuncts; the clash of the semantic [±HUMAN] features seems to override morphological resolution, whilst still disallowing default agreement, as illustrated in (17).

(17) ∗*Ma-shodu* *le* *ma-phela* ***a/ba/di*** *ka* *kamore-ng*
6-thief and 6-cockroach **SM6/SM2/SM8/10** LOC room-LOC
Int.: 'The thieves and the cockroaches are in the room.'

In (17), both conjuncts carry [PLURAL] and [NC6] features, so theoretically SM6 should be able to appear on the verb. For some reason, however, the conflicting animacy features on the conjuncts prevent this from happening.

Interestingly, the word *lephela* (*cockroach*) can be used metaphorically to denote a shady individual, with the same sort of connotations as *scumbag*, and when interpreted with this meaning, *lephela can* be coordinated with another [+HUMAN] DP. Agreement resolution in such instances will then follow the pattern followed by all other &Ps consisting of two [+HUMAN] conjuncts, namely using morphological agreement when the conjuncts are plural and NC-balanced (18a), and using default agreement in all other cases, such as (18b).

(18) a. *Ma-shodu* *le* *ma-phela* ***a*** *ka* *kamore-ng*
6-thief and 6-cockroach **SM6** LOC room-LOC
'The thieves and the scumbags are in the room.'

b. *Le-shodu* *le* *le-phela* ***ba*** *ka* *kamore-ng*
5-thief and 5-cockroach **SM2** LOC room-LOC
'The thief and the scumbag are in the room.'

5.2.5 NC4 & NC6 examples

There is one subset of data which is totally inconsistent, and does not fit the generalizations laid out above: When a conjunct from NC4 is coordinated with a conjunct from NC6 something strange happens. One would expect it to follow the same pattern as other plural NC-unbalanced conjuncts and take default agreement, but this does not happen. Instead, the consultants alternated between using SM4 and SM6, with no perceivable pattern in when which subject marker is used, as can be seen from (19) – (21). The data in (19) – (21) is all from one consultant; no two consultants answered in exactly the same way for these sentences, although they did all alternate between SM4 and SM6.

(19) a. *Me-lomo* *le* *ma-tsoho* ***a*** *bohloko*
4-mouth and 6-hand **SM6** sore
'The mouths and hands are sore.'

b. *Ma-tsoho le me-lomo* ***a*** *bohloko*
6-hand and 4-mouth **SM6** sore
'The hands and mouths are sore.'

(20) a. *Me-ropa le ma-tata* ***e*** *ka kamor-eng*
4-drum and 6-duck **SM4** LOC room-LOC
'The drums and the ducks are in the room.'

b. *Ma-tata le me-ropa* ***e*** *ka kamor-eng*
6-duck and 4-drum **SM4** LOC room-LOC
'The ducks and the drums are in the room.'

(21) a. *Me-ropa le ma-phela* ***a*** *ka kamor-eng*
4-drum and 6-cockroach **SM6** LOC room-LOC
'The drums and the cockroaches are in the room.'

b. *Ma-phela le me-ropa* ***e*** *ka kamor-eng*
6-cockroach and 4-drum **SM4** LOC room-LOC
'The cockroaches and the drums are in the room.'

At times it seemed as though the consultant favoured SM6 over SM4, regardless of the order of the conjuncts (as in (19a&b)), while at other times she seemed to favour SM4 instead (as in 20a&b)). At other times still, she seemed to favour agreement with the closest conjunct (as in (21a&b)). All that can be said is that when [-HUMAN] conjuncts from NC4 are coordinated with [-HUMAN] conjuncts from NC6, some form of partial agreement is used.[2] As this data is inconsistent, and no generalization can be made to capture the patterns of agreement in it, the analysis in §5.3 will not attempt to account for it.

5.2.6 The generalizations

Putting aside the issue of the ineffability of mixed [±HUMAN] coordination, and the erratic resolution strategies employed when an NC4 noun is coordinated with an NC6 noun, agreement resolution mechanisms in Sesotho can by and large be described by the following generalizations.

[2]The generalization that the exception applies only to [-HUMAN] conjuncts, rather than [+HUMAN] conjuncts comes from the following: NC3/4 does not typically include [+HUMAN] nouns, so there is no data for the coordination of [+HUMAN] conjuncts from NC4 with [+HUMAN] conjuncts from NC6, and as mixed [±HUMAN] coordination is ineffable in Sesotho, the trend could not be checked with reference to a [-HUMAN] conjunct from NC4 with a [+HUMAN] conjunct from NC6.

(22)
1. When the conjuncts are **NC-unbalanced**, regardless of whether they are singular or plural, default agreement is used (SM2 for [+HUMAN] conjuncts, and SM8/10 for [-HUMAN] conjuncts).
2. When both conjuncts are **singular** and **NC-balanced**, default agreement is used.

3a. When both conjuncts are **plural** and **NC-balanced**, morphological agreement is used.

3b. Exception: **NC-balanced [+HUMAN] conjuncts from NC8 or NC10** trigger default agreement (i.e. SM2).

5.3 An OT analysis

The generalizations listed in (22) show that the [NUMBER], [NOUN CLASS] and [±HUMAN] features of the conjuncts are all important factors in determining which subject marker will be used. The [±HUMAN] feature affects the choice of default SM. The [NUMBER] feature is responsible for the difference between generalizations 2 and 3a. Finally, contrasting generalization 3a with generalization 3b shows that the [NOUN CLASS] feature also has the potential to affect the agreement resolution strategy used.

The constraints introduced in the previous chapter – RES#, MAXNC, MAX[+H], and MAX[-H] – each make reference to one of these features, and so go some way to accounting for the generalization listed above. However, this constraint set must be supplemented with an additional constraint to account for generalization 3b. This generalization basically amounts to a prohibition from using morphological agreement when doing so would result in the [-HUMAN] default SM being used with [+HUMAN] conjuncts.

(23) lists the constraints which are necessary to account for the Sesotho data.

(23) Constraint definitions

1. MAX[+HUMAN] (MAX[+H]): The semantic [+HUMAN] feature on each conjunct must be expressed morphologically by the SM.
2. MAX[-HUMAN] (MAX[-H]): The semantic [-HUMAN] feature on each conjunct must be expressed morphologically by the SM.

3. RESOLVENUMBER (RES#): Coordinated DPs in subject position trigger plural agreement on the predicate.
4. MAXNOUNCLASS (MAXNC): The noun class feature on *each conjunct* should correspond to the noun class feature on the SM.
5. DEP[-HUMAN] (DEP[-H]): The [-HUMAN] specification of the SM must be expressed semantically by *at least one* of the conjuncts.

Constraints 1 to 4 behave here exactly as they do in the Xitsonga analysis: MAX[+H] stipulates that if there is a [+HUMAN] feature in the input, it must be reflected in the output, through the use of a subject marker which carries a [+HUMAN] feature. Similarly, MAX[-H] stipulates that if there is a [-HUMAN] feature in the input, it must parsed by the subject marker in the output. RES# penalizes any SM which carries a [SINGULAR] number feature, as this clashes with the semantic plurality of the &P.

MAXNC requires all NC-features in the input to be parsed by the subject marker. As explained on page 70, MAXNC evaluates subject markers in a binary fashion: A subject marker satisfies this constraint if, and only if, it parses *all* the NC-features in the input; if it parses none or only one of the NC-features, the subject marker is penalized. The violation incurred by a SM which only parses one of the NC-features in the input carries equal weight as the violation incurred by a SM which parses none of the NC-features in the input.

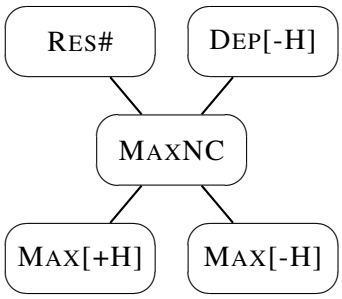
DEP[-H] is a new constraint, introduced in this chapter. It differs from MAX[-H] in two crucial ways: Firstly, MAX[-H] works by first examining the nouns in the input for [-HUMAN] features, and then looking for a corresponding [-HUMAN] feature on the SM (therefore operating in the direction of input to output), whereas DEP[-H] operates in the opposite direction (output to input), by first examining the SM for a [-HUMAN] feature, and *then* looking for a corresponding feature on the subject nouns. What this boils down to is that as *all* nouns carry either a [+HUMAN] or [-HUMAN] feature, MAX[+H] and MAX[-H] will always require a SM which explicitly carries a [+HUMAN] or [-HUMAN] feature; conversely, not all subject markers carry one of these features: As table (2) on page 90 shows, some subject markers, such as SM4 and SM6, are under-specified with regards to humanness ([ØHUMAN]). Thus DEP[-H] will treat a [ØHUMAN] subject marker as neutral, and it will not be penalized, regardless of whether the subject nouns are [+HUMAN] or [-HUMAN]. Only SMs which specifically carry a [-HUMAN] feature will be penalized if there is no

corresponding [-HUMAN] feature in the input.

The second way in which DEP[-H] differs from MAX[-H] is that the latter demands that *both* conjuncts match the SM in terms of the [-HUMAN] feature. This is analogous with the traditional use of MAX in phonology, where it stipulates 'No Deletion': Here no [±HUMAN] feature in the input is allowed to be 'lost', it must be reflected in the output. By contrast, DEP[-H] just requires that the [-HUMAN] feature on the SM must be somewhere in the input, even if only on *one* of the conjuncts. Again, this is somewhat analogous with the traditional use of DEP: No attention is paid to whether all the features in the input have a corresponding feature in the output, but instead the emphasis is on ensuring that the features which *do* appear in the output have a corresponding feature in the input.

The following ranking, (24), best captures the generalizations laid out in (22):

(24) Ranking of constraints in Sesotho



As MAX[+H] and MAX[-H] apply in entirely different contexts (namely exclusively human conjuncts for the former, and exclusively nonhuman conjuncts for the latter), the two constraints operate independently of each other. The only context in which these two constraints directly interact with each other is when the &P contains both human and nonhuman conjuncts. However, as such coordination is ineffable in Sesotho (see §5.2.4), no ranking relation can be established between these two constraints. MAX[+H] and MAX[-H] are therefore ranked in the same stratum.

MAXNC must dominate MAX[+H] and MAX[-H]; the latter two constraints are responsible for default agreement in the output, and if they were ranked above MAXNC, morphological agreement would never occur. MAXNC dominating MAX[+H] and MAX[-H] ensures that morphological agreement happens when the conjuncts are plural and NC-balanced, as can be seen in (25).

(25) *Ma-shodu le ma-polesa **a** lwana*
6-thief and 6-policeman **SM6** fight
'The thieves and the policemen are fighting.'

6[+H] & 6[+H]	RES#	DEP[-H]	MAXNC	MAX[+H]	MAX[-H]
a. ba (2)			*!		
b. ☞ a (6)				*	
c. di (8/10)		*!	*	*	
d. o (1)	*!		*		
e. le (5)	*!		*	*	

In (25b) MAXNC must crucially dominate MAX[+H], as the latter favours the non-optimal candidate (a) over the optimal candidate (b), because SM2 parses the [+HUMAN] feature of both conjuncts, whereas SM6 is underspecified for the [±HUMAN] feature. By contrast, MAXNC penalizes candidate (a) (as well as candidates (c)-(e)) because it does not express the NC6 feature which is found on both conjuncts, and favours candidate (b), because it *does* carry an NC feature which matches the NC feature on the conjuncts. Thus is MAX[+H] dominated MAXNC, the suboptimal candidate (a) would be favoured over the optimal candidate (b).

For similar reasons, MAXNC must also dominate MAX[-H]. When both conjuncts are nonhuman, MAX[-H] will always penalize any candidate which does not parse the [-HUMAN] feature – this constraint therefore favours the [-HUMAN] default SM. As morphological agreement is used when both conjuncts are plural and NC-balanced, a constraint which disqualifies the default SM must be ranked above MAX[-H]. As before, MAXNC does precisely this, as can be seen in (26): Only candidate (b) does not violate MAXNC, as only SM4 matches the NC-feature on both conjuncts in the input.

(26) *Me-nwana le me-lala **e** bohloko*
4-finger and 4-neck **SM4** sore
'The fingers and the necks are sore.'

4[-H] & 4[-H]	RES#	DEP[-H]	MAXNC	MAX[+H]	MAX[-H]
a. ba (2)			∗!		∗
b. ☞ e (4)					∗
c. di (8/10)			∗!		
d. u (3)	∗!		∗		∗
e. se (7)	∗!		∗		

The dominance of MAXNC over both MAX[+H] and MAX[-H] thus accounts for generalization (3a) in (22) on page 101, which states that when both conjuncts are plural and NC-balanced, morphological agreement will be used. Ranking DEP[-H] over MAXNC in turn accounts for the exception to this, which is listed as generalization (3b) – if the plural NC-balanced conjuncts carry [+HUMAN] features, as well as NC8 *or* NC10 features, then morphological agreement is *not* used, but instead the [+HUMAN] default SM is used, as can be seen in (27). In this example, both conjuncts come from NC8, but SM2, *ba*, is used, rather than SM8/10, *di*.

(27) *Di-tsibi le di-ra* ***ba (/∗di)*** *bua*
8-expert and 8-enemy **SM2 (/∗SM8/10)** talk
'The experts and the enemies are talking.'

8[+H] & 8[+H]	RES#	DEP[-H]	MAXNC	MAX[+H]	MAX[-H]
a. ☞ ba (2)			∗		
b. a (6)			∗	∗!	
c. di (8/10)		∗!		∗	
d. le (5)	∗!		∗	∗	

As explained above, DEP[-H] penalizes any candidate which has a [-HUMAN] feature on it if the conjuncts do not also carry a [-HUMAN] feature. This is precisely what happens with candidate (c). Although candidate (c) is the only candidate which parses the NC-feature on each conjunct, and is therefore the only candidate which satisfies MAXNC, it violates the more highly ranked DEP[-H], as it carries a [-HUMAN] feature when there is no [-HUMAN] feature in the input. Candidates (a) and (b) are thus the only candidates which are considered as possible options at this point.[3] MAX[+H] then evaluates the [±HUMAN]

[3]Candidate (d) is no longer considered as an option as it violates RES# – see below.

features on these candidates against those on the conjuncts. Candidate (b) carries a [ØHUMAN] feature, and so does not parse the [+HUMAN] feature on the conjuncts; it therefore violates MAX[+H]. By contrast, candidate (a) *does* carry a [+HUMAN] feature, and therefore satisfies this constraint, thereby making *ba* the optimal candidate. DEP[-H] therefore explains why morphological agreement cannot happen when the morphologically agreeing subject marker carries a [-HUMAN] feature which clashes with the [+HUMAN] features on the conjuncts (generalization (3b)).

The other context where a subject marker which would agree morphologically with both conjuncts exists but is not used, is when the conjuncts are *singular* and NC-balanced. In such instances, the appropriate [±HUMAN] default agreement is used instead (generalization (2)). The dominance of RES# over MAXNC accounts for this. This is because when both conjuncts are singular and NC-balanced, the only candidate subject marker which carries a NC-feature which matches those of the conjuncts, and therefore satisfies MAXNC, will also carry a [SINGULAR] feature; this violates the RES# constraint, which demands that the SM carry a [PLURAL] feature when the subject consists of conjoined DPs.

This is demonstrated in (28), where candidate (d) is the only candidate which is not penalized by MAXNC, and yet it is not the optimal candidate, as it violates the higher ranked RES# constraint. All of the other candidates violate MAXNC, and thus it is up to MAX[+H] to evaluate the [±HUMAN] features on the remaining candidates against the [+HUMAN] features in the input (MAX[-H] will not distinguish between any of the candidates in this example, as the input contains no [-HUMAN] features which must be parsed – all the candidates thus vacuously satisfy this constraint); of the remaining candidates, only (a) parses the [+HUMAN] feature, so it is the only one which satisfies MAX[+H]. Candidate (a) therefore rendered the optimal candidate.

(28) *Le-shodu* *le* *le-polesa* ***ba*** *lwana*
5-thief and 5-policeman **SM2** fight
'The thief and the policeman are fighting.'

5[+H] & 5[+H]	RES#	DEP[-H]	MAXNC	MAX[+H]	MAX[-H]
a. ☞ ba (2)			*		
b. a (6)			*	*!	
c. di (8/10)		*!	*	*	
d. o (1)	*!		*		
e. le (5)	*!			*	

Because RES# rules out the possibility of morphological agreement with NC-balanced conjuncts when they are singular, and results in default agreement being used instead, the situation arises where one can have two different inputs which each consist of conjuncts with exactly the same singular NC-features, but the optimal candidate will be different in each case, because the two inputs differ in terms of their [±HUMAN] specifications. For example, in (29), the input consists of two conjoined NC5 nouns, just like in (28), but unlike (28), the nouns in (29) both carry a [-HUMAN] feature, and thus the optimal candidate is now in fact (c), as it is the only subject marker with the [-HUMAN] feature (in this case MAX[+H] plays no role, as there are no [+HUMAN] features in the input against which to evaluate the candidates' features).

(29) *Le-tata le le-phela **di** ya tsamaya*
5-duck and 5-cockroach **SM8/10** PRES walk
'The duck and the cockroach are walking.'

5[-H] & 5[-H]	RES#	DEP[-H]	MAXNC	MAX[+H]	MAX[-H]
a. ba (2)			*		*!
b. a (6)			*		*!
c. ☞ di (8/10)			*		
d. le (5)	*!				*

Note that RES# and DEP[-H] never interact with each other, and thus no ranking relation between the two constraints can be determined.

This ranking also accounts for generalization (1), which states that when the conjuncts are NC-unbalanced (i.e. carry different NC- features), the appropriate [±HUMAN] default will be used, based on the [±HUMAN] features on the conjuncts. This applies regardless of whether the conjuncts are singular, as in (30), or plural, as in (31).

(30) *Mo-rena le le-kgoba* ***ba*** *ja*
1-king and 5-slave **SM2** eat
'The king and the slave are eating.'

1[+H] & 5[+H]	RES#	DEP[-H]	MAXNC	MAX[+H]	MAX[-H]
a. ☞ ba (2)			*		
b. a (6)			*	*!	
c. di (8/10)		*!	*	*	
d. u (1)	*!		*	*	
e. le (5)	*!		*	*	

In (30b), candidates (d) and (e) each agree morphologically with one of the conjuncts in the input, but they both carry a [SINGULAR] feature, and thereby violate RES#. Candidate (c) carries a [-HUMAN] feature for which there is no corresponding feature in the input; it is therefore penalized by DEP[-H]. Candidate (b) does not reflect the [+HUMAN] features found in the input, and so violates MAX[+H]. This therefore leaves (a) as the optimal candidate.

The tableau looks very similar when the conjuncts both denote non-human entities, with the main difference being that it is MAX[-H] which is responsible for the final differentiation between a non-optimal candidate and the optimal candidate. Like in (30), in (31), RES# penalizes candidates (d) and (e), as they are both singular subject markers. None of the candidates violate DEP[-H]: Only candidates (c) and (e) have a [-HUMAN] feature against which to evaluate the features in the input, but as these match, the two candidates satisfy DEP[-H]; the other candidates satisfy DEP[-H] vacuously. Candidates (a) - (c) all violate MAXNC equally, as none of these subject markers can simultaneously express NC6 and NC8; likewise, none of these candidates violates MAX[+H], as there are no [+HUMAN] features in the input which require expression in the output. Finally, MAX[-H] penalizes candidates (a) and (b), because they do not parse the [-HUMAN] features in the input. This renders (c) as the optimal candidate.

(31) *Mo-ropa le se-lepe* ***di*** *nyametsa*
4-drum and 8-axe **SM8/10** disappear.CAUS
'The drums and the axes are missing.'

6[-H] & 8[-H]	RES#	DEP[-H]	MAXNC	MAX[+H]	MAX[-H]
a. ba (2)			*		*!
b. e (4)			*		*!
c. ☞ di (8/10)			*		
d. u (3)	*!		*		*
e. se (7)	*!		*		*

As the analysis stands now, it might go some way towards explaining the ineffability of coordination involving mixed [±HUMAN] conjuncts. Consider the tableau in (32), for example:

(32) **Mo-nna le n-tja* ***ba/di*** *ya teropo-ng*
1-man and 9-dog **SM2/SM10** go town-LOC
Int.: 'The man and the dog are going to town.'

1[+H] & 9[-H]	RES#	DEP[-H]	MAXNC	MAX[+H]	MAX[-H]
a. ba (2)			*		*
b. a (6)			*	*	*!
c. di (8/10)			*	*	
d. u (1)	*!		*		
e. i (9)	*!		*		

In (32), candidates (d) and (e) both carry a [SINGULAR] feature, thus violating RES#. None of the remaining candidates violate DEP[-H], but they all violate MAXNC. Candidates (a) - (c) are therefore all available for evaluation under MAX[+H] and MAX[-H], which are equally ranked. Candidates (b) and (c) violate MAX[+H], as neither parses the [+HUMAN] feature found on the first conjunct. Candidates (a) and (b) violate MAX[-H], as neither parses the [-HUMAN] feature found on the second conjunct. With a total of two violations under the last two constraints, candidate (b) is definitively eliminated as a possible output. However, with one violation apiece, there is nothing to promote the choice of either candidate (a) or candidate (c) as the optimal candidate. It is possibly for this reason that coordination involving a human and a non-human conjunct may be considered as ineffable.

5.4 Concluding remarks

This chapter has shown how the constraints proposed in chapter 4 for Xitsonga, along with one new constraint, DEP[-H], can account for the agreement resolution strategies used in Sesotho. As in Xitsonga, MAX[+H] and MAX[-H] are the constraints which give rise to default agreement conditioned by the 'humanness' of the conjuncts. However, MAXNC's dominance over both MAX[+H] and MAX[-H] accounts for the use of morphological agreement with plural NC-balanced conjuncts, whilst RES#'s dominance over MAXNC accounts for the restriction against morphological agreement with *singular* NC-balanced conjuncts. Finally, DEP[-H]'s dominance over MAXNC accounts for the use of default agreement (instead of morphological agreement) with NC-balanced [+HUMAN] conjuncts from either NC8 or NC10.

The next chapter goes on to show how the constraint set used in this chapter can be further supplemented to account for the agreement resolution strategies used in isiXhosa. It should also be noted that one of the constraints introduced in the next chapter will go some way towards explaining *why* Sesotho exhibits some anomalies in agreement with &Ps coordinating [-HUMAN] conjuncts from NC4 with [-HUMAN] conjuncts from NC6 – this is discussed in the final chapter, on pages 164–167.

Chapter 6

Coordination in isiXhosa

The last chapter examined the patterns of agreement resolution in Sesotho and provided an Optimality Theoretic analysis to account for them, and this chapter will now do the same for isiXhosa. Like Xitsonga and Sesotho, default agreement conditioned by the [±HUMAN] feature on the verb is a common resolution strategy in isiXhosa; however, isiXhosa also employs a resolution strategy not found systematically in either of the other two languages: partial agreement. What is particularly interesting is that both closest conjunct agreement and furthest conjunct agreement are used. The data and descriptive generalizations are laid out in §6.2, followed by an OT analysis in §6.3. Predictions which follow from the analysis are then explored in section 6.4, and the analysis is adapted accordingly.

6.1 Background information

Part of the Nguni group, isiXhosa (S41) is the second largest of South Africa's 11 official languages. With 8.2 million L1 speakers, it is the mother tongue of 16% of South Africans. These speakers are most densely concentrated in the Eastern Cape (5.1 million speakers), and the neighbouring Western Cape (1.4 million speakers) (Statistics South Africa 2012). Outside of South Africa, an estimated 27 300 speakers can be found in Lesotho and Swaziland (Lewis 2009).

All of the consultants used for the isiXhosa study were born and raised in the Eastern Cape; two were from Butterworth, one from Mthatha, and one from King William's Town. While there has been some familial migration from town to town over the generations, all

of the consultants said their families had lived in the Eastern Cape for several generations. Two had studied isiXhosa at school, while the other two had not.

There are thirteen noun classes in isiXhosa: Classes 1-11, and 14-15. NC14 consists primarily of abstract nouns, and NC15 consists of verbs and nouns derived from verbs; initially, I attempted to include these classes in the study, but then had to omit them, as the consultants informed me that coordination involving words from these classes was pragmatically strange. For this reason, this study has focused on noun classes 1-11, the subject markers for which are given in table (1).

(1) The Subject Markers in isiXhosa

Singular NC	Subject Marker	Plural NC	Subject Marker
1 & 1a	u	2 & 2a	ba
3	u	4	i
5	li	6	a
7	si	8	zi
9	i	10	zi
11	lu		

As with Xitsonga and Sesotho, I assume that each subject marker carries a combination of features regarding number, noun class, and 'humanness'.

(2) Features associated with the various subject markers

Singular SMs	Features	Plural SMs	Features
SM1 (u)	[SG, NC1, +H]	SM2 (ba)	[PL, NC2, +H]
SM3 (u)	[SG, NC3, ØH]	SM4 (i)	[PL, NC4, ØH]
SM5 (li)	[SG, NC5, ØH]	SM6 (a)	[PL, NC6, ØH]
SM7 (si)	[SG, NC7, -H]	SM8/10 (zi)[1]	[PL, NC8/10, -H]
SM9 (i)	[SG, NC9, -H]		
SM11 (lu)	[SG, NC11, -H]		

The assumptions about the [±HUMAN] features are derived from the data laid out in §6.2 below. As SM2 appears to be the default subject marker for conjuncts which are [+HUMAN], I have assumed that SM2 carries a [+HUMAN] feature. Similarly, as NC1 is traditionally the

[1] SM8 and SM10 are homophonous: Both are realised as *zi*. As in the Sesotho chapter, I assume that SM8 and SM10 are in fact the same subject marker in isiXhosa, and I thus use the label SM8/10.

singular noun class for NC2 nouns, I have assumed that SM1 also carries a [+HUMAN] feature. By contrast, SM8/10 is used as the default for [-HUMAN] conjuncts, therefore I have assumed that SM8/10 carries a [-HUMAN] feature. At this point, the issue of assigning an explicit [±HUMAN] feature becomes slightly inelegant: As the singular noun classes which traditionally correspond with NC8 and NC10 are NC7 and NC9, I have also assigned a [-HUMAN] feature to SM7 and SM9, however, as NC11 also takes its plural in NC10, I have also assigned a [-HUMAN] feature to SM11; while this may be inelegant, it does not affect the analysis below, as singular subject markers are always ruled out fairly early in the tableaux. All the remaining subject markers are assumed to be under-specified with regard to the [HUMAN] feature. As noted in the Xitsonga and Sesotho chapters, the assumptions about the [HUMAN] features only apply to the subject markers, not to the nouns within the corresponding noun classes.

6.2 Descriptive generalizations for isiXhosa data

This section briefly describes the general patterns of agreement resolution used, and shows that there are two main resolution strategies: Default agreement and closest conjunct agreement. The choice of strategy appears to be determined by the number features on the conjuncts; for this reason the coordination of singular conjuncts (§6.2.1) and plural conjuncts (§6.2.2) are discussed in separate subsections. A few exceptions are discussed in §6.2.3, and the issue of mixed [±HUMAN] coordination is examined in §6.2.4. The patterns discussed in §6.2.1, §6.2.2 and §6.2.4 are then captured by three resolution rules in §6.2.5.

6.2.1 Singular conjuncts

In isiXhosa the coordination of two singular nouns triggers the use of default agreement, regardless of whether the conjuncts are NC-balanced (§6.2.1.1) or NC-unbalanced (§6.2.1.2). There are two default subject markers, the choice of which is conditioned by the value of the [±HUMAN] feature on the conjuncts.

6.2.1.1 NC-balanced singular conjuncts

NC-balanced singular conjuncts appear to behave exactly the same in isiXhosa as they do in Xitsonga and Sesotho. For example, when they are both [+HUMAN], they take the default subject marker *ba* (SM2), as example (3) shows; in each case both conjuncts belong to the same noun class (NC5 in (a), and NC7 in (b)), but the subject markers used do not correspond to these noun classes, nor do they correspond to the plural noun classes associated with NC5 or NC7 (NC6 and NC8, respectively).

(3) a. *I-polisa* *ne-gqwetha* ***ba**-ya-sebenza*
5-policeman and.5-lawyer **SM2**-PRES-work
'The policeman and the lawyer are working.'

b. *Isi-tyebi* *nesi-bhanxa* ***ba**-ya-funda*
7-rich.person and.7-fool **SM2**-PRES-study
'The rich person and the fool are studying.'

Similarly, when neither conjunct is human, as in (4), the subject marker used will also be a default, instead of a SM which corresponds to the noun class of the conjuncts (NC3 for (a), and NC5 for (b)), or one which corresponds to the plural noun class associated with them (NC4 and NC6, respectively). As before, the SM used for nonhuman conjuncts differs from that used for human ones; in isiXhosa this default is *zi*, SM8/10.

(4) a. *Um-pu* *nom-bhobho* ***zi**-se* *rumi-ni*
3-gun and.3-pipe **SM8/10**-LOC room-LOC
'The gun and the pipe are in the room.'

b. *I-sele* *ne-dada* ***zi**-ya-qubha*
5-frog and.5-duck **SM8/10**-PRES-swim
'The frog and the duck are swimming.'

Comparing (4a) with (4b) shows that *zi* is used as a default, regardless of whether the conjuncts denote animal or inanimate entities. This thus seems to indicate that, at least *grammatically*, isiXhosa does not encode animacy in terms of 'human', 'animal' and 'inanimate', but rather in terms of [±HUMAN]. *Zi* is therefore the default subject marker for [-HUMAN] conjuncts.

6.2.1.2 NC-unbalanced singular conjuncts

These default subject markers are also used when the conjuncts are not NC-balanced. In (5), the conjuncts do not carry the same NC features as each other: (5a) coordinates nouns from NC1 and NC5, whilst in (5b) the conjuncts are from NC5 and NC7. However, in each case, the conjuncts both carry a [+HUMAN] feature, and thus, just as in (3) above, SM2 is used to agree with this feature.

(5) a. *Um-fundi* *ne-polisa* ***ba**-ya-baleka*
1-student and.5-policeman **SM2**-PRES-run
'The student and the policeman are running.'

b. *I-gqwetha* *nesi-bhanxa* ***ba**-ya-lwa*
5-lawyer and.7-fool **SM2**-PRES-fight
'The lawyer and the fool are fighting.'

Similarly, in (6), the NC features on the conjuncts do not match each other – with conjuncts from NC3 and NC5 being coordinated in (6a), and conjuncts from NC5 and NC7 in (6b) – but as in (4), each conjunct carries a [-HUMAN] feature, and SM8/10 is thus used to agree with this feature.

(6) a. *Um-nqathe* *ne-qanda* ***zi**-se* *tafile-ni*
3-carrot and.5-egg **SM8/10**-LOC table-LOC
'The carrot and the egg are on the table.'

b. *I-dada* *nesi-khova* ***zi**-ya-bhabha*
5-duck and.7-owl **SM8/10**-PRES-fly
'The duck and the owl are flying.'

Thus when singular conjuncts are coordinated, agreement indexes a [±HUMAN] distinction, irrespective of whether the conjuncts are NC-balanced or NC-unbalanced. In this respect the isiXhosa data is similar to the Xitsonga and Sesotho data in the previous chapters.

The consistent use of SM2 to agree with conjuncts with a [+HUMAN] feature is fairly compelling evidence upon which I base the assumption that SM2 itself must have a [+HUMAN] feature. Likewise, the consistent use of SM8/10 with [-HUMAN] conjuncts seems to indicate that this subject marker must carry a [-HUMAN] feature. It is on the basis of this that the [±HUMAN] features in table 2 on page 112 were assigned.

6.2.2 Plural conjuncts

When both conjuncts are plural, isiXhosa diverges from the Xitsonga pattern, in that default agreement is not used at all. At first glance, however, the trend seems similar to one of the trends followed by Sesotho, as morphological agreement appears to be used when the plural conjuncts are NC-balanced (§6.2.2.1). However, comparing this data with instances of coordination involving plural NC-unbalanced conjuncts reveals that there is a somewhat different agreement resolution mechanism at work here (§6.2.2.2).

6.2.2.1 NC-balanced plural conjuncts

When the plural conjuncts are NC-balanced the agreement resolution patterns appear to emulate those of Sesotho, as they seem to use morphological agreement, regardless of whether the conjuncts are [-HUMAN], as in (7), or [+HUMAN], as in (8). In each sentence in (7) and (8), the first conjunct carries the same NC feature as the second conjunct, and in each case that same NC feature can be found on the subject marker.

(7) a. *Imi-pu nemi-bhobho **i**-se rumi-ni*
4-gun and.4-pipe **SM4**-LOC room-LOC
'The guns and the pipes are in the room.'

b. *Ama-sele nama-dada **a**-ya-qubha*
6-frog and.6-duck **SM6**-PRES-swim
'The frogs and the ducks are swimming.'

(8) a. *Ama-polisa nama-gqwetha **a**-ya-sebenza*
6-policeman and.6-lawyer **SM6**-PRES-work
'The policemen and the lawyers are working.'

b. *Izi-tyebi nezi-bhanxa **zi**-ya-funda*
8-rich.person and.8-fool **SM8/10**-PRES-study
'The rich people and the fools are studying.'

Example (8b), however, shows a slight deviation from the Sesotho pattern. In Sesotho, when [+HUMAN] conjuncts from noun classes 8 or 10 are conjoined, morphological agreement is *not* used; instead default agreement is used, resulting in the use of SM2 (c.f. (12) on page 96). Thus whereas isiXhosa seems to use morphological agreement consistently for the coordination of NC-balanced conjuncts, Sesotho appears to remain sensitive to animacy

in the same contexts. This signals a more prominent role played by animacy in Sesotho, while isiXhosa seems to prefer formal mechanisms of agreement instead.

6.2.2.2 NC-unbalanced plural conjuncts

More differences from the Sesotho pattern appear when one examines agreement with NC-unbalanced plural conjuncts. Careful attention must be paid to the data, for these differences to emerge, as examples such as those in (9) might lead one to believe that isiXhosa, like Sesotho, uses default agreement in these contexts.

(9) a. *Izi-caka* *naba-pheki* ***ba**-ya-pheka*
8-servant and.2-cook **SM2**-PRES-cook
'The servants and the cooks are cooking.'

b. *Ama-dada* *nezi-khova* ***zi**-ya-bhabha*
6-duck and.8-owl **SM8/10**-PRES-fly
'The ducks and the owls are flying.'

In (9a), the conjuncts are from NC8 and NC2 respectively, and both carry a [+HUMAN] feature, and SM2 is is used. By contrast, in (9b), the conjuncts are from NC6 and NC8 respectively, both carry a [-HUMAN] feature, and SM8/10 is used on the verb. It thus might appear as though the subject marker is chosen on the basis of the [±HUMAN] feature found on the conjuncts.

However, the data in (10), where the order of the conjuncts has been switched, shows that the data in (9) is in fact misleading.

(10) a. *Aba-pheki* *nezi-caka* ***zi**-ya-pheka*
2-cooks and.8-servants **SM8/10**-PRES-cook
'The cooks and the servants are cooking.'

b. *Izi-khova* *nama-dada* ***a**-ya-bhabha*
8-owl and.6-duck **SM6**-PRES-fly
'The owls and the ducks are flying.'

The examples in (10) show that it is actually the conjunct linearly closest to the verb which controls the agreement: In (10a) SM8/10 is used, as it agrees with the noun class of the closest conjunct, NC8. Similarly, in (10b), the closest conjunct comes from NC6, and thus SM6 is used on the verb. The examples in (9) thus merely appear to be using default

agreement, because the closest conjunct in each case come from a noun class whose subject marker is used as the default for the relevant [±HUMAN] distinction.

We are now in a position to reinterpret the examples in (7) and (8): Closest conjunct agreement could also account for the agreement patterns of these NC-balanced plural conjuncts – it is impossible to tell whether those patterns are the result of morphological resolution, or closest conjunct agreement, as both strategies would result in the same subject marker being used. However, the principle of Ockham's Razor states that simple explanations should be favoured over more complicated ones, and thus if two related sets of data can be explained by the same phenomenon, this is preferable to explaining them with reference to two different phenomena. For this reason, this chapter will analyse the coordination of both NC-balanced and NC-unbalanced plural conjuncts as being resolved by means of closest conjunct agreement in isiXhosa.

Still, the question arises as to why closest conjunct agreement only happens when the conjuncts are plural, and not when they are in the singular. The answer is quite simple: When the conjuncts are in the singular, closest conjunct agreement would result in a subject marker with a [SINGULAR] feature, and this would clash with the semantic plurality inherent to &P. This is accounted for in the OT analysis in §6.3.

6.2.3 Exceptions to these trends

There are two groups of exceptions to the generalizations described above: One involving NC11 nouns which take their plural in NC6, and one involving non-human nouns from NC1a/2a. It should, however, be noted that not all nouns from these classes behave in the deviant manner described below.

6.2.3.1 Exceptions involving NC11 nouns which take their plurals in NC6

Most NC11 nouns take their plural form in NC10, but there is a small set which take their plural in NC6 instead. When this happens, coordination involving these plural forms follows a different pattern from that described above. Rather notably, the way in which they deviate from the trend differs according to the [±HUMAN] status of the conjuncts.

Thus far, the generalization has been that when both conjuncts are plural, closest conjunct agreement is used, and no attention is paid to the [±HUMAN] feature on the nouns.

However, when one of these plural conjuncts is from NC6 with its singular in NC11, closest conjunct agreement is not always used. In fact, when the conjuncts both denote humans, CCA is not used at all, and instead the [+HUMAN] default subject marker is used, as can be seen in (11).[2]

(11) a. *Izi-gulane nama-sana* ***ba****-ya-dlala*
8-patient and.6-baby **SM2**-PRES-play
'The patients and the babies are playing.'

b. *Iin-kawu nama-sana* ***ba****-ya-dlala*
10-albino and.6-baby **SM2**-PRES-play
'The albinos and the babies are playing.'

c. *Ama-sana nezi-gulane* ***ba****-ya-dlala*
6-baby and.8-patient **SM2**-PRES-play
'The babies and the patients are playing.'

d. *Ama-sana neen-kawu* ***ba****-ya-dlala*
6-baby and.10-albino **SM2**-PRES-play
'The babies and the albinos are playing.'

Interestingly, this avoidance of CCA happens regardless of whether the NC6 noun is the conjunct closest to the verb, as in (11a&b), or not, as in (11c&d). It is worth noting that the only [+HUMAN] noun from NC11 that takes its plural in NC6 that I could find was *usana*, so this *might* be an anomaly associated specifically with that word. However, the fact that more anomalies involving NC11→NC6 nouns appear when both conjuncts are [-HUMAN] leads me to believe otherwise.

Due to the fact that these anomalous nouns trigger the [+HUMAN] default SM when the conjuncts denote human entities, one might expect them to trigger the [-HUMAN] default when the conjuncts denote non-human entities, but this does *not* happen. When both conjuncts are [-HUMAN] an altogether different strategy is employed: NC6 agreement is employed, regardless of whether the NC11→NC6 noun is the first conjunct or the second conjunct, as illustrated in (12). [3]

(12) a. *Ama-khuko neen-lokwe* ***a****-phantsi*
6-mat and.10-dress **SM6**-down
'The mats and the dresses are on the floor.'

[2]The singular form of the NC6 noun *amasana* is *unsana*, which is in NC11.

[3]The singular forms of the NC6 nouns *amakhuko* and *amahadi* are *ukhuko* and *uhadi*, respectively; both are NC11 nouns.

b. *Ama-hadi nezi-dlalisi **a**-phezu kwe-tafile*
6-harp and.8-toy **SM6**-on LOC-table
'The harps and the toys are on the table.'

c. *Iin-lokwe nama-khuko **a**-phantsi*
10-dress and.6-mat **SM6**-down
'The dresses and the mats are on the floor.'

d. *Izi-dlalisi nama-hadi **a**-phezu kwe-tafile*
8-toy and.6-harp **SM6**-on LOC-table
'The toys and the harps are on the table.'

In (c) and (d), the NC6 noun is the conjunct closest to the verb, and thus the use of SM6 is not particularly remarkable – it merely appears to be following the trend of using closest conjunct agreement when both conjuncts are plural. By contrast, in (a) and (b), the NC6 noun is the first conjunct, with nouns from NC10 and NC8, respectively, closest to the verb, and yet SM8/10 is not used – instead SM6 is used, apparently agreeing with the furthest conjunct. I can offer no explanation for these anomalies, other than to suggest that the divergent NC features born by NC11→NC6 nouns are somehow responsible.

6.2.3.2 Exceptions involving [-HUMAN] NC1a/2a nouns

The second group of exceptions involves [-HUMAN] conjuncts from NC1a/2a. Traditionally, NC1 and NC2 are reserved exclusively for [+HUMAN] nouns; however, NC1a and NC2a are special subsets of NC1 and NC2 (NC2a even has its own prefix: *oo-* instead of *aba-*), and in these subgroups can be found a small selection of [-HUMAN] nouns, such as *umatshini (machine), uloliwe (train), unonkala (crab)* and *udyakalashe (jackal)*. Several exceptions which involve such nouns arose, but unfortunately the exceptions were not entirely systematic.

When singular conjuncts are NC-balanced, the trend thus far has been to use the appropriate [±HUMAN] default. However, when the conjuncts are [-HUMAN] from NC1a, the [-HUMAN] default is not used; instead SM2 appears on the predicate (13).

(13) a. *U-loliwe no-matshini **ba**-se galaji-ni*
1a-train and.1a-machine **SM2**-LOC garage-LOC
'The train and the machine are in the garage.'

b. *U-nomadukudwane no-nokala **ba**-ya-tya*
1a-scorpian and.1a-crab **SM2**-PRES-eat
'The scorpian and the crab are eating.'

c. *U-dyakalashe no-matshini **ba**-se gadi-ni*
1a-jackal and.1a-machine **SM2**-LOC garden-LOC
'The jackal and the machine are in the garden.'

This could be interpreted as either the [+HUMAN] default SM being used, or as coordinated NC-balanced singular conjuncts taking the subject marker of the plural noun class traditionally associated with them. I do not know which is the correct interpretation: If it is the latter, why does this not happen with other singular NC-balanced conjuncts? And while one might argue that the very presence of these nouns in NC1a might mean that they have an honourary [+HUMAN] status, which would trigger the [+HUMAN] default SM, the evidence which I present below seems to indicate that NC1a/2a nouns are deviant in a way which cannot be accounted for by assuming a [+HUMAN] feature.

When the conjuncts are plural, the trend has been to use closest conjunct agreement. While this is sometimes the case when the closest conjunct is from NC2a, as in (14), at other times closest conjunct agreement seems to be actively avoided in favour of furthest conjunct agreement when the closest conjunct is from NC2a, as in (15).

(14) a. *Izin-ja noo-dyakalashe **ba**-ya-baleka*
10-dog and.2a-jackal **SM2**-PRES-run
'The dogs and the jackals are running.'

b. *Imi-bhobho noo-matshini **ba**-se galaji-ni*
4-pipe and.2a-machine **SM2**-LOC garage-LOC
'The pipe and the machines are in the garage.'

(15) a. *Ama-sele noo-nonkala **a**-hleli e-dame-ni*
6-forg and.2a-crab **SM6**-stay.PERF LOC-dam-LOC
'The frogs and the crab stay at the dam.'

b. *Iin-cwadi noo-matshini **zi**-lahlek-ile*
10-book and.2a-machine **SM8/10**-get.lost-STAT
'The books and the machines are missing.'

In (14) closest conjunct agreement has been used, resulting in the use of SM2, whilst in (15) closest conjunct agreement seems to be avoided when it would result in the use of SM2. This shows that the exceptions are not systematic; as a result, the analysis in §6.3 will not be able to account for them.

6.2.4 Mixed [±HUMAN] conjuncts

As in Xitsonga, there was some disagreement amongst the participants about whether or not [+HUMAN] nouns may be coordinated with [-HUMAN] nouns. Two of the consultants felt that it was highly ungrammatical to coordinate conjuncts which clash in terms of their [±HUMAN] feature, whilst the other two expressed no such concern, even when specifically asked about the grammaticality of the data. The data in this section comes specifically from these two consultants.

I managed to elicit 100 examples of mixed [±HUMAN] coordination: 50 examples coordinating two singular nouns, and 50 coordinating two plural nouns. This data is somewhat messy, as there is no one set pattern which the consultants followed, but there are noticeable trends which emerged.

6.2.4.1 Both conjuncts are singular

For the most part (36/50 examples), when the conjuncts are both singular, the consultants said that the subject marker which seemed most grammatical was SM2. This was regardless of whether the [+HUMAN] came first, as in (16), or second, as in (17).

(16) a. *Isi-gulane* *nesi-tye* ***ba**-se* *kitshi-ni*
7-patient and.7-dish **SM2**-LOC kitchen-LOC
'The patient and the dish are in the kitchen.'

b. *I-gwala* *nom-vundla* ***ba**-zimele*
5-coward and.3-hare **SM2**-hide
'The coward and the hare are hiding.'

(17) a. *Isi-tya* *nesi-gulane* ***ba**-se* *kitshi-ni*
7-dish and.7-patient **SM2**-LOC kitchen-LOC
'The dish and the patient are in the kitchen.'

b. *Um-vundla* *ne-gwala* ***ba**-zimele*
3-hare and 5-coward **SM2**-hide
'The hare and the coward are hiding.'

The above data show that NC-balancedness (see (a)-examples) and NC-unbalancedness (see (b)-examples) do not affect the choice of subject marker, nor does the animacy of the

non-human conjunct. However, this generalization accounts for only 72% of the 50 sentences. The remaining 14 sentences do the complete opposite, and take *zi* (SM8/10), as (18) and (19) show.

(18) a. *In-doda nen-ja **zi**-lele*
9-man and.9-dog **SM8/10**-sleep
'The man and the dog are sleeping.'

b. *In-doda nom-nqathe **zi**-se gadi-ni*
9-man and.3-carrot **SM8/10**-LOC garden-LOC
'The man and the carrot are in the garden.'

(19) a. *In-ja nen-doda **zi**-lele*
9-dog and.9-man **SM8/10**-sleep
'The dog and the man are sleeping.'

b. *Um-nqathe nen-doda **zi**-se gadi-ni*
3-carrot and.9-man **SM8/10**-LOC garden-LOC
'The carrot and the man are in the garden.'

Again, the order of the [+HUMAN] conjunct with respect to the [-HUMAN] conjunct seems to have no effect on the choice of subject marker, nor does the animacy of the non-human conjunct. The only commonality to be found in the 14 sentences which use SM8/10 instead of SM2, is that in each case [+HUMAN] conjunct comes from either NC7 or NC9. However, not all instances of mixed [±HUMAN] coordination involving a [+HUMAN] conjunct from NC7 or NC9 use SM8/10, as (16a) above shows. In fact, the same conjuncts in a different sentence do not necessarily take the same SM: Contrast (20a) with (18a), and (20b) with (19a) – all four sentences were produced by the same speaker.

(20) a. *In-doda nen-ja **ba**-ya e-doloph-ini*
9-man and.9-dog **SM2**-go LOC-town-LOC
'The man and the dog are going to town.'

b. *In-ja nen-doda **ba**-ya e-doloph-ini*
9-dog and.9-man **SM2**-go LOC-town-LOC
'The dog and the man are going to town.'

In sentences in (20) have exactly the same conjuncts as those in (18a) and (19a), and yet SM2 has been used in these instances, instead of SM8/10 as before. The erratic switching between the two different defaults is perhaps indicative of the fact that although the two

consultants allowed for mixed [±HUMAN] coordination, it may be somewhat dispreferred. Nonetheless, the overwhelming tendency is to use the [+HUMAN] default, possible indicating that the [+HUMAN] feature is cognitively more salient.

6.2.4.2 Both conjuncts are plural

One would expect there to be fewer inconsistencies when both conjuncts are plural; thus far the [±HUMAN] feature on the DPs has had no effect the choice of subject marker in these contexts – instead, closest conjunct agreement has been used consistently. For the most part, this continues to be the strategy of choice when the conjuncts differ in terms of their [±HUMAN] specification. In (21a), the conjunct closest to the verb carries a NC6 feature, thus triggering the use of SM6 on the verb. When the order of the conjuncts is reversed in (21b), resulting in the conjunct from NC2 being closer to the verb, the subject marker which carries the NC2 feature is used. The same trend of closest conjunct agreement can be observed in (22).

(21) a. *Aba-ntwana* *nama-dada* ***a**-se* *gadi-ni*
2-child and.6-duck **SM6**-LOC garden-LOC
'The children and the ducks are in the garden.'

b. *Ama-dada* *naba-ntwana* ***ba**-se* *gadi-ni*
6-duck and.2-child **SM2**-LOC garden-LOC
'The ducks and the children are in the garden.'

(22) a. *Ama-doda* *nemi-nqathe* ***i**-se* *gadi-ni*
6-man and.4-carrot **SM4**-LOC garden-LOC
'The men and the carrots are in the garden.'

b. *Emi-nqathe* *nama-doda* ***a**-se* *gadi-ni*
4-carrot and.6-man **SM6**-LOC garden-LOC
'The carrots and the men are in the garden.'

In the above examples, the order of the [+HUMAN] conjunct with respect to the [-HUMAN] conjunct does not affect the choice of the subject marker (contrast the (a) and (b) examples), nor does the animacy of the non-human conjunct (contrast (21) with (22)). However, CCA only accounts for 35 of the 50 coordinated mixed [±HUMAN] plural conjuncts.

Of the remaining 15 sentences, 12 of these employed furthest conjunct agreement, instead of closest conjunct agreement.[4] There is no rhyme or reason to be found in the sentences which resolve agreement clashes this way.

(23) a. *Aba-fundi* *nezi-khova* ***ba****-se* *hlati-ni*
2-student and.8-owl **SM2**-LOC forest-LOC
'The students and the owls are in the forest.'

b. *Emi-vundla* *naba-pheki* ***i****-se* *kitshi-ni*
4-hare and.2-cook **SM4**-LOC kitchen-LOC
'The hares and the cooks are in the kitchen.'

c. *Ama-hashe* *nezi-tyebi* ***a****-se* *bale-ni*
6-horse and.8-rich.man **SM6**-LOC field-LOC
'The horses and the rich men are in the field.'

d. *Izi-gulane* *nem-vundla* ***zi****-ya-dlala*
8-patient and.4-hare **SM8/10**-PRES-play
'The patients and the hares are playing.'

In these exceptions, the furthest conjunct agreement does not appear to be conditioned by the furthest conjunct being either [+HUMAN], as in (23a&d), or [-HUMAN], as in (23b&c). Nor does it appear to be conditioned by a particular noun class.

Although the coordination of mixed [±HUMAN] conjuncts does not yield an entirely consistent pattern, 71% of the examples can be accounted for by the following two generalizations:

1. When the mixed [±HUMAN] conjuncts are singular, SM2 is used.

2. When the mixed [±HUMAN] conjuncts are plural, closest agreement is used.

6.2.5 The generalizations

Putting aside the few exceptions discussed in 6.2.3 and 6.2.4, the agreement patterns of isiXhosa can be captured fairly neatly by the following generalizations:

[4]The remaining 3 sentences will not be discussed at length – they all involved coordination of *amadoda (men)* (NC6) and *izinja (dogs)* (NC10), and SM2, *ba*, was used in these three instances. However, it is interesting to note that *amadoda* is an anomalous noun: In the singular it falls into NC9, however, instead of going into NC10 in the plural, it instead goes into NC6. Thus perhaps a correlation can be drawn between this anomalous behaviour, and that of the NC11→NC6 nouns discussed in §6.2.3.

(24)
1. Coordinated singular DPs, regardless of whether they are NC-balanced or unbalanced, take default agreement: SM2, *ba*, if they are human, or SM8/10, *zi*, if they are nonhuman.
2. When the singular conjuncts differ in terms of their [±HUMAN] feature, SM2, *ba*, is the preferred subject marker.
3. Coordinated plural DPs take closest conjunct agreement.

6.3 An OT Analysis for the isiXhosa data

As isiXhosa uses some of the same agreement resolution strategies employed by Xitsonga and Sesotho, the same constraint set can be used to account for some of the isiXhosa data. However, none of the constraints make reference to the order of the conjuncts. An additional constraint is therefore needed in order to account for closest conjunct agreement; this constraint, AGREECLOSEST, requires that the subject marker must reflect the NC-feature of the conjunct linearly closest to the verb.

(25) Constraint definitions

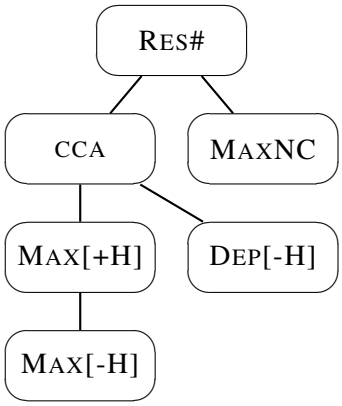
1. MAX[+HUMAN] (MAX[+H]): The semantic [+HUMAN] feature on each conjunct must be expressed morphologically by the SM.
2. MAX[-HUMAN] (MAX[-H]): The semantic [-HUMAN] feature on each conjunct must be expressed morphologically by the SM.
3. RESOLVENUMBER (RES#): Coordinated DPs in subject position trigger plural agreement on the predicate.
4. MAXNOUNCLASS (MAXNC): The noun class feature on *each conjunct* should correspond to the noun class feature on the SM.
5. DEP[-HUMAN] (DEP[-H]): The [-HUMAN] specification of the SM must be expressed semantically by *at least one* of the conjuncts.
6. AGREECLOSEST (CCA): The SM should parse the NC-feature on the conjunct which is linearly closest to it.[5]

[5]In the tableaux which follow, this constraint is abbreviated as 'CCA' for reasons of space. However, to disambiguate this from the phenomenon of closest conjunct agreement, the constraint name is written out in full as AGREECLOSEST when not in tableaux or ranking diagrams.

The first five constraints work in the same way here as they do in the analysis in chapters 4 and 5. If one/both of the conjuncts denote/s a human entity, MAX[+H] penalizes any SM which does not parse the [+HUMAN] feature.[6] MAX[-H] penalizes any SM which does not parse the [-HUMAN] feature if one/both of the conjuncts denote/s a non-human entity.[7] RES# penalizes any SM which does not carry a [PLURAL] feature when the subject consists of coordinated DPs. MAXNC requires the subject marker to carry the an NC-feature which agrees with all the NC-features in the input – it therefore penalizes any SM which does not parse the NC- feature of both conjuncts simultaneously. DEP[-H] penalizes any subject marker which carries a [-HUMAN] feature if there is no corresponding [-HUMAN] feature in the input. Finally, the new constraint, AGREECLOSEST, penalizes any SM which does not parse the NC-feature of the conjunct lineally closest to the verb.

These constraints interact with each other in such a way that the ranking in (26) best captures the generalizations in (24). MAXNC is not all that relevant to the isiXhosa analysis, as any of the generalizations which it accounts for can also be accounted for by AGREECLOSEST. However, as AGREECLOSEST assigns violations in a subset of the environments in which MAXNC does, these two constraints have been ranked equally. Likewise, DEP[-H] is not vital to this analysis either, as it does not account for any of the generalizations, but as it assigns violations to a subset of the candidates penalized by MAX[+H],[8] these two constraint have been ranked equally.

(26) Ranking of constraints in isiXhosa



[6] In isiXhosa, only SM1 and SM2 carry a [+HUMAN] feature.

[7] In isiXhosa, only SM7-11 carry a [-HUMAN] feature.

[8] When the conjuncts denote human entities, MAX[+H] penalizes all candidates which do not carry a [+HUMAN] feature, whilst DEP[-H] will only penalized the candidates which explicitly carry a [-HUMAN] feature.

The ranking of RES# above AGREECLOSEST accounts for the fact that closest conjunct agreement only happens when the conjuncts are plural. When they are singular, the conjunct closest to the verb will come from an NC with a [SINGULAR] feature, and thus the subject marker which carries the same NC feature will also bear the [SINGULAR] feature. This subject marker will therefore violate RES#, and as RES# dominates AGREECLOSEST, this SM will no longer be considered a viable candidate, as can be seen from (27).

(27) *Um-fundi ne-polisa* ***ba****-ya-baleka*
1-student and.5-policeman **SM2**-PRES-run
'The student and the policeman are running.'

1[+H] & 5[+H]	RES#	CCA	MAXNC	MAX[+H]	DEP[-H]	MAX[-H]
a. ☞ ba (2)		*	*			
b. a (6)		*	*	*!		
c. zi (8/10)		*	*	*!	*	
d. u (1)	*!	*	*			
e. li (5)	*!		*	*		

In (27) candidate (e) is the only candidate which satisfies AGREECLOSEST, but it, along with candidate (d), violates the more highly ranked RES#, and is therefore no longer available for consideration. Candidates (a)-(c) satisfy RES#, but all violate AGREECLOSEST and MAXNC equally. Thus the second stratum of constraints cannot distinguish between these three candidates. All three are therefore available for evaluation in the third stratum, where MAX[+H] picks the optimal candidate. As candidate (a) is the only remaining candidate which bears the [+HUMAN] feature, it is the only candidate which satisfies this constraint. It is therefore chosen as the optimal candidate.

By contrast, when the conjuncts are plural, the subject marker which agrees with the NC feature of the closest conjunct will also carry the [PLURAL] feature, therefore satisfying RES#, and making it available for evaluation by AGREECLOSEST, as can be seen in (28).

(28) *Aba-fundi nama-polisa* ***a****-ya-baleka*
2-student and.6-policeman **SM6**-PRES-run
'The students and the policemen are running.'

2[+H] & 6[+H]	RES#	CCA	MAXNC	MAX[+H]	DEP[-H]	MAX[-H]
a. ba (2)		*!	*			
b. ☞ a (6)			*	*		
c. zi (8/10)		*!	*	*	*	
d. u (1)	*!	*	*			
e. li (5)	*!	*	*	*		

Here, the two subject markers (candidates (d) and (e)) with a [SINGULAR] feature violate RES#. Only SM6 (candidate (b)) satisfies AGREECLOSEST, and all of the subject markers violate MAXNC. This leaves candidate (b) as the optimal candidate, with only one violation in the second stratum, as opposed to two, like candidates (a) and (c).

This ranking also accounts for the fact that NC-balanced conjuncts behave in the same way as NC-unbalanced conjuncts, using the appropriate [±HUMAN] default when the conjuncts are both singular, and closest conjunct agreement when they are both plural. When the NC-balanced conjuncts are both singular (as in (29)), a subject marker exists which would satisfy both AGREECLOSEST and MAXNC (candidate (d) in this case), but just as in (27), this SM will have a [SINGULAR] feature, and will therefore violate the more highly ranked RES#. Candidate (d) is therefore suboptimal.

(29) *Um-pu* *nom-bhobho* ***zi-se*** *rumi-ni*
3-gun and.3-pipe **SM8/10**-LOC room-LOC
'The gun and the pipe are in the room.'

3[-H] & 3[-H]	RES#	CCA	MAXNC	MAX[+H]	DEP[-H]	MAX[-H]
a. ba (2)		*	*			*!
b. i (4)		*	*			*!
c. ☞ zi (8/10)		*	*			
d. u (3)	*!					*

None of the remaining candidates satisfy AGREECLOSEST or MAXNC, and thus they must all be evaluated under MAX[+H]. As the conjuncts in (29) both denote non-human entities, there is no [+HUMAN] feature in the input which requires expression in the output – as a result, no violations are assigned under this constraint. Similarly, Dep[-H] does

not assign any violations, as candidates (a) and (b) do not carry a [-HUMAN] feature, and therefore do not make any assumptions about there being a [-HUMAN] feature in the input. By contrast, SM8/10 (candidate (c)) does carry a [-HUMAN] feature, and therefore requires this feature to be input if it is to satisfy DEP[-H]; as the conjuncts are nonhuman, this condition is met.

MAX[-H], on the other hand, starts by registering the presence of the [-HUMAN] feature in the input, and examines each candidate to see if it parses this feature. It assigns one violation each to candidates (a) and (b), as these subject markers do not reflect the [-HUMAN] feature of the conjuncts, while candidate (c) does. This renders (c) as the optimal candidate.

When the NC-balanced conjuncts are plural, the only significant way in which the tableau differs from a tableau for plural NC-unbalanced conjuncts (such as (28)), is that the optimal candidate will satisfy MAXNC in addition to AGREECLOSEST, as can be seen in (30).

(30) *Imi-pu* *nemi-bhobho* ***i**-se* *rumi-ni*
4-gun and.4-pipe **SM4**-LOC room-LOC
'The guns and the pipes are in the room.'

4[-H] & 4[-H]	RES#	CCA	MAXNC	MAX[+H]	DEP[-H]	MAX[-H]
a. ba (2)		*	*!			*
b. ☞ i (4)						*
c. zi (8/10)		*	*!			
d. u (3)	*!	*	*			*

In (30), the only constraints which have the option of picking the winning candidate are AGREECLOSEST, MAXNC, and MAX[-H]. As MAX[-H] favours a non-optimal candidate (c), it must crucially be ranked below AGREECLOSEST.

The dominance relationship between MAX[+H] and MAX[-H] can be established by looking at mixed [±HUMAN] coordination. As established in §6.2.4, when the mixed [±HUMAN] conjuncts are both singular, the predominant trend is for SM2 to be used. This means that the constraint requiring the output to faithfully express the [+HUMAN] feature present in the input (i.e. MAX[+H]) must dominate the constraint requiring the output to faithfully express the [-HUMAN] feature of a conjunct (i.e. MAX[-H]). This can be seen in (31).

(31) *In-tombi* *nesi-khova* ***ba****-se* *gadi-ni*
9-girl and.7-owl **SM2**-LOC garden-LOC
'The girl and the owl are in the garden.'

9[+H] & 7[-H]	RES#	CCA	MAXNC	MAX[+H]	DEP[-H]	MAX[-H]
a. ☞ ba (2)		*	*			*
b. i (4)		*	*	*!		*
c. zi (8/10)		*	*	*!		
d. i (9)	*!	*	*	*		*
e. si (7)	*!		*	*		

In (31), candidates (d) and (e) each agree with the NC-feature of one of the conjuncts, but as these SMs both have a [SINGULAR] feature, they both violate RES#. This therefore means that even though candidate (e) is the only candidate which satisfies AGREECLOSEST, it is not considered optimal. None of the remaining candidates can satisfy MAXNC, as there is no subject marker which can simultaneously parse the NC7 and NC9 features. MAX[+H] favours candidate (a), while MAX[-H] favours candidate (c); as (a) is in fact the optimal candidate, MAX[+H] must dominate MAX[-H].

As mentioned above, DEP[-H] does not play a significant role in the analysis of isiXhosa. The context in which it was important in Sesotho was to explain why morphological agreement is *not* used when the plural NC-balanced conjuncts are both [+HUMAN] and come from NC8 or NC10; in these contexts, DEP[-H] would penalize SM8/10, as it carries a [-HUMAN] feature with no corresponding [-HUMAN] feature in the input. However as morphological agreement is allowed in precisely these contexts in isiXhosa, it is possible to tell that DEP[-H] must be ranked lowly here – crucially, it must be dominated by AGREECLOSEST and MAXNC, because these two constraints favour the optimal candidate, whilst DEP[-H] penalizes it, as SM8/10 carries a [-HUMAN] feature which is not reflected in the input, as can be seen in (32).

(32) *Izi-tyebi* *nezi-bhanxa* ***zi****-ya-funda*
8-rich.person and.8-fool **SM8/10**-PRES-study
'The rich people and the fools are studying.'

8[+H] & 8[+H]	RES#	CCA	MAXNC	MAX[+H]	DEP[-H]	MAX[-H]
a. ba (2)		*!	*!			*
b. i (4)		*!	*!	*		*
c. ☞ zi (8/10)				*	*	
d. si (7)	*!	*	*	*	*	

In truth, the analysis would not be affected if DEP[-H] were ranked lower, either on par with, or below MAX[-H]; however, I have applied the principle of only demoting constraint X from one stratum to the next if there is evidence that one of the other constraints in the former stratum *must* dominate constraint X. In this case, there is no evidence which requires DEP[-H] to be dominated by either MAX[+H] or MAX[-H].

6.4 Predictions

The analysis, as it stands now, makes some interesting predictions about a set of data which had not yet been collected when the analysis was constructed, namely the coordination of number-unbalanced conjuncts (i.e. singular conjuncts being coordinated with plural conjuncts). In particular, the analysis makes two predictions. The first is that when the first conjunct is singular, and the second conjunct is plural, closest conjunct agreement will be used, as can be seen in (33) (note that no data accompanies the tableau, as at this point it is merely a prediction).

(33) Input: DP[SG, NC5, +H] & DP[PL, NC8, +H]

5[+H] & 8[+H]	RES#	CCA	MAXNC	MAX[+H]	DEP[-H]	MAX[-H]
a. ba (2)		*!	*			
b. a (6)		*!	*	*		
c. ☞ zi (8/10)			*	*	*	
d. li (5)	*!	*	*	*		

As the closest conjunct is plural, the subject marker which carries the agreeing NC feature (in this case, candidate (c)) also carries a PLURAL feature. This means that it does not violate RES#, and is therefore available to be chosen as the optimal candidate by AGREECLOSEST.

However, when the first conjunct is plural, and the second conjunct is singular, the prediction changes. The analysis now predicts that default agreement will be used – SM2 if both conjuncts denote humans (34), and SM8/10 if both conjuncts denote non-humans (35).

(34) Input: DP[PL, NC8, +H] & DP[SG, NC5, +H]

8[+H] & 5[+H]	RES#	CCA	MAXNC	MAX[+H]	DEP[-H]	MAX[-H]
a. ☞ ba (2)		*	*			
b. a (6)		*	*	*!		
c. zi (8/10)		*	*	*!	*!	
d. li (5)	*!		*	*		
e. o (1)	*!		*			

(35) Input: DP[PL, NC6, -H] & DP[SG, NC9, -H]

6[-H] & 9[-H]	RES#	CCA	MAXNC	MAX[+H]	DEP[-H]	MAX[-H]
a. ba (2)		*	*			*!
b. a (6)		*	*			*!
c. ☞ zi (8/10)		*	*			
d. i (9)	*!		*			

In both these tableaux, the subject marker which satisfies AGREECLOSEST (candidate (d) in each case), carries a [SINGULAR] feature, and so is disqualified from consideration by the most highly ranked constraint, RES#. It is therefore left to MAX[+H] and MAX[-H] to evaluate the remaining subject markers' faithfulness to the [±HUMAN] features in the input. In (34), the conjuncts both carry a [+HUMAN] feature; only candidates (a) and (e) parse the [+HUMAN] feature, and so only these two candidates satisfy MAX[+H]. However, candidate (e) also carries a singular feature, and so violates RES#; the analysis therefore predicts that SM2 will be the optimal candidate in this instance. By contrast, in (35) the conjuncts both denote non-human entities, so none of the candidates violate MAX[+H] (or DEP[-H]). MAX[-H] then penalizes candidates (a) and (b), as neither parses the [-HUMAN] feature. The analysis therefore predicts that SM8/10 will be the optimal candidate in instances such as this.

In order to check these predictions, I scheduled additional interviews with the consultants, and set up new questionnaires, designed specifically elicit to sentences coordinating

number-unbalanced conjuncts.[9]

6.4.1 Checking the Predictions

When the first conjunct is singular, and the second conjunct is plural, agreement works in exactly the way the analysis above predicted it would: The SM agrees with the conjunct closest to the verb, as illustrated in (36).

(36) a. *I-gwetha nezi-denge **zi**-ya-lwa*
5-lawyer and.8-fool **SM8/10**-PRES-fight
'The lawyer and the fools are fighting.'

b. *Isi-denge nama-gwetha **a**-ya-lwa*
7-fool and.6-lawyer **SM6**-PRES-fight
'The fool and the lawyers are fighting.'

c. *In-komo nama-hashe **a**-tya ingca*
9-cow and.6-horse **SM6**-eat grass
'The cow and the horses are eating grass.'

d. *I-hashe neen-komo **zi**-tya ingca*
5-horse and.10-cow **SM8/10**-eat grass
'The horse and the cows are eating grass.'

In (36a), the conjunct closest to the verb comes from NC8, and SM8(/10) is used. Similarly, in (36b&c), the second conjunct comes from NC6, and SM6 is used on the verb. Finally, in (36d) the second conjunct comes from NC10, and again, this NC-feature shows up on the SM. Thus the analysis's first prediction is correct.

The second prediction – that default agreement will be used when the first conjunct is plural, and the second singular – does not, however, hold up to scrutiny. As the data in (37) show, default agreement is *not* used in this context.

(37) a. *Ama-gwetha nesi-denge **a**-ya-lwa*
6-lawyer and.7-fool **SM6**-PRES-fight
'The lawyers and the fool are fighting.'

b. *Izi-denge ne-gwetha **zi**-ya-lwa*
8-fool and.5-lawyer **SM8/10**-PRES-fight
'The fools and the lawyer are fighting.'

[9]It was also at this time that I started collecting data for number-unbalanced conjuncts for Sesotho and Xitsonga.

c. *Ama-hashe nen-komo **a**-tya ingca*
6-horse and.9-cow **SM6**-eat grass
'The horses and the cow are eating grass.'

d. *Iin-komo ne-hashe **zi**-tya ingca*
10-cow and.5-horse **SM8/10**-eat grass
'The cows and the horse are eating grass.'

Contrary to the predictions made by (34), (37a) and (37b) do not use SM2, nor does (37c) use SM8/10, as predicted by (35). And while it may at first glance seem as though default agreement is being employed in (37d), this is not what is actually happening. Instead, in each of the above sentences, the subject marker is determined by the noun class of the *first* conjunct.

Thus partial agreement is always used when one of the conjuncts is plural. If the closest conjunct is plural, the partial agreement takes the form of closest conjunct agreement. If the closest conjunct is singular, but the furthest conjunct is plural, then the partial agreement takes the form of furthest conjunct agreement.

6.4.2 Tweaking the Analysis

The constraints set in (25) on page 126 make no provision for furthest conjunct agreement. The data in (37) suggests the need for a new constraint which can licence agreement with the furthest conjunct. This constraint is DEPNC, which requires that for any NC-feature present on the subject marker, there must be a corresponding NC-feature somewhere in the input.

(38) DEPNOUNCLASS (DEPNC): The noun class feature on the subject marker should reflect the noun class feature of at least one of the conjuncts.

This means that if the conjuncts are NC-unbalanced, there will be two candidates which will satisfy DEPNC – one which agrees with the first conjunct, and one which agrees with the second. This is in contrast to AGREECLOSEST, which will only ever be satisfied by one candidate.

DEPNC must be ranked below RES#, otherwise it would favour subject markers with a [SINGULAR] feature when both conjuncts are singular (see discussion of (43) on page 138). Furthermore, as DEPNC prevents default agreement from being used in contexts such as

(37), it must clearly dominate the two constraints which favour default agreement, namely MAX[+H] and MAX[-H]. However, no dominance relationship can be established between DEPNC and AGREECLOSEST; this is because these two constraints are in a stringency relation, and it is not possible to derive a direct ranking argument when two constraints are in this sort of superset-subset relationship (McCarthy 2008:32). Thus these two constraints have been left unranked with respect to each other.

The addition of this constraint now explains how furthest conjunct agreement can be licensed in isiXhosa. As can be seen in (39), the subject markers which agree with the first conjunct, candidate (c), and the second conjunct, candidate (d), both satisfy DEPNC, while all other candidates violate it. However, as (d) violates the more highly ranked RES#, (c) is rendered as the optimal candidate.

(39) *Izi-denge* *ne-gwetha* ***zi**-ya-lwa*
8-fool and.5-lawyer SM**8/10**-PRES-fight
'The fools and the lawyer are fighting.'

8[+H] & 5[+H]	RES#	DEPNC	CCA	MAXNC	MAX[+H]	DEP[-H]	MAX[-H]
a. ba (2)		∗!	∗	∗			
b. a (6)		∗!	∗	∗	∗		
c. ☞ zi (8/10)			∗	∗	∗	∗	
d. li (5)	∗!			∗	∗		

Similarly, in (40) DEPNC assigns violations to candidates (a) and (c), as there is no NC2 or NC8/10 feature in the input; candidates (b) and (d), however, satisfy this constraint, as the conjuncts are from NC6 and NC9, respectively. However, while AGREECLOSEST favours (d) over (b), (d) is in fact not a viable candidate, as it violates RES#. SM6 is therefore considered the optimal candidate, even though it does not agree with the closest conjunct.

(40) *Ama-hashe* *nen-komo* ***a**-tya* *ingca*
6-horse and.9-cow **6**-eat grass
'The horse and the cows are eating grass.'

6[-H] & 9[-H]	RES#	DEPNC	CCA	MAXNC	MAX[+H]	DEP[-H]	MAX[-H]
a. ba (2)		*!	*	*			*
b. ☞ a (6)			*	*			*
c. zi (8/10)		*!	*	*			
d. i (9)	*!			*			

The addition of DEPNC does not affect the optimal candidate in other contexts. When the number-unbalanced conjuncts are in the order of singular first, plural second, RES# will again remove the singular SM from consideration early on (such as (d) in (41)). The SM which agrees with the NC of the plural conjunct (candidate (c)) will then satisfy both DEPNC and AGREECLOSEST, making it the optimal candidate.

(41) *I-gwetha nezi-denge **zi**-ya-lwa*
5-lawyer and.8-fool **SM8/10**-PRES-fight
'The lawyer and the fools are fighting.'

5[+H] & 8[+H]	RES#	DEPNC	CCA	MAXNC	MAX[+H]	DEP[-H]	MAX[-H]
a. ba (2)		*!	*!	*			
b. a (6)		*!	*!	*	*		
c. ☞ zi (8/10)				*	*	*	
d. li (5)	*!		*	*	*		

The generalizations in (24) on page 126 still hold true with this tweaked version of the analysis. When both conjuncts are singular, the candidates which would satisfy DEPNC will again be penalized by RES# (candidates (d) and (e) in (42)). Candidates (a), (b), and (c) all accrue equal violations under DEPNC, AGREECLOSEST, and MAXNC, and no violations under MAX[+H] and DEP[-H]. MAX[-H] favours candidate (c), as it is the only one of the remaining SMs which parses the [-HUMAN] feature in the input. Thus the addition of DEPNC does not prohibit default agreement from happening when both conjuncts are singular.

(42) *I-khaphetsu nes-onka **zi**-se tafile-ni*
5-cabbage and.7-bread **SM8/10**-LOC table-LOC
'The cabbage and the bread are on the table.'

5[-H] & 7[-H]	RES#	DEPNC	CCA	MAXNC	MAX[+H]	DEP[-H]	MAX[-H]
a. ba (2)		*	*	*			*!
b. a (6)		*	*	*			*!
c. ☞ zi (8/10)		*	*	*			
d. li (5)	*!		*	*			*
e. si (7)	*!			*			

This example illustrates why DEPNC must be ranked *below* RES#: If DEPNC dominated RES#, as it does in (43), it would penalize candidates (a)-(c), making them unavailable for evaluation under subsequent constraints; only candidates (d) and (e) would be left as possible optima. RES# would not distinguish between these two candidates, penalizing them equally. It would then be up to AGREECLOSEST to distinguish between the two; this constraint would pick (e), as it agrees with the NC feature of the closest conjunct. This is clearly not the optimal candidate – therefore RES# *must* dominate DEPNC.

(43)

5[-H] & 7[-H]	DEPNC	RES#	CCA	MAXNC	MAX[+H]	DEP[-H]	MAX[-H]
a. ba (2)	*!		*	*			*
b. a (6)	*!		*	*			*
c. ☞ zi (8/10)	*!		*	*			
d. li (5)		*	*!	*			*
e. ☛ si (7)		*		*			

The addition of DEPNC does not cause problems when both conjuncts are plural either, as all of the plural SMs satisfy RES#, and are thus available for evaluation under DEPNC and AGREECLOSEST. Of the plural SMs, only those which agree with at least one of the conjuncts will satisfy DEPNC (candidates (b) and (c) in (44)), and only the candidate which specifically agrees with the closest conjunct will satisfy AGREECLOSEST (candidate (b) here); it will therefore be the candidate which satisfies AGREECLOSEST that has no violations at this point, rendering it the optimal candidate.

(44) *Izi-onka nama-khaphetsu **a**-se tafile-ni*
8-bread and.6-cabbage **SM6**-LOC table-LOC
'The bread [multiple loaves] and the cabbages are on the table.'

8[-H] & 6[-H]	RES#	DEPNC	CCA	MAXNC	MAX[+H]	DEP[-H]	MAX[-H]
a. ba (2)		*!	*!	*			*
b. ☞ a (6)				*			*
c. zi (8/10)			*!	*			
d. li (5)	*!	*	*	*			*
e. si (7)	*!	*	*	*			

Tableaux (42) and (44) thus show that the addition of DEPNC poses no problems for the analysis of number-balanced conjuncts.

6.5 Concluding remarks

This chapter has shown how the constraints proposed in the previous two chapters need to be supplemented with two additional constraints (AGREECLOSEST and DEPNC) in order to account for the isiXhosa data. The dominance of MAXNC, AGREECLOSEST and DEPNC over MAX[+H], MAX[-H], and DEP[-H] accounts for the use of syntactic agreement (rather than default agreement) with plural conjuncts. In turn, RES#'s dominance over the three NC-related constraints accounts for the prohibition of syntactic agreement with singular conjuncts, thereby necessitating the use of default agreement instead. Finally, MAX[+H]'s dominance over MAX[-H] explains the preference for parsing the [+HUMAN] feature over the [-HUMAN] feature in contexts where both are present in the input.

The next chapter explores the factorial typology which can be generated through various orderings of these constraints, and examines some interesting trends which emerge.

Chapter 7

A Typology of agreement resolution strategies

In the previous chapters I have proposed a total of seven constraints, and shown how three different rankings of these constraints can account for the different sets of resolution rules in Xitsonga, isiXhosa, and Sesotho. But these three rankings are merely the tip of the proverbial iceberg: With seven constraints, there are 5040 different possible linear rankings. Using OTWorkplace (Prince et al. 2014), I generated a factorial typology using a carefully selected set of inputs and candidates, and ascertained that all these different rankings only give rise to 159 distinct sets of outputs.

However, if one assumes that RESOLVENUMBER is universally undominated, resulting in plural agreement with a subject &P, one can generate a sub-typology by restricting the output to 'languages' in which RES# is undominated. This leaves only 47 language types. A description of what each of these typologies would entail is given in appendix A.

§7.1 describes the process used to generate the factorial typology, the inputs and candidates used, and the reasoning behind the choice of these inputs and candidates. The rest of the chapter then highlights a few of the implications of the typology, especially two predictions of universal trends (§7.2), and some unexpected effects which result from certain constellations of constraints (§7.3). Finally, §7.4 uses the Chichewa data from Corbett and Mtenje (1987) to show how the factorial typology can be applied to other studies on agreement with subject &Ps in Bantu.

7.1 The basis for the typology

In generating the factorial typology, it is very important to ensure that the various inputs and candidates used are sufficient to capture all the variations possible. This section briefly describes types of inputs and candidates used, and explores the reasoning behind these choices.

The factorial typology was generated in OTWorkplace (Prince et al. 2014), using a subset of the violation tableaux from the isiXhosa data. However, as this data set consisted of over 500 inputs, many of which employed the same resolution strategy, I pared down the number of inputs by working out that these inputs could be divided into groups of inputs which will always trigger the same type of agreement strategy. There are 22 groupings of inputs, each of which could feasibly trigger different agreement resolution strategies. One input from each of these groups was then used to generate the typology. These groups are listed in (1).[1]

(1) Groups of inputs for the full factorial typology

Number Balanced, [±HUMAN] Balanced Conjuncts

Both Conjuncts [+HUMAN]

1. NC-balanced singular conjuncts
2. NC-unbalanced singular conjuncts
3. NC-balanced plural conjuncts, *not* NC8
4. NC-balanced plural conjuncts, *specifically* NC8
5. NC-unbalanced plural conjuncts, where *neither* conjunct is NC2, and second conjunct is *not* NC8
6. NC-unbalanced plural conjuncts, where first conjunct *is* NC2
7. NC-unbalanced plural conjuncts, where second conjunct *is* NC8

Both Conjuncts [-HUMAN]

8. NC-balanced singular conjuncts
9. NC-unbalanced singular conjuncts
10. NC-balanced plural conjuncts
11. NC-unbalanced plural conjuncts, where *neither* conjunct is NC8
12. NC-unbalanced plural conjuncts, where first conjunct *is* NC8

[1] In this list, as well as the rest of this chapter and appendix A, NC2 is used to refer to the noun class associated with the [+HUMAN] default subject marker, and NC8 is used to refer to the noun class associated with the [-HUMAN] default subject marker.

Mixed [±HUMAN] Conjuncts

13. NC-balanced singular conjuncts
14. NC-unbalanced singular conjuncts
15. NC-balanced plural conjuncts
16. NC-unbalanced plural conjuncts, where neither conjunct is NC2 or NC8
17. NC-unbalanced plural conjuncts, where first conjunct is NC2
18. NC-unbalanced plural conjuncts, where first conjunct is NC8

Mixed Number Conjuncts

19. Both conjuncts human: PL & SG
20. Both conjuncts human: SG & PL
21. Both conjuncts nonhuman: PL & SG
22. Both conjuncts nonhuman: SG & PL

The reasoning behind this choice of inputs is that each group of inputs has the potential to behave differently. Chapters 4, 5 and 6 have shown that the resolution strategies used will vary according to whether the conjuncts are NC-balanced/unbalanced, [±HUMAN]-balanced/unbalanced and number-balanced/unbalanced. Additionally, the singular/plural and human/nonhuman distinctions can affect the choice of resolution strategy, as may the order of the conjuncts when they are number-unbalanced.

Furthermore, using a conjunct which comes from the noun class associated with the [±HUMAN] default subject markers can obscure certain information. For example, for the data "NC6[+H] & NC2[+H] → SM2", it is not clear whether closest conjunct agreement or default agreement is being used; likewise, for "NC8[-H] & NC4[-H] → SM8", it is not clear whether it is default agreement, or furthest conjunct agreement being licensed by a ranking such as DEPNC»MAX[-H]»CCA (see §7.2.2). For this reason, it was important to include coordinated pairs which do not include the noun classes associated with the default subject markers (this accounts for inputs 5, 11, and 16), and then to also include pairs where one of the conjuncts (preferably the first) does come from those noun classes (inputs 6, 12, 17, and 18). Additionally, as §7.3.4 will show, NC8 [+HUMAN] conjuncts sometimes behave differently from other [+HUMAN] conjuncts; it is therefore important to include coordination involving [+HUMAN] NC8 conjuncts as separate groups which need to be represented in the typology (inputs 4 and 7).

The candidate set used to generate the typology included, *at a minimum*, the following candidates:

a. The [+HUMAN] default SM (represented by SM2)

b. the [-HUMAN] default SM (represented by SM8)

c. The SM which agreed with NC-feature of the first conjunct

d. The SM which agreed with NC-feature of the second conjunct

Additionally, if the conjuncts were singular, the candidate set also included the SMs of the plural noun classes those conjuncts would fall into. Furthermore, if the input included a [+HUMAN] feature, SM1 was included in the candidate set, on the assumption that SM1 carries the same [+HUMAN] feature as SM2; likewise, SM7 was included if the input included a [-HUMAN] feature. However, in the full factorial typology SM1 and SM7 are almost always harmonically bounded, unless they happen to meet the criteria of (c) or (d) above; in the sub-typology, which only includes languages in which RES# is undominated, SM1 and SM7 are *always* harmonically bounded.

The remainder of this chapter explores some of the facets of the typology which is based on the inputs and candidate sets described above.

7.2 Predictions of universality

As already mentioned, this chapter focuses solely on the sub-typology in which RES# is undominated. By systematically going through each of the 47 language types in this sub-typology, one can pick up trends which appear to be entirely uniform in their manifestation. This section examines two such trends, showing that the typology predicts that 1) default agreement will always be used when both conjuncts are singular (§7.2.1), and 2) furthest conjunct agreement cannot occur in a grammar unless closest conjunct agreement is also used in the same grammar (§7.2.2).

7.2.1 Agreement with singular conjuncts

The first prediction made is a very concrete one: When RES# is undominated, the coordination of two **singular** conjuncts will always give rise to default [+HUMAN] or [-HUMAN] agreement, depending on the [±HUMAN] feature on the conjuncts. This is because, with the

exception of RES#, all of the constraints demand faithfulness to either the NC-features in the input, or the [±HUMAN] features in the input – yet any subject marker which is faithful to the NC-feature of a singular conjunct would violate RES#.[2] As agreement with the NC-features of the conjuncts is no longer possible, this leaves agreement with the [±HUMAN] features of the conjuncts as the only option.

This can be seen by looking at tableaux (2) and (3), which show how various candidates fare against the constraints when both conjuncts are singular. To abstract away from any given language – and therefore any assumptions about specific subject markers and rankings of constraints – I've used features as the inputs and candidate sets, and left all the constraints except RES# unranked. In both (2) and (3), the input is NC-balanced, as both singular conjuncts carry the same NC feature (X). The only difference between the the two tableaux is that in (2), the conjuncts both denote humans, while in (3), they both denote nonhumans.

(2)

[NCX, +H, SG] & [NCX, +H, SG]	RES#	DEPNC	MAXNC	CCA	DEP[-H]	MAX[-H]	MAX[+H]
a. ☞ SM: [NCY, +H, PL]		*	*	*			
b. SM: [NCZ, -H, PL]		*	*	*	*!		*!
c. SM: [NCX, ØH, SG]	*!						*
d. SM: [NCX2, ØH, PL]		*	*	*			*!

(3)

[NCX, -H, SG] & [NCX, -H, SG]	RES#	DEPNC	MAXNC	CCA	DEP[-H]	MAX[-H]	MAX[+H]
a. SM: [NCY, +H, PL]		*	*	*		*!	
b. ☞ SM: [NCZ, -H, PL]		*	*	*			
c. SM: [NCX, ØH, SG]	*!					*	
d. SM: [NCX2, ØH, PL]		*	*	*		*!	

Candidate (a) represents the [+HUMAN] default (such as SM2), while candidate (b) represents the [-HUMAN] default (such as SM8). Candidate (c) represents the subject marker which agrees with the noun class of both conjuncts, and candidate (d) represents the subject marker of the plural noun class typically associated with the singular noun class X.[3]

In both cases, candidate (c) is the only SM which does not violate DEPNC, MAXNC or CCA, but it is not a viable candidate, as it violates the highest ranked constraint, RES#. Even

[2] As the conjuncts are singular, the subject marker which parses the NC-feature would also have to have a [SINGULAR] feature, which conflicts with RES#'s requirement for plural agreement.

[3] i.e.: If NCX was NC3, NCX2 would be NC4, and if NCX was NC5, then NCX2 would be NC6, etc.

though candidate (d) represents the SM of the plural NC typically associated with NCX, it does not carry the NC-features required to satisfy any of the noun class related constraints, and thus it fares as badly as the two defaults do on these constraints. Only the constraints which address faithfulness to the [±HUMAN] feature can distinguish between candidates (a), (b) and (d). When both conjuncts are [+HUMAN], as in (2), MAX[+H] will favour the [+HUMAN] default, candidate (a). When both conjuncts are [-HUMAN], as in (3), MAX[-H] will favour the [-HUMAN] default, candidate (b).

Importantly, the theory predicts that two singular conjuncts of a particular noun class will never take the agreement of the plural noun class of that gender – for example, two conjuncts from NC5 will not take NC6 agreement, as this subject marker will be harmonically bounded: No ranking of the constraints will favour that candidate. The exception to this would be when the the singular conjuncts happen to be in the same gender as the required [±HUMAN] default: Thus two nonhuman conjuncts from NC7 would take NC8 agreement, and two human conjuncts from NC1 would take NC2 agreement. This, however, is still the result of resolution to the [±HUMAN] default, rather than resolution to the plural associated with a particular gender.

While this accords largely with findings in other languages, such as Chichewa (Corbett and Mtenje 1987), Setswana (Roberts and Wolentis 1974) and Chathu [4] (Mous and Mreta 2004), it is not entirely unproblematic. Marten (2000) cites Krifka (1995) and Schadeburg (1992) as saying that in Kiswahili singular conjuncts from the same two-class gender will take the agreement associated with the plural of that gender. The only example of this which is given, however, involves the coordination of two NC1 [+HUMAN] conjuncts taking NC2 agreement, and this could be an example of [+HUMAN] default agreement. Nonetheless, Kiswahili is not the only language for which there have been reports of this strategy.

Simango (2012) provides the following example, in (4), of coordination involving singular NC-balanced conjuncts in ciNsenga, where both conjuncts carry NC12 and [-HUMAN] features. As the analysis predicts, agreement is resolved to the [-HUMAN] default, SM8, rather than the plural subject marker of gender 12/13.

[4]There is one exception in Chathu: When both conjuncts are non-human, and the closest conjunct is from NC9, then SM10 may be used. This is not a typical resolution strategy, however, as Mous and Mreta (2004) emphasize that this method of agreement *only* happens when the closest conjunct is from NC9.

(4) *ka-mbwili na ka-temo* ***v(/∗tw)-a-sowa***
12-hoe and 12-axe **SM8(∗SM13)**-PAST-miss
'The hoe and the axe are missing.' (ciNsenga; Simango 2012:179)

However, Simango (personal communication) also provides the following example, noting that when the singular NC-balanced conjuncts both denote human entities, then the plural SM from that particular gender is more likely to be used, as illustrated in (5).

(5) *N-thombi na m-fumu zu-lawizh-ana*
9-girl and 9-chief **SM10**-talk-RECIP
'The girl and the chief are talking to each other.' (ciNsenga)

To accommodate languages which allow for in-gender resolution the OT analysis would have to be tweaked, perhaps through the addition of a constraint such as MAXGENDER, which would favour any subject marker which expressed the two-class gender of both conjuncts. However, I leave this to future research.

7.2.2 Partial agreement

An interesting implication of the analysis proposed in this thesis is the predictions it makes about partial agreement (i.e. agreement with only one conjunct). Whilst the analysis allows for both closest conjunct agreement (CCA) and furthest conjunct agreement (FCA), it predicts that FCA will *never* be a preferred strategy for agreement resolution. In short: 1) FCA *cannot* occur in a grammar which *does not* also allow CCA, but 2) CCA *can* occur occur in a grammar which does not use FCA. Thus, of the languages which allow partial agreement, the partial agreement can either manifest as a) both CCA and FCA, or b) only CCA, but never as c) only FCA.

This asymmetry is the result of the stringency relationship between AGREECLOSEST[5] and DEPNC. These are the two constraints which are responsible for the effects of partial agreement. DEPNC favours any candidate SM which agrees with the NC-feature on either conjunct, regardless of whether it is the closest conjunct or the furthest conjunct. AGREECLOSEST, on the other hand, only favours one candidate SM, namely the one which

[5]Please note that while this constraint has been abbreviated as CCA in the tableaux in this thesis, in the text of the thesis (§7.2.2) I refer to the constraint in its unabbreviated form, AGREECLOSEST, to disambiguate it from the phenomenon of closest conjunct agreement (CCA). Thus when the abbreviation CCA used in-text, it refers to closest conjunct agreement.

agrees with the NC-feature of the conjuncts linearly closest to the verb. When considered in isolation, these two constraints show a noticeable bias towards closest conjunct agreement: If AGREECLOSEST dominates DEPNC, they will conspire to produce CCA, as can be seen in (6); likewise if DEPNC dominates AGREECLOSEST, CCA will again prevail over FCA, as (7) shows.

(6)

NC8 & NC6	CCA	DEPNC
a. ☞ SM6		
b. SM8	∗!	

(7)

NC8 & NC6	DEPNC	CCA
a. ☞ SM6		
b. SM8		∗!

This two tableaux show that if partial agreement is allowed in a language, the CCA will be chosen over FCA. FCA will only be used when the closest conjunct violates some higher-ranked constraint, which the furthest conjunct satisfies. One obvious example of this is when the closest conjunct is singular, and the furthest conjunct is plural – in such instances, the SM which agrees with the closest conjunct would violate RES#, leaving agreement with the furthest conjunct as more optimal than agreement with the closest conjunct (8).[6]

(8)

NC4 & NC5	RES#	CCA	DEPNC
a. ☞ SM4		∗	
b. SM5	∗!		
c. SM6		∗	∗!

In (8), candidate (c) does not agree with the NC-feature of either conjunct, and thus violates both AGREECLOSEST and DEPNC. In contrast, (b) satisfies both AGREECLOSEST and DEPNC, yet it is not optimal – even though it agrees with the NC-feature of the closest conjunct – as it violates RES#. It is in precisely this instance which FCA may surface: Even

[6]Note that this is specifically for languages which **do** license FCA – those which do not license it would have DEPNC in the bottom stratum of constraints, not dominating any other constraint.

though candidate (a) violates AGREECLOSEST, it is the only candidate which satisfies both RES# and DEPNC.

In the above example, FCA is licensed by the fact that the closest conjunct seems to "repel" CCA – that is to say, it has a feature (in this case [SINGULAR]) which actively prohibits agreement with that conjunct. Another example where the closest conjunct could be considered as "repelling" CCA would be in languages in which DEP[-H] is highly ranked. In such languages, the SM which is used as the [-HUMAN] default may not be used when both conjuncts denote humans. Thus even if such a language licenses CCA, this agreement strategy would not be allowed when the closest conjunct carries [+HUMAN, NC8] features,[7] as can be seen in (9).

(9)

6[+H] & 8[+H]	DEP[−H]	DEPNC	CCA	MAX[+H]
a. SM2		∗!	∗	
b. ☞ SM6			∗	∗
c. SM8	∗!			∗

In this example the relative ranking of AGREECLOSEST and DEPNC does not affect the outcome; for this reason I have left them unranked relative to each other. Here, candidate (c), which satisfies AGREECLOSEST, is not allowed as it violates the more highly ranked DEP[-H]. This rules out the use of CCA as an agreement strategy in this instance, and the speaker is left with a choice between two other strategies: Either use [+HUMAN] default SM, candidate (a), or agree with the NC-feature of the furthest conjunct, candidate (b). If DEPNC dominates MAX[+H], as it does in (9), FCA will be the strategy chosen.

The two examples above show FCA being licensed by a feature on the closest conjunct "repelling" agreement with that conjunct. Another way of licensing FCA is for the furthest conjunct to actively "attract" agreement, as is illustrated below.

The tableaux in (10) show the relative rankings of a subset of constraints in a hypothetical language. In this language, DEPNC » MAX[+H] » CCA, which typically results in closest conjunct agreement, as can be seen in (10a) where two human conjuncts, from NC6 and NC8 respectively, have been conjoined, and SM8 appears on the verb. Although both conjuncts are human, and MAX[+H] dominates AGREECLOSEST, the [+HUMAN] default

[7] I assume here that SM8 is the [-HUMAN] default for this hypothetical language

(candidate (a)) is *not* used, and CCA *is* used. This is because DEPNC dominates MAX[+H], and candidate (a) violates DEPNC, making it unavailable for evaluation under MAX[+H]. Thus, although candidates (b) and (c) violate MAX[+H], there is no other available candidate which satisfies this constraint, resulting in (b) and (c) being available for evaluation by AGREECLOSEST. AGREECLOSEST then favours candidate (c), as the conjunct closest to the verb has a NC8 feature.

(10) a.

6[+H] & 8[+H]	DEPNC	MAX[+H]	CCA
a. SM2	∗!		∗
b. SM6		∗	∗!
c. ☞ SM8		∗	

b.

2[+H] & 8[+H]	DEPNC	MAX[+H]	CCA
a. ☞ SM2			∗
b. SM6	∗!	∗	∗
c. SM8		∗!	

In tableau (10b), however, where two human conjuncts from NC2 and NC6, respectively, have been conjoined, the same ranking yields a different result: Here, CCA is *not* used, and instead SM2 appears on the verb. On the surface, it might seem as though default [+HUMAN] agreement is being used. While this is *partially* true, it is not the whole truth: If this were the case, then why not use SM2 in tableau (10a)? In both cases, after all, the conjuncts are both [+HUMAN]. The answer to this lies in the fact that DEPNC and AGREECLOSEST are crucially ranked differently with respect to MAX[+H].

In (10a), SM2 was eliminated from consideration by DEPNC, as neither conjunct was from NC2. By contrast, in (10b) one of the conjuncts has a NC2 feature, and thus candidate (a) satisfies DEPNC, making it available for evaluation under MAX[+H], along with candidate (c), which agrees with the NC-feature of the closest conjunct. As MAX[+H] dominates AGREECLOSEST, the candidate which parses the [+HUMAN] feature on the conjuncts (i.e. SM2) is preferred over the candidate which parses the NC-feature of the closest conjunct (i.e. SM8, in this example). Thus, in this instance, FCA is licensed, as the furthest conjunct satisfies two requirements: 1) it agrees with the NC feature of at least on of the conjuncts,

and 2) it expresses the requisite [+HUMAN] feature.

Thus candidate (a) attracts agreement with more 'force' than candidate (c), despite the fact that (c) satisfies the standard resolution strategy employed by plural conjuncts in the language.

An additional example of FCA being licensed by the furthest conjunct attracting agreement in a language which normally employs CCA is a ranking very similar to that in (10) above, but with MAX[-H] instead of MAX[+H]: DEPNC»MAX[-H]»AGREECLOSEST. Under this ranking, just as in (10a), CCA will be the normal resolution strategy for agreement with plural conjuncts, as illustrated in (11a). However, when both conjuncts are [-HUMAN], and the furthest conjunct carries the NC-feature which is associated with the [-HUMAN] default marker, then the furthest conjunct will attract agreement more strongly than the closest conjunct, as illustrated in (11b).

(11) a.

6[-H] & 4[-H]	DEPNC	MAX[-H]	CCA
a. ☞ SM4		*!	
b. SM6		*!	*
c. SM8	*!		*!

b.

8[-H] & 4[-H]	DEPNC	MAX[–H]	CCA
a. SM4		*!	
b. SM6	*!	*!	*
c. ☞ SM8			*

This section has therefore shown that when partial agreement is allowed in a grammar, the preference will be biased towards closest conjunct agreement, rather than furthest conjunct agreement. FCA will only be licensed when asymmetries in either [NUMBER] or [±HUMAN] features trump the asymmetry of distance from the verb.

7.3 Surface effects of specific constellations of constraints

While systematically going through the 47 language types in appendix A, and surmising what descriptive generalizations are predicted for each language, some patterns emerged as

constant under certain constellations of constraints. These patterns are interesting to note, as they reveal that certain descriptive generalizations are always the result of a specific configuration of two or three constraints. This section briefly describes four of these configurations and their effects. §7.3.1 looks at the interaction of noun class oriented constraints with the [±HUMAN] constraints, while §7.3.2 discusses the configurations of constraints involved in the resolution of mixed [±HUMAN] conjuncts. §7.3.3 then looks at the configurations of constraints which determine the resolution strategies used with number-unbalanced conjuncts. Finally, §7.3.4 discusses the interaction of Dep[-H] with the NC oriented constraints.

7.3.1 The interaction of NC constraints with [±HUMAN] constraints

One observation is that plural[8] [+HUMAN] and [-HUMAN] conjuncts do not always utilise the same agreement resolution strategies, as was seen in variety two of Xitsonga (see chapter 4). This can be accounted for by the fact that anytime MAX[+H] and MAX[-H] are ranked differently in respect to any NC constraint (i.e. MAXNC, DEPNC, or AGREECLOSEST) human and non-human conjuncts will behave differently. If MAX[αH][9] is ranked above a particular NC constraint, and MAX[$\neg\alpha$H] is ranked below it, then [$\neg\alpha$HUMAN] conjuncts will be affected by the NC constraint, whilst [αHUMAN] conjuncts will not be.

This can give rise to some unexpected agreement resolution patterns. For the ranking MAX[αH]»MAXNC»MAX[$\neg\alpha$H], conjuncts with the [$\neg\alpha$HUMAN] feature would use morphological resolution when they are NC-balanced, whilst conjuncts with the [αHUMAN] feature would instead use default agreement. Likewise, if the intervening NC constraint is AGREECLOSEST or DEPNC, then partial agreement could be used with [$\neg\alpha$HUMAN] conjuncts, but not with [αHUMAN] conjuncts.

7.3.2 Resolution strategies for mixed [±HUMAN] conjuncts

As noted in §7.2.1, when two singular conjuncts are conjoined, the resulting SM will always be either the [+HUMAN] or [-HUMAN] default, depending on the [±HUMAN] specification of the conjuncts. The use of the [±HUMAN] default still holds true when the two singular conjuncts differ in their [±HUMAN] features, but the choice of default is not automatically

[8] Singular conjuncts always take default agreement, as discussed in §7.2.1.

[9] Where α represents either a positive or negative sign.

apparent, as both the [+HUMAN] or the [-HUMAN] default SM would satisfy the demands of faithfulness to the [±HUMAN] features in the input. Thus what ultimately decides which default will be used is the ranking of MAX[+H] and MAX[-H] relative to one-another. If MAX[+H] dominates MAX[-H], then the [+HUMAN] default will be used, as in (12a). By contrast, if MAX[-H] dominates MAX[+H], then the [-HUMAN] default will be used, as can be seen in (12b).

(12) a.

5[+H] & 3[-H]	MAX[+H]	MAX[-H]
a. ☞ SM2		*
b. SM8	*!	

b.

5[+H] & 3[-H]	MAX[-H]	MAX[+H]
a. SM2	*!	
b. ☞ SM8		*

When the mixed [±HUMAN] conjuncts are plural, there are more resolution strategies available. There are three possibilities: One can either

(1) use the same [+HUMAN] or [-HUMAN] default as was used with the singular conjuncts,

(2) use morphological agreement if the plural conjuncts are NC-balanced (i.e. NC4 & NC4 → SM4), and then default agreement with NC-unbalanced conjuncts (again, the default SM would be the same as the one used when both conjuncts are singular), *or*

(3) use the SM which agrees with the NC-feature of closest conjunct.

If the first strategy is used, we know that the MAX[αH] constraint which chooses the default must be ranked above all of the NC-constraints (i.e. MAX[αH]»DEPNC, CCA, MAXNC). If the second strategy is used, MAXNC must dominate the MAX[αH] constraint responsible for choosing the default, and MAX[αH] must in turn dominate AGREECLOSEST and DEPNC (i.e. MAXNC»MAX[αH]»CCA, DEPNC). Finally, if the third strategy is used, then AGREECLOSEST must dominate both MAX[+H] and MAX[-H]; the ranking of MAXNC and DEPNC cannot be determined without further data.

To summarise: When the mixed [±HUMAN] conjuncts are singular, one of the default subject markers will be used; the matter of *which* default is used is determined exclusively

by the ranking of MAX[+H] and MAX[-H] relative to one another. However, when the mixed [±HUMAN] conjuncts are plural, the ranking of the highest MAX[αH] constraint relative to the three NC constraints must be taken into consideration.

7.3.3 Resolution strategies for mixed number conjuncts

When the two conjuncts differ in their number feature, the order of the conjuncts can affect the resolution strategy used. If the first conjunct is singular, and the second conjunct is plural, the resolution strategy used should be the same as the one used for NC-unbalanced plural conjuncts with the same [±HUMAN] feature. Thus if CCA is used for NC-unbalanced plural [αHUMAN] conjuncts, it will also be used when the &P consists of a singular [αHUMAN] conjunct coordinated with a plural [αHUMAN] conjunct (in that order).

However, when the order of the conjuncts is reversed, and the singular conjunct is placed second, the matter is not as simple as merely using the same resolution strategy as was used for singular conjuncts. Instead there are two possible choices: Either 1) use the appropriate [±HUMAN] default SM, or 2) agree with the NC-feature of the plural conjunct (i.e. use furthest conjunct agreement). If default agreement is used, the relevant MAX[±H] constraint dominates DEPNC, whereas the opposite is true if FCA is used.

Note that just as with number balanced conjuncts, [+HUMAN] and [-HUMAN] mixed number conjuncts do not always employ the same resolution strategy: One set might use default agreement, whilst the other set might trigger furthest conjunct agreement being used. As discussed in §7.3.1, all that is required for this to happen is for MAX[+H] and MAX[-H] to be ranked differently with respect to DEPNC. Thus a ranking of MAX[αH]»DEPNC» MAX[$\neg\alpha$H] will result in default agreement being used for [αHUMAN] conjuncts, but FCA for [$\neg\alpha$HUMAN] conjuncts.

To summarise: Number-unbalanced &Ps with the order ‘SG & PL’ will use the same agreement resolution strategy as plural NC-unbalanced &Ps with the same [±HUMAN] specification. By contrast, when the order of the conjuncts is reversed (i.e. ‘PL & SG’), the agreement resolution strategy is determined by the ranking of DEPNC relative to MAX[+H] and MAX[-H].

7.3.4 The interaction of DEP[-H] with NC oriented constraints

Some languages, like Sesotho (see chapter 5), do not allow the [-HUMAN] default SM to be used with [+HUMAN] conjuncts, even if the SM in question agrees with the noun class of the conjuncts. In such languages, if the [-HUMAN] default is SM8, then the use of SM8 is not allowed with an &P consisting of NC8[+H] & NC8[+H] conjuncts. This is the result of the DEP[-H] constraint dominating both MAXNC and AGREECLOSEST: DEP[-H] penalizes a SM if it expresses [-HUMAN] if there is no [-HUMAN] conjunct in the input, and thus if DEP[-H] dominates MAXNC and AGREECLOSEST, it overrides the requirements for faithfulness to noun class features.

An interesting quirk arises *in languages which allow closest conjunct agreement*, but in which DEP[-H] dominates only AGREECLOSEST, and not MAXNC (i.e. the three constraints occur in the following configuration: MAXNC»DEP[-H]»CCA). In such languages, the use of SM8 would be allowed with NC-balanced [+HUMAN] conjuncts from NC8, as (13a) shows, but it *not* be used with NC-unbalanced conjuncts where the closest conjunct is a [+HUMAN] noun from NC8, as illustrated in (13b) – this despite the fact that closest conjunct agreement is allowed with other NC-unbalanced plural agreement, as shown in (13c).

(13) a.

8[+H] & 8[+H]	MAXNC	DEP[-H]	CCA
a. SM2	*!		*
b. SM6	*!		*
c. ☞ SM8		*	

b.

6[+H] & 8[+H]	MAXNC	DEP[-H]	CCA
a. ☞ SM2	*		*
b. ☞ SM6	*		*
c. SM8	*	*!	

c.

2[+H] & 6[+H]	MAXNC	DEP[-H]	CCA
a. SM2	*		*!
b. ☞ SM6	*		
c. SM8	*	*!	

In (13a), MAXNC's dominance over DEP[-H] allows for SM8 to be the optimal output for an input of '8[+H]&8[+H]'. However, in (13b), SM8 does not parse *both* NC-features in the input, and so violates MAXNC just as equally as SM2 and SM6 do. DEP[-H] then penalizes SM8, as it carries a [-HUMAN] feature and there is no [-HUMAN] feature in the input. Thus even though SM8 is the only candidate to satisfy AGREECLOSEST, this candidate is suboptimal; either candidate (a) or (b) will be optimal, depending on the ranking of the constraints which follow.

There are two possible ways to resolve the agreement dilemma in (13b): Either use the [+HUMAN] default SM, or agree with the NC-feature of the first conjunct. If the default is used, then we can tell that MAX[+H] must dominate DEPNC; the inverse must be true if FCA is used (i.e. DEPNC»MAX[+H]).

Thus in languages which allow closest conjunct agreement with plural [+HUMAN] conjuncts, if MAXNC and AGREECLOSEST have different rankings relative to DEP[-H], there will be an asymmetry in the grammar regarding syntactic agreement with NC8 [+HUMAN] conjuncts: Morphological agreement resulting in SM8 will be licensed, but closest conjunct agreement resulting in SM8 will be prohibited. On the other hand, if all of the NC oriented constraints have the same ranking as each other relative to DEP[-H], there will be no asymmetry in the agreement with NC8 [+HUMAN] conjuncts – if they all dominate DEP[-H], syntactic agreement will be allowed in such instances, but if DEP[-H] dominates MAXNC, DEPNC and AGREECLOSEST, semantic agreement will be favoured with NC8 [+HUMAN] conjuncts.

§7.3 has thus shown that certain surface effects will always be the result of the interaction between two or three specific constraints. §7.4 will now go on to illustrate the application of the analysis proposed in this thesis to a previous study on agreement with subject &Ps in Bantu.

7.4 Applying the analysis to a previous study

One of the benefits of generating a factorial typology is that, theoretically, it should generate any possible combination of resolution strategies that a noun class language might actually use. After systematically working out what descriptive generalizations would apply to each of the rankings generated, I examined the 47 language types to see if any of them matched the combination of resolution strategies described for Chichewa by Corbett and Mtenje (1987). These descriptive generalizations are summarised in (14).[10]

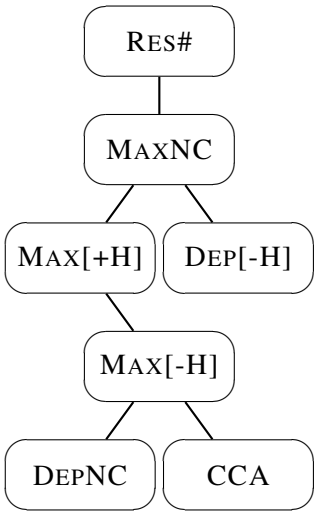
(14) Descriptive generalizations of agreement resolution in Chichewa

- NC-unbalanced conjuncts (regardless of number features):
  1. If conjuncts are [±HUMAN]-balanced: use appropriate [±HUMAN] default (SM2 for [+HUMAN] conjuncts, SM8/10 for [-HUMAN] conjuncts).
  2. If conjuncts are [±HUMAN]-unbalanced: coordination is ineffable (i.e. no SM is satisfactory).
- NC-balanced conjuncts:
  3. If conjuncts are *singular* and [±HUMAN]-balanced: use appropriate [±HUMAN] default.
  4. If conjuncts are *singular* and [±HUMAN]-unbalanced: coordination is ineffable.
  5. If both conjuncts are *plural*, regardless of whether they are [±HUMAN]-balanced or [±HUMAN]-unbalanced: use morphological agreement (i.e. use the SM which agrees with the NC of both conjuncts).

There was one very close match: "Language 45" (see appendix), which has the ranking given in (15), best matches the resolution patterns of Chichewa.

[10] In these generalizations, I use the revised Bleek-Meinhoff noun class system, rather than the Watkins system used by Corbett and Mtenje (1987). Please note that in Chichewa the subject markers for NC8 and NC10 are homophonous; for convenience sake, I will refer to these subject markers as SM8/10.

(15) Ranking diagram for 'Language 45'



The dominance of MAXNC over MAX[+H] and MAX[-H] accounts for the use of morphological agreement with plural NC-balanced conjuncts (as in (16)), while the dominance of RES# over MAXNC prohibits the use of morphological agreement with singular conjuncts; thus when the input consists of singular conjuncts, the optimal SM is chosen by MAX[+H] or MAX[-H] instead (as in (17)). Finally, the the dominance of MAX[+H] and MAX[-H] over DEPNC and AGREECLOSEST results in plural NC-unbalanced conjuncts triggering default agreement, rather than partial agreement, as (18) shows.[11]

(16) *Ma-lalanje ndi ma-samba **a**-ku-bvunda*
6-orange and 6-leaf **SM6**-PRES-rot
'The oranges and the leaves are rotting.' (Corbett and Mtenje 1987:20)

6[-H] & 6[-H]	RES#	MAXNC	MAX[+H]	DEP[–H]	MAX[–H]	DEPNC	CCA
a. SM2		*!			*	*	*
b. ☞ SM6					*		
c. SM8/10		*!				*	*

[11]The glosses in the examples have been changed to reflect the Bleek-Meinhoff NC-system, rather than the Watkins gender system.

(17) *Lalanje ndi tsamba* ***a****-ku-bvunda*
5.orange and 5.leaf **SM6**-PRES-rot
'The orange and the leaf are rotting.' (Corbett and Mtenje 1987:19)

5[-H] & 5[-H]	RES#	MAXNC	MAX[+H]	DEP[–H]	MAX[–H]	DEPNC	CCA
a. SM2					*!	*	*
b. SM6					*!	*	*
c. ☞ SM8/10						*	*
d. SM5	*!	*			*		

(18) *Ma-ukonde ndi mi-peni* ***zi****-li uko*
6-net and 4-knife **8/10**-be there
'The nets and the knives are there.' (Corbett and Mtenje 1987:22)

6[-H] & 4[-H]	RES#	MAXNC	MAX[+H]	DEP[–H]	MAX[–H]	DEPNC	CCA
a. SM2		*			*!	*	*
b. SM4		*			*!		
c. SM6		*			*!		*
d. ☞ SM8/10		*				*	*

Unfortunately, the ranking in (15) fails to capture the ineffability of NC-unbalanced mixed [±HUMAN] conjuncts. As MAX[+H] dominating MAX[-H] is only responsible for the use of the [+HUMAN] default SM with [±HUMAN]-unbalanced conjuncts in "language 45", and this does not happen in Chichewa, one could argue that there is no evidence for these two constraints to be ranked with respect to each other in Chichewa. In fact, by removing this dominance relationship, and ranking these two constraints in the same stratum (i.e. RES# » MAXNC » MAX[+H], MAX[-H], DEP[-H] » DEPNC, CCA), one produces a ranking in which there is no one optimal candidate for mixed [±HUMAN] conjuncts, as illustrated in (19), unless they are plural and NC-balanced, as in (20).[12]

(19) **Mu-nthu ndi ng'ombe* ***a/zi****-ku-yenda*
1-person and 9-cow **SM2/SM8/10**-PRES-walk
Int: 'A person and a cow are walking.' (Corbett and Mtenje 1987:34)

[12]NC13 is the plural of NC12; these are the diminutive noun classes in Chichewa, and the noun prefixes attach to the stem which includes the NC prefix of the word in its non-diminutive form.

1[+H] & 9[-H]	RES#	MAXNC	MAX[+H]	DEP[–H]	MAX[–H]	DEPNC	CCA
a. SM2		*			*!	*	*
b. SM8/10		*	*!			*	*
c. 1	*!	*			*		*
d. 9	*!	*	*				

(20) *ti-a-na* *ndi* *ti-mi-phika* ***ti****-li* *apo*
13-2-child and 13-4-pot **SM13**-be there
'The infants and the small pots are there.' (Corbett and Mtenje 1987:33)

13[+H] & 13[-H]	RES#	MAXNC	MAX[+H]	DEP[–H]	MAX[–H]	DEPNC	CCA
a. SM2		*!			*	*	*
b. SM4		*!	*		*	*	*
c. SM8/10		*!	*			*	*
d. ☞ SM13			*		*		

In (19), two singular conjuncts from different noun classes are conjoined. The two singular subject markers, candidates (c) and (d), both violate RES#, and so are non-optimal. All of the candidates violate MAXNC. The remaining possible candidates, (a) and (b), each violate one constraint in the third stratum of constraints, and they both violate the two constraints on the next stratum, and therefore they are equally 'bad'. Thus there is no optimal candidate.

By contrast, when the [±HUMAN]-unbalanced conjuncts are plural and NC-balanced, as in (20), the SM which agrees with the NC both conjuncts will win under MAXNC; this accounts for the fact that coordination of [±HUMAN] unbalanced conjuncts is allowed when the conjuncts are plural and NC-balanced.

Thus by ranking MAX[+H] and MAX[-H] in the same stratum of constraints, one can account for the grammaticality of coordination involving plural NC-balanced nouns which differ in their [±HUMAN] specification, but the ineffability of all other &Ps consisting of mixed [±HUMAN] conjuncts in Chichewa.

In this section I have shown how the analysis proposed in this thesis can be extended to other studies of agreement with &Ps in Bantu languages. Unfortunately, although there

are several other such studies, many do not contain representative inputs of enough of the input groups described in (1) on page 141; without more data one cannot determine where the languages described in those studies fit within the typology.

7.5 Concluding remarks

This chapter discussed the generation of a typology of agreement resolution strategies. I showed that the Optimality Theoretic analysis proposed in this chapter predicts that 1) &Ps consisting of singular conjuncts will always trigger default agreement in noun class languages (§7.2.1), and that 2) Furthest conjunct agreement will never be used in a grammar which does not *also* use closest conjunct agreement (§7.2.2).

I then went on to show that certain surface effects are always the result of the interaction between two or three constraints. Firstly, when MAX[+H] and MAX[-H] are ranked differently relative to one or more of the NC constraints, there will be an asymmetry in the resolution strategies employed by human and nonhuman conjuncts (§7.3.1). Secondly, when the two plural conjuncts differ in their [±HUMAN] specification, there are three resolution options available – the resolution strategy ultimately used is determined by the ranking of the highest MAX[±H] constraint relative to the three NC-constraints (§7.3.2). Thirdly, when the conjuncts differ in their NUMBER feature, the order of the conjuncts affects the choice of resolution strategy – if the plural conjunct is closest to the verb, one will use same resolution strategy as one uses for NC-unbalanced conjunct; if, however, the singular conjunct is closest to the verb, the ranking of DEPNC relative to MAX[+H] and MAX[-H] will determine whether FCA or default agreement is used (§7.3.3). Finally, languages which allow closest conjunct agreement, but in which MAXNC and AGREECLOSEST are ranked differently relative to Dep[-H], will display an asymmetry with regards to syntactic agreement with [+HUMAN] conjuncts from NC8 (§7.3.4).

§7.4 then showed that the OT analysis proposed in this thesis can be extended to account for the data in other studies on of agreement with subject &Ps in Bantu, provided the data covers a wide range of representative inputs.

Chapter 8

Concluding Remarks

This thesis set out to study agreement with coordinated subjects in Xitsonga, Sesotho and isiXhosa. It addressed the following questions:

1. What strategies are used to resolve agreement with coordinated preverbal subjects in Xitsonga, Sesotho and isiXhosa?
2. a) Are these strategies restricted to certain agreement feature contexts?
   b) If so, which agreement feature contexts trigger the use of which strategies?
3. a) Are these strategies ontological primitives, or merely the surface manifestation of a conspiracy between underlying constraints?
   b) Can the restriction of certain agreement strategies to specific agreement feature contexts be accounted for by the hierarchical ranking of such constraints within an Optimality Theoretic framework?
4. What insight can be gleaned from generating and examining a factorial typology of agreement strategies, based on the constraints proposed in (3)?

Section 8.1 discusses how this thesis answered these research questions, and section 8.2 discusses issues for further research.

8.1 Answering the research questions

Xitsonga, Sesotho and isiXhosa display, to varying degrees, use of attested agreement resolution strategies such as morphological agreement, partial agreement and default (semantic) agreement. However, I showed that these agreement resolution strategies can not be used

interchangeably, and that the strategy used will be determined by the combination of features present on the conjuncts. Moreover, the combination of features which trigger specific agreement resolution strategies is not consistent cross-linguistically. This is grounds for arguing that these strategies *are not* ontological primitives, but rather the surface manifestations of the interactions of various constraints.

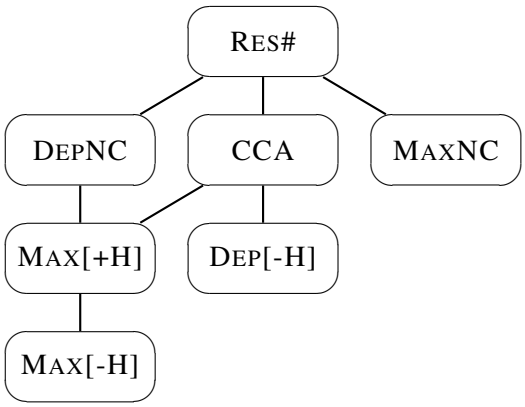
For the most part, these constraints address faithfulness to features on the conjuncts, although one constraint (RES#) addresses the issue of singular agreement with an &P being marked. The constraints requiring faithfulness to the [±HUMAN] features on the conjuncts (namely MAX[+H], MAX[-H], and DEP[-H]) are responsible for the manifestation of 'semantic' agreement , while the constraints which address faithfulness to NC-features (namely MAXNC, DEPNC, and AGREECLOSEST) are responsible for the various manifestations of 'syntactic agreement' – i.e. morphological agreement and partial agreement. Partial agreement, in particular, is shown to be licensed by DEPNC and AGREECLOSEST. Admittedly, the latter of these constraints performs the same function as the common strategy of closest conjunct agreement, and thus one might be tempted to argue that this strategy *could* be an ontological primitive. However, it is reasonable to argue that the strategy arises as a result of the constraint, on the basis of two issues: 1) AGREECLOSEST is not the only constraint to license partial agreement, but merely appears to be a locality version of DEPNC, and 2) AGREECLOSEST's interaction with other constraints results in restrictions on the agreement feature contexts where closest conjunct agreement is allowed.

One other agreement resolution strategy described in previous studies is that of phonological agreement; Voeltz (1971) attributed the licensing of agreement with conjuncts from NC8 and NC10 in isiXhosa to the fact that SM8 and SM10 are homophonous in this language. However, this effect, along with morphological agreement with plural NC-balanced conjuncts, can also be attributed to closest conjunct agreement – a strategy used with all plural conjuncts. Phonological agreement therefore played no real role in the data collected for this study.

Chapters 4, 5 and 6 discussed the specifics of which agreement feature contexts trigger which agreement strategies in Xitsonga, Sesotho, and isiXhosa. The constraints responsible for producing the effects of these strategies were introduced as they became relevant. This resulted in only the analysis for isiXhosa showing the complete ranking of all con-

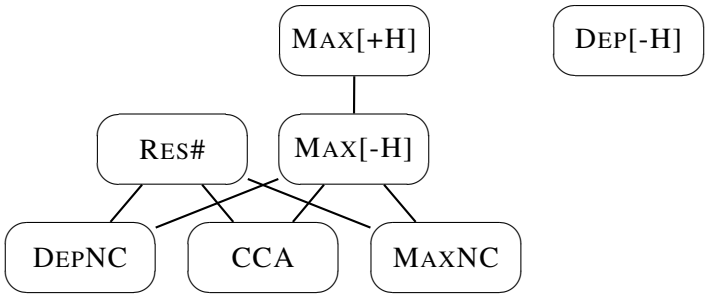
straints. The most noticeable trait of the isiXhosa ranking, given in (1), is that all of the constraints requiring faithfulness to NC-features dominate all of the constraints requiring faithfulness to [±HUMAN] features. This is results in syntactic agreement being preferred over semantic agreement. Semantic agreement is licensed only when syntactic agreement would result in a [SINGULAR] feature appearing on the subject marker; this is because SMs with a [SINGULAR] feature are penalized by RES#, and RES# dominates the constraints which would favour the syntactically agreeing subject marker.

(1) Ranking of constraints in isiXhosa



The complete ranking for the first variety of Xitsonga, showing the place of *all* of the constraints within the hierarchy, shows the reverse trend from that of isiXhosa: All of the constraints requiring faithfulness to [±HUMAN] features dominate all of the constraints requiring faithfulness to NC-features, as (2) shows. This results in semantic agreement being favoured over syntactic agreement. Moreover, as the default markers are both plural SMs, the use of these subject markers will never violate RES#, and thus syntactic agreement will never be required as a 'fall back' plan (as semantic agreement is in isiXhosa).

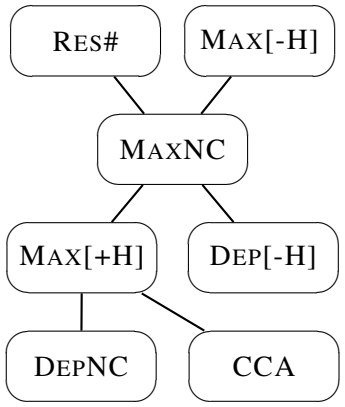
(2) Ranking of constraints in Xitsonga variety 1



Note that in (2) DEP[-H] does not seem to interact with any of the other constraints. The main surface effect of a highly ranked Dep[-H] constraint is the prohibition against syntactic agreement in contexts where syntactic agreement would result in the [-HUMAN] default subject marker being used with [+HUMAN] conjuncts (see discussion of (27) on page 105) – as syntactic agreement is not actually used in this variety of Xitsonga, there is no evidence of DEP[-H] interacting with any of the other constraints. DEP[-H] therefore neither dominates, nor is dominated by, other constraints.

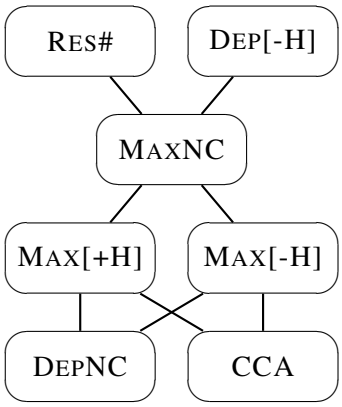
In the second variety of Xitsonga, semantic agreement is always preferred for non-human conjuncts, but under certain circumstances syntactic agreement is allowed with human conjuncts. This asymmetry is the result of MAX[-H] and MAX[+H] being ranked differently relative to one of the constraints favouring syntactic agreement, namely MAXNC. Partial agreement is not licensed at all, as all the constraints favouring semantic agreement dominate DEPNC and AGREECLOSEST, as shown in (3).

(3) Ranking of constraints in Xitsonga Variety 2



In Sesotho, on the other hand, MAX[+H] and MAX[-H] have the same ranking as each other relative to the constraints favouring syntactic agreement, which results in human and non-human conjuncts almost always using the same resolution strategies in the same agreement feature contexts. The exception to this (plural NC-balanced [+HUMAN] conjuncts from NC8 or NC10 triggering default agreement, rather than morphological agreement) is the result of DEP[-H] dominating MAXNC. The two constraints which license partial agreement, DEPNC and AGREECLOSEST, which are introduced in chapter 8, at first seem to play no active role in Sesotho, and so they fall at the bottom of the hierarchy, as shown in (4).

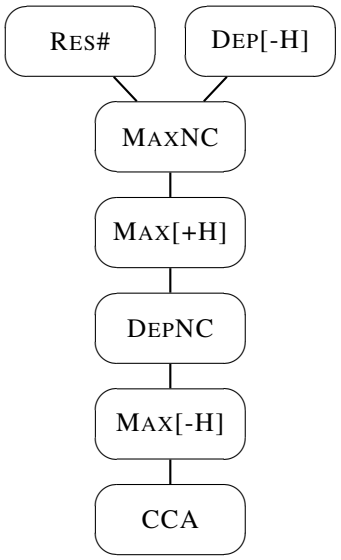
(4) Ranking of constraints in Sesotho



However, it is worth noting that there was some evidence of partial agreement in Sesotho (see § 5.2.5 on page 99). The problem is that 1) the partial agreement did not systematically manifest as either closest conjunct agreement or furthest conjunct agreement, but rather seemed to switch haphazardly between the two, and 2) it only ever occurred when [-HUMAN] conjuncts from NC4 and NC6 were coordinated with each other. Although no explanation can be offered as to why the form of partial agreement used is not consistent, the of DEPNC may offer some clue as to why partial agreement only happens there is coordination between [-HUMAN] conjuncts from NC4 and NC6.

As partial agreement only happens with [-HUMAN] conjuncts, we know that MAX[+H] and MAX[-H] must be ranked differently in respect to at least one of the constraints which favour partial agreement. As the form of partial agreement used is not systematically closest conjunct agreement, the constraint in question must be DEPNC, which favours agreement with either conjunct. Now, consider the implications for plural conjuncts if one revises the ranking in (4), such that DEPNC is ranked below MAX[+H], but above MAX[-H], as in (5).

(5) Revised ranking of constraints in Sesotho



As we already know, MAXNC dominating MAX[+H] and MAX[-H] results in morphological agreement with plural NC-balanced conjuncts. If MAX[+H] in turn dominates DEPNC, one would expect semantic agreement with plural NC-unbalanced [+HUMAN] conjuncts; this is exactly what one finds in Sesotho – thus the rankings in (4) and (5) do not differ in this regard. On the other hand, if DEPNC dominates MAX[-H], one would expect partial agreement to be used with all plural NC-unbalanced [-HUMAN] conjuncts.[1] As the generalization thus far has been that semantic agreement is used in such instances, this ranking appears to make the wrong prediction. However, consider the following facts:

1. There are only five plural noun classes in Sesotho: NC2, NC4, NC6, NC8, and NC10.
2. NC2 is typically reserved for [+HUMAN] conjuncts.
3. Thus there are only four plural noun classes which can contain [-HUMAN] conjuncts.
4. These four noun classes can be combined as unordered pairs in only five different combinations: 4&6, 4&8, 4&10, 6&8 and 6&10.
5. Of the five pairs, only one pair (4&6) does not include either NC8 or NC10.
6. When the conjuncts are NC-unbalanced DEPNC always favours two subject markers: one for each conjunct.
7. This means that for four of the five pairs, SM8/10 will always be one of the SMs favoured by DEPNC, and that only with the 4&6 pair is SM8/10 penalized.

[1] As already discussed, the high ranking of RES# prevents syntactic agreement with singular conjuncts, as this would result in using a SM with a [SINGULAR] feature.

If one of the conjuncts is from NC8 or NC10, as in (6), SM8/10 will be one of the two candidates favoured by DEPNC; in this instance, the other candidate favoured is (b), SM4, and the remaining candidates incur fatal violations. MAX[-H] then evaluates candidates (b) and (d) as to whether they parse the [-HUMAN] features in the input. As SM4 is underspecified for the [±HUMAN] feature, it violates MAX[-H]. On the other hand, SM8/10 parses the [-HUMAN] feature on the conjuncts, and therefore satisfies this constraint. SM8/10 thus remains the optimal candidate in such instances of coordinating NC-unbalanced plural [-HUMAN] conjuncts, even when DEPNC dominates MAX[-H].

(6)

8[-H] & 4[-H]	RES#	DEP[-H]	MAXNC	MAX[+H]	DEPNC	MAX[-H]	CCA
a. ba (2)			*		*!	*	*
b. e (4)			*			*!	
c. a (6)			*		*!	*	*
d. ☞ di (8/10)			*				*

By contrast, when neither of the NC-unbalanced plural [-HUMAN] conjuncts is from NC8/10, SM8/10 agrees with neither conjunct's NC feature, and so will incur a fatal violation under DEPNC, as can be seen in (7). This means that SM8/10 is not available to be evaluated by the MAX[-H], the constraint which favours default agreement with non-human conjuncts. Instead, only SM4 and SM6 are left as viable candidates. DEPNC's dominance over MAX[-H] could thus account for the lack of default agreement with NC-unbalanced conjuncts from NC4 and NC6, and the preference for using either SM4 or SM6 on the verb.

(7)

4[-H] & 6[-H]	RES#	DEP[-H]	MAXNC	MAX[+H]	DEPNC	MAX[-H]	CCA
a. ba (2)			*		*!	*	*
b. e (4)			*			*	*
c. ☞ a (6)			*			*	
d. di (8/10)			*		*!		*

Unfortunately, however, this does not fully account for the strange agreement behaviour observed in such instances: The presence of AGREECLOSEST would then suggest that closest conjunct agreement should be preferred with these conjuncts, but in truth no systematic use of closest conjunct agreement is displayed.

The final research question was addressed in chapter 7, where I showed that not only can certain surface effects in agreement be accounted for by the interactions of specific constellations of constraints, but also that two predictions follow from the typology generated: 1) When RES# is undominated, the coordination of singular conjuncts should also trigger semantic agreement, and never 'in-gender' agreement, and 2) Furthest conjunct agreement should never be a preferred agreement strategy – it will only be used in languages which allow closest conjunct agreement, but only in contexts where closest conjunct agreement is prohibited for some reason.

8.2 Avenues for future research

The analysis presented in this thesis can be extended in many ways. Firstly, it would be interesting to see to what degree these constraints can account for agreement resolution patterns in other Bantu languages, and what, if any, modifications need to be made to capture the generalizations for those languages.

Secondly, the scope of agreement with &Ps studied in this thesis was restricted to preverbal subjects. As evidenced in Kiswahili, postverbal subject &Ps trigger different agreement strategies (Marten 2000); as yet, my analysis cannot account for why this may be.

Finally, agreement with non-subject &Ps was also outside the scope of this study. It still remains to be seen whether the analysis presented here would also account for the agreement resolution strategies used in other agreement contexts, such as object agreement, adjectival agreement, pronominal agreement, and relative agreement.

Appendix A

More on the Factorial Typology

This appendix consists of two sections. The first section, A1, consists of a subsection of the factorial typology generated in OTworkplace (Prince et al. 2014). As discussed in chapter 7, only the 47 output languages in which RES# is undominated are shown here. The second section, A2, gives the relevant ranking diagrams for each of the languages shown in the table in A1; additionally, to abstract away from specific inputs and outputs, a set of descriptive generalizations accompanies each ranking diagram.

A.1 Inputs and their various outputs in each 'language'

This section consists of a table showing inputs and their various outputs across the different languages. Each input is labeled as belonging to one of the groups listed in (1) on page 141 – this list is repeated overpage, for ease of reference. Please note that for convenience's sake, SM2 is used to refer to the [+HUMAN] default, and that SM8 is used to refer to the [-HUMAN] default. Similarly, NC2 and NC8, respectively, are used to refer to the noun classes associated with the relevant default subject markers. For any language which uses different subject markers as the defaults, please substitute accordingly.

The first few rows of the table which follows indicate the input groups being represented, along with an example of such an input. The rows below that then indicate what subject marker would be the output for the given input, according to a particular language type; the language type is labeled in the left-most column.

For reasons of space, in the 'example input' row in the table, the [+HUMAN] feature is

indicated by 'h' appearing after the NC-feature, and the [-HUMAN] feature is indicated by 'i.' appearing after the NC-feature.

Number Balanced, [±HUMAN Balanced Conjuncts]

Both Conjuncts [+HUMAN]

1. NC-balanced singular conjuncts
2. NC-unbalanced singular conjuncts
3. NC-balanced plural conjuncts, *not* NC8
4. NC-balanced plural conjuncts, *specifically* NC8
5. NC-unbalanced plural conjuncts, where *neither* conjunct is NC2, and second conjunct is *not* NC8
6. NC-unbalanced plural conjuncts, where first conjunct *is* NC2
7. NC-unbalanced plural conjuncts, where second conjunct *is* NC8

Both Conjuncts [-HUMAN]

8. NC-balanced singular conjuncts
9. NC-unbalanced singular conjuncts
10. NC-balanced plural conjuncts
11. NC-unbalanced plural conjuncts, where *neither* conjunct is NC8
12. NC-unbalanced plural conjuncts, where first conjunct *is* NC8

Mixed [±HUMAN] Conjuncts

13. NC-balanced singular conjuncts
14. NC-unbalanced singular conjuncts
15. NC-balanced plural conjuncts
16. NC-unbalanced plural conjuncts, where neither conjunct is NC2 or NC8
17. NC-unbalanced plural conjuncts, where first conjunct is NC2
18. NC-unbalanced plural conjuncts, where first conjunct is NC8

Mixed Number Conjuncts

19. Both conjuncts human: PL & SG
20. Both conjuncts human: SG & PL
21. Both conjuncts nonhuman: PL & SG
22. Both conjuncts nonhuman: SG & PL

	Both Conjuncts [+HUMAN]							Both Conjuncts [-HUMAN]					[±HUMAN]-Unbalanced Conjuncts						#-Unbalanced Conjuncts			
Group	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**	**20**	**21**	**22**
e.g. input	**5h& 5h**	**5h& 9h**	**6h& 6h**	**8h& 8h**	**6h& 10h**	**2h& 6h**	**6h& 8h**	**5i& 5i**	**3i& 5i**	**6i& 6i**	**4i& 6i**	**8i& 6i**	**5h& 5i**	**5h& 9i**	**6h& 6i**	**6h& 4i**	**2h& 6i**	**8i& 6h**	**6h& 7h**	**7h& 6h**	**6i& 9i**	**9i& 6i**
Lg#1	SM2	SM2	SM6	SM2	SM10	SM6	SM2	SM8	SM8	SM6	SM6	SM6	SM2	SM2	SM6	SM4	SM6	SM6	SM2	SM6	SM6	SM6
Lg#2	SM2	SM2	SM6	SM2	SM10	SM6	SM2	SM8	SM8	SM6	SM6	SM6	SM2	SM2	SM6	SM4	SM6	SM6	SM2	SM6	SM8	SM6
Lg#3	SM2	SM2	SM6	SM2	SM10	SM6	SM2	SM8	SM8	SM6	SM6	SM6	SM2	SM2	SM6	SM4	SM6	SM6	SM6	SM6	SM6	SM6
Lg#4	SM2	SM2	SM6	SM2	SM10	SM6	SM2	SM8	SM8	SM6	SM6	SM6	SM8	SM8	SM6	SM4	SM6	SM6	SM2	SM6	SM8	SM6
Lg#5	SM2	SM2	SM6	SM2	SM10	SM6	SM2	SM8	SM8	SM6	SM6	SM6	SM8	SM8	SM6	SM4	SM6	SM6	SM6	SM6	SM6	SM6
Lg#6	SM2	SM2	SM6	SM2	SM10	SM6	SM2	SM8	SM8	SM6	SM6	SM6	SM8	SM8	SM6	SM4	SM6	SM6	SM6	SM6	SM8	SM6
Lg#7	SM2	SM2	SM6	SM2	SM10	SM6	SM2	SM8	SM8	SM6	SM6	SM8	SM8	SM8	SM6	SM4	SM6	SM8	SM6	SM6	SM6	SM6
Lg#8	SM2	SM2	SM6	SM2	SM10	SM6	SM2	SM8	SM8	SM6	SM8	SM8	SM8	SM8	SM6	SM8	SM8	SM8	SM2	SM6	SM8	SM8
Lg#9	SM2	SM2	SM6	SM2	SM10	SM6	SM2	SM8	SM8	SM6	SM8	SM8	SM8	SM8	SM6	SM8	SM8	SM8	SM6	SM6	SM8	SM8
Lg#10	SM2	SM2	SM6	SM2	SM10	SM6	SM2	SM8	SM8	SM8	SM8	SM8	SM8	SM8	SM8	SM8	SM8	SM8	SM2	SM6	SM8	SM8
Lg#11	SM2	SM2	SM6	SM2	SM10	SM6	SM2	SM8	SM8	SM8	SM8	SM8	SM8	SM8	SM8	SM8	SM8	SM8	SM6	SM6	SM8	SM8
Lg#12	SM2	SM2	SM6	SM8	SM10	SM6	SM8	SM8	SM8	SM6	SM6	SM6	SM2	SM2	SM6	SM4	SM6	SM6	SM2	SM6	SM6	SM6
Lg#13	SM2	SM2	SM6	SM8	SM10	SM6	SM8	SM8	SM8	SM6	SM6	SM6	SM2	SM2	SM6	SM4	SM6	SM6	SM2	SM6	SM8	SM6
Lg#14	SM2	SM2	SM6	SM8	SM10	SM6	SM8	SM8	SM8	SM6	SM6	SM6	SM2	SM2	SM6	SM4	SM6	SM6	SM6	SM6	SM6	SM6
Lg#15	SM2	SM2	SM6	SM8	SM10	SM6	SM8	SM8	SM8	SM6	SM6	SM6	SM8	SM8	SM6	SM4	SM6	SM6	SM2	SM6	SM8	SM6
Lg#16	SM2	SM2	SM6	SM8	SM10	SM6	SM8	SM8	SM8	SM6	SM6	SM6	SM8	SM8	SM6	SM4	SM6	SM6	SM6	SM6	SM6	SM6
Lg#17	SM2	SM2	SM6	SM8	SM10	SM6	SM8	SM8	SM8	SM6	SM6	SM6	SM8	SM8	SM6	SM4	SM6	SM6	SM6	SM6	SM8	SM6
Lg#18	SM2	SM2	SM6	SM8	SM10	SM6	SM6	SM8	SM8	SM6	SM6	SM8	SM8	SM8	SM6	SM4	SM6	SM8	SM6	SM6	SM6	SM6
Lg#19	SM2	SM2	SM6	SM8	SM10	SM6	SM2	SM8	SM8	SM6	SM8	SM8	SM8	SM8	SM6	SM8	SM8	SM8	SM2	SM6	SM8	SM8
Lg#20	SM2	SM2	SM6	SM8	SM10	SM6	SM6	SM8	SM8	SM6	SM8	SM8	SM8	SM8	SM6	SM8	SM8	SM8	SM6	SM6	SM8	SM8
Lg#21	SM2	SM2	SM6	SM8	SM10	SM6	SM8	SM8	SM8	SM8	SM8	SM8	SM8	SM8	SM8	SM8	SM8	SM8	SM2	SM6	SM8	SM8
Lg#22	SM2	SM2	SM6	SM8	SM10	SM6	SM8	SM8	SM8	SM8	SM8	SM8	SM8	SM8	SM8	SM8	SM8	SM8	SM6	SM6	SM8	SM8
Lg#23	SM2	SM2	SM6	SM2	SM10	SM2	SM6	SM8	SM8	SM6	SM6	SM6	SM2	SM2	SM6	SM4	SM2	SM6	SM6	SM6	SM6	SM6
Lg#24	SM2	SM2	SM6	SM2	SM10	SM2	SM6	SM8	SM8	SM6	SM6	SM8	SM2	SM2	SM6	SM4	SM2	SM8	SM6	SM6	SM6	SM6
Lg#25	SM2	SM2	SM6	SM2	SM10	SM2	SM6	SM8	SM8	SM6	SM6	SM8	SM8	SM8	SM6	SM4	SM2	SM8	SM6	SM6	SM6	SM6

	Both Conjuncts [+HUMAN]							Both Conjuncts [-HUMAN]					[±HUMAN]-Unbalanced Conjuncts						#-Unbalanced Conjuncts			
Group	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**	**20**	**21**	**22**
e.g. input	**5h& 5h**	**5h& 9h**	**6h& 6h**	**8h& 8h**	**6h& 10h**	**2h& 6h**	**6h& 8h**	**5i& 5i**	**3i& 5i**	**6i& 6i**	**4i& 6i**	**8i& 6i**	**5h& 5i**	**5h& 9i**	**6h& 6i**	**6h& 4i**	**2h& 6i**	**8i& 6h**	**6h& 7h**	**7h& 6h**	**6i& 9i**	**9i& 6i**
Lg#26	SM2	SM2	SM6	SM2	SM10	SM2	SM6	SM8	SM8	SM6	SM8	SM8	SM8	SM8	SM6	SM8	SM8	SM8	SM6	SM6	SM8	SM8
Lg#27	SM2	SM2	SM6	SM2	SM10	SM2	SM6	SM8	SM8	SM8	SM8	SM8	SM8	SM8	SM8	SM8	SM8	SM8	SM6	SM6	SM8	SM8
Lg#28	SM2	SM2	SM6	SM8	SM10	SM2	SM6	SM8	SM8	SM6	SM6	SM6	SM2	SM2	SM6	SM4	SM2	SM6	SM6	SM6	SM6	SM6
Lg#29	SM2	SM2	SM6	SM8	SM10	SM2	SM6	SM8	SM8	SM6	SM6	SM8	SM2	SM2	SM6	SM4	SM2	SM8	SM6	SM6	SM6	SM6
Lg#30	SM2	SM2	SM6	SM8	SM10	SM2	SM6	SM8	SM8	SM6	SM6	SM8	SM8	SM8	SM6	SM4	SM2	SM8	SM6	SM6	SM6	SM6
Lg#31	SM2	SM2	SM6	SM8	SM10	SM2	SM6	SM8	SM8	SM6	SM8	SM8	SM8	SM8	SM6	SM8	SM8	SM8	SM6	SM6	SM8	SM8
Lg#32	SM2	SM2	SM6	SM8	SM10	SM2	SM6	SM8	SM8	SM8	SM8	SM8	SM8	SM8	SM8	SM8	SM8	SM8	SM6	SM6	SM8	SM8
Lg#33	SM2	SM2	SM2	SM2	SM2	SM2	SM2	SM8	SM8	SM6	SM6	SM6	SM2	SM2	SM2	SM2	SM2	SM2	SM2	SM2	SM6	SM6
Lg#34	SM2	SM2	SM2	SM2	SM2	SM2	SM2	SM8	SM8	SM6	SM6	SM6	SM2	SM2	SM2	SM2	SM2	SM2	SM2	SM2	SM8	SM6
Lg#35	SM2	SM2	SM2	SM2	SM2	SM2	SM2	SM8	SM8	SM6	SM8	SM8	SM2	SM2	SM2	SM2	SM2	SM2	SM2	SM2	SM8	SM8
Lg#36	SM2	SM2	SM2	SM2	SM2	SM2	SM2	SM8	SM8	SM8	SM8	SM8	SM2	SM2	SM2	SM2	SM2	SM2	SM2	SM2	SM8	SM8
Lg#37	SM2	SM2	SM2	SM2	SM2	SM2	SM2	SM8	SM8	SM8	SM8	SM8	SM8	SM8	SM8	SM8	SM8	SM8	SM2	SM2	SM8	SM8
Lg#38	SM2	SM2	SM6	SM2	SM2	SM2	SM2	SM8	SM8	SM6	SM6	SM6	SM2	SM2	SM6	SM2	SM2	SM2	SM2	SM2	SM6	SM6
Lg#39	SM2	SM2	SM6	SM2	SM2	SM2	SM2	SM8	SM8	SM6	SM6	SM6	SM2	SM2	SM6	SM2	SM2	SM2	SM2	SM2	SM8	SM6
Lg#40	SM2	SM2	SM6	SM2	SM2	SM2	SM2	SM8	SM8	SM6	SM8	SM8	SM2	SM2	SM6	SM2	SM2	SM2	SM2	SM2	SM8	SM8
Lg#41	SM2	SM2	SM6	SM2	SM2	SM2	SM2	SM8	SM8	SM6	SM8	SM8	SM8	SM8	SM6	SM8	SM8	SM8	SM2	SM2	SM8	SM8
Lg#42	SM2	SM2	SM6	SM2	SM2	SM2	SM2	SM8	SM8	SM8	SM8	SM8	SM8	SM8	SM8	SM8	SM8	SM8	SM2	SM2	SM8	SM8
Lg#43	SM2	SM2	SM6	SM8	SM2	SM2	SM2	SM8	SM8	SM6	SM6	SM6	SM2	SM2	SM6	SM2	SM2	SM2	SM2	SM2	SM6	SM6
Lg#44	SM2	SM2	SM6	SM8	SM2	SM2	SM2	SM8	SM8	SM6	SM6	SM6	SM2	SM2	SM6	SM2	SM2	SM2	SM2	SM2	SM8	SM6
Lg#45	SM2	SM2	SM6	SM8	SM2	SM2	SM2	SM8	SM8	SM6	SM8	SM8	SM2	SM2	SM6	SM2	SM2	SM2	SM2	SM2	SM8	SM8
Lg#46	SM2	SM2	SM6	SM8	SM2	SM2	SM2	SM8	SM8	SM6	SM8	SM8	SM8	SM8	SM6	SM8	SM8	SM8	SM2	SM2	SM8	SM8
Lg#47	SM2	SM2	SM6	SM8	SM2	SM2	SM2	SM8	SM8	SM8	SM8	SM8	SM8	SM8	SM8	SM8	SM8	SM8	SM2	SM2	SM8	SM8

A.2 Descriptive generalizations of Languages 1-47

This section provides a ranking diagram for each of the language types listed in the table in the previous section, as well as a brief description of resolution patterns which surface as a result of the rankings. The following abbreviations are use.

Key of Abbreviations used:

CCA: closest conjunct agreement

FCA: furthest conjunct agreement

CC: closest conjunct

FC: furthest conjunct

def.: Default agreement

[+H] def: default SM used with conjuncts with [+human] features

[-H] def: default SM used with conjuncts with [-human] features

[+H]: (In isolation) conjuncts with [+human] features

[-H]: (In isolation) conjuncts with [-human] features

Xh: [+human] conjunct, with any NC-feature

8h: [+human] conjunct, with NC8 feature

SM2: default SM used with conjuncts with [+human] features

SM8: default SM used with conjuncts with [-human] features

Xi: [-human] conjunct, with any NC-feature

8i: [-human] conjunct, with NC8 feature

(B): conjuncts are NC-balanced

(U): conjuncts are NC-unbalanced

h: human conjuncts

i: nonhuman conjuncts

(Bh): [+human] conjuncts are NC-balanced

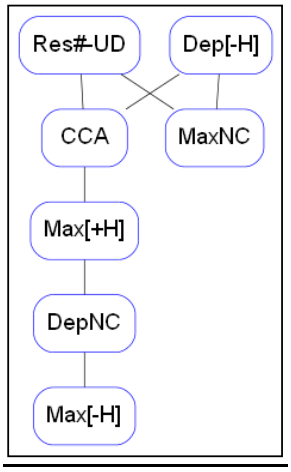
(Bi): [-human] conjuncts are NC-balanced

(Uh): [+human] conjuncts are NC-unbalanced

(Ui): [-human] conjuncts are NC-unbalanced

morph: Morphological agreement

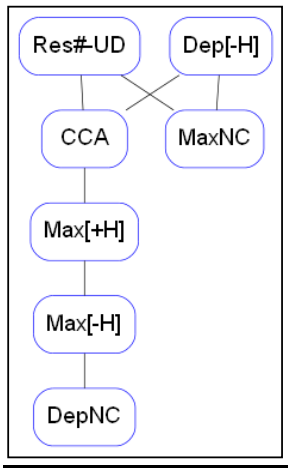
Lang1



[±H]-Bal	[±H]-Unbal	#-Unbal
SG & SG→ def. PL & PL → CCA BUT: Xh&8h→ SM2	SG & SG→ SM2 PL & PL → CCA	PL & SG → def. for [+H] → FCA for [-H] SG & PL → CCA

Notes: 1) Due to high ranking of AgreeClosest, issue of NC-bal/unbal plays no role. 2) High ranking of Dep[-H] yields avoidance of [-H] default when CC is 8[+H]. 3) Max[+H] >> DepNC >> Max[-H] responsible for different resolution strategies in PL&SG contexts.

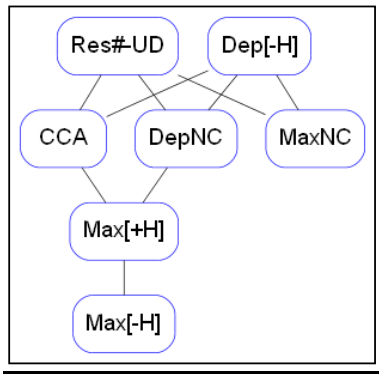
Lang2



[±H]-Bal	[±H]-Unbal	#-Unbal
SG & SG→ def. PL & PL → CCA BUT: Xh&8h→ SM2	SG & SG→ SM2 PL & PL → CCA	PL & SG→ def. SG & PL → CCA

Notes: 1) Due to high ranking of AgreeClosest, issue of NC-bal/unbal plays no role. 2 High ranking of Dep[-H] yields avoidance of [-H] default when CC is 8[+H]. 3) Low ranking of DepNC results in PL&SG always using default.

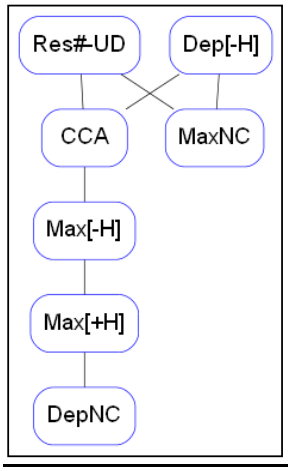
Lang3



[±H]-Bal	[±H]-Unbal	#-Unbal
SG & SG→ def. PL & PL → CCA BUT: Xh&8h→ SM2	SG & SG→ SM2 PL & PL → CCA	PL & SG→ FCA SG & PL → CCA

Notes: 1) Due to high ranking of AgreeClosest, issue of NC-bal/unbal plays no role. 2) High ranking of Dep[-H] yields avoidance of [-H] default when CC is 8H. 3) High DepNC yields FCA in PL&SG

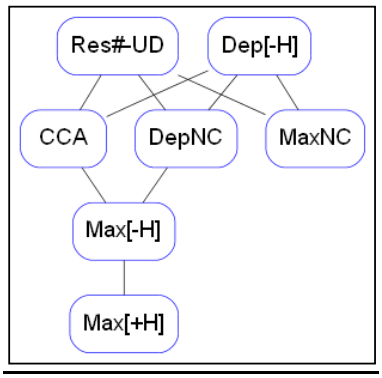
Lang4



[±H]-Bal	[±H]-Unbal	#-Unbal
SG & SG→ def. PL & PL → CCA BUT: Xh&8h→SM2	SG & SG→ SM8 PL & PL → CCA	PL & SG→ def. SG & PL → CCA

Notes: Only difference from Lang2 is that Max[-H] >> Max[+H], so mixed [±Human] coordination yields [-H] default.

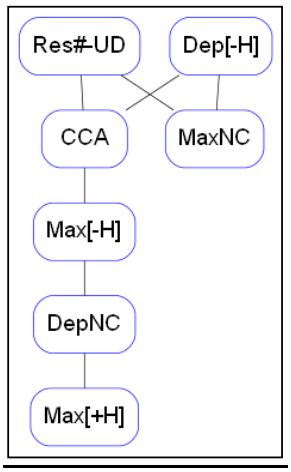
Lang5



[±H]-Bal	[±H]-Unbal	#-Unbal
SG & SG→ def. PL & PL → CCA BUT: Xh&8h→SM2	SG & SG→ SM8 PL & PL → CCA	PL & SG→ FCA SG & PL → CCA

Notes: Only difference from Lang3 is that Max[-H] >> Max[+H], so mixed [±Human] coordination yields [-H] default.

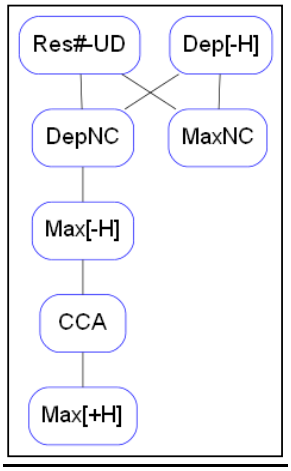
Lang6



[±H]-Bal	[±H]-Unbal	#-Unbal
SG & SG→ def. PL & PL → CCA BUT: Xh&8h→ SM2	SG & SG→ SM8 PL & PL → CCA	PL & SG → def. for [-H], →FCA for [+H] SG & PL → CCA

Notes: Similar to Lang1, except order of Max[-H] and Max[+H] reversed, so 1) mixed [±Human] coordination yields [-H] default, and 2) DepNC below **Max[-H],** but above **Max[+H],** responsible for different resolution strategies in PL&SG contexts.

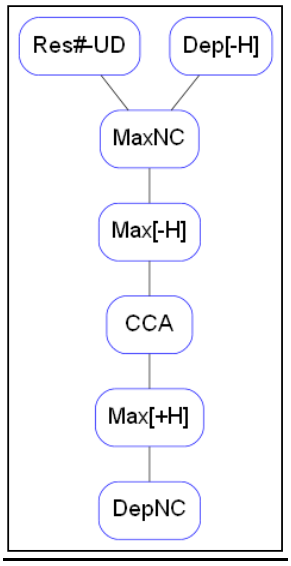
Lang7



[±H]-Bal	[±H]-Unbal	#-Unbal
SG & SG→ def. PL & PL → CCA BUT: Xh&8h→SM2 and: 8i&Xi→SM8	SG & SG→ SM8 PL & PL → CCA BUT 8&X→SM8	PL & SG→ FCA. SG & PL → CCA

Notes: Ranking AgreeClosest below Max[-H] means that if NC8 is in the input, SM8 will be favoured over CCA – BUT because Dep[-H] is highly ranked, SM8 is only favoured if at least one of the conjuncts is non-human.

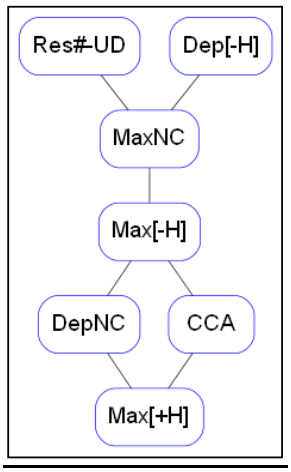
Lang8



[±H]-Bal	[±H]-Unbal	#-Unbal
SG & SG→ def. PL & PL → CCA (B) → CCA (Uh) →8 (Ui) BUT: Xh&8h→ SM2	SG & SG→ SM8 PL & PL → SM8 (U) → morph (B)	PL & SG→ def. SG & PL → CCA (h) → def (i)

Notes: High MaxNC, without equal/higher NC constraint, means that there's a difference between NC-bal and NC-unbal conjuncts – but only for [-Human] conjuncts, because there's another NC constraint above Max[+H] , which prevents [+H] def. Max[-H] >> AgreeClosest; DepNC explains lack of CCA in mixed [±Human] PLs.

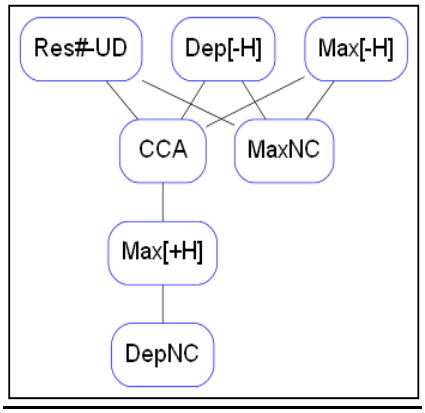
Lang9



[±H]-Bal	[±H]-Unbal	#-Unbal
SG & SG→ def. PL & PL → CCA (B) → CCA (Uh) →8 (Ui) BUT: Xh&8h→ 2	SG & SG→ SM8 PL & PL → SM8 (U) → morph (B)	PL & SG→ FCA(h) → def. (I) SG & PL → CCA (h) → def. (i)

Notes: Like Lang8, except DepNC >> Max[+H] licences FCA for Human PL&SG.

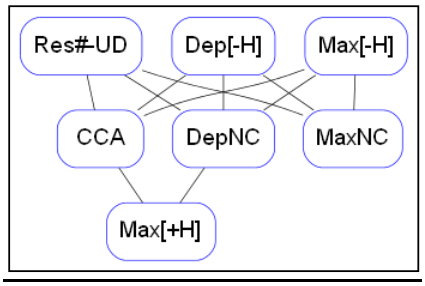
Lang10



[±H]-Bal	[±H]-Unbal	#-Unbal
SG & SG→ def. PL & PL → CCA (h) → def. (i) BUT: Xh&8h→ SM2	SG & SG→ SM8 PL & PL → SM8	PL & SG→ def. SG & PL → CCA (h) → def. (i)

Notes: Like Lang8, except Max[-H] >> AgreeClosest means that morphological agreement is never used when the conjuncts are [-Human].

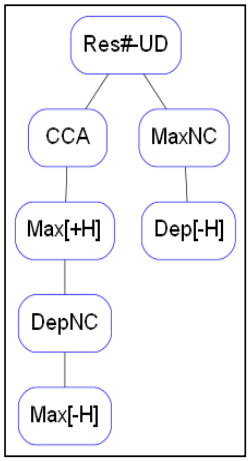
Lang11

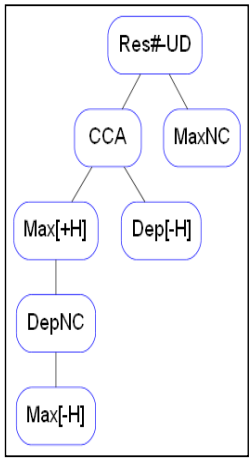


[±H]-Bal	[±H]-Unbal	#-Unbal
SG & SG→ def. PL & PL → CCA (H) → def. (I) BUT: Xh&8h→ SM2	SG & SG→ SM8 PL & PL → SM8	PL & SG→ FCA(h) → def. (I) SG & PL → CCA (h) → 8 (I)

Notes: Like Lang9, except Max[-H] above AgreeClosest means that morphological agreement is never used when the conjuncts are [-H].

Lang12

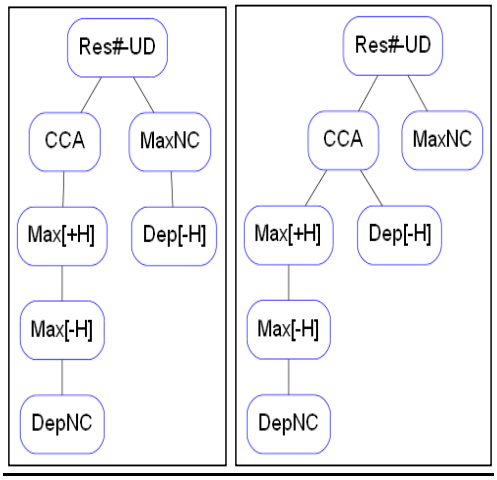




[±H]-Bal	[±H]-Unbal	#-Unbal
SG & SG→ def. PL & PL → CCA	SG & SG→ SM2 PL & PL → CCA	PL & SG→ def (h) → FCA (i) SG& PL → CCA

Notes: 1) Due to high ranking of AgreeClosest, issue of NC-bal/unbal plays no role. 2) Ranking Dep[-H] lower than MaxNC /AgreeClosest allows for "[-H] def" when CC is 8H. 3) Max[+H] >>DepNC>>Max[-H] yields def. in PL&SG for humans but FCA for nonhumans.

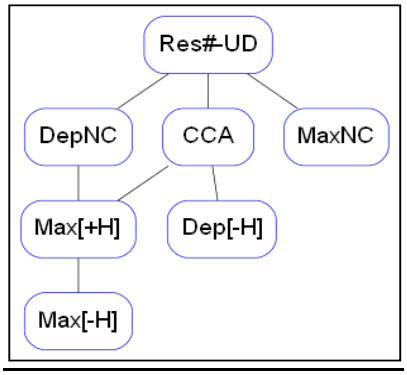
Lang13

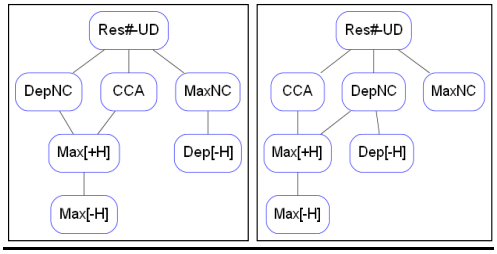


[±H]-Bal	[±H]-Unbal	#-Unbal
SG & SG→ def.	SG & SG→ SM2	PL & SG→ def
PL & PL → CCA	PL & PL → CCA	SG& PL → CCA

Notes: 1) Due to high ranking of AgreeClosest, issue of NC-bal/unbal plays no role. 2) Ranking Dep[-H] lower than MaxNC /AgreeClosest allows for "[-H] def" when CC is 8H.

Lang14

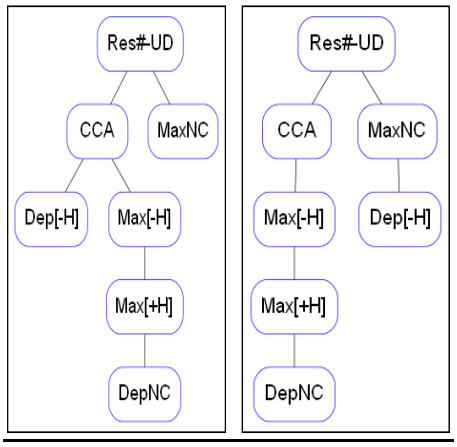




[±H]-Bal	[±H]-Unbal	#-Unbal
SG & SG→ def.	SG & SG→ SM2	PL & SG→ FCA
PL & PL → CCA	PL & PL → CCA	SG& PL → CCA

Notes: All of the NC constraints dominate all of the [±Human] constraints, favouring syntactic agreement with all plural conjuncts. This is the ranking which isiXhosa uses.

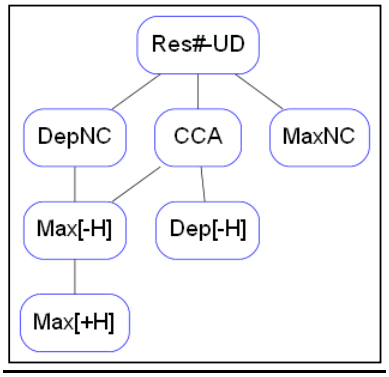
Lang15



[±H]-Bal	[±H]-Unbal	#-Unbal
SG & SG→ def.	SG & SG→ SM8	PL & SG→ def.
PL & PL → CCA	PL & PL → CCA	SG& PL → CCA

Notes: Like Lang13, except Max[-H] >> Max[+H] , so SM8 is used for mixed [±Human] conjuncts.

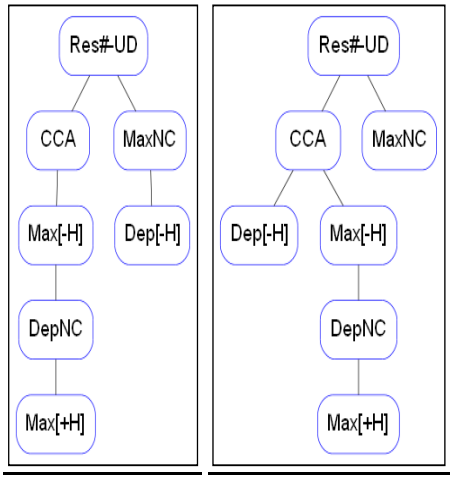
Lang16



[±H]-Bal	[±H]-Unbal	#-Unbal
SG & SG→ def.	SG & SG→ SM8	PL & SG→ FCA
PL & PL → CCA	PL & PL → CCA	SG& PL → CCA

Notes: Like Lang14, except Max[-H] >> Max[+H], so SM8 is used for mixed [±Human] conjuncts.

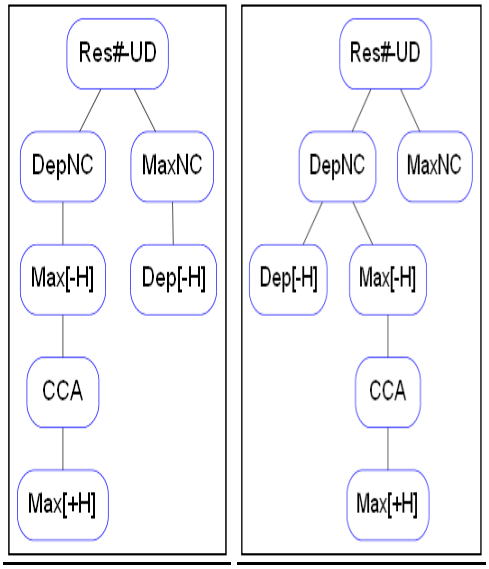
Lang17



[±H]-Bal	[±H]-Unbal	#-Unbal
SG & SG→ def.	SG & SG→ SM8	PL & SG→ def (i) → FCA (h)
PL & PL → CCA	PL & PL → CCA	SG& PL → CCA

Notes: Max[-H] >> DepNC >> Max[+H] responsible for different resolution strategies in PL&SG contexts.

Lang18

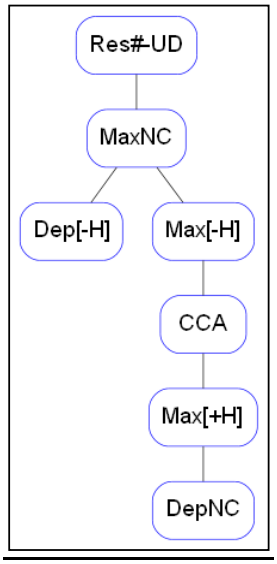


[±H]-Bal	[±H]-Unbal	#-Unbal
SG & SG→ def.	SG & SG→ SM8	PL & SG→ FCA
PL & PL → CCA	PL & PL → CCA	SG& PL → CCA

CCA for plurals, EXCEPT when HC is 8i: Then SM8 is used.

NOTE: For plurals, DepNC favours 2 candidates, which will be decided between by AgreeClosest. BUT if one of the conjuncts is 8i, then Max[-H] chooses SM8 over the other candidate, which could result in FCA, instead of CCA. For mixed [±Human] conjuncts, the NC8 needn't be [-H], as any [-H] conjunct will licence SM8.

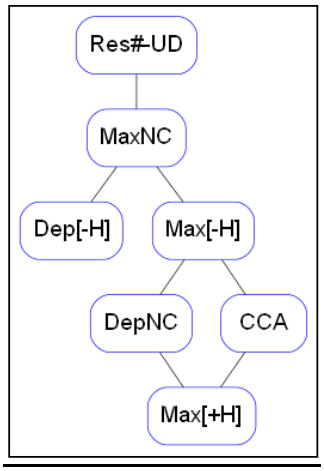
Lang19



[±H]-Bal	[±H]-Unbal	#-Unbal
SG & SG→ def. PL & PL → CCA (B) → CCA (Uh) →SM8 (Ui) BUT Xh&8h →SM2	SG & SG→ SM8 PL & PL → SM8 (U) → morph (B)	PL & SG→ def. SG & PL → CCA (h) → SM8 (i)

Notes: 1) MaxNC >> Dep[-H] allows 8h&8h→SM8, but Dep[-H] >> AgreeClosest prohibits SM8 with NC-unbal. conjuncts when CC is 8h. 2) Low DepNC results in default agreement always being used with PL&SG.

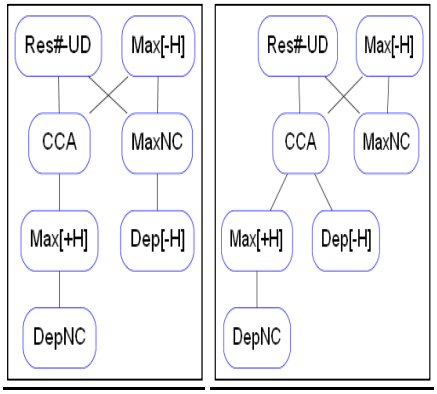
Lang20



[±H]-Bal	[±H]-Unbal	#-Unbal
SG & SG→ def. PL & PL → CCA (B) → CCA (Uh) → SM8 (Ui) BUT xh&8h →FCA	SG & SG→ SM8 PL & PL → SM8 (U) → morph (B)	PL & SG→ def (i) → FCA (h) SG & PL → CCA (h) → SM8 (i)

Notes: 1) MaxNC >> Dep[-H] allows 8h&8h→SM8, but Dep[-H] >> AgreeClosest prohibits SM8 with NC-unbal. conjuncts when CC is 8h 2) DepNC >> Max[+H] allows for FCA.

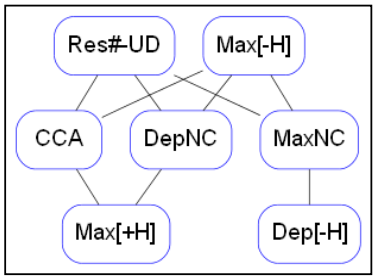
Lang21

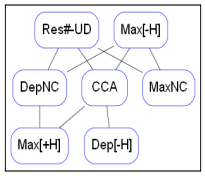


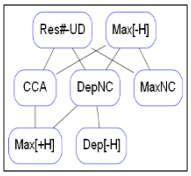
[±H]-Bal	[±H]-Unbal	#-Unbal
SG & SG→ def. PL & PL → CCA (h) →8 (i)	SG & SG→ SM8 PL & PL → SM8	PL & SG→ def SG & PL → CCA (h) → SM8 (i)

Notes: 1) High Max[-H] means default for any coordination involving [-H] conjuncts. NC-constraints over Max[+H] means that syntactic agreement is favoured with plural [+H] conjuncts.

Lang22



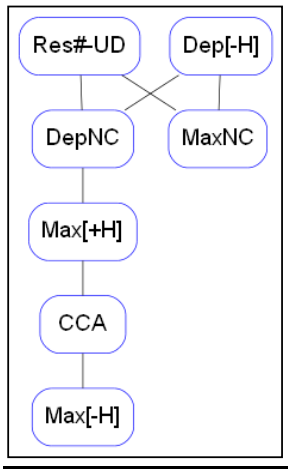




[±H]-Bal	[±H]-Unbal	#-Unbal
SG & SG→ def. PL & PL → CCA (h) →SM8 (i)	SG & SG→ SM8 PL & PL → SM8	PL & SG→ def (i) → FCA (h) SG & PL → CCA (H) → SM8 (i)

Notes: 1) High Max[-H] means default for any coordination involving [-H] conjuncts. All NC-constraints over Max[+H] means that syntactic agreement is favoured with plural [+H] conjuncts. NB: AgreeClosest and DepNC must dominate Max[+H], doesn't matter which NC constraint dominates Dep[-H] (to allow SM8 with [+H] conjuncts)

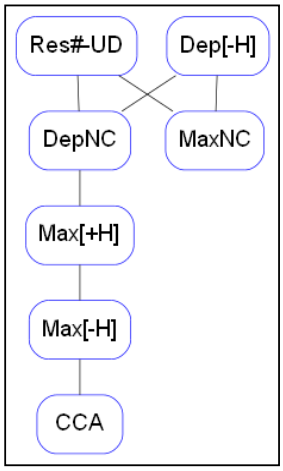
Lang23



[±H]-Bal	[±H]-Unbal	#-Unbal
SG & SG→ def. PL & PL → CCA BUT: 8h&8h→SM2 xh&8h→SMX 2h&Xh→SM2	SG & SG→ SM2 PL & PL → CCA (unless HC is NC2: then SM2)	PL & SG→ FCA SG & PL → CCA

Notes: 1) High Dep[-H] means 8h&8h→SM2, and FCA in Xh&8h (because DepNC is higher than Max[+H]); 2) DepNC >> Max[+H] >> AgreeClosest has result that some sort of NC-agreement is always preferred, and will usually be CCA, unless HC is NC2. AgreeClosest>Max[-H] means that HC NC8 has no effect.

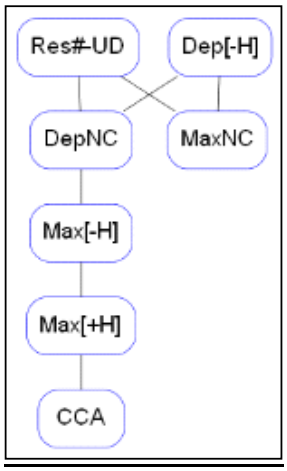
Lang24



[±H]-Bal	[±H]-Unbal	#-Unbal
SG & SG→ def. PL & PL → CCA BUT: 8h&8h→SM2 xh&8h→SMX 2h&Xh→SM2 8i&Xi→SM 8	SG & SG→ SM2 PL & PL → CCA (unless HC is NC2/8: then SM2/8)	PL & SG→ FCA SG & PL → CCA

Notes: 1) High Dep[-H] means 8h&8h→2 and FCA in Xh&8h 2) DepNC >> Max[±H]> > AgreeClosest has result that some sort of NC agreement is always preferred, and will usually be CCA, unless HC is NC of default for the relevant [±H] status.

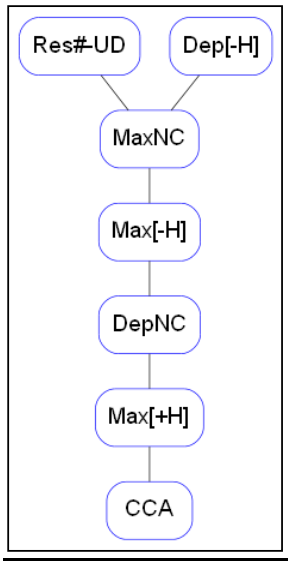
Lang25



[±H]-Bal	[±H]-Unbal	#-Unbal
SG & SG→ def. PL & PL → CCA BUT: 8h&8h→SM2 xh&8h→SMX 2h&Xh→SM2 8i&Xi→ SM8	SG & SG→ SM8 PL & PL → CCA (unless HC is NC2/8: then SM2/8)	PL & SG→ FCA SG & PL → CCA

Notes: Basically same as lg24, except Max[-H] dominates Max[+H], and that means [-H] default in mixed [±human] coordination.

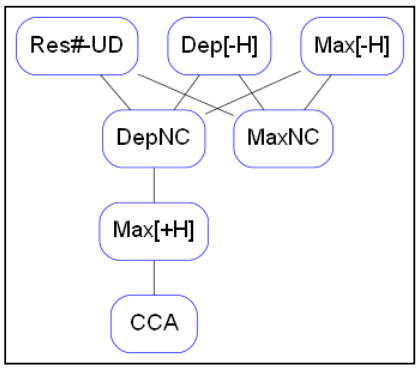
Lang26



[±H]-Bal	[±H]-Unbal	#-Unbal
SG & SG→ def. PL & PL → CCA (H) → morph (Bi) → def (Ui) BUT: 8h&8h→SM2 xh&8h→SMX 2h&Xh→SM2	SG & SG→ SM8 PL & PL → SM8 (U) → morph (B)	PL & SG→ FCA (h) → def (i) SG & PL → CCA (h) → def (i)

Notes: 1) Max[-H] above DepNC means no partial agreement for non-humans, but partial agreement still licensed for human conjuncts.

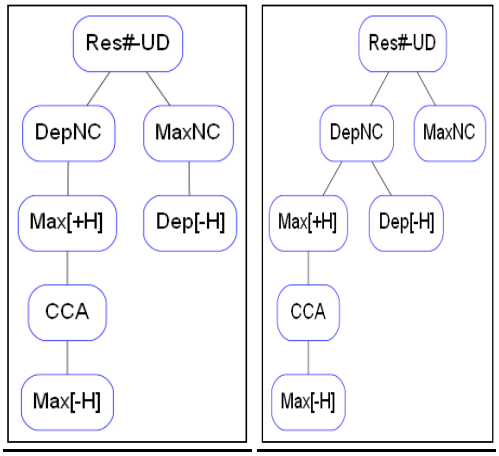
Lang27



[±H]-Bal	[±H]-Unbal	#-Unbal
SG & SG→ def. PL & PL → CCA (h) → def (i) BUT: 8h&8h→SM2 xh&8h→SMX 2h&Xh→SM2	SG & SG→ SM8 PL & PL → SM8	PL & SG→ FCA (h) → def (i) SG & PL → CCA (h) → def (i)

Notes: 1) Max[-H] above NC constraints means no NC agreement for non-humans, but NC agreement still licensed for human conjuncts..

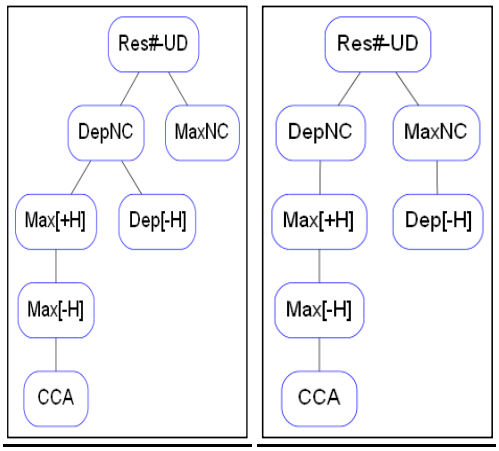
Lang28



[±H]-Bal	[±H]-Unbal	#-Unbal
SG & SG→ def. PL & PL → CCA BUT: xh&8h→SMX 2h&Xh→SM2	SG & SG→ SM2 PL & PL → CCA (unless HC is NC2: then SM2)	PL & SG→ FCA SG & PL → CCA

Notes: 1) Like lg23, except that MaxNC, DepNC >Dep[-H] allows for 8h&8h→SM8; DepNC>Max[+H], Dep[-H] means that when CC is NC8h, SM8 is disallowed, and FCA is favoured over [+H] default.

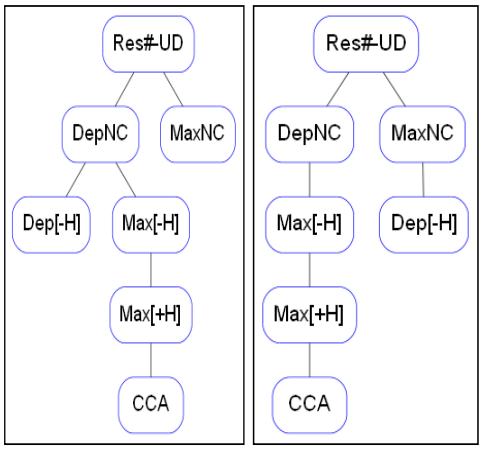
Lang29



[±H]-Bal	[±H]-Unbal	#-Unbal
SG & SG→ def. PL & PL → CCA BUT: xh&8h→SMX 2h&Xh→SM2 8i&Xi→SM8	SG & SG→ SM2 PL & PL → CCA (unless HC is NC2/8: then SM2/8)	PL & SG→ FCA SG & PL → CCA

Notes: Like Lang28, except now Max[-H]>AgreeClosest, so when HC is NC8i, SM8 is favoured over CCA.

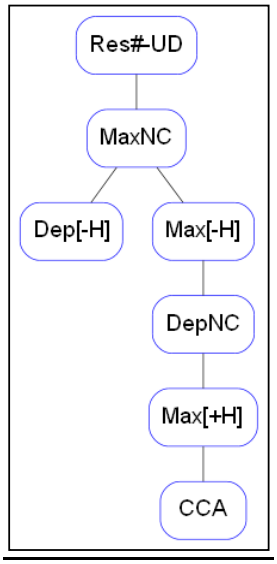
Lang30



[±H]-Bal	[±H]-Unbal	#-Unbal
SG & SG→ def. PL & PL → CCA BUT: xh&8h→SMX 2h&Xh→SM2 8i&Xi→SM8	SG & SG→ SM8 PL & PL → CCA (unless HC is NC2/8: then SM2/8)	PL & SG→ FCA SG & PL → CCA

Notes: Like Lang29, except now Max[-H]>Max[+H], so non-human default chosen in Mixed[±H] coordination.

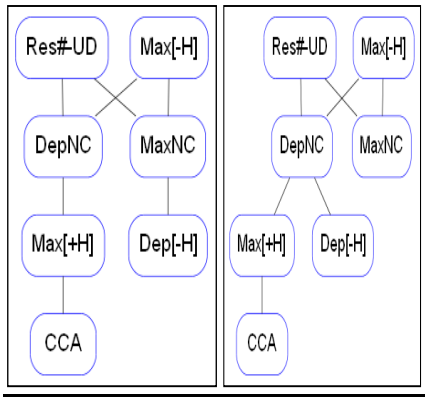
Lang31



[±H]-Bal	[±H]-Unbal	#-Unbal
SG & SG→ def. PL & PL → CCA (h) → morph (Bi) → def (Uh) BUT: xh&8h→SMX 2h&Xh→SM2	SG & SG→ SM8 PL & PL → SM8 (U) → morph (B)	PL & SG→ FCA (h) → def (i) SG & PL → CCA (h) → def (i)

Notes: 1) Max[-H] above DepNC means no partial agreement for non-humans, but partial agreement still licensed for human conjuncts.

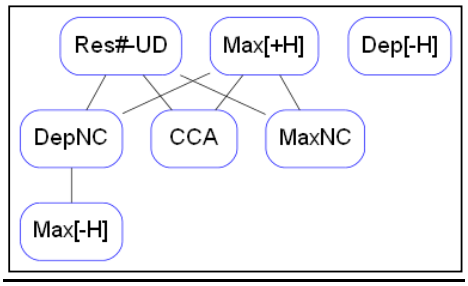
Lang32



[±H]-Bal	[±H]-Unbal	#-Unbal
SG & SG→ def. PL & PL → CCA (h) → def (i) BUT: xh&8h→SMX 2h&Xh→SM2	SG & SG→ SM8 PL & PL → SM8	PL & SG→ FCA (h) → def (i) SG & PL → CCA (h) → def (i)

Notes: 1) Max[-H] above NC constraints means no NC agreement for non-humans.

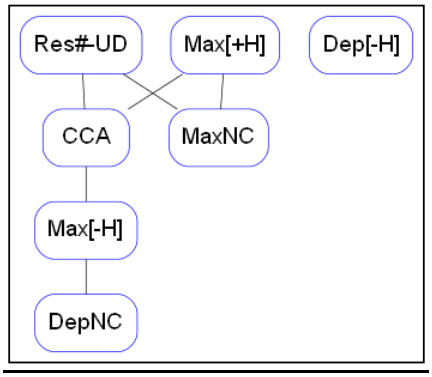
Lang33



[±H]-Bal	[±H]-Unbal	#-Unbal
SG & SG→ def PL & PL → def (h) → CCA (i)	SG & SG→ SM2 PL & PL → SM2	PL & SG→ FCA (i) → def (h) SG & PL → CCA (i) → def (h)

Notes: 1) Max[+H] above NC constraints means no NC agreement for humans; 2) Low Max[-H] means that non-human default is a last resort. 3) High Max[+H] means that Dep[-H] is kind of irrelevant.

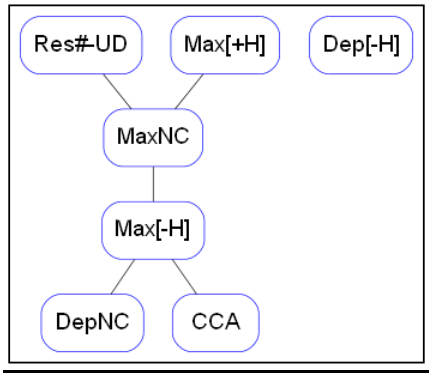
Lang34



[±H]-Bal	[±H]-Unbal	#-Unbal
SG & SG→ def PL & PL → def (h) → CCA (i)	SG & SG→ SM2 PL & PL → SM2	PL & SG→ def SG & PL → CCA (i) → def (h)

Notes: 1) Max[+H] above NC constraints means no NC agreement for humans; 2) High Max[+H] means that Dep[-H] is kind of irrelevant. 3) Low DepNC means that FCA is never licensed.

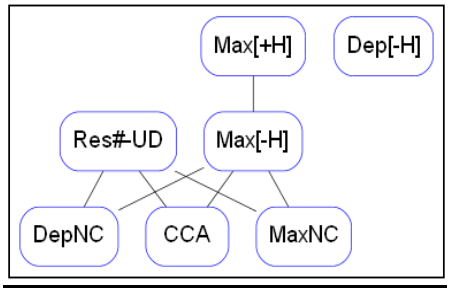
Lang35



[±H]-Bal	[±H]-Unbal	#-Unbal
SG & SG→ def PL & PL → def (H) →Morph (Bi) → def (Ui)	SG & SG→ SM2 PL & PL → SM2	PL & SG→ def SG & PL→ def

Notes: 1) Max[+H] above NC constraints means no NC agreement for humans; 2) High Max[+H] means that Dep[-H] is kind of irrelevant. 3) Partial agreement never licensed.

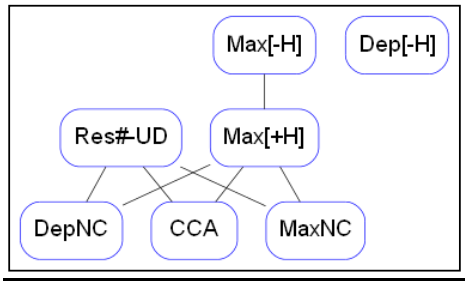
Lang36



[±H]-Bal	[±H]-Unbal	#-Unbal
SG & SG→ def PL & PL → def	SG & SG→ SM2 PL & PL → SM2	PL & SG→ def SG & PL→ def

Notes: All the [±H] constraints dominate all the NC constraints resulting in all default agreement. This is Tsonga variety1.

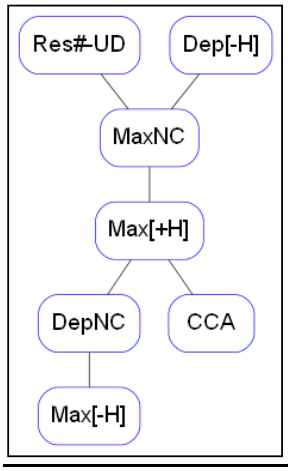
Lang37



[±H]-Bal	[±H]-Unbal	#-Unbal
SG & SG→ def PL & PL → def	SG & SG→ SM8 PL & PL → SM8	PL & SG→ def SG & PL→ def

Notes: Like Lang36, except that Max[-H]>>Max[+H], so non-human default is used in mixed [±H] coordination.

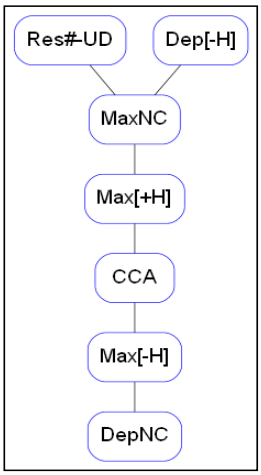
Lang38



[±H]-Bal	[±H]-Unbal	#-Unbal
SG & SG→ def PL & PL → 2 (Uh) →CCA (Ui) →Morph (B) BUT: 8h&8h→SM2	SG & SG→ SM2 PL & PL → SM2 (U) → morph (B)	PL & SG→ def (h) →FCA (i) SG & PL→ def (h) →CCA (i)

Notes: 1) High MaxNC allows for morphological agreement in balanced coordination (except for 8h&8h). 2) Max[+H]> DepNC, AgreeClosest > Max[-H], means default for [+H], but partial agreement for [-H].

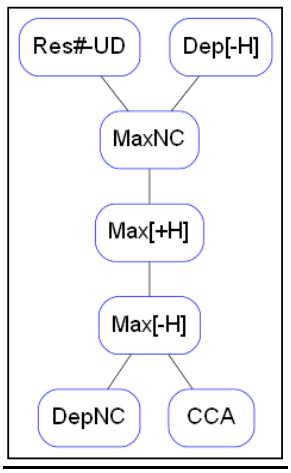
Lang39



[±H]-Bal	[±H]-Unbal	#-Unbal
SG & SG→ def PL & PL → 2 (Uh) →CCA (Ui) →Morph (B) BUT: 8h&8h→SM2	SG & SG→ SM2 PL & PL → SM2 (U) → morph (B)	PL & SG→ def SG & PL→ def (h) →CCA (i)

Notes: Same as Lang38, except no FCA in PL&SG for [-H] nouns now.

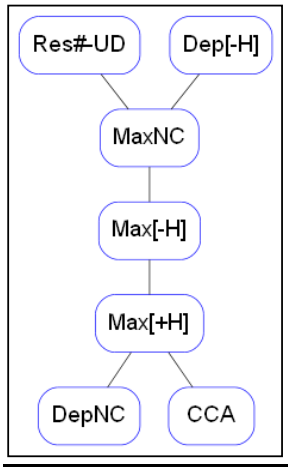
Lang40



[±H]-Bal	[±H]-Unbal	#-Unbal
SG & SG→ def PL & PL → def (U) →Morph (B) BUT: 8h&8h→SM2	SG & SG→ SM2 PL & PL → SM2 (U) → morph (B)	PL & SG→ def SG & PL→ def

Notes: So basically, morphological agreement in PL NC-Bal. coordination, and default everywhere else.

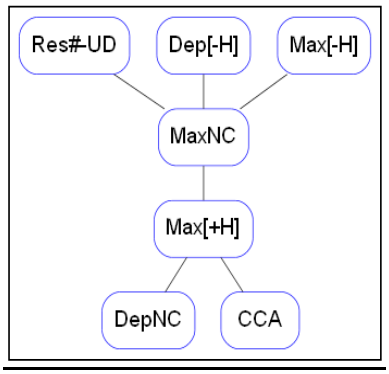
Lang41



[±H]-Bal	[±H]-Unbal	#-Unbal
SG & SG→ def PL & PL → def (U) →Morph (B) BUT: 8h&8h→SM2	SG & SG→ SM8 PL & PL → SM8 (U) → morph (B)	PL & SG→ def SG & PL→ def

Notes: Basically same as Lang40, except Max[-H]>Max[+H], which means SM8 default in mixed [±Human] coordination.

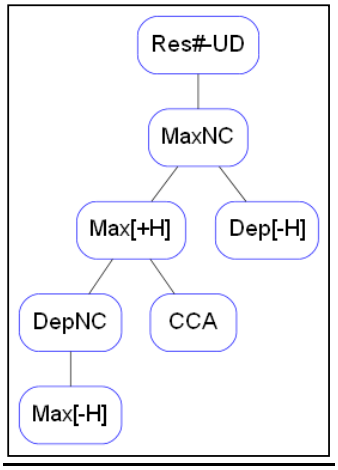
Lang42



[±H]-Bal	[±H]-Unbal	#-Unbal
SG & SG→ def PL & PL → def (U) → def (Bi) →Morph (Bh) BUT: 8h&8h→SM2	SG & SG→ SM8 PL & PL → SM8	PL & SG→ def SG & PL→ def

Notes: Max[-H] dominates all NC constraints, so [-H] conjuncts always use default; NC agreement only ever used with balanced [+H] conjuncts (except for 8H).

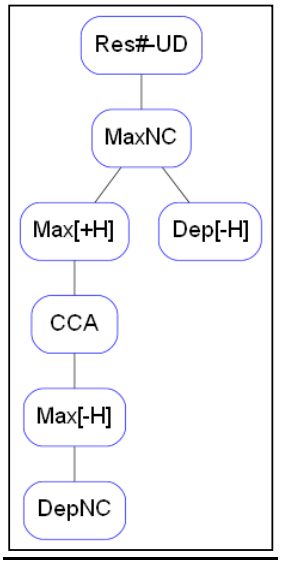
Lang43



[±H]-Bal	[±H]-Unbal	#-Unbal
SG & SG→ def PL & PL → CCA (Ui) → def (Uh) →Morph (B)	SG & SG→ SM2 PL & PL → SM2 (U) →morph (B)	PL & SG→ def (h) →FCA (i) SG & PL→ def (h) →CCA (i)

Notes: [+H] conjuncts only ever use NC agreement when balanced. [-H] default only a last resort, with Max[-H] being so low.

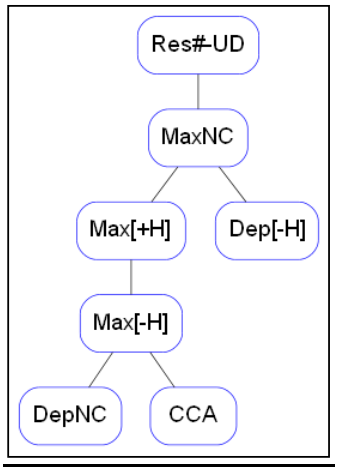
Lang44



[±H]-Bal	[±H]-Unbal	#-Unbal
SG & SG→ def PL & PL → CCA (Ui) → def (Uh) →Morph (B)	SG & SG→ SM2 PL & PL → SM2 (U) →morph (B)	PL & SG→ def SG & PL→ def (h) →CCA (i)

Notes: Basically same as Lang43, except Max[-H]>DepNC does not license FCA in PL&SG contexts.

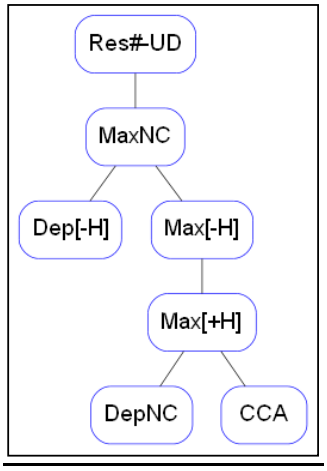
Lang45



[±H]-Bal	[±H]-Unbal	#-Unbal
SG & SG→ def PL & PL → def (U) →Morph (B)	SG & SG→ SM2 PL & PL → SM2 (U) → morph (B)	PL & SG→ def SG & PL→ def

Notes: Basically like Lang40, except that 8h&8h→SM8 due to MaxNC>Dep[-H].

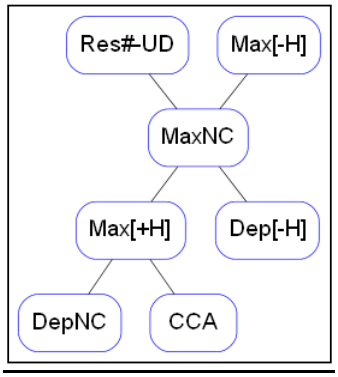
Lang46



[±H]-Bal	[±H]-Unbal	#-Unbal
SG & SG→ def PL & PL → def (U) →Morph (B)	SG & SG→ SM8 PL & PL → SM8 (U) → morph (B)	PL & SG→ def SG & PL→ def

Notes: Basically like Lang45, except Max[-H]>Max[+H], which means [-H] default in mixed [±human] coordination.

Lang47



[±H]-Bal	[±H]-Unbal	#-Unbal
SG & SG→ def	SG & SG→ SM8	PL & SG→ def
PL & PL → def (U) →Morph (Bh) →Def (Bi)	PL & PL → SM8	SG & PL→ def

Notes: Basically like lg46, except Max[-H]>MaxNC (as well as other NC Constraints), which means [-H] default anytime there's a [-H] conjuncts. ALSO: this is the version of Lang42 where 8h&8h→8.

Bibliography

Aoun, J., Benmamoun, E., and Sportiche, D. (1999). Further remarks on first conjunct agreement. *Linguistic Inquiry*, 30(4):669–681.

Aronoff, M. and Xu, Z. (2010). A realization Optimality-Theoretic approach to affix order. *Morphology*, 20(2):381–411.

Badecker, W. (2007). A feature principle for partial agreement. *Lingua*, 117:1541–1565.

Baković, E. (2006). Partial identity avoidance as cooperative interaction. In Alcazar, A., Martinez, T., and Hernandez, R., editors, *Proceedings of the Thirty-Third Western Conference on Linguistics*, volume 16, pages 13–25, Fresno, CA. WECOL 2004.

Baumbach, E. J. M. (1987). *Analytical Tsonga grammar*. University of South Africa, Pretoria, South Africa.

Bokamba, E. G. (1985). Verbal agreement as a non-cyclic rule in Bantu. In Goyvaerts, D. L., editor, *African linguistics: Essays in memory of M. W. K. Semikenke*, pages 9–54. Benjamins, Amsterdam, Netherlands.

Bosch, S. (1988). Aspects of subject conjunction in Zulu. *South African Journal of African Languages*, 8(3):70–74.

Bresnan, J. (2000). Optimal syntax. In Dekkers, J., van der Leeuw, F., and van der Weijer, J., editors, *Optimality Theory: Phonology, syntax, and acquisition*. Oxford University Press, Oxford, UK.

Bresnan, J. (2001). Explaining morphosyntactic competition. In Baltin, M. and Collins, C., editors, *The handbook of contempory syntactic theory*, pages 11–44. Blackwell, Oxford, UK.

Buell, L. (2012). Class 17 as a non-locative noun class in Zulu. *Journal of African Languages and Linguistics*, 33(1):1–35.

Chatelain, W. and Junod, H. A. (1909). *Pocket dictionary of Thonga (Shangaan) - English, English - Thonga (Shangaan)*. George Bridel & Co, Lausanne, Switzerland.

Chinque, G. (1988). On *si* constructions and the theory of *Arb*. *Linguistic Inquiry*, 19:521–582.

Chomsky, N. (1964). Current issues in linguistic theory. In Fodor, J. and Katz, J., editors, *The structure of language*. Prentice-Hall, Englewood Cliffs, NJ.

Chomsky, N. (1965). *Aspects of the theory of syntax*. MIT Press, Cambridge, MA.

Chomsky, N. (1986). *Knowledge of language: Its nature, origin, and use*. Praeger, New York, NY.

Chomsky, N. (1995). *The Minimalist Program*. MIT Press, Cambridge, MA.

Chomsky, N. (2000). Minimalist inquiries: The framework. In Martin, R., Michaels, D., and Uriagereka, J., editors, *Step by step: Essays on Minimalist Syntax in honour of Howard Lasnik*, pages 89–155. MIT Press, Cambridge, MA.

Chomsky, N. and Halle, M. (1968). *The sound pattern of English*. Harper & Row, New York, NY.

Corbett, G. (1983a). *Hierarchies, targets and controllers: Agreement patterns in Slavic*. Croom Helm, London, UK.

Corbett, G. (1983b). Resolution rules: Agreement in person, number, and gender. In Gazdar, G., Klein, E., and Pullum, G., editors, *Order, concord and constituency*, pages 175–206. Foris Publications, Dordrecht, Netherlands.

Corbett, G. (1985). Gender and noun classes. In Shopen, T., editor, *Language typology and syntactic description III: Grammatical categories and the lexicon*, pages 241–279. Cambridge University Press, Cambridge, UK, 2007 edition.

Corbett, G. (1991). *Gender*. Cambridge University Press, Cambridge, UK.

Corbett, G. (2006). *Agreement*. Cambridge University Press, Cambridge, UK.

Corbett, G. and Mtenje, A. (1987). Gender agreement in Chichewa. *Studies in African Linguistics*, 18(1):1–38.

Crabtree, W. (1923). *Elements of Luganda grammar*. S.P.C.K., London, UK, 2nd edition.

Cuenod, R. (1967). *Tsonga-English dictionary*. Swiss Mission in South Africa, Johannesburg, South Africa.

de Swart, P., Lamers, M., and Lestrade, S. (2008). Animacy, argument structure, and argument encoding. *Lingua*, 118:131–140.

de Vos, M. and Mitchley, H. (2012). Subject marking and preverbal coordination in Sesotho: A perspective from Optimality Theory. *South African Linguistics and Applied Language Studies*, 30(2):155–170.

Dimmendaal, G. J. (1983). *The Turkana language*. Foris Publications, Dordrecht, Netherlands.

Dixon, R. M. W. (1972). *The Dyirbal language of North Queensland*. Cambridge University Press, Cambridge, UK.

Dowling, T. (1998). *Speak Xhosa with Us: Beginners to Advanced Course*. Mother Tongues Multimedia Development, Cape Town, South Africa.

Duarte, F. B. (2011). Tense encoding, agreement patterns, definiteness and relativization strategies in Changana. In Bokamba, E. G., Shosted, R. K., and Ayalew, B. T., editors, *Selected Proceedings of the 40th Annual Conference on African Linguistics: African Languages and Linguistics Today*, pages 80–94, Somerville, MA. Cascadilla Proceedings Project.

Givón, T. (1970). The resolution of gender conflicts in Bantu conjunction: When syntax and semantics clash. In *Papers from the Sixth Regional Meeting, Chicago Linguistics Society*, pages 250–261. Chicago Linguistics Society.

Givón, T. (1971). Some historical changes in the noun class system of Bantu, their possible causes and wider implications. In Kim, C. W. and Stahlke, H., editors, *Papers in African Linguistics*, pages 33–54. Linguistic Research, Inc, Edmonton, Canada.

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

Grimshaw, J. (1997b). Projection, heads, and optimality. *Linguistic Inquiry*, 28:373–422.

Grimshaw, J. (2001). Optimal clitic positions and the lexicon in Romance clitic systems. In Legendre, G., Grimshaw, J., and Vikner, S., editors, *Optimality-Theoretic syntax*, pages 205–240. MIT Press, Cambridge, MA.

Guthrie, M. (1948). *The classification of the Bantu languages*. International African Institute, London, UK.

Halle, M. and Marantz, A. (1993). Distributed morphology and the pieces of inflection. In Hale, K. and Jay Keyser, S., editors, *The View from Building 20*, pages 111–176. MIT Press, Cambridge, MA.

Hristov, B. (2011). Animacy and agreement with conjoined nouns in Bulgarian. In Absi, Z., Althobaiti, M., Heyer, V., Ogawa, M., Patterson, C., and Soultatis, T., editors, *Language at the University of Essex (LangUE) 2010 Proceedings*, pages 51–63. Department of Language and Linguistics, University of Essex.

Johannessen, J. (1996). Parial agreement and coordination. *Linguistic Inquiry*, 27:661–676.

Johannessen, J. (1998). *Coordination*. Cambridge University Press., Cambridge.

Kager, R. (1999). *Optimality Theory*. Cambridge University Press, Cambridge, UK.

Kisseberth, C. (1970). On the functional unity of phonological rules. *Linguistic Inquiry*, 1:291–306.

Kriel, T. J. (1950). *The new Sesotho-English dictionary*. APB Publishers and booksellers, Johannesburg, South Africa.

Krifka, M. (1995). Swahili. In Jacobs, J., von Stechow, A., Sternefeld, W., and Vennemann, T., editors, *Syntax: Ein internationales Handbuch zeitgenoessischer Forschung [Syntax: An international handbook of contemporary research]*, volume 2, pages 1397–1418, Berlin, Germany. Walter de Gruyter.

Leben, W. (1973). *Suprasegmental phonology.* PhD thesis, Massachusetts Institute of Technology.

Lewis, M. P., editor (2009). *Ethnologue: Languages of the world*. SIL International, Dallas, TX, sixteenth edition. Online version: http://www.ethnologue.com/16; accessed January 2014.

Maho, J. P. (2009). The online version of the new updated Guthrie list, a referential classification of the Bantu languages. http://goto.glocalnet.net/mahopapers/nuglonline.pdf.

Marten, L. (2000). Agreement with conjoined noun phrases in Swahili. *Afrikanistische Arbeitspapiere: Swahili Forum VII*, 64:75–96.

Marten, L. (2005). The dynamics of agreement and conjunction. *Lingua*, 115(4):527–547.

Marušič, F. and Nevins, A. (2009). Two types of neuter: Closest-conjunct agreement in the presence of '5 and ups'. In *Formal Approaches to Slavic Linguistics*, volume 18, Ithaca NY. Cornell University, http://nrs.harvard.edu/urn-3:HUL.InstRepos:3427430.

McCarthy, J. (1986). OCP effects: Gemination and antigemination. *Linguistic Inquiry*, pages 207–263.

McCarthy, J. (2008). *Doing Optimality Theory: Applying theory to data*. Blackwell, Malden, MA, & Oxford.

Milroy, L. (1987). *Language and social networks*. Blackwell, Oxford.

Mokoena, A. (1998). *Sesotho made easy: A step-by-step guide*. African Languages Made Easy Series. J L van Schaik Publishers, Pretoria, South Africa.

Moloto, A. N. (1992). Agreement and concord with conjoined noun phrases in Northern Sotho. Master's thesis, University of Stellenbosch, Stellenbosch, South Africa.

Moshi, L. (1998). Word order in multiple object constructions in Kivunjo-Chaga. *Journal of African Languages and Linguistics*, 19(2):137–152.

Mous, M. and Mreta, A. (2004). The comitative/conjunctive preposition *na* in Chathu (Pare), a Bantu language of Tanzania. *South African Journal of African Languages*, 4:219–229.

Nabe, H., Dreyer, P., and Kakana, G., editors (1976). *Educum Xhosa Dictionary*. Educum, Johannesburg, South Africa.

Newell, H. (2005). A consideration of feminine default gender. Master's thesis, University of Cincinnati, Cininnati.

Noyer, R. (1997). *Features, positions, and affixes in autonomous morphological structure*. Garland, New York. Revised version of 1992 MIT Doctoral Dissertation.

Nyembezi, S. (1970). *Learn more Zulu*. Shuter and Shooter, Pietermaritzburg, South Africa.

Prince, A. and Smolensky, P. (1993/2004). *Optimality Theory: Constraint interaction in generative grammar*. Blackwell, Malden, MA, & Oxford. [Revision of 1993 technical report, Rutgers University Center for Cognitve Science. ROA-537].

Prince, A., Tesar, B., and Merchant, N. (2014). Otworkplace X69. Freeware. https://sites.google.com/site/otworkplace/.

Pullum, G. and Zwicky, A. (1986). Phonological resolution of syntactic feature conflict. *Language*, pages 751–773.

Reynierse, C., editor (1991). *South African Multi-Language Dictionary and Phrase Book*. The Reader's Digest Association South Africa, Cape Town.

Riedel, K. (2009). *The Syntax of Object Marking in Sambaa: A Comparative Bantu Perspective*. PhD thesis, Leiden University, LOT, Utrecht.

Roberts, L. and Wolentis, M. (1974). Conjunction and concord in Bantu. In Voeltz, E., editor, *Third Annual Conference in African Linguistics*, volume 7 of *Indiana University Publications, African Series*, pages 231–242, Bloomington. Indiana University, Research Center for the Language Sciences.

Russell, K. (1995). Morphemes and candidates in optimality theory. *ROA-44*.

Saussure, F. d. (1916). *Cours de linguistique générale*. Payot, Paris.

Schadeburg, T. (1992). *A sketch of Swahili morphology*. R. Köppe, Köln.

Short, D. (1993). Czech. In Comrie, B. and Corbett, G., editors, *The Slavonic languages*, pages 455–532. Routledge, London, UK.

Simango, S. R. (2012). Subject marking, coordination and noun classes in ciNsenga. *South African Linguistics and Applied Language Studies*, 30(2):171–183.

Statistics South Africa (2012). Census 2011: Census in brief. Technical Report 03-01-41, Statistics South Africa, Pretoria.

Steel, S. (1978). *Universals of human language IV: Syntax*, pages 585–623. Stanford University Press, Stanford.

Steinmetz, D. (2006). Gender shifts in Germanic and Slavic: Semantic motivation for neuter? *Lingua*, 116:1418–1440.

Stump, G. (2001). *A theory of paradigm structure*. CUP, Cambridge.

Trommer, J. (2001). *Distributed Optimality*. PhD thesis, University of Potsdam, Potsdam.

Trommer, J. (2002). Modularity in OT-morphosyntax. In Fanselow, G. and Féry, C., editors, *Resolving conflicts in grammar: Optimality Theory in syntax, morphology and phonology*, Special Issue 11 of Linguistische Berichte, pages 83–117. Helmut Buske Verlag, Hamburg.

van Koppen, M. (2005). *One probe - two goals: Aspects of agreement in Dutch dialects*. PhD thesis, Leiden University, LOT, Utrecht.

van Wyk, E. B. (1957). An augmentative noun class in Tsonga. *African Studies*, 16(1):25–36.

Voeltz, E. (1971). Surface constraints and agreement resolution: Some evidence from Xhosa. *Studies in African Linguistics*, 2(1):37–60.

Webb, V. (2008). Overview of issues at stake. In Lafon, M. and Webb, V., editors, *The standardisation of African languages: Language political realities*, pages 7–21, Johannesburg, South Africa. CentRePol and IFAS.

Woolford, E. (1999). Animacy hierarchy effects on object agreement. In Kotey, P., editor, *New Dimensions in African Linguistics and Languages (Trends in African Linguistics 3)*, pages 203–216, Trenton, NJ. Africa World Press.

Xu, Z. (2007). *Inflectional morphology in Optimality Theory*. PhD thesis, Stony Brook University, New York.

Xu, Z. (2011). Optimality theory and morphology. *Language and Linguistics Compass*, 5(7):466–484.

Yip, M. (1998). Identity avoidance in phonology and morphology. In Lapointe, S., Brentari, D., and Farrell, P., editors, *Morphology and its relation to phonology and syntax*, pages 216–246. CSLI, Stanford.

Zhang, N. (2009). *Coordination in syntax*. Cambridge University Press, New York.

Zoerner, E. (1995). *Coordination: The syntax of &P*. PhD thesis, University of California, Irvine, CA.