Feature Indeterminacy and Feature Resolution

Mary Dalrymple; Ronald M. Kaplan

Language, Vol. 76, No. 4. (Dec., 2000), pp. 759-798.

Stable URL:
http://links.jstor.org/sici?sici=0097-8507%28200012%2976%3A4%3C759%3AFIAFR%3E2.0.CO%3B2-R

Language is currently published by Linguistic Society of America.

FEATURE INDETERMINACY AND FEATURE RESOLUTION

MARY DALRYMPLE AND RONALD M. KAPLAN

Xerox Palo Alto Research Center

Syntactic features like CASE, PERSON, and GENDER are often assumed to have simple atomic values that are checked for consistency by the standard predicate of equality. The CASE feature has values such as NOM or ACC, and values like MASC and FEM are assumed for the feature GENDER. But such a view does not square with some of the complex behavior these features exhibit. It allows no obvious account of FEATURE INDETERMINACY (how a particular form can satisfy conflicting requirements on a feature like CASE), nor does it give an obvious account of FEATURE RESOLUTION (how PERSON and GENDER features of a coordinate noun phrase are determined on the basis of the conjuncts). We present a theory of feature representation and feature checking that solves these two problems, providing a straightforward characterization of feature indeterminacy and feature resolution while sticking to structures and standard interpretations that have independent motivation. Our theory of features is formulated within the LFG framework, but we believe that similar solutions can be developed within other syntactic approaches.*

INTRODUCTION. Well-known examples of FEATURE INDETERMINACY, such as 1, show that a single phrase can appear to simultaneously fulfill conflicting syntactic requirements (Dyla 1984, Zaenen & Karttunen 1984, Pullum & Zwicky 1986, Zwicky 1991).

(1) Kogo Janek lubi a Jerzy nienawidzi? (Polish)
who Janek likes and Jerzy hates
? OBJ CASE = ACC OBJ CASE = GEN
'Who does Janek like and Jerzy hate?'

In ex. 1, the pronoun *kogo* is the object of two coordinated verbs; the first verb requires accusative case for its object, while the second requires genitive case. *Kogo* can fill both of these requirements, as the accusative and genitive forms for this particular pronoun are the same. It is not possible to characterize the pronoun in 1 as bearing either accusative or genitive case, since it must satisfy both requirements simultaneously. Example 2 (Groos & van Reimsdijk 1979, Ingria 1990, Bayer 1996) shows that such examples can also appear in noncoordinate structures.

(2) Ich habe gegessen was übrig war. (German)
I have eaten what left was
OBJ CASE = ACC ? SUBJ CASE = NOM
'I ate what was left.'

On the basis of such examples it has been argued (Ingria 1990, Johnson & Bayer 1995) that Boolean combinations of equality constraints over atomic values like those standardly used in constraint-based grammatical theories are not useful for enforcing

* For helpful discussion and assistance, we are grateful to Mark Aronoff, Kristjan Arnason, Sam Bayer, Jim Blevins, Joan Bresnan, Miriam Butt, Greville Corbett, Chris Culy, Amy Dahlstrom, Mark Donohue, Anette Frank, Mark Johnson, Lauri Karttunen, Stefan Kaufmann, Martin Kay, Paul Kiparsky, Tracy Holloway King, Terry Langendoen, John Maxwell, Stefan Müller, Paula Newman, María-Eugenia Niño, Toshiki Osada, Christian Rohrer, Louisa Sadler, Ivan Sag, Hadar Shemtov, Petr Sgall, Peggy Speas, Nigel Vincent, Steve Wechsler, Annie Zaenen, Arnold Zwicky, two anonymous referees, and audiences at the Linguistics Colloquium at Stanford University and at the LFG97 Conference in San Diego. A preliminary version of this work appeared as Dalrymple & Kaplan 1997.

feature-matching requirements such as these, and even that constraint-based frameworks in general are incapable of handling such phenomena. Some authors have suggested instead that additional devices of various sorts are needed to provide the necessary grammatical characterizations.

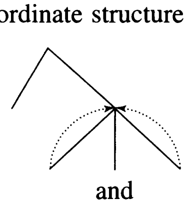
We agree with the particular claim that equality constraints over atomic feature values cannot adequately account for these patterns of dependency, but we take issue with the general conclusion that this calls for a wholesale revision to the underlying theory of grammar. Rather, we see these and other recalcitrant patterns of agreement as problems that can be solved within a constraint-based architecture by means of only minor adjustments to representations and mechanisms. In fact, we can overcome these difficulties by a very simple change of representation: we allow feature values to be SETS of items as well as individual symbols. The CASE value for an ordinary nominative noun is represented as the singleton set {NOM} instead of the symbol NOM, and an accusative noun has the value {ACC}. A word such as *was* in 2, which can satisfy both nominative and accusative requirements, has {NOM,ACC} as its CASE value, where the different elements indicate the different possibilities of realization. Finally, the contextual case-agreement requirements are imposed by set-membership assertions instead of equality conditions. We present more examples of feature indeterminacy and our solution for this kind of complex behavior in §4.

The phenomenon of FEATURE RESOLUTION in coordinate noun phrases raises a different set of challenges for atomic-value-with-equality approaches. The problem is to give a principled account of how certain features of conjoined elements are resolved to determine the agreement features of the coordination as a whole. For example, when one of the conjuncts in a coordinate noun phrase in Slovak is first person, verb agreement shows that the person of the coordinate phrase is also first person:

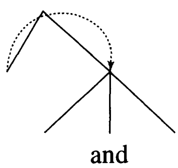
(3) Ja a ty sme bratia. (Slovak)
I and you are.1PL brothers
'You and I are brothers.'

This problem and the analogous problem for gender have been studied extensively by Corbett (1983a,b, 1991), but many formal accounts of agreement address this issue only partially, if at all. If the PERSON feature is represented as an atomic value such as 1 or 2, it is not clear how the rules for determining the person of coordinate phrases can be formulated in a non-ad-hoc way for examples like 3, since there is no straightforward way of operating on the features 1 and 2 of the individual conjuncts in 3 to obtain the feature 1 for the coordinate subject noun phrase.

We investigate the feature resolution problem in a range of languages and propose a general and uniform account of feature resolution in coordinate constructions. Here as well we make use of set-valued features. We assign set values to features of the individual conjuncts, {S,H} for first person and {H} for second person. We then take the union of the conjunct sets, in this case {S,H}, to be the feature value for the coordinate structure (4). The features required by the verb must be compatible with the features of the coordinate structure as a whole (5).

(4) 

Feature resolution (person, gender)

(5) 

Feature checking in coordination

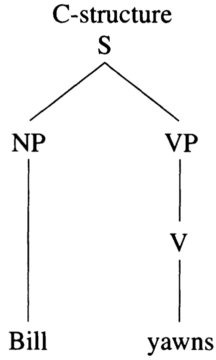
Our general approach to coordinate feature resolution is outlined in §5, our account of resolution of the PERSON feature in §6 and of the GENDER feature in §7.

1. BACKGROUND: LEXICAL FUNCTIONAL GRAMMAR. Our theory of feature indeterminacy and feature resolution is formulated within the framework of LEXICAL FUNCTIONAL GRAMMAR. We provide a brief summary of the general framework. For a more complete explication of the syntactic assumptions of LFG, see Bresnan 1982, Dalrymple et al. 1995, and the references cited there.

LFG postulates two syntactic levels of representation. Constituent structure (C-STRUCTURE) encodes phrasal dominance and precedence relations and is represented as a phrase structure tree. Functional structure (F-STRUCTURE) encodes syntactic predicate-argument structure and is represented as an attribute-value matrix. The c-structure and abbreviated f-structure for sentence 6 are given in 7.

(6) Bill yawns.

(7) C-structure



F-structure

$$\begin{bmatrix} \text{PRED} & \text{`YAWN'} \\ \text{TENSE} & \text{PRES} \\ \text{SUBJ} & [\text{PRED `BILL'}] \end{bmatrix}$$

An f-structure is a function from attributes such as PRED, SUBJ, and TENSE to values like PRES, semantic forms like 'YAWN' or 'BILL', f-structures, or to sets of such values. Coordinate structures can be represented as sets of f-structures, as proposed by Kaplan & Bresnan 1982, Bresnan et al. 1985 and Kaplan & Maxwell 1988; complex values also can be represented as sets.

The properties of f-structures that correspond to c-structure nodes are specified by annotations on phrase structure rules and lexical entries. The annotated phrase-structure rule in 8 is used in the analysis of sentence 6.

$$(8)\ \ S \rightarrow \underset{(\uparrow \text{SUBJ}) = \downarrow}{\text{NP}} \quad \underset{\uparrow = \downarrow}{\text{VP}}$$

In each rule or lexical entry, the ↑ metavariable refers to the f-structure corresponding to the mother c-structure node, and the ↓ metavariable refers to the f-structure corresponding to the nonterminal labeled by the assertion (Kaplan & Bresnan 1982:183). The assertions on the rule in 8 indicate that the f-structure for the S (↑ in the annotation on the NP node) has a SUBJ attribute whose value is the f-structure for the NP daughter (↓ in the annotation on the NP node), and that the S node corresponds to an f-structure that is the same as the f-structure for the VP daughter. When the phrase-structure rule for S is used in the analysis of a given sentence, the metavariables ↑ and ↓ are instantiated to particular f-structures placed in correspondence with nodes of the c-structure. The analysis is acceptable if those f-structures satisfy all the requirements that the phrasal annotations specify as well as the assertions associated with lexical entries.

Lexical entries also use the metavariables ↑ and ↓ to encode information about the f-structures of the preterminal nodes that immediately dominate them. Lexical entries for the words *Bill* and *yawns* are:

(9) *Bill* NP (↑ PRED) = 'BILL'
yawns V (↑ PRED) = 'YAWN'
(↑ TENSE) = PRES

For example, the assertion (↑ PRED) = 'BILL' states that the preterminal node immediately dominating the terminal symbol *Bill* has an f-structure whose value for the attribute PRED is 'BILL'.

The feature assertions in 8 and 9 are satisfied by the f-structure in 7 but also by all larger f-structures that contain other combinations of attributes and values besides the ones shown in 7. These larger structures have properties that are not in any way related to the words or phrases of the sentence, and, as discussed in Kaplan & Bresnan 1982, they are excluded from grammatical consideration. The f-structure assigned to an utterance is the SMALLEST f-structure that satisfies all the assertions that come from its words and phrases. This is the MINIMAL MODEL for the collection of feature assertions, since it satisfies all lexical and phrasal constraints but contains no additional attributes or values.

The assertions presented above are taken to DEFINE the minimal model for the f-structure of an utterance. LFG also provides for CONSTRAINING EQUATIONS, which require the minimal model for the defining equations of an utterance to have certain properties. Constraining equations are distinguished from other assertions by the presence of a subscript *c*, as in 10.

(10) (↑ NUM) $=_c$ PL

We will make use of constraining equations in our analysis of feature resolution in coordination.

2. TRADITIONAL APPROACHES TO FEATURE REPRESENTATION. It is well-known that some words, such as verbs, can place constraints on features of their arguments, in particular on the person, number, gender, or case values. In the following Hindi example (McGregor 1972), the verb requires its subject to be third person, singular, masculine, and nominative.

(11) Ram calegaa. (Hindi)
Ram.MASC will go.MASC.3SG
'Ram will go.'

We represent the agreement relationship by an arrow: Here, the verb *calegaa* agrees with the subject *Ram*. Under traditional LFG assumptions about the form of agreement features, the f-structure for example 11 is 12. In LFG the agreement constraints contributed by the lexical entries for *Ram* and *calegaa* are instantiated as 13.

(12)
$$f: \begin{bmatrix} \text{PRED 'GO'} & \\ \text{TENSE PRES} & \\ \text{SUBJ } g: & \begin{bmatrix} \text{PRED 'RAM'} \\ \text{PERSON 3} \\ \text{NUM SG} \\ \text{GENDER MASC} \\ \text{CASE NOM} \end{bmatrix} \end{bmatrix}$$

(13) a. calegaa (f SUBJ PERSON) = 3
(f SUBJ NUM) = SG
(f SUBJ GENDER) = MASC
(f SUBJ CASE) = NOM

b. Ram (g PERSON) = 3
(g GENDER) = MASC
(g NUM) = SG
(g CASE) = NOM

We can deduce from annotations on the Hindi phrase structure rules that (f SUBJ) = g, so we conclude that the agreement features in this example are consistent and the sentence is correctly predicted to be well formed. In contrast, the lexical entry for *calogii* in the ungrammatical string **Ram calogii* provides the equations in 14.

(14) calogii (f SUBJ PERSON) = 2
(f SUBJ GENDER) = FEM

Given the substitution and transitivity axioms of equality, there is a chain of deduction from these premises that leads to the conclusions 2 = 3 and MASC = FEM, inconsistencies between atomic feature values that account for the unacceptability of this string.

On the face of it, agreement seems to be a problem that has a straightforward solution: Two items in a sentence covary systematically, and this covariance is nicely encoded in LFG by means of atomic feature-value assignments and deductions using the standard properties of the equality relation. Essentially equivalent treatments have been provided in other constraint-based syntactic theories, including HPSG (Pollard & Sag 1994).

3. INDETERMINACY: A PROBLEM FOR TRADITIONAL APPROACHES. It is possible for a noun phrase to be ambiguous or vague between two case features. In some instances, the noun phrase can appear to simultaneously fulfill conflicting syntactic requirements. This phenomenon is sometimes referred to as RESOLUTION, by Pullum and Zwicky (1986) for example; here we use the terms FEATURE CHECKING and FEATURE INDETERMINACY, reserving the term resolution for the process of assigning person and gender features to coordinate phrases,[1] described in §§6 and 7.

Groos and van Reimsdijk (1979) point out that in German free relatives, the case of the relative pronoun must match the case required by the matrix verb as well as by the verb in the relative clause.[2]

[1] Pullum and Zwicky (1986) use the term PRINCIPLED RESOLUTION to refer to the process by which the features of a coordinate noun phrase are determined.

[2] In fact, German speakers vary as to whether case agreement in the matrix clause is required. For many speakers, case agreement is required only within the relative clause, while for other speakers, matrix case agreement is not required unless the free relative clause appears in matrix subject position (Arnold Zwicky, p.c.). The following example is grammatical for these speakers:

(i) %Ich nehme,	[wem du	vertraust].	(German)
I take	who you	trust	
ACC $\in$ OBJ CASE {DAT}		DAT $\in$ OBJ CASE	

'I take who(ever) you trust.'

(15) Wer nicht gefördert, wird, muss klug sein. (German)
Who not supported is must clever be
NOM SUBJ CASE = NOM SUBJ CASE = NOM
'Who isn't supported must be clever.'

If case requirements conflict, the result is ungrammatical:

(16) *Wer nicht geholfen wird, muss klug sein. (German)
Who not helped is must clever be
NOM SUBJ CASE = DAT SUBJ CASE = NOM
'Who isn't helped must be clever.'

These results follow from the traditional atomic-value-and-equality account of agreement. As Groos and van Reimsdijk (1979) note, and as we mentioned above, this simple approach fails in some instances. In 17, the matrix verb requires an accusative argument and the verb in the relative clause requires a nominative argument. The form *was* can separately satisfy the requirements of both NOM and ACC verbs, and the grammaticality of this example shows that it can appear as the single target of two conflicting demands.

(17) Ich habe gegessen was übrig war.
I have eaten what left was
OBJ CASE = ACC ? SUBJ CASE = NOM
'I ate what was left.'

Examples such as these have been widely discussed—Zaenen & Karttunen 1984, Pullum & Zwicky 1986, Ingria 1990, Johnson & Bayer 1995, Bayer 1996, Müller 1999—since they pose a direct challenge for equality-based treatments of agreement. We briefly describe the two most obvious ways of assimilating the treatment of these examples to the traditional approach and show why these proposals are not viable.

3.1. INDETERMINACY AS DISJUNCTION? It may seem that we can obtain the desired result by assuming that *was* is specified as a Boolean disjunction, so that it is either NOM or ACC.

(18) was: ($\uparrow$ CASE) = NOM $\vee$ ($\uparrow$ CASE) = ACC

This approach does not succeed, however, under the usual axioms of Boolean logic that conjunction and disjunction distribute over each other. Any formula can be converted to a disjunctive normal form in which disjunctions have the widest possible scope, so that formally there is no difference between the disjunctive statement in 18 and a list of homophones.

(19) was: ($\uparrow$ CASE) = NOM
was: ($\uparrow$ CASE) = ACC

Thus, no matter which value of CASE we choose for *was*, it conflicts with the requirements of one of the verbs in the construction:

In the following discussion, we consider only the judgments of speakers of the more restrictive dialect. We are grateful to Arnold Zwicky for discussion of these examples.

(20) Ich habe gegessen was übrig war.
I have eaten what left was
OBJ CASE = ACC ACC SUBJ CASE = NOM
'I ate what was left.'

(21) Ich habe gegessen was übrig war.
I have eaten what left was
OBJ CASE = ACC NOM SUBJ CASE = NOM
'I ate what was left.'

We see that disjunctively specifying the case of *was* does not predict the grammaticality of 17.

3.2. INDETERMINACY AS UNDERSPECIFICATION? Alternatively, we might assume that *was* is completely unspecified for case:

(22) was: no case at all

But this approach also does not account for the grammaticality of the German free relative. This situation can be diagrammed with a variable x standing for the case of the pronoun rather than the two different values in the illustration above:

(23) Ich habe gegessen was übrig war.
I have eaten what left was
OBJ CASE = ACC x SUBJ CASE = NOM
'I ate what was left.'

The unknown value x for the case of *was* is required by the matrix verb to be ACC, and by the verb of the relative clause to be NOM. Even though no particular value is known, the transitivity of equality still implies the undesired conclusion that NOM = ACC.

(24) $x = \text{NOM} \wedge x = \text{ACC} \Rightarrow \text{NOM} = \text{ACC}$ (!)
(transitivity of equality)

Thus a solution based on feature underspecification fares no better than the disjunction proposal: The agreement requirements of both verbs still are not simultaneously satisfiable. Underspecification has a separate difficulty in that it fails to account for the fact that *was* is not compatible with requirements for cases other than NOM or ACC. For example, *was* cannot appear in contexts requiring dative or genitive case, but such a situation cannot be ruled out if no feature value is specified.

Other approaches to indeterminacy have been proposed and are discussed briefly in §4.6. Some researchers (Johnson & Bayer 1995, Müller 1999) have proposed solutions that apply only to the case of German reduced relatives, but do not extend to the coordinate examples discussed in section 4.3. Other researchers (Sag et al. 1985, Shieber 1989) have made proposals to handle feature indeterminacy in coordinate structures, but these solutions do not extend to noncoordinate examples like 23. There have also been proposals to augment the standard constraint-based architecture with additional evaluative operations (Ingria 1990) or to abandon such architectures altogether (Johnson & Bayer 1995). It turns out, however, that a simple and general solution is available within the existing LFG framework. This solution accounts for all of the examples we have presented and does not require fundamental changes to the architecture of the theory.

4. A SET-BASED THEORY OF INDETERMINATE FEATURE VALUES AND REQUIREMENTS. We provide a general solution to the problem of agreement with indeterminate features by representing features as sets of atomic values rather than as individual symbols. The elements of a set encode the alternative values that an indeterminate feature can agree with, and agreement requirements are imposed by assertions of set membership.

(25) Sets encode indeterminate feature possibilities.

We illustrate our approach to feature indeterminacy by considering the differing patterns of agreement exemplified by case in German free relatives, case in Polish coordination, and noun class in Xhosa coordination.

4.1. THE CASE OF GERMAN FREE RELATIVES. We take the value of the feature CASE for noun phrases always to be a set whose elements are the atomic symbols (NOM, ACC, etc.) of the traditional approach. Often the value is a singleton set, as with the unambiguously nominative German relative pronoun *wer*. For relative pronouns like *was*, which can be used in both nominative and accusative contexts, the set has more than one member.

(26) wer: $(\uparrow \text{CASE}) = \{\text{NOM}\}$
was: $(\uparrow \text{CASE}) = \{\text{NOM, ACC}\}$

Case checking by verbs consists in testing for the appropriate members in the CASE set (27).

(27) gegessen: $\text{ACC} \in (\uparrow \text{OBJ CASE})$
übrig: $\text{NOM} \in (\uparrow \text{SUBJ CASE})$

Thus we assign the functional structure in 29 to German free relatives like 28, where a line between feature positions indicates that exactly the same value appears in both places (some details omitted).

(28) Ich habe gegessen was übrig war.
I have eaten what left was
ACC ∈ OBJ CASE {NOM, ACC} NOM ∈ SUBJ CASE
'I ate what was left.'

(29)
$$\left[\begin{array}{ll} & \text{'EAT'} \\ \text{TENSE} & \text{PAST} \\ \text{SUBJ} & \left[\begin{array}{ll}\text{PRED} & \text{'PRO'}\\ \text{NUM} & \text{SG} \\ \text{PERSON} & 1\end{array}\right] \\ \text{OBJ } o\text{:} & \left[\begin{array}{ll}\text{PRED} & \text{'WHAT'} \\ \text{CASE} & \{\text{NOM, ACC}\} \\ \text{RELMODS} & \left\{\left[\begin{array}{ll}\text{PRED} & \text{'LEFT'}\\ \text{SUBJ } s\text{:} & \left[\begin{array}{l}\text{PRED}\\ \text{CASE}\end{array}\right]\end{array}\right]\right\}\end{array}\right]\end{array}\right]$$

The matrix verb *gegessen* 'eaten' requires that ACC be a member of its OBJ's case value, and the embedded predicate *übrig war* 'was left' requires that NOM be a member of its SUBJ's case value. These requirements are embodied in the assertions in 30.

(30) was: $(o\ \text{CASE}) = \{\text{NOM, ACC}\}$
gegessen: $\text{ACC} \in (o\ \text{CASE})$
übrig: $\text{NOM} \in (s\ \text{CASE})$

In a free relative construction the PRED and CASE values of the relative phrase are shared between the matrix sentence and the relative clause, so for this example we also know that

(31) (o CASE) = (s CASE)

The conditions in 30 and 31 are mutually consistent: the case requirements for both verbs are satisfied and the sentence is marked as grammatical.

In contrast, 32 is not acceptable.[3]

(32) *Wem du vertraust muss klug sein. (German)
who you trust must clever be
{DAT} DAT ∈ OBJ CASE NOM ∈ SUBJ CASE
'Who(ever) you trust must be clever.'

The relative pronoun *wem* 'who' is unambiguously dative, and so its value for the CASE feature is the singleton set {DAT}. The requirements of the verb *vertraust* 'trust' are satisfied, since it requires a dative object. However, the matrix verb *muss* 'must' requires a nominative subject, and this constraint is not met. Letting d and n stand respectively for the f-structures of the dative object and nominative subject, the relevant constraints for this example are given in 33.

(33) (d CASE) = (n CASE)
wem: (d CASE) = {DAT}
vertraust: DAT ∈ (d CASE)
muss: NOM ∈ (n CASE) [not satisfied]

4.2. SETS AND THEIR DESCRIPTIONS. Our use of set representations as the formal encoding of feature values is a conservative solution to the problem of feature indeterminacy. It is not an extension to LFG's domain of f-structure entities, since sets have been part of the LFG f-structure ontology from the beginning. Kaplan & Bresnan 1982 introduced sets of f-structures to represent collections of adjuncts and modifiers, and for the set of relative clause modifiers, RELMODS in ex. 29, and Bresnan et al. 1985 and Kaplan & Maxwell 1988 used sets to represent the conjoined elements of coordinate constructions. Thus, LFG's functional description language has always included set membership ∈ as a primitive relation, and this is the relation by which the verbal lexical entries above check that their set-valued argument features meet the appropriate requirements.

The assignment of set values for the pronoun case features does involve a minor extension to the original f-description notation. The CASE value for *was* is defined by the equation in 34.

(34) was: (↑ CASE) = {NOM,ACC}

The expression {NOM,ACC} in this equation is a SET DESIGNATOR, a notational innovation indicating that the feature value is a set and also providing an EXHAUSTIVE ENUMERATION of the set's elements. This captures some of the disjunctive intuitions of feature indeterminacy: If x stands for some unknown value, we know that

(35) $x \in \{\text{NOM,ACC}\} \Rightarrow \text{NOM} = x \bigvee \text{ACC} = x$ (implicit disjunction)
(n.b., $\text{GEN} \neq x$)

Whereas explicit disjunction as denoted by the Boolean operator $\bigvee$ has the undesired

[3] We are grateful to Stefan Kaufmann for providing the German example given in 32.

wide-scope interpretation, elements in a set are disjunctive only with respect to specific assertions of set membership. In this way, our set representations can be viewed as formally clarifying the intuitions that underlie the disjunctive values of Zaenen & Karttunen 1984. Conversely, we also know that if y is defined by a set designator, then it does in fact contain all of the designated elements:

(36) $y = \{\textsc{nom}, \textsc{acc}\} \Rightarrow \textsc{nom} \in y \wedge \textsc{acc} \in y$ (implicit conjunction)
(n.b., $\textsc{gen} \notin y$)

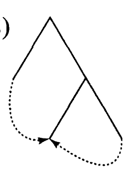
This is why the conjunction of differing verbal requirements is satisfiable with indeterminate pronoun specifications. This combination of disjunctive and conjunctive characteristics is the basis of our analysis of indeterminacy.

The fact that NOM and ACC belong to the set could also be encoded without the use of set designators, by the pair of assertions in 37.

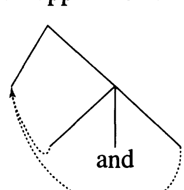
(37) was: $\textsc{nom} \in (\uparrow \textsc{case})$
$\textsc{acc} \in (\uparrow \textsc{case})$

But this is not only more cumbersome than the single assertion above, it also does not express the important fact that the set contains no other members. Thus, unlike the set designator in 36, the assertions in 37 would have the undesired property of being consistent with a further assertion that $\textsc{gen} \in (\uparrow \textsc{case})$.

4.3. INDETERMINACY IN COORDINATION. The German free relative example discussed above can be schematically represented as 38.

(38) 
Shared values (German CASE)

Other examples of feature indeterminacy as well as feature determination involve an argument that is shared between two conjuncts. Again, as indicated in 39, the shared argument can appear to satisfy conflicting requirements in the two conjuncts.

(39) 

Case features in Polish coordinate constructions provide evidence for this type of three-way indeterminate agreement, and our set-based solution also applies to these examples. Dyła (1984:701) discusses the following Polish example (40), in which a noun phrase is shared across a coordinate structure and different requirements are placed on its case feature in each conjunct.

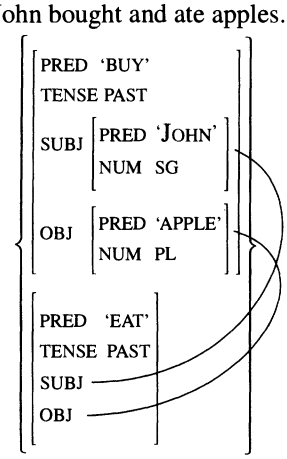
(40) Kogo Janek lubi a Jerzy nienawidzi? (Polish)
who Janek likes and Jerzy hates
{ACC, GEN} ACC ∈ OBJ CASE GEN ∈ OBJ CASE
'Who does Janek like and Jerzy hate?'

This sentence is grammatical because the pronoun *kogo* 'who' satisfies both agreement requirements, in contrast to the ungrammaticality of 41 in which *kogo* 'who' is replaced by *co* 'what'.

(41) *Co Janek lubi a Jerzy nienawidzi? (Polish)
who Janek likes and Jerzy hates
{NOM, ACC} ACC ∈ OBJ CASE GEN ∈ OBJ CASE
'Who does Janek like and Jerzy hate?'

Our use of set-valued features interacts with the LFG theory of coordination to account for this pattern of agreement. As originally proposed by Bresnan and colleagues (1985) and discussed by Kaplan and Maxwell (1988), the f-structure for a coordinate structure is a set containing the f-structures for the conjuncts.[4] Thus the f-structure for a sentence like *John bought and ate apples* is 42.

(42) John bought and ate apples.



This example of verb coordination is described by the phrase structure rule 43.

(43) $\underset{\downarrow\in\uparrow}{\mathrm{V}} \rightarrow \mathrm{V}\ \ \mathrm{CONJ}\ \ \underset{\downarrow\in\uparrow}{\mathrm{V}}$

Given this rule and given the fact that the coordinated V is the head of the VP and S, the f-structure for the whole sentence is a set containing an f-structure corresponding to each conjunct. The NP *John* contributes the subject of this f-structure set by virtue of the simple function-application schema (↑ SUBJ) = ↓, and the NP *apples* provides its object. The original definition of function application (Kaplan & Bresnan 1982), however, did not provide a value for such an expression when ↑ denotes a set of f-structures instead of a single f-structure. Thus, it did not produce the desired effect of distributing the subject and object to each of the conjunct f-structures. Bresnan et al. 1985 outlined a general theory of feature distribution in coordinate constructions by specifying the conditions under which a set can be treated as an f-structure with attributes and values.

(44) For any attribute-value property P and set s: $P(s)$ iff $\forall f \in s.P(f)$

According to this definition, if the NP *John* is the subject of the coordinate structure, it is distributed to be the subject of each of the conjuncts.[5]

[4] For more discussion of coordination in LFG, in particular of nonconstituent coordination, see Maxwell and Manning 1996.

[5] This definition differs from the later formulation in Kaplan & Maxwell 1988. They proposed that when s is a set, $(s\ a) = v$ iff v is mathematically the generalization of all of the elements of s applied to a.

For our Polish example 40 the effect of this definition is to distribute *kogo* 'who' so that it becomes the object in each of the conjunct f-structures.

(45) Kogo Janek lubi a Jerzy nienawidzi (Polish)
who Janek likes and Jerzy hates
'Who does Janek like and Jerzy hate?'

$$\left\{\begin{array}{l} \left[\begin{array}{ll} \text{FOCUS } w\text{:} & \left[\begin{array}{ll} \text{PRED} & \text{'WHO'} \\ \text{CASE} & \{\text{ACC, GEN}\} \end{array}\right] \\ \text{PRED} & \text{'LIKE'} \\ \text{SUBJ} & [\text{PRED 'JANEK'}] \\ \text{OBJ} & \end{array}\right] \\ \left[\begin{array}{ll} \text{FOCUS} & \\ \text{PRED} & \text{'HATE'} \\ \text{SUBJ} & [\text{PRED 'JERZY'}] \\ \text{OBJ} & \end{array}\right] \end{array}\right\}$$

The agreement pattern for examples such as 45 now follows from the same kind of feature markings we used for the noncoordinate German sentences: Polish *kogo* 'who' has a CASE feature whose value is {ACC, GEN}. The verb *lubi* 'likes' requires its object's case to contain ACC, and the verb *nienawidzi* 'hates' requires its object's case to contain GEN. The form *kogo* is the object of both of these verbs and satisfies both of these requirements, and the sentence is well formed.

(46)
kogo $(w \text{ CASE}) = \{\text{ACC,GEN}\}$
lubi $\text{ACC} \in (w \text{ CASE})$
nienawidzi $\text{GEN} \in (w \text{ CASE})$

Examples of indeterminacy can be found with features other than case. Pullum and Zwicky (1986:761) discuss constraints on the form of the verb used in a verb phrase complement. They present examples such as 47-48 and note that they are ill formed.[6]

(47) *I certainly will, and you already have, clarify the situation with respect to the budget.

(48) *I certainly will, and you already have, clarified the situation with respect to the budget.

The auxiliary *will* requires a bare infinitive complement, while auxiliary *have* requires the past participle form. There is no form of the verb *clarify* that has both of these features, and so the example is ill formed.

Example 49, in contrast, is perfectly acceptable.

This reformulation, however, makes unwanted predictions involving constraining as opposed to defining properties.

[6] Some speakers allow for alternative agreement strategies in which only the requirements of the nearest conjunct need be satisfied. For those speakers, example 48 is acceptable. See §5 for more discussion of alternative agreement strategies involving coordinate phrases.

(49) I certainly will, and you already have, set the record straight with respect to the budget.

This follows if the value of the VFORM feature, like the value of the CASE feature, is a set which may contain more than one element. Since the verb form *set* represents the bare infinitive and the past participle form, the VFORM value of the verb *set* contains both values.

(50) *set* V $(\uparrow \text{PRED}) = \text{`SET'}$
$(\uparrow \text{VFORM}) = \{\text{BASE, PPART}\}$

This form satisfies the requirements of both *will* and *have* and allows for the grammaticality of 49.

4.4 INDETERMINATE AGREEMENT REQUIREMENTS. The data explored in the previous section show that the case of a noun phrase can be indeterminate and can independently satisfy each of two different case requirements imposed by different verbs. Conversely, a verb may be indeterminate in the requirement it imposes on an argument, and this requirement might be satisfied in two different ways in a conjunction of two differently specified noun phrases. Such a situation is illustrated by noun-class agreement in the Bantu languages as discussed for Xhosa in Voeltz 1971 and Pullum & Zwicky 1986,[7] and for Chicheŵa in Corbett 1991.

In Xhosa, as in other Bantu languages, nouns belong to several noun classes. We follow standard practice in referring to these noun classes with numbers. Generally, the singular form is a member of one class while the plural form is a member of a second class, and so the noun class of a noun is represented by a pair of numbers such as 1/2, indicating that the singular is in class 1, the plural in class 2.

Voeltz (1971) shows that Xhosa verbs agree with their subjects in noun class and that coordinating subject noun phrases of the same noun class is permissible, as illustrated in 51 and 52.

(51) umfana nomfazi bayagoduka. (Xhosa)
young.man and.young.woman go.home
1/2 1/2 SUBJ CLASS = 1/2
'The young man and the young woman are going home.'

(52)
$$\begin{bmatrix} \text{PRED} & \text{`GO.HOME'} \\ \text{SUBJ} & \left\{ \begin{matrix} \begin{bmatrix} \text{PRED} & \text{`YOUNG.MAN'} \\ \text{CLASS} & 1/2 \end{bmatrix} \\ \begin{bmatrix} \text{PRED} & \text{`YOUNG.WOMAN'} \\ \text{CLASS} & 1/2 \end{bmatrix} \end{matrix} \right\} \end{bmatrix}$$

The noun-class requirement is met by each conjunct of the coordinate subject noun phrase: Each conjunct noun phrase is class 1/2.

It is generally impossible to coordinate subject noun phrases of different noun classes.

[7] The noun class system of Bantu is sometimes referred to as a gender system. We use the term 'noun class' here to distinguish this system from the gender systems in languages like French or Hindi, which behave differently in terms of feature resolution.

(53) a. *Igqira nesanuse ayagoduka. (Xhosa)
doctor and.diviner go.home
5/6 7/8 SUBJ CLASS = 5/6

b. *Igqira nesanuse ziyagoduka. (Xhosa)
doctor and.diviner go.home
5/6 7/8 SUBJ CLASS = 7/8
'The doctor and diviner are going home.'

When verb agreement is not involved, such coordinate structures are unproblematic; the problem with the examples in 53 is that the subject noun-class requirement imposed by the verb is distributed to each conjunct of the coordinate subject phrase but is not satisfied by all of them.

Voeltz notes that an interesting exception to this generalization is provided by verbs that can take subjects in either of two noun classes.[8] In such cases, the verb may take a coordinate subject with nouns from these different classes, and each conjunct separately satisfies the requirements imposed by the verb. In 54, the prefix *zi* is the agreement form for both 7/8 and 9/10 plurals.

(54) Izandla neendlebe zibomvu. (Xhosa)
hands and.ears are.red
7/8 9/10 SUBJ CLASS ∈{7/8,9/10}
'The hands and the ears are red.'

We again use set representations to encode an indeterminate pattern of agreement. In this instance, the set appears as a part of the subject noun-class requirement of the verb, not as a property of the nouns as in the German and Polish examples. In Xhosa, an indeterminate verb such as *zibomvu* checks the noun class of its subject by checking for membership in a set of possible noun classes. In 54, the subject's noun class must be either 7/8 or 9/10. This indeterminacy is distributed to the conjuncts of the subject noun phrase, and each conjunct satisfies one of the requirements. Thus, a conjunction of nouns in these classes is permitted.

(55)
$$f: \left[\begin{array}{ll} \text{PRED} & \text{'ARE.RED'} \\ \text{SUBJ} & \left\{ \begin{array}{l} h: \left[\begin{array}{ll} \text{PRED} & \text{'HANDS'} \\ \text{CLASS} & 7/8 \end{array}\right] \\ e: \left[\begin{array}{ll} \text{PRED} & \text{'EARS'} \\ \text{CLASS} & 9/10 \end{array}\right] \end{array} \right\} \end{array}\right]$$

(56) izandla (h CLASS) = 7/8
neendlebe (e CLASS) = 9/10
zibomvu (f SUBJ CLASS) ∈ {7/8,9/10}

[8] Roberts and Wolontis (1974) give additional evidence to support Voeltz's claims, and also present a transformationally based analysis similar in some respects to the one presented here: A verb like *zifikile* 'arrived' is marked as both 7/8 and 9/10, and either or both features can be matched by the subject noun phrase.

Interestingly, though DuPlessis and Visser (1992) do not discuss agreement with conjoined subjects in Xhosa, they give rules for Xhosa clitic/object agreement that are different from the generalizations noted by Voeltz: They claim that clitic/object agreement is semantically governed, not determined by the syntactic processes that Voeltz describes and we have analyzed.

The agreement requirements of verbs with determinate subject prefixes, such as *bayago-duka*, are still stated with the equality relation.

Facts similar to the Xhosa data are reported by Corbett (1991:276 ff.) for Chichewa, based on work by Corbett and Mtenje (1987), where the verb agreement form for plural nouns of the 1/2 and 5/6 genders are both marked with the prefix *a-*.

(57) ma-lalanje ndi ma-samba a-kubvunda. (Chichewa)
orange and leaf are.rotting
6 6 SUBJ CLASS $\in \{2, 6\}$
'The oranges and the leaves are rotting.'

(58) a-mphaka ndi a-galu a-kuthamanga. (Chichewa)
cat and dog are.running
2 2 SUBJ CLASS $\in \{2, 6\}$
'The cats and the dogs are running.'

Coordinate structures with mixtures of class 2 and 6 nouns are also allowed, as, for example, in 59.

(59) a-mphaka ndi ma-lalanje a-li uko. (Chichewa)
cat and orange be there
2 6 SUBJ CLASS $\in \{2, 6\}$
'The cats and the oranges are there.'

Another example of indeterminate agreement requirements involves the person feature and right-node raising in German. The following observations were originally presented by Eisenberg (1973) and have since been discussed by Ingria (1990) and Pullum and Zwicky (1986:757).

(60) weil [wir das Haus kaufen] und [die Müllers den Garten kaufen] (German)
because we the house buy and the Müllers the garden buy
1 3
'because we buy the house and the Müllers buy the garden'

(61) weil [wir das Haus] und [die Müllers den Garten] kaufen
because we the house and the Müllers the garden buy
1 3 SUBJ $= 1/3$
'because we [buy] the house and the Müllers buy the garden'

Example 60 shows that the German verb *kaufen* 'buy' agrees with both 1 and 3 person subjects. In example 61 the subject of the first conjunct is a 1 person form, the subject of the second conjunct is a 3 person form, and the verb is compatible with both of these. This follows from a simple indeterminate constraint in the lexical entry for this form of *kaufen*.

(62) kaufen ($\uparrow$ SUBJ PERSON) $\in \{1, 3\}$

When the verb is right-node raised out of a conjunction of clauses, this constraint is distributed across the elements of the resulting set of conjuncts and thus applied independently to the subject of each conjunct. The sentence is grammatical because each conjunct satisfies the verb's requirement.

(63)
$$\left\{ \begin{bmatrix} \begin{bmatrix} \text{PRED} & \text{`BUY'} \\ \text{SUBJ} & \begin{bmatrix} \text{PRED} & \text{`PRO'} \\ \text{PERSON} & 1 \\ \text{NUM} & \text{PL} \end{bmatrix} \\ \text{OBJ} & \begin{bmatrix} \text{PRED} & \text{`HOUSE'} \end{bmatrix} \end{bmatrix} \\ \begin{bmatrix} \text{PRED} & \\ \text{SUBJ} & \begin{bmatrix} \text{PRED} & \text{`MÜLLERS'} \\ \text{PERSON} & 3 \\ \text{NUM} & \text{PL} \end{bmatrix} \\ \text{OBJ} & \begin{bmatrix} \text{PRED} & \text{`GARDEN'} \end{bmatrix} \end{bmatrix} \end{bmatrix} \right\}$$

Pullum and Zwicky (1986:771) and Zaenen and Karttunen (1984) discuss more complicated examples where combinations of more than one feature are involved in the indeterminacy. Example 64 involves both the PERSON and NUM features.[9]

(64) weil [Ihr das Haus] und [Franz den Garten] kauft (German)
because you the house and Franz the garden buy
2PL 3SG SUBJ = 2PL/3SG
'because you buy the house and Franz buys the garden'

Our analysis of the more interesting indeterminacy of correlated agreement requirements resembles the treatment we have given for the Xhosa, Chicheŵa, and German examples above. For the verb *kauft* the two combinations of correlated features each contribute to an element of the set representing the indeterminacy:

(65) kauft $(\uparrow \text{SUBJ}) \in \{x, y\}$
$(x \text{ PERSON}) = 2$
$(x \text{ NUM}) = \text{PL}$
$(y \text{ PERSON}) = 3$
$(y \text{ NUM}) = \text{SG}$

These equations assert that the subject of *kauft* is either a second person plural phrase (the combination of requirements represented by x) or a third person singular phrase (the requirements represented by y). In example 64, these requirements are distributed to, and satisfied separately by, the subject of each conjunct, just as they were in the simpler examples involving only one feature.

Most previous treatments of feature indeterminacy do not properly account for examples involving indeterminacy of correlated features; indeed, besides our own account, only Bayer 1996 successfully handles such cases. The underlying problem for most accounts (Pullum & Zwicky 1986, Johnson & Bayer 1995, as discussed by Bayer 1996)

[9] Pullum and Zwicky found that example 64 was ungrammatical for all of the German speakers they consulted, but at least some of our informants find it grammatical, as do some speakers consulted by Zaenen and Karttunen (p.c. to Pullum and Zwicky).

is that they are framed in terms of standard notions of conjunction or disjunction (or analogous operators) that are distributive and associative and thus allow for incorrect combinations of features.

4.5. SPEAKER VARIATION. As examples such as 64 show, some instances of correlated feature indeterminacy are acceptable for only some speakers. There are also many examples of speaker variation where just a single feature is involved. Zaenen and Karttunen (1984) discuss 66, an example of case indeterminacy in Icelandic, noting that only some speakers find it acceptable.

(66) %Hann stal og borðaði köku. (Icelandic)
he stole and ate cookie
OBJ CASE = DAT OBJ CASE = ACC {ACC, DAT}
'He stole and ate the cookie.'

The noun phrase *köku* appears as the object of two coordinated verbs, the first of which requires its object to bear dative case and the second of which requires accusative case. Following Zaenen and Karttunen, we assume that such examples are treated as indeterminate by speakers who accept them, that is, that the case of *köku* is analyzed as the set {ACC,DAT}. The single form can then satisfy both requirements. For speakers who find these examples ungrammatical, *köku* must be disjunctively specified for case (($\uparrow$ CASE) = {ACC} $\vee$ ($\uparrow$ CASE) = {DAT}). Of course, if the different feature possibilities correlate with some other clearly disjunctive ambiguity, for example an ambiguity of sense, a disjunctive specification is required. From this observation follows what Zaenen & Karttunen 1984 calls the ANTI-PUN ORDINANCE: a word that has a seemingly indeterminate value for CASE or some other syntactic feature may not be used in two different senses at the same time.

Examples which are analyzed by some speakers as indeterminate and by others as disjunctive are fairly common. Zaenen and Karttunen (1984) also discuss the sentence in 67, an example originally considered by Kayne (1975). They note that 'some prescriptive grammarians condemn' such examples.

(67) %Il m'a frappé et donné des coups de pied. (French)
he me hit and gave kicks
{ACC, DAT} ACC $\in$ OBJ CASE DAT $\in$ OBJ CASE
'He hit and kicked me.'

In the same vein, Pullum and Zwicky (1986:765) claim that examples like 68 and 69 are acceptable for all German speakers, but the speakers we have checked with do not find these examples acceptable.

(68) %Sie hat Karl gefunden und geholfen. (German)
she has Karl found and helped
{ACC, DAT} ACC $\in$ OBJ CASE DAT $\in$ OBJ CASE
'She has found and helped Karl.'

(69) %Er findet und hilft Frauen. (German)
he finds and helps women
ACC $\in$ OBJ CASE DAT $\in$ OBJ CASE {ACC, DAT}
'He finds and helps women.'

These examples reinforce the point that different speakers may analyze the same forms either indeterminately or disjunctively.

Pullum and Zwicky (1986) are troubled by the flexibility allowed by approaches such as ours and Zaenen and Karttunen's. They reject the possibility that different speakers may analyze a given form in different ways, and propose instead a more restrictive treatment of idiolectal variation. They claim that the grammar for each speaker can be characterized by a particular integer that specifies the maximum number of features that can be indeterminate in any phonological form. For example, if the number for a speaker is zero, then that speaker will not accept any instances of indeterminacy. If the number is one, the speaker will accept examples such as 69 but not example 64 above, in which two correlated features are involved. However, some German speakers we consulted accept the one-feature indeterminacy in example 61, with the indeterminate verb *kaufen*, but reject 69, which also has only one indeterminate feature. This is a counterexample to Pullum and Zwicky's particular proposal. We note that their concerns about the range of idiolectal variation may be addressed in other ways. A lexicalist theory such as LFG, for example, offers other devices for expressing generalizations across the lexicon, and these may be the basis for a restricted set of possible idiolects.

4.6. COMPARISON WITH EARLIER APPROACHES. To our knowledge, Zaenen and Karttunen (1984) were the first to observe that equality constraints over atomic values are inadequate to characterize case agreement phenomena in their full generality. They identified many of the empirical and formal problems considered in later work, but they presented only a sketch of a coherent treatment of the problem. Subsequent proposals for dealing with case agreement in German free relatives and other examples of indeterminacy have involved more or less radical changes or augmentations to the architecture of constraint-based grammatical theories.

Ingria (1990) proposes handling indeterminate agreement by means of a nondistinctness check rather than by either a unification operation or an equality constraint. A relative pronoun such as German *was* is disjunctively marked for case, bearing the case NOM $\vee$ ACC. Verbs check for consistency with this feature without instantiating it to either one of the disjuncts: The nondistinctness check ensures only that the case value required by the verb is consistent with the value of the relative pronoun. Ingria's account resembles ours in that it supplies complex feature values for indeterminate features and checks these features in a special way, though our account expresses the properties of different features in different constructions in terms of simple, standard, and independently motivated mathematical relations and structures.

Johnson and Bayer (1995) also discuss the interaction of underspecification with case checking, presenting additional problems for certain approaches to case checking based on unification or equality constraints. They offer an alternative framework based on deduction in Lambek categorial grammar, arguing that unification-based approaches cannot handle the full range of feature indeterminacy and that a deductive approach is therefore preferable. Their primary argument against a constraint-based account is aimed at proposals for feature determination in coordination by Sag et al. (1985) and Shieber (1989). Those analyses assume that the features of a coordinate structure are the generalization of the features of the conjuncts (Sag et al.), or that the features of the coordinate structure must subsume the features of the conjuncts (Shieber). As Johnson and Bayer point out, coordinating two transitive verbs that require different cases for their objects—for example, a verb that requires an accusative object and a verb that requires a dative object—will produce a coordinate structure compatible with an object with either accusative or dative case, the wrong result. Like the original Bresnan et al. 1985 approach, the account we present here does not suffer from such incorrect predictions.

Though our approach and the Johnson/Bayer account are set on very different formal foundations, they have substantial overlap in their empirical coverage, at least for the phenomena examined here.[10] These approaches differ in two main respects. First, as noted in §4.4, our account handles examples in which multiple correlated features are indeterminate. Bayer (1996) points out that the Johnson/Bayer approach cannot deal with these constructions, for what seem to be fundamental reasons. Second, their solution for indeterminacy in coordination (§4.3) does not extend to indeterminacy in noncoordinate constructions such as German free relatives (§4.1). To handle these cases, Johnson and Bayer propose a completely separate mechanism that relies on two separate values for the CASE of the relative pronoun: One of these is checked in the matrix clause, and the other is checked in the relative clause. Müller (1999) offers a similar two-feature analysis for German free relatives, but he does not also offer a solution for coordination. We prefer a solution that treats both types of indeterminacy in the same way, since the phenomena are clearly related.

Müller (1999) and Anette Frank (p.c.) point out that since our set-based account of indeterminacy is designed to eliminate the undesired consequences of transitivity of equality, it would lead to more complicated treatments if any phenomena can be found in which indeterminacies must be resolved in correlated ways. According to Müller, the sentence in 70 exemplifies this situation.

(70) weil Kasparov [seine Gegner] [einen nach dem anderen] geschlegen
because Kasparov his opponents one after the other beaten
{NOM, ACC} {ACC} ACC $\in$ OBJ CASE
*[einer nach dem anderen]
{NOM}
hat
has
'because Kasparov beat his opponents one after the other' (German)

Müller asserts that the adverbial phrase *einen nach dem anderen* 'one after the other'

[10] In this work, we confine our analysis to functional features, features that are represented at the level of functional structure. Using the same formal devices as the Johnson/Bayer account, Bayer (1996) provides a detailed analysis of examples in which phrase structure constituents of unlike category are coordinated, assimilating these examples to his account of features such as CASE:

(i) Jermaine is boring and a fool. (AP and NP)
(ii) Pat is healthy and of sound mind. (AP and PP)

In LFG, phrase structure category information is represented separately from functional information. This modularity reflects the general architecture of LFG, which assumes that subsystems of grammar, such as phrasal and functional structure, exhibit NEAR-DECOMPOSABILITY, with major interactions within a module and minor interactions across modules. In most instances, functional information is sufficient to characterize subcategorization requirements. Occasionally, however, the phrase structure category of the subcategorized element is also at issue, and it is these examples which Bayer's analysis treats. We can handle these instances of category indeterminacy within our approach by using the CAT predicate, introduced by Kaplan and Maxwell (1996). This imposes constraints on the category labels of the phrase structure nodes that an f-structure corresponds to. We do not discuss such examples further in this article.

shows CASE agreement with its antecedent, and that this must be consistent with the verb's assignment of CASE to its OBJ. This assumption accounts for the ungrammaticality of the variant in which the adverbial bears NOM case. In fact, though, the correct analysis of these examples is that the CASE value of the adverb is dependent on the antecedent's grammatical function; the adverb must be ACC because its antecedent is an OBJ.[11] Examples of this sort therefore do not undermine our approach.

To summarize, we agree with Zaenen & Karttunen 1984, Ingria 1990 and Johnson & Bayer 1995 that simple equality constraints with atomic feature values such as NOM or ACC are inappropriate for characterizing feature indeterminacy and feature checking. However, our solution shows that these phenomena do not necessitate the addition of new processes for satisfying special kinds of constraints (as proposed by Ingria) or the abandonment of constraint-based frameworks for linguistic description (as suggested by Johnson and Bayer).

5. FEATURE RESOLUTION WITH COORDINATE NOUN PHRASES: DISTRIBUTIVE AND NONDISTRIBUTIVE PROPERTIES. Thus far we have examined cases in which features of the individual conjuncts in a coordinate structure are checked separately to determine well-formedness. We discussed examples of WH-fronting in coordinate structures in Polish (§4.3), where the case of the fronted object is checked independently in each conjunct. In §4.4 we discussed Xhosa examples in which the noun class of each conjunct in a coordinate noun phrase is checked independently by the verb. In these situations, the coordinate structure as a whole does not have its own features distinct from the features of the conjuncts.

However, some features of a coordinate noun phrase may be different from the features of the conjuncts. The first person plural form of the verb is appropriate, as shown in 71, even though *José* is a third person singular noun phrase.

(71) $\boxed{\text{José y yo}}$ hablamos/*habláis/*hablan (Spanish)
José and I speak-1PL/*speak-2PL/*speak-3PL

Intuitively, a coordinate noun phrase has its own person, number, and gender, just as a pronoun does.

To make this intuition explicit, we follow a proposal by John Maxwell (p.c.) for distinguishing the different kinds of properties a coordinate structure can have. Person, number and gender features of a coordinate noun phrase are properties of the coordinate structure as a whole. We refer to these features as NONDISTRIBUTIVE features, since they do not distribute to the individual conjuncts. Instead, they are properties of the set of f-structures representing the coordinate phrase. Other syntactic features, such as case or noun class, are DISTRIBUTIVE features: properties of the individual conjunct noun phrases and not of the coordinate phrase. They are checked by inspecting the features of each conjunct, as in the Xhosa examples described above. Maxwell's proposal was foreshadowed in early work on coordination in LFG by Peterson (1982), who depicts coordinate structures as sets with attributes and values as well as elements. Andrews (1983) also discussed the NUM and GENDER features of coordinate noun phrases within an LFG setting, and considered the problem of determining these features on the basis of the feature values of the individual conjuncts of the phrase.

On this view, f-structures for coordinate structures are HYBRID OBJECTS: structures with both elements and nondistributive features. This is shown in 72.

[11] See Berman 1999 for discussion of the correlation between case and grammatical function in German.

(72) 'José y yo':

$$\left[\begin{array}{l} \left\{\begin{array}{l} \left[\begin{array}{ll}\text{PRED} & \text{'JOSÉ'}\\ \text{PERSON} & 3\\ \text{NUM} & \text{SG}\end{array}\right] \\ \left[\begin{array}{ll}\text{PRED} & \text{'PRO'}\\ \text{PERSON} & 1\\ \text{NUM} & \text{SG}\end{array}\right]\end{array}\right\} \\ \text{PERSON}\ \ 1 \\ \text{NUM}\ \ \text{PL}\end{array}\right]$$

The coordinate structure differs from its conjuncts in its values for the person and number features. While the coordinate phrase as a whole is 1PL, neither conjunct has these features; one conjunct is 1SG, and the other is 3SG.

Formally, we refine the definition of function application given in 44 to take into account the distinction between distributive and nondistributive properties of sets:

(73) For any DISTRIBUTIVE property P and set s, $P(s)$ iff $\forall f \in s.P(f)$.
For any NONDISTRIBUTIVE property P and set s, $P(s)$ iff P holds of s itself.

If PERSON is a nondistributive feature, a constraint such as ($\uparrow$ SUBJ PERSON) = 1 on a verb is satisfied by a first person noun phrase or a coordinate noun phrase with the 1 value for the PERSON feature, even if one or more of its conjuncts bears some other PERSON value.

Coordinate phrases bear nondistributive PERSON NUM, and GENDER features just as noncoordinate phrases do, and we require a means of determining what these features are. For some features, such as the NUM feature, the determination is primarily semantic: A plural noun phrase is one whose referent includes more than one individual. We have little more to say about the semantically determined NUM feature. Our analysis focuses on features, in particular PERSON and GENDER, that are related in a systematic way to the corresponding features of the conjuncts by a syntactic process of FEATURE RESOLUTION. Thus, on our view, agreement features fall into the classes shown in 74.

(74)

Distributive	Nondistributive	
	Syntactic resolution	Semantically determined
CASE, Bantu noun class	PERSON, GENDER	NUM

We develop a set-based representation for the values of resolving features that permits a simple formal statement of the way they behave.

Our theory of feature resolution can be tested by observing patterns of agreement with coordinate phrases, but care must be taken in the use of this diagnostic. Corbett points out that languages often exhibit alternative agreement strategies in which agreement with coordinate phrases depends on the person and gender features of just one of its conjuncts (Corbett 1991; see also Steinberg & Caskey 1988). In the following Swahili examples, the verb shows noun-class agreement with the closest conjunct of the coordinate subject phrase (Corbett 1991:265):

(75) a. ki-ti na m-guu wa meza u-mevunjika. (Swahili)
7-chair and 3-leg of table 3-be.broken
'The chair and the leg of the table are broken.'

b. m-guu wa meza na ki-ti ki-mevunjika.
3-leg of table and 7-chair 7-be.broken
'The leg of the table and the chair are broken.'

Peterson (1986) and, more recently, Sadler (1999) provide extensive discussion of these issues within LFG. The issues raised by these examples are orthogonal to the treatment of resolution, and are therefore not examined further here.

6. PERSON RESOLUTION. The Spanish example given in 71 shows that a first person plural verb form is used with first person plural coordinate subjects. Slovak (Corbett 1983a:178) provides another illustration of the same dependency (76).

(76) $\boxed{\text{Ja a ty}}$ sme bratia. (Slovak)
I and you are.1PL brothers
'You and I are brothers.'

This phenomenon is commonly described in terms of a hierarchy of values for the person feature: if a phrase with first person agreement is coordinated with any other phrase, the coordinate structure bears first person features; if the coordinate structure contains a second person pronoun but no first person pronoun, the coordinate structure bears second person features, and so on.[12] Our goal is to permit person features to be represented so that the person feature of a coordinate structure can be computed systematically on the basis of the person features of its conjuncts. We accomplish this by again using set values for agreement attributes, but not in a way that encodes indeterminacy. For person, gender, and other RESOLVING features:

(77) Sets encode complex values, not indeterminacy.

In particular, we refine the representation of values such as 1 for the PERSON feature and take them to be sets whose elements are drawn from the markers S (for speaker) and H (for hearer). The advantage of using sets to represent these complex values is that they provide a simple method for resolving the person feature in coordination: The value of the person feature for a coordinate structure is resolved to be the UNION of the person features of the conjuncts. Initially, we can view these sets as reflecting the referent of the coordinate phrase: The union of the person values of the conjuncts represents the membership of the set of individuals referred to. This correlation between reference and syntactic markers breaks down in certain cases, however, as we discuss in §6.2.

6.1. INCLUSIVE/EXCLUSIVE PRONOUNS. Our inventory of markers S and H gives four possible values for the PERSON feature. We assign the following interpretations to these four combinations.[13]

{S}: first person singular, first person exclusive nonsingular
{S,H}: first person inclusive nonsingular
{H}: second person
{}: third person

[12] As observed by Corbett (1991:263), this hierarchy matches the hierarchy for pronominal reference proposed by Zwicky (1977:717); see also Silverstein 1976:

1. Use the first person (I) pronoun *we* for any reference set with the referential element 1;
2. Otherwise, use the second person (II) pronoun *you* for any reference set with the referential element 2;
3. Otherwise, use the third person (III) pronoun *they*.

[13] The referent of an inclusive first person pronoun includes the hearer, while the referent for an exclusive first person pronoun excludes the hearer.

Fula (Forchheimer 1953, Arnott 1970) is a language with an inclusive/exclusive distinction in the first person plural pronominal system.[14] Fula exhibits the following pattern of person resolution, where the ampersand denotes coordination.

(78)

1.SG & 2	=	1INC.PL
1.SG & 2 & 3	=	1INC.PL
1.SG & 3	=	1EXC.PL
1EXC.PL & 3	=	1EXC.PL
2 & 3	=	2
3 & 3	=	3

In Fula it is possible to front noun phrases, including pronominal forms, thereby putting them in focus. An incorporated pronoun then appears on the verb and shows agreement with the fronted subject. The forms of the verb meaning 'live' with incorporated pronominal subjects are those in 79.

(79)

	SINGULAR	PLURAL
1.	mi-djodi	*inclusive*: djodu-dèn
		exclusive: mèn-djodi
2.	djodu-da	djodu-don
3.	wo-djodi	bè-djodi

In example 80, the first person singular pronoun *min* is fronted, and the incorporated pronoun shows first person singular agreement with it.

(80) min kö Afriki mi-djodi. (Fula)
I in Afrika live
1 1
'Me, I live in Africa.'

This construction provides a way to examine resolution of the PERSON feature, since conjoined phrases can also undergo fronting. Fronted coordinate phrases containing a second person conjunct and a third person conjunct show second person agreement, as is typically found in other languages.

(81) an e Bill kö Afriki djodu-don. (Fula)
you and Bill in Afrika live
2 3 2
'You and Bill, you live in Africa.'

Fronted coordinate phrases containing only third person conjuncts show third person agreement (82). More interesting examples involve first person pronouns. Coordinate phrases with a first person singular pronoun and a second person pronoun show first person inclusive plural agreement (83, 84).

(82) Bill e George kö Afriki bè-djodi. (Fula)
Bill and George in Afrika live
3 3 3
'Bill and George, they live in Africa.'

[14] The variety of Fula described here is spoken in Guinea. We are grateful to Aicha Kone for providing the data, and to Chris Culy for help with the analysis.

(83) an e min kö Afriki djodu-dèn. (Fula)
you and I in Afrika live
2 1 1INCL
'You and I, we live in Africa.'

(84) an e Bill e min kö Afriki djodu-dèn. (Fula)
you and Bill and I in Afrika live
2 3 1 1INCL
'You and Bill and I, we live in Africa.'

When the coordinate phrase contains a first person pronoun and a third person noun phrase (but no second person phrase), first person exclusive plural agreement is found (85). Finally, when a first person exclusive pronoun is coordinated with a third person noun phrase, first person exclusive plural agreement is found (86).

(85) Bill e min kö Afriki mèn-djodi. (Fula)
Bill and I in Afrika live
3 1 1EXCL
'Bill and I, we live in Africa.'

(86) Bill e mènèn kö Afriki mèn-djodi. (Fula)
Bill and we-Excl in Afrika live
3 1EXCL 1EXCL
'Bill and us, we live in Africa.'

These facts are predicted by our union analysis if we assign the sets of person agreement features in 87.[15]

(87)

	SINGULAR	PLURAL
1.	min: {S}	*inclusive*: enèn: {S,H}
		exclusive: mènèn: {S}
2.	an: {H}	onön: {H}
3.	ö: {}	kombè:{}

Assuming that the PERSON feature of a coordinate noun phrase is the union of the PERSON features of the conjuncts predicts exactly this pattern; the table in 88 recapitulates the facts given in 78. For example, if a second person pronoun and a third person pronoun are coordinated, the person value for the coordinate noun phrase is the union of the second and third person features. Similarly, if a first person singular pronoun and a second person pronoun are coordinated, the person value for the coordinate noun phrase is again the union of the first and second person features.

[15] Nigel Vincent pointed out to us that our representation of the third person feature as the empty set {} can be taken to reflect the unmarked status of this value on a markedness hierarchy (for more discussion of markedness and the representation of agreement features, see Blevins 2000).

(88) Fula

$$
\begin{aligned}
\{S\}\ (1\text{SG}) \cup \{H\}\ (2) &= \{S, H\}\ (1\text{INC.PL})\\
\{S\}\ (1\text{SG}) \cup \{H\}\ (2) \cup \{\}\ (3) &= \{S, H\}\ (1\text{INC.PL})\\
\{S\}\ (1\text{SG}) \cup \{\}\ (3) &= \{S\}\ (1\text{EXC.PL})\\
\{S\}\ (1\text{EXC.PL}) \cup \{\}\ (3) &= \{S\}\ (1\text{EXC.PL})\\
\{H\}\ (2) \cup \{\}\ (3) &= \{H\}\ (2)\\
\{\}\ (3) \cup \{\}\ (3) &= \{\}\ (3)
\end{aligned}
$$

The union operation produces f-structures like 89, which satisfy constraints such as the ones given in 90.

(89)

$$
c: \left[\begin{array}{l}
\left\{\begin{array}{l}
i: \left[\begin{array}{l}\text{PRED 'PRO'}\\ \text{NUM SG}\\ \text{PERSON } \{S\}\end{array}\right]\\
y: \left[\begin{array}{l}\text{PRED 'PRO'}\\ \text{NUM SG}\\ \text{PERSON } \{H\}\end{array}\right]
\end{array}\right\}\\
\text{NUM PL}\\
\text{PERSON } \{S, H\}
\end{array}\right]
$$

(90)

$$
\begin{aligned}
(i\ \text{PERSON}) &= \{S\}\\
(y\ \text{PERSON}) &= \{H\}\\
(c\ \text{PERSON}) &= \{S\} \cup \{H\} = \{S, H\}
\end{aligned}
$$

6.2. FEWER PRONOMINAL DISTINCTIONS. Many languages, for example English, Spanish, and Slovak, do not make all the distinctions described above in their pronominal systems, but instead have a single first person plural pronominal form that collapses the inclusive/exclusive distinction. This raises a question: Do these languages make the full set of syntactic distinctions of Fula, thus preserving the correlation between marker sets and reference? Or is the inclusive/exclusive distinction neutralized in the syntax of languages like English, Spanish, and Slovak, with a concomitant weakening of the relation between reference and marker sets?

Consider the first alternative: English, like Fula, has two first-person plural values, {S} for first person exclusive, and {S, H} for first person inclusive. The coordinate phrase *you and I* has the person value {S, H}, while the phrase *Bill and I* has a different person value, {S}. As noted, this preserves the referential values of the person marker sets, but at the cost of introducing a distinction that is unattested by any agreement phenomena in the language.

The second alternative is to collapse the inclusive/exclusive distinction in English, and to assume that coordinate phrases like *Bill and I* and *you and I* have the same value for the person feature. Since the second person introduces the H marker, this common value must be {S, H}. And since third person phrases introduce no marker, the first person singular pronoun must itself be associated with the marker set {S, H}. This alternative does not introduce internally unmotivated syntactic distinctions into the syntax of languages like English and Spanish, at the cost of associating the set {S, H} with the first person singular pronoun and thus breaking the connection between marker sets and pronominal reference.

The choice between these two feature encodings thus depends on the relative importance assigned to two competing principles, the principle that a syntactic distinction is assumed only if it is attested versus the principle that marker sets should uniformly reflect the referent of a phrase. Despite its intuitive appeal, however, other data suggest that the referential principle cannot be strictly maintained. Mark Aronoff pointed out to us that the English second person pronominal form can also be used for impersonal third person reference while still requiring second-person agreement, and the French third person singular pronominal form *on* can be used for first person plural colloquial reference. This breakdown in the referential correlation undermines the argument for including unattested distinctions in the English marker system. Thus, we prefer to give up the direct referential correspondence in favor of an encoding that makes only the attested syntactic distinctions.

For English, Spanish, Slovak, and similar languages, then, we suggest that only three of the four possible combinations of markers S and H are realized, with a single set being used for all first person forms:

{S, H}: first person
{H}: second person
{}: third person

Agreement patterns for such languages are shown in 91.

(91) English, Spanish, Slovak, . . .:
1 & 2 = 1
1 & 3 = 1
2 & 3 = 2
3 & 3 = 3

The table in 92 represents the same information as in 91, and shows that treating resolution as the union of conjunct sets again makes the correct predictions.

(92) $\{S, H\}\ (1) \cup \{H\}\ (2) = \{S, H\}\ (1)$
$\{S, H\}\ (1) \cup \{\}\ (3) = \{S, H\}\ (1)$
$\{H\}\ (2) \cup \{\}\ (3) = \{H\}\ (2)$
$\{\}\ (3) \cup \{\}\ (3) = \{\}\ (3)$

6.3. MORE COMPLICATED PERSON SYSTEMS? Given our inventory of person markers, S and H, no language can make more than four distinctions in its person system for purposes of feature resolution. These four possibilities are fully attested in languages with distinctions between first person inclusive, first person exclusive, second person, and third person. There has been speculation about the existence of more complicated pronominal systems: Boas (1911:35) contemplates a system in which one of the first person plural pronouns includes the speaker and hearer or hearers only (the 'limited inclusive'), another includes the speaker and excludes the hearer or hearers (the exclusive), and still another includes the speaker, the hearer, and other persons not present (the 'general inclusive'); the terminology comes from Foster & Foster 1948. Such a system would be more elaborate than the one we have proposed, requiring, for example, a third marker O (for 'other') in addition to S and H to make all the necessary distinctions.

As far as we know, there is in fact no language that makes these additional distinctions. The only apparent support for the necessity of a more complex formal system comes from the affix pronouns of Sierra Popoluca (Foster & Foster 1948, Forchheimer 1953, Elson 1960),[16] which are reported by Foster and Foster (1948) to distinguish limited and general inclusive first person forms. This claim, however, has been dismissed by Noyer (1992), who shows convincingly that the form referred to by Foster

[16] We are grateful to María-Eugenia Niño for assistance with the Sierra Popoluca data.

and Foster as the 'limited inclusive' is actually a first person inclusive DUAL form, referring to the speaker and a single hearer. If reference to multiple hearers is involved, the 'general inclusive'/inclusive plural form is used.

Zwicky (1977) also contemplates the existence of more complicated pronominal systems, but again these systems are not attested. Thus, there is no evidence to justify a more elaborate set of markers than the S, H system we have proposed.

6.4. FORMAL SPECIFICATION OF FEATURE RESOLUTION. We have seen that we obtain the appropriate value for the nondistributive PERSON feature of a coordinate phrase if we represent PERSON values as sets and define the PERSON value for a coordinate structure to be the union of the PERSON values of all the conjuncts. LFG allows statements of relations between the f-structures of mother and daughter nodes, via the $\uparrow$ and $\downarrow$ metavariables. We can provide a simple formalization of feature resolution, based on the proposition in 93.

(93) $x \cup y$ is the smallest set z such that $x \subseteq z \wedge y \subseteq z$

Thus, the union will result if we take the SMALLEST set that satisfies a collection of separately stated MOTHER-DAUGHTER SUBSET assertions. Both these conditions are easily achieved within the LFG architecture, since, as noted in §1, LFG f-structures are the minimal models for the functional description of an utterance. Thus, the smallest set is naturally picked out as the mother's PERSON value. The necessary subset relations are established by annotations on the LFG phrase structure rule for coordinate noun phrases.

(94)
$$\begin{array}{llll} \text{NP} \longrightarrow & \text{NP} & \text{CONJ} & \text{NP} \\ & \downarrow\in\uparrow & & \downarrow\in\uparrow \\ & (\downarrow \text{PERSON}) \subseteq (\uparrow \text{PERSON}) & & (\downarrow \text{PERSON}) \subseteq (\uparrow \text{PERSON}) \end{array}$$

The annotations on each conjunct use only the $\uparrow$ and $\downarrow$ metavariables to assert that the PERSON markers for the conjunct are contained in the person value for the coordinate structure as a whole.

Example 95 shows that a Spanish coordinate structure with a second person and a third person conjunct takes second person agreement; the f-structure for this example is 96.

(95) $\boxed{\text{José y tu}}$ habláis (Spanish)
José and you speak.2PL

(96) José y tu habláis

$$f: \begin{bmatrix} \text{PRED} & \text{'SPEAK'} \\ \text{SUBJ } c: & \begin{bmatrix} \left\{ \begin{array}{l} j: \begin{bmatrix} \text{PRED 'JOSÉ'} \\ \text{NUM SG} \\ \text{PERSON } \{\} \end{bmatrix} \\ t: \begin{bmatrix} \text{PRED 'PRO'} \\ \text{NUM SG} \\ \text{PERSON } \{\text{H}\} \end{bmatrix} \end{array} \right\} \\ \text{NUM PL} \\ \text{PERSON } \{\text{H}\} \end{bmatrix} \end{bmatrix}$$

Rule 94 asserts that the PERSON value for each of the conjuncts is a subset of the PERSON value of the coordinate structure as a whole.

(97) (j PERSON) $\{\} \subseteq$ (c PERSON)
(t PERSON) $\{H\} \subseteq$ (c PERSON)

The smallest set that satisfies these conditions is the set {H}, the second person value, as desired; the nonminimal first person set {S,H}, which also satisfies the conditions in 97, is not allowed. Only the smaller set appears in the minimal model for the coordinate NP.

We can now account for patterns of verb agreement. The first person verb *hablamos* has a lexical entry with the specifications in 98.

(98) hablamos: ($\uparrow$ SUBJ NUM) = PL
($\uparrow$ SUBJ PERSON) $=_c$ {S, H}

The constraining equation ($=_c$) requires that the PERSON feature in the minimal model be the first person set {S,H}. However, the PERSON set for the coordinate subject is {H}, which does not satisfy this requirement. Therefore, this verb cannot be used if its subject is *José y tu* 'José and you'.

The second person verb in example 95, *habláis*, imposes the following requirements:

(99) habláis: (f SUBJ NUM) = PL
(f SUBJ PERSON) $=_c$ {H}

The constraining equation for this verb asserts that the PERSON of the verb's SUBJ is the set {H}. This condition is satisfied for example 95.

6.5. INTERSECTION OR UNION? Sag et al. 1985 also makes use of sets to encode values for the person feature, but provides a different means for determining the person value for a coordinate structure. The authors propose that the feature on a coordinate NP is the intersection, not the union, of features on the conjuncts. They specify the sets of markers in 100.[17]

(100) 1 {}
2 {+XSP}
3 {+XSP, +THP}

This gives the desired results for English, Slovak, and Spanish:

(101)
$$\begin{aligned} \{\}\ (1) \cap \{+XSP\}\ (2) &= \{\}\ (1) \\ \{\}\ (1) \cap \{+XSP, +THP\}\ (3) &= \{\}\ (1) \\ \{+XSP\}\ (2) \cap \{+XSP, +THP\}\ (3) &= \{+XSP\}\ (2) \\ \{+XSP, +THP\}\ (3) \cap \{+XSP, +THP\}\ (3) &= \{+XSP, +THP\}\ (3) \end{aligned}$$

This particular set of values does not work, however, for languages like Fula that have different verb agreement forms for inclusive and exclusive first person plurals. In Fula the coordinate phrase corresponding to *you and I* takes first person plural INCLUSIVE agreement, while the coordinate phrase corresponding to *Bill and I* takes first person plural EXCLUSIVE agreement. Sag et al.'s analysis fails because its value for the first person number feature is the empty set, and intersecting the empty set with any other set gives the empty set. Thus it is in principle impossible to distinguish different kinds of coordination involving a first person pronoun.

It turns out however, that any analysis in which features are combined by union

[17] Karttunen (1984) suggests a variant of the Sag et al. approach, which makes the same predictions and also suffers from the same problems (outlined below) with languages that distinguish inclusive and exclusive forms of the first person plural pronoun, such as Fula. Farkas and Ojeda (1983) also provide a discussion of number and gender resolution in coordination, noting issues that arise within the GPSG framework, but they do not make a specific proposal for handling person or gender resolution.

can be transformed into a corresponding analysis in which features are combined by intersection. This dual analysis is obtained by applying De Morgan's laws: intersection is used instead of union, and the sets are replaced by their complements. Applying this transformation to our previous account of Fula gives 102.

(102) Fula, alternative analysis involving complement sets

SINGULAR		PLURAL
1. min: {H}	*inclusive*:	enèn: {}
	exclusive:	mènèn: {H}
2. an: {S}		onön: {S}
3. ö: {S, H}		kombè: {S, H}

These person values seem unintuitive, but it is not difficult to reinterpret them in a more natural way: interpret H as indicating that the hearer is ABSENT from the referent set, and S as indicating that the speaker is absent. It is easy to verify that these assignments give the correct results when the PERSON value for the coordinate structure is obtained by intersecting the PERSON values of its conjuncts:

(103)
{H} (1SG) ∩ {S} (2) = {} (1INC.PL)
{} (1INC.PL) ∩ {S, H} (3) = {} (1INC.PL)
{H} (1SG) ∩ {S, H} (3) = {H} (1EXC.PL)
{S} (2) ∩ {S, H} (3) = {S} (2)
{S, H} (3) ∩ {S, H} (3) = {S, H} (3)

Whether union or intersection should be used to combine sets is, then, not an issue of whether an analysis is mathematically possible or impossible: exactly the same range of facts can be covered with either resolution operator. But an intersection analysis is not easy to state within the standard LFG architecture. Expressing this analysis by means of equations appearing on the individual conjunct daughters would require the PERSON value of the coordinate structure to be the largest/most fully specified subset of the PERSON values of the individual conjuncts. This is incompatible with the LFG principle of taking the minimal solution to a collection of constraints. In less restrictive theories, such as HPSG, the operations of set union and intersection may already be available as explicit operators on grammatical structures, and schemata may be used to refer directly to the arguments of the union and intersection operations. In those theories, since the operations may be explicitly stated, there may be no formal motivation to choose union over intersection to characterize feature resolution.

7. GENDER RESOLUTION. As we have noted, there is a near-iconic relation between the sets representing the PERSON value and the referents of the phrases bearing these features: The mnenomic names S and H, for speaker and hearer, very nearly represent the semantic contribution of the pronouns bearing these features. In contrast, gender systems vary much more widely across languages than person systems: Many languages have no syntactic gender distinctions, or have gender only in certain limited paradigms; other languages have two, three, or more syntactic genders. Languages also vary in their patterns of gender resolution. This variability is related to the fact that syntactic gender systems often do not reflect 'natural' or semantic gender; for instance, the German noun *Mädchen* 'girl' is syntactically neuter, though semantically feminine. Thus, compared to the PERSON feature, the marker sets for the values of the syntactic GENDER feature are less susceptible to an intuitive motivation on the basis of semantic facts.

The complex relationship between syntactic and semantic gender is evident in patterns of gender resolution in many languages. Corbett (1983b) points out that in many languages, gender agreement with coordinate phrases is semantically based, depending on referential, nonsyntactic properties of the conjuncts such as animacy or humanness rather than syntactic gender. Wechsler (1999) also explores this phenomenon, claiming that although agreement with conjoined inanimates usually depends on syntactic resolution rules, agreement with conjoined animates often depends only on semantic factors (104).

(104) [La personne avec la barbe] et Marie sont idiots/?*idiotes
[the person with the beard]-FEM and Marie-FEM are idiots-MASC/?*FEM
'The person with the beard and Marie are idiots.' (French)

In this example, the predicate shows MASC agreement with the conjoined phrase *la personne avec la barbe et Marie* based on semantic gender (one conjunct denotes a male), even though each conjunct is syntactically feminine. Such examples do not involve syntactic feature resolution and thus are not relevant to the present discussion.

There are certain common and recurring characteristics of syntactic gender resolution:

(105) a. A coordinate phrase behaves as if it has the same gender as some noncoordinate element in the language. That is, there are no special gender agreement markers that appear only with coordinate phrases.
b. A coordinate structure with same-gender elements behaves as if it has that same gender.
c. A coordinate structure with mixed-gender elements may or may not behave as if it has the gender of one of its conjuncts (but it does behave as if it has a gender that is independently found in the language).

These characteristics follow if we treat the GENDER feature exactly like the PERSON feature: Values for the GENDER feature are encoded as sets of markers, and the GENDER of a coordinate structure is the union of the GENDER features of the individual conjuncts. This generalization is expressed in the annotated rule 106.

(106) NP $\rightarrow$

NP	CONJ	NP
$\downarrow \in \uparrow$		$\downarrow \in \uparrow$
$(\downarrow \text{PERSON}) \subseteq (\uparrow \text{PERSON})$		$(\downarrow \text{PERSON}) \subseteq (\uparrow \text{PERSON})$
$(\downarrow \text{GENDER}) \subseteq (\uparrow \text{GENDER})$		$(\downarrow \text{GENDER}) \subseteq (\uparrow \text{GENDER})$

As with the PERSON feature, the GENDER value of the coordinate phrase will be the smallest set containing the GENDER values of the individual conjuncts. To illustrate how this general rule operates, we first examine instances in which the pattern in 105 holds. Then we discuss a departure from the pattern in 105b, showing that it can be analyzed as involving a minor variation on the general theme.

7.1. TWO GENDERS: HINDI. A simple pattern of gender resolution is found in Hindi (McGregor 1972:20), which (like Spanish, French, and many other languages) has two genders, masculine and feminine. When a masculine and a feminine noun phrase are coordinated, the coordinate phrase has masculine gender (107). A coordinate phrase has feminine gender only when all of its conjuncts are feminine (108).

(107)

meraa kuttaa aur merii billii	mere saath g^h^ar mẽ rahte hãĩ.	(Hindi)
my dog and my cat	with me house LOC live	
MASC FEM	SUBJ GENDER = MASC	

'My dog and my cat live with me in the house.'

(108) $\boxed{\text{yah laṛki aur uski mãã / this girl and her mother / FEM FEM}}$ dilli mẽ rahtii hãĩ. (Hindi)

yah	laṛki	aur	uski	mãã	dilli	mẽ	rahtii	hãĩ.
this	girl	and	her	mother	Delhi	LOC	live	
	FEM			FEM			SUBJ GENDER = FEM	

'This girl and her mother live in Delhi.'

The rules for gender resolution in Hindi can be expressed informally in the following way (Corbett 1983a, 1991):

(109) Hindi
- If the conjuncts include at least one masculine noun phrase, the masculine form is used.
- Otherwise the feminine form is used.

The observations in 109 can be expressed in tabular form, with the ampersand again representing coordination.

(110) Hindi
MASC & MASC = MASC
MASC & FEM = MASC
FEM & FEM = FEM

We see that, in accordance with the generalization given in 105, coordination of like genders preserves that gender, and coordination of mixed genders behaves as a masculine phrase. The distinctions of this two-way gender system can be represented as in 111, using a single primitive marker M.

(111) Hindi
MASC {M}
FEM {}

Our assumption that the gender feature for a coordinate structure is obtained by taking the union of the gender features of the conjuncts gives the right result, as 112 shows. The cases it illustrates are parallel to the table in 110.[18]

(112) Hindi
$\{\text{M}\}$ (MASC) $\cup$ $\{\text{M}\}$ (MASC) = $\{\text{M}\}$ (MASC)
$\{\text{M}\}$ (MASC) $\cup$ $\{\}$ (FEM) = $\{\text{M}\}$ (MASC)
$\{\}$ (FEM) $\cup$ $\{\}$ (FEM) = $\{\}$ (FEM)

Our union analysis provides the gender value {M} for the coordinate structure in 107 and the value {} for the coordination in 108. These are compatible with the verbal agreement constraints in 113.

(113) a. rahte: $(\uparrow \text{ SUBJ GENDER}) =_c \{\text{M}\}$
b. rahtii: $(\uparrow \text{ SUBJ GENDER}) =_c \{\}$

7.2. THREE GENDERS: ICELANDIC. Icelandic exhibits a slightly more complicated pattern of gender resolution. It has three genders, and Corbett (1983b) shows that the following generalization holds:

[18] The marker sets for the person feature (see n. 15) seem to reflect an independently motivated markedness hierarchy, with the unmarked third person value represented as the empty set. It is tempting to speculate that GENDER values might also be assigned to correlate with a markedness hierarchy. Corbett (1991), however, observes that there is no consistent relationship across languages between markedness and gender resolution: A mixed-gender coordinate phrase does not always resolve to the unmarked gender.

(114) Icelandic
- If the conjuncts are all masculine, the masculine form is used.
- If the conjuncts are all feminine, the feminine form is used.
- Otherwise the neuter form is used.

Again, like genders are preserved, and any structure involving coordination of mixed genders has neuter gender.[19]

(115) [drengurinn og telpan] eru þreytt. (Icelandic)
the.boy and the.girl are tired
MASC FEM SUBJ GENDER = NEUT
'The boy and the girl are tired.'

(116) [maðurinn og barnið] eru þreytt. (Icelandic)
the.man and the.baby are tired
MASC NEUT SUBJ GENDER = NEUT
'The man and the baby are tired.'

(117) ég sá [á og lamb] bæði svört. (Icelandic)
I saw a.ewe and a.lamb both black
FEM NEUT SUBJ GENDER = NEUT
'I saw a ewe and a lamb, both black.'

These results can be restated in tabular form in 118.

(118) Icelandic

MASC	& MASC	=	MASC
FEM	& FEM	=	FEM
NEUT	& NEUT	=	NEUT
MASC	& FEM	=	NEUT
MASC	& NEUT	=	NEUT
FEM	& NEUT	=	NEUT

We assign the following sets for MASC, FEM, and NEUT gender in Icelandic:

(119) Icelandic
MASC {M}
FEM {F}
NEUT {M, F}

Example 120 parallels the table in 118 and illustrates that these assignments make the correct predictions.

(120) Icelandic

$$\begin{aligned}
\{\text{M}\}\ (\text{MASC}) \cup \{\text{M}\}\ (\text{MASC}) &= \{\text{M}\}\ (\text{MASC})\\
\{\text{F}\}\ (\text{FEM}) \cup \{\text{F}\}\ (\text{FEM}) &= \{\text{F}\}\ (\text{FEM})\\
\{\text{M, F}\}\ (\text{NEUT}) \cup \{\text{M, F}\}\ (\text{NEUT}) &= \{\text{M, F}\}\ (\text{NEUT})\\
\{\text{M}\}\ (\text{MASC}) \cup \{\text{F}\}\ (\text{FEM}) &= \{\text{M, F}\}\ (\text{NEUT})\\
\{\text{M}\}\ (\text{MASC}) \cup \{\text{M, F}\}\ (\text{NEUT}) &= \{\text{M, F}\}\ (\text{NEUT})\\
\{\text{F}\}\ (\text{FEM}) \cup \{\text{M, F}\}\ (\text{NEUT}) &= \{\text{M, F}\}\ (\text{NEUT})
\end{aligned}$$

[19] We are grateful to Kristjan Arnason for discussion of these examples.

Given the features in 119 and appropriate agreement constraints, our general rule 106 accounts for this pattern.

7.3. UNEXPECTED AGREEMENT WITH SAME-GENDER COORDINATION: SLOVENE. Some languages exhibit a phenomenon that violates the generalization in 105b: coordinate structures with conjuncts of the same gender unexpectedly require agreement of a different gender. For example, in the South Slavonic language Slovene (Corbett 1983a, 1991), a neuter plural verb form exists, but it is not used for coordinated phrases in which the conjuncts are neuter; instead, the masculine plural verb form is used.[20]

Corbett claims that the pattern of gender agreement in 121 holds for Slovene.

(121) Slovene (Corbett 1983a:186)
- If all conjuncts are feminine, then the feminine form is used.
- Otherwise the masculine form is used.

This pattern resembles the Icelandic pattern discussed in §7.2 in all but two respects: the masculine form rather than the neuter form is used for mixed-gender coordination, and, surprisingly, masculine agreement is required when all conjuncts are neuter. We can begin to account for this pattern by postulating the following sets:

(122) Slovene

MASC {F, N}
FEM {F}
NEUT {N}

These sets have essentially the same structure as the Icelandic sets, with the MASC set playing a role analogous to the NEUT set of Icelandic. However, the fact that coordination of neuter conjuncts takes masculine agreement is not yet accounted for.[21]

(123)

To drevo in	gnezdo	na njem	mi	bosta	ostala	v spominu.
that.tree and	the.nest	on.it	to.me	will	remain	in memory
NEUT	NEUT				SUBJ GENDER = MASC	

'That tree and the nest on it will remain in my memory.' (Slovene)

(124) Slovene

MASC & MASC	=	MASC
MASC & FEM	=	MASC
MASC & NEUT	=	MASC
MASC & FEM & NEUT	=	MASC
FEM & NEUT	=	MASC
FEM & FEM	=	FEM
NEUT & NEUT	=	MASC

We present two possible accounts of this phenomenon. The first account assumes

[20] Corbett (1983a, 1991) provides an extensive discussion of gender resolution in Slavic, pointing out that West Slavic languages exhibit interestingly different patterns of gender resolution with neuter conjuncts as compared to South Slavonic languages like Slovene. The West Slavic language Czech is particularly well studied in this regard; see, for example, Sgall et al. 1986.

[21] Corbett points out that there is actually some degree of controversy over resolution patterns for gender in Slovene. According to work by Bajec, cited by Corbett (1983a, 212), coordination of neuter forms requires NEUT rather than MASC agreement, a pattern that is directly captured by the sets given in 122 and raises no further issues.

that gender resolution takes place in the expected way, given the sets in 122, and that some other property of the coordinate neuter phrase is responsible for its peculiar agreement behavior. That is, a coordinate structure with neuter conjuncts is in fact neuter plural. However, coordinate noun phrases must also have some special property unrelated to gender marking which noncoordinate phrases do not share. The neuter plural verb form forbids its subject to have this property, and as a result, this verb form can appear with a noncoordinate neuter plural subject but not a coordinate one. On the other hand, masculine plural verbs are marked so as to allow noncoordinate masculine plural phrases as well as coordinate masculine and neuter phrases. Given this account, we would expect at least some phenomena other than verb agreement to treat coordinated neuter phrases as neuter, since they do in fact bear that gender.

In contrast, it is possible that a coordinate phrase with NEUT conjuncts actually carries the features of a masculine noun phrase. Evidence from Lama (Yu 1988) supports the general possibility that a feature value of a coordinate phrase can be determined independently of conjunct values. Lama noun phrases belong to different classes, as in Xhosa. Yu shows that all coordinate phrases in Lama belong to noun class 2, regardless of the noun class of their conjuncts. Unlike the analysis of Xhosa presented in §4.4, we would analyze noun class in Lama as a nondistributive feature. This feature is not resolved (determined on the basis of the features of the conjuncts) but is always set to the value 2 as an additional annotation on the CONJ in the general coordination rule (125).

(125) CONJ
$(\uparrow \text{CLASS}) = 2$

In Lama, then, the coordinate structure does not just contribute to but completely determines the noun class of the coordinate phrase.

We can provide an analysis of Slovene along similar lines: the conjunction contributes a constraint affecting the gender of the coordinate phrase. We specialize the general coordination rule in 106 so that it adds the element F to the set representing the gender of the coordinate phrase:

(126) CONJ
$\text{F} \in (\uparrow \text{GENDER})$

With this addition to the rule, the pattern in 124 is obtained for Slovene. In the table in 127 we have indicated the contribution of the coordination with an ampersand.

(127) Slovene

$$\begin{aligned}
\{\text{F, N}\}\ (\text{MASC}) \cup \{\text{F, N}\}\ (\text{MASC}) \cup \{\text{F}\}\ (\&) &= \{\text{F, N}\}\ (\text{MASC})\\
\{\text{F, N}\}\ (\text{MASC}) \cup \{\text{F}\}\ (\text{FEM}) \cup \{\text{F}\}\ (\&) &= \{\text{F, N}\}\ (\text{MASC})\\
\{\text{F, N}\}\ (\text{MASC}) \cup \{\text{N}\}\ (\text{NEUT}) \cup \{\text{F}\}\ (\&) &= \{\text{F, N}\}\ (\text{MASC})\\
\{\text{F, N}\}\ (\text{MASC}) \cup \{\text{F}\}\ (\text{FEM}) \cup \{\text{N}\}\ (\text{NEUT}) \cup \{\text{F}\}\ (\&) &= \{\text{F, N}\}\ (\text{MASC})\\
\{\text{F}\}\ (\text{FEM}) \cup \{\text{N}\}\ (\text{NEUT}) \cup \{\text{F}\}\ (\&) &= \{\text{F, N}\}\ (\text{MASC})\\
\{\text{N}\}\ (\text{NEUT}) \cup \{\text{N}\}\ (\text{NEUT}) \cup \{\text{F}\}\ (\&) &= \{\text{F, N}\}\ (\text{MASC})\\
\{\text{F}\}\ (\text{FEM}) \cup \{\text{F}\}\ (\text{FEM}) \cup \{\text{F}\}\ (\&) &= \{\text{F}\}\ (\text{FEM})
\end{aligned}$$

Unlike our first analysis, this account asserts that coordinated NEUT phrases actually become MASC PL and should behave as masculine phrases in every circumstance. We have not been able to find evidence of other Slovene phenomena that would allow us to determine which of our two accounts is correct.

8. INDETERMINACY IN GENDER AND NUMBER. We now briefly explore the question of whether the kind of indeterminacy discussed in §4 can be found with other features.

Examples like 128 suggest that the value of the GENDER feature cannot be indeterminate. The Hindi third person pronoun *wah* is ambiguously masculine and feminine, but it cannot be treated as masculine and feminine at the same time.

(128) *wah pahũcaa aur pahučĩi. (Hindi)
he/she arrived.MASC and arrived.FEM
MASC/FEM SUBJ GENDER = MASC SUBJ GENDER = FEM
'He/she arrived (masculine) and arrived (feminine).'

But other examples seem to indicate the contrary: that some forms are indeterminately specified for the GENDER feature. Eisenberg (1973) and Pullum and Zwicky (1986) discuss examples like 129.

(129) der oder die älteste Abgeordnete (German)
the.MASC or the.FEM oldest representative
'the oldest representative (masculine) or representative (feminine)'

Here, the common gender noun *Abgeordnete* 'representative' is shared between two determiners specified for different genders. This is not possible for unambiguously masculine nouns like *Lehrer* or unambiguously feminine nouns like *Lehrerin*:

(130) *der oder die Lehrer/ Leherin (German)
the.MASC or the.FEM teacher.MASC/ teacher.FEM
'the teacher or teacher'

Patterns similar to these can also involve the number feature. Even where one form can be used sometimes as singular and sometimes as plural, it is not possible for that form to be used as singular and plural at the same time (Zaenen & Karttunen 1984, Pullum & Zwicky 1986), as 131–132 illustrate.

(131) *Ich habe den Dozenten gesehen und geholfen. (German)
I have the docent seen and helped
ACC SG/DAT PL SUBJ CASE = ACC SUBJ CASE = DAT
'I have seen and helped the docent/docents.'

(132) *Sie singt und singen. (German)
she/they sings and sing
3SG/3PL SUBJ = 3SG SUBJ = 3PL
'She/they sings and sing.'

But Eisenberg 1973 and Zaenen & Karttunen 1984 present examples showing that an unmarked noun can be shared between two determiners, one singular and one plural, as in 133.

(133) der Antrag des oder der Dozenten (German)
the petition of the-SG or the-PL docents
'the petition of the docent or docents'

Analogous examples in English have been judged marginally grammatical by Zaenen and Karttunen (1984), though we do not find examples like 134 acceptable.

(134) %this and these sheep

English examples with coordinated singular and plural nouns and a shared determiner, however, are fully acceptable.

(135) a. *this representatives or representative
b. *these representatives or representative
c. the representatives or representative

To account for these facts, Ingria (1990) and Bayer (1996) draw a distinction between semantically 'potent' features like number and semantically 'impotent' features like case; we assume that they would also classify gender and person as semantically potent like the number feature. They claim that phenomena involving feature indeterminacy can occur only with semantically impotent features, thus accounting for the unacceptability of examples 128, 131, and 132. According to their analysis, these examples are ungrammatical because they involve the semantically potent and therefore determinate features NUM and GENDER. To account for the acceptability of examples like 133, they introduce an exception to this general claim. They propose that the number feature can in fact be indeterminate, although only in limited circumstances: subconstituents of NP, but not NP itself, can be indeterminately specified for number. We suppose that 129 would lead them to the same conclusion for the GENDER feature.

Example 136, however, shows that this proposal is incorrect.[22]

(136) *der Antrag des Dozenten, die gesprochen haben
the petition of the-SG docent that-PL spoken have-PL
'the petition of the docent that have spoken'

Although the agreement involved in this example is NP-internal, the example is still ungrammatical. There are also cases of ungrammaticality involving NP-internal GENDER agreement (137).[23]

(137) *Dieser Abgeordnete, die sich im Spiegel sah (German)
this.MASC representative that.FEM self in mirror saw
'this (masculine) representative who saw herself in the mirror'

The Hindi examples in 138 show that the same noun cannot be modified by adjectives of different genders.[24]

(138) a. wah saccaa aur acchaa pradhaanmãtrii thaa.
(s)he honest.MASC and good.MASC prime minister was.MASC
'He was an honest and good prime minister.'
b. wah saccii aur acchii pradhaanmãtrii thĩĩ.
(s)he honest.FEM and good.FEM prime minister was.FEM
'She was an honest and good prime minister.'
c. *wah saccii aur acchaa pradhaanmãtrii thaa/thĩĩ.
(s)he honest.FEM and good.MASC prime minister was.MASC/FEM
'He/she was an honest and good prime minister.'

We see that these examples cannot be explained by appealing to a distinction in the way features behave inside and outside NP. Instead, the distinguishing property of the acceptable examples is that they involve either determiner or nominal coordination, as in 129, 133, and 135.

[22] We are grateful to Thomas Breuel for providing the German example 136.

[23] We are grateful to Christian Rohrer for providing the German example 137.

[24] We are grateful to Vineet Gupta for providing the Hindi examples in 138.

Given these facts, we could simply refine the stipulation that Ingria and Bayer make, proposing that the gender and number features can be indeterminate only in the more restrictive cases of determiner or noun coordination. But this begs the question of why they should behave differently in just these constructions. Our framework suggests a different approach, one that assumes that these features are never indeterminate. Rather, the difference in behavior follows from the LFG representation of coordinate structures. The conjuncts in a coordinate structure correspond to different f-structures, and those can take on different values for the GENDER feature in the acceptable coordinate cases 129 and 133.

(139) der oder die Abgeordnete (German)
the.MASC or the.FEM representative
'the representative (masculine) or representative (feminine)

$$\left\{ \begin{array}{l} \left[\begin{array}{ll} \text{SPEC} & \text{THE} \\ \text{GENDER} & \{\text{M}\} \\ \text{PRED} & \text{`REPRESENTATIVE'} \end{array} \right] \\ \left[ \begin{array}{ll} \text{SPEC} & \text{THE} \\ \text{GENDER} & \{\} \\ \text{PRED} & \text{`REPRESENTATIVE'} \end{array} \right] \end{array} \right\}$$

Though we do not provide an explicit treatment here, we believe that a successful analysis of these examples will emerge from a careful consideration of the interactions between coordination and feature distribution.

In cases not involving determiner or noun coordination, such as examples 128, 136, 137, and 138c, we expect to observe the effects of feature clashes within a single f-structure. For example, we represent the value of the GENDER feature for an ambiguous pronoun like *wah* in example 138 as a (wide-scope) disjunction and not as an indeterminate set:

(140) wah: $(\uparrow \text{GENDER}) = \{\text{M}\}$
$\vee\ (\uparrow \text{GENDER}) = \{\}$

This accounts for the fact that, as shown in example 128, *wah* cannot be used as a masculine and feminine pronoun at the same time: it is ambiguous, not indeterminate. Both of the sentences in 141 are acceptable.

(141) a. wah pahũcaa. (Hindi)
he/she arrived.MASC
MASC/FEM SUBJ GENDER = MASC
'He arrived.'

b. wah pahũcii. (Hindi)
he/she arrived.FEM
MASC/FEM SUBJ GENDER = FEM
'She arrived.'

Similarly, the determiner *des* in 136 requires its noun phrase to be singular, while the relative clause is acceptable only as a modifier of a plural phrase.

Thus, examples in which a resolving feature appears to be indeterminate turn out to involve structures with coordinate nouns or coordinate determiners in which the feature is specified in different ways in different conjuncts. We conclude that resolving features like PERSON and GENDER are not indeterminate. This state of affairs is exactly what our theory of feature representation predicts: An indeterminate resolving feature would have to be a set of possible PERSON or GENDER sets, and our union analysis would not operate properly on such a complex value.

Given our conclusion that features like PERSON and GENDER are not indeterminate, it may be worthwhile to recall our treatment of examples such as 61, repeated here for convenience.

(142) (=61)
weil [wir das Haus] und [die Müllers den Garten] kaufen (German)
because we the house and the Müllers the garden buy
1 3 SUBJ = 1/3
'because we buy the house and the Müllers buy the garden'

This is not a counterexample to our claim that the PERSON of the subject of each conjunct is determinately specified. As noted by Ingria (1990) and discussed in §4.4, the verb can place an indeterminate constraint on any of its subject's features, even those that are themselves determinate. This indeterminate constraint is satisfied separately in each conjunct.

9. CONCLUSION. We have provided a general theory of feature representation that easily accounts for feature indeterminacy and feature resolution within the LFG framework. Our set representations, together with minimal extensions to LFG's description language, enable a simple and intuitive characterization of both these phenomena. We represent indeterminate values by sets containing the alternative possibilities. This approach to indeterminacy makes the right predictions but does not require large-scale formal revisions to avoid the difficulties that previous proposals encountered. We represent resolving features as sets of markers that are combined by union in coordinate structure feature resolution. This provides an explanation for the seemingly complicated behavior of the person and gender features in coordinate structures. Although we have used the conceptual tools and techniques of LFG in formulating our theory, we believe that similar solutions could be framed within other syntactic approaches.

REFERENCES

ANDREWS, AVERY, III. 1983. Constituent coordination in LFG. Canberra: Australian National University, MS.
ARNOTT, D. W. 1970. The nominal and verbal systems of Fula. Oxford: Clarendon Press.
BAYER, SAMUEL. 1996. The coordination of unlike categories. Language 72.579–616.
BERMAN, JUDITH. 1999. Does German satisfy the subject condition? On-line proceedings of the LFG99 Conference, ed. by Miriam Butt and Tracy Holloway King. http://csli-publications.stanford.edu/LFG/4/lfg99.html.
BLEVINS, JAMES P. 2000. Markedness and agreement. Transactions of the Philological Society 98.233-62.
BOAS, FRANZ. 1911. Introduction to the Handbook of American Indian languages. Washington: Government Printing Office. Republished in 1991.
BRESNAN, JOAN (ed.) 1982. The mental representation of grammatical relations. Cambridge, MA:MIT Press.
——; RONALD M. KAPLAN; and PETER PETERSON. 1985. Coordination and the flow of information through phrase structure. Palo Alto, CA: Xerox PARC, MS.
CORBETT, GREVILLE G. 1983a. Hierarchies, targets, and controllers: Agreement patterns in Slavic. London: Croom Helm.
——. 1983b. Resolution rules: Agreement in person, number, and gender. Order, concord, and constituency, ed. by Gerald Gazdar, Ewan Klein, and Geoffrey K. Pullum, 175–206. Dordrecht: Foris.
——. 1991. Gender. Cambridge: Cambridge University Press.

——, and ALFRED D. MTENJE. 1987. Gender agreement in Chicheŵa. Studies in African Linguistics 18.1–38.

DALRYMPLE, MARY, and RONALD M. KAPLAN. 1997. A set-based approach to feature resolution. On-line proceedings of the LFG97 Conference, ed. by Miriam Butt and Tracy Holloway King.http://csli-publications.stanford.edu/LFG/2/lfg97.html.

——; ——; JOHN T. MAXWELL III; and ANNIE ZAENEN (eds.) 1995. Formal issues in lexical-functional grammar. Stanford University: CSLI Publications.

DUPLESSIS, J.A., and M. VISSER. 1992. Xhosa syntax. Pretoria: Via Afrika.

DYLA, STEFAN. 1984. Across-the-board dependencies and case in Polish. Linguistic Inquiry 15.701–5.

EISENBERG, PETER. 1973. A note on 'identity of constituents'. Linguistic Inquiry 4.417–20.

ELSON, BENJAMÍN. 1960. Gramatica Popoluca de la Sierra. Xalapa, Mexico: Universidad Veracruzana.

FARKAS, DONKA, and ALMERINDO OJEDA. 1983. Agreement and coordinate NPs. Linguistics 21.659–73.

FORCHHEIMER, PAUL. 1953. The category of person in language. Berlin: Walter de Gruyter.

FOSTER, MARY L., and GEORGE M. FOSTER. 1948. Sierra Popoluca speech. Smithsonian Institution. Institute of Social Anthropology, publication 8.

GROOS, ANNEKE, and HENK VAN REIMSDIJK. 1979. Matching effects in free relatives: A parameter of core grammar. Theory of markedness in generative grammar: Proceedings of the 1979 GLOW Conference, ed. by Adriana Belletti, Luciana Brandi, and Luigi Rizzi, 171-216. Pisa: Scuola Normale Superiore de Pisa.

INGRIA, ROBERT J. P. 1990. The limits of unification. Proceedings of the 28th annual meeting of the ACL, Pittsburgh, 194–204.

JOHNSON, MARK, and SAMUEL BAYER. 1995. Features and agreement in Lambek categorial grammar. Proceedings of the Formal Grammar Workshop.

KAPLAN, RONALD M., and JOAN BRESNAN. 1982. Lexical-functional grammar: A formal system for grammatical representation. The mental representation of grammatical relations, ed. by Joan Bresnan, 173–281. Cambridge, MA: MIT Press. Reprinted in Dalrymple et al. 1995. 29–130.

——, and JOHN T. MAXWELL. 1988. Constituent coordination in Lexical-Functional Grammar. Proceedings of the 12th International Conference on Computational Linguistics (COLING88), Budapest, vol. 1, 303–5. Reprinted in Dalrymple et al. 1995. 199–210.

——, ——. 1996. LFG grammar writer's workbench. Technical report, Xerox Palo Alto Research Center. URL ftp://ftp.parc.xerox.com/pub/lfg/lfgmanual.ps.

KARTTUNEN, LAURI. 1984. Features and values. Proceedings of the 10th International Conference on Computational Linguistics (COLING84), Stanford, 28–33.

KAYNE, RICHARD S. 1975. French Syntax: The transformational cycle. Cambridge, MA: The MIT Press.

MAXWELL, JOHN T., III, and CHRISTOPHER D. MANNING. 1996. A theory of non-constituent coordination based on finite-state rules. On-line proceedings of the 1st LFG conference, Rank Xerox, Grenoble, August 26–28, 1996, ed. by Miriam Butt and Tracy Holloway King. URL http://csli-publications.stanford.edu/LFG/1/lfg1.html.

MCGREGOR, R. S. 1972. Outline of Hindi grammar. Oxford and Delhi: Oxford University Press.

MÜLLER, STEFAN. 1999. An HPSG-analysis for free relative clauses in German. Grammars 2.53–105.

NOYER, ROLF. 1992. Features, positions, and affixes in autonomous morphological structure. Cambridge, MA: MIT dissertation.

PETERSON, PETER G. 1982. Conjunction in LFG. Cambridge, MA and New South Wales, Australia: MIT and University of Newcastle, MS.

——. 1986. Establishing verb agreement with disjunctively conjoined subjects: Strategies vs. principles. Australian Journal of Linguistics 6.231–49.

POLLARD, CARL, and IVAN A. SAG. 1994. Head-driven phrase structure grammar. Chicago: University of Chicago Press.

PULLUM, GEOFFREY K., and ARNOLD M. ZWICKY. 1986. Phonological resolution of syntactic feature conflict. Language 62.751–73.

ROBERTS, LINDA, and MARJORIE WOLONTIS. 1974. Conjunction and concord in Bantu. Third annual conference on African linguistics, ed. by Erhard Voeltz, 231-42. Indiana University, 7–8 April 1972.

SADLER, LOUISA. 1999. Non-distributive features and coordination in Welsh. On-line Proceedings of the LFG99 Conference, ed. by Miriam Butt and Tracy Holloway King. http://csli-publications.stanford.edu/LFG/4/lfg99.html.

SAG, IVAN A.; GERALD GAZDAR; THOMAS WASOW; and STEVEN WEISLER. 1985. Coordination and how to distinguish categories. Natural Language and Linguistic Theory 3.117–71.

SGALL, PETR; EVA HAJIČOVÁ; and JARMILA PANEVOVÁ. 1986. The meaning of the sentence in its semantic and pragmatic aspects. Dordrecht: D. Reidel.

SHIEBER, STUART M. 1989. Parsing and type inference for natural and computer languages. Palo Alto, CA: Stanford University dissertation.

SILVERSTEIN, MICHAEL. 1976. Hierarchy of features and ergativity. Grammatical categories in Australian languages, ed. by R. M. W. Dixon, 112–71. Canberra: Australian Institute for Aboriginal Studies.

STEINBERG, ELISA, and ALEXANDER F. CASKEY. 1988. The syntax and semantics of gender (dis)agreement: An autolexical approach. Chicago Linguistic Society 24.291–303.

VOELTZ, ERHARD. 1971. Surface constraints and agreement resolution: Some evidence from Xhosa. Studies in African Linguistics 2.37–60.

WECHSLER, STEPHEN. 1999. 'Elsewhere' in gender resolution. Austin: University of Texas, MS.

YU, ELLA OZIER. 1988. Agreement in left dislocation of coordinate structures. Chicago Linguistic Society 24.2.322–36.

ZAENEN, ANNIE, and LAURI KARTTUNEN. 1984. Morphological non-distinctiveness and coordination. Eastern States Conference on Linguistics, 309–20.

ZWICKY, ARNOLD M. 1977. Hierarchies of person. Chicago Linguistic Society 13.714–33.

——. 1991. Systematic versus accidental phonological identity. Paradigms: The economy of inflection, ed. by Frans Plank, 113–31. Berlin: Mouton de Gruyter.

Xerox Palo Alto Research Center
Palo Alto, CA 94304
[dalrymple@parc.xerox.com]
[kaplan@parc.xerox.com]

[Received 1 July 1999;
accepted 12 August 1999]