

A set-based approach to feature resolution

Mary Dalrymple and Ronald M. Kaplan
Xerox Palo Alto Research Center/Stanford University

Proceedings of the LFG97 Conference

University of California, San Diego

Miriam Butt and Tracy Holloway King (Editors)


1997

CSLI Publications

<http://www-csli.stanford.edu/publications/>

1 Traditional approaches to agreement

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 English example, the verb is marked to indicate that its subject is third person and singular:

(1)  Bill arrives.

We represent the agreement relationship by an arrow: here, the verb *arrives* agrees with the subject *Bill*. Under traditional LFG assumptions about the form of agreement features, the f-structure for example (1) is:

$$(2) \quad f : \left[\begin{array}{l} \text{PRED} \quad \text{'ARRIVE'} \\ \text{SUBJ} \quad g : \left[\begin{array}{l} \text{PRED} \quad \text{'BILL'} \\ \text{PERSON} \quad 3 \\ \text{NUM} \quad \text{SG} \end{array} \right] \end{array} \right]$$

In LFG the agreement constraints contributed by the lexical entries for *Bill* and *arrives* are usually instantiated as:

$$(3) \quad \begin{array}{l} \text{arrives:} \quad (f \text{ SUBJ PERSON}) = 3 \\ \quad \quad \quad (f \text{ SUBJ NUM}) = \text{SG} \\ \\ \text{Bill:} \quad (g \text{ PERSON}) = 3 \\ \quad \quad \quad (g \text{ NUM}) = \text{SG} \end{array}$$

We can deduce from annotations on the English phrase structure rules that $(f \text{ SUBJ}) = g$, so we conclude that the agreement features in this example are consistent and the sentence is correctly predicted to be well-formed. On the other hand, the lexical entry for *arrive* in the ungrammatical string **Bill arrive* provides the equations

$$(4) \quad \begin{array}{l} \text{arrive:} \quad (f \text{ SUBJ PERSON}) = 3 \\ \quad \quad \quad (f \text{ SUBJ NUM}) = \text{PL} \end{array}$$

Given the substitution and transitivity axioms of equality, there is a chain of deduction from these premises that leads to the conclusion $\text{PL}=\text{SG}$, an inconsistency between atomic feature values that accounts 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 feature-based syntactic theories, including HPSG (Pollard and Sag, 1994). However, it has been frequently noted that this simple and traditional solution does not cover the whole range of agreement phenomena that natural languages exhibit, particularly where agreement relations hold among three or more items. It does not make correct predictions in situations where one or more of the agreeing items is ambiguous or indeterminate, and without further elaboration it does not predict the patterns of agreement that coordinate structures enter into.

In this paper we present a new account of feature agreement aimed at overcoming the known difficulties of the traditional approach. Our approach represents feature values as sets of atomic symbols and makes use of the logic of simple set theory to enforce the appropriate patterns of covariation. We first consider the problem of feature indeterminacy: we illustrate some of the shortcomings of the traditional approach and then present the essential elements of our solution. The later sections of the paper give a set-based account of feature agreement under coordination.

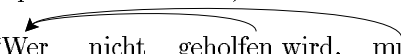
2 Indeterminacy: a problem for the traditional approach

It is possible for a noun phrase to be ambiguous or vague between two case features. In some instances, it can appear to simultaneously fulfill conflicting syntactic requirements (Zaenen and Karttunen, 1984; Pullum and Zwicky, 1986). We will use the terms ‘feature checking’ and ‘feature indeterminacy’ in our discussion of this type of feature resolution.

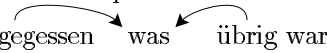
Groos and van Reimsdijk (1979) point out that in German free relatives, the case of the relative pronoun must match the case required by the matrix verb as well as by the verb in the relative clause:

- (5) 
 Wer nicht gefördert wird, muss klug sein.
 Who not supported is must clever be
 NOM SUBJ=NOM SUBJ=NOM
 ‘Who isn’t supported must be clever.’

If case requirements conflict, the result is ungrammatical:

- (6) 
 *Wer nicht geholfen wird, muss klug sein.
 Who not helped is must clever be
 NOM SUBJ=DAT SUBJ=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, however, this simple approach fails in some instances. In the following example, 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:

- (7) 
 Ich habe gegessen was übrig war
 I have eaten what was left
 OBJ=ACC ? SUBJ=NOM
 ‘I ate what was left.’

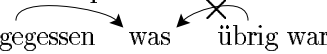
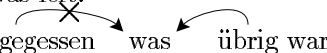
Examples such as these have been widely discussed (Zaenen and Karttunen, 1984; Pullum and Zwicky, 1986; Ingria, 1990; Johnson and Bayer, 1995), since they pose a direct challenge for equality-based treatments of agreement. We briefly discuss the two most obvious ways of extending the traditional approach and show why these proposals are not viable.

2.1 Indeterminacy as disjunction?

It may seem that we can obtain the desired result by assuming that *was* is specified as a wide-scope disjunction, so that it is either NOM or ACC:

- (8) $was: (\uparrow \text{ CASE}) = \text{NOM} \vee (\uparrow \text{ CASE}) = \text{ACC}$

This approach does not succeed, however, under the usual assumptions 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. Thus, no matter which value of CASE we choose for *was*, it will conflict with the requirements of one of the verbs in the construction:

- (9) 
 Ich habe gegessen was übrig war
 I have eaten what was left
 OBJ=ACC ACC SUBJ=NOM
 ‘I ate what was left.’
- (10) 
 Ich habe gegessen was übrig war
 I have eaten what was left
 OBJ=ACC NOM SUBJ=NOM
 ‘I ate what was left.’

We see that disjunctively specifying the case of *was* does not predict the grammaticality of (7).

2.2 Indeterminacy as underspecification?

Alternatively, we might assume that *was* is completely unspecified for case:

(11) *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:

(12) Ich habe gegessen was übrig war
 I have eaten what was left
 OBJ=ACC x SUBJ=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:

(13) $x = \text{NOM} \wedge x = \text{ACC} \Rightarrow \text{NOM} = \text{ACC}$ (!)
 (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 a genitive case, but such a situation cannot be ruled out if no feature value is specified.

3 A set-based theory of feature indeterminacy and feature checking

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

3.1 The case of German free relatives

We propose that the value of the feature CASE for noun phrases is always a set whose elements are the atomic symbols (NOM, ACC, etc.) of the traditional approach. Often the value will be a singleton set, as with the unambiguously nominative relative pronoun *wer*. For relative pronouns like *was*, which can be used in both nominative and accusative contexts, the set will have more than one member:

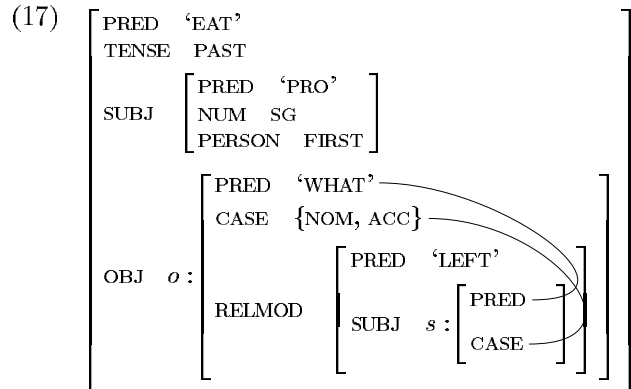
(14) *wer*: $(\uparrow \text{CASE}) = \{\text{NOM}\}$
was: $(\uparrow \text{CASE}) = \{\text{NOM}, \text{ACC}\}$

Case checking by verbs consists in testing for the appropriate members in the CASE set:

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

We assign the functional structure in (17) to German free relatives like (16) (some details omitted):

(16) Ich habe gegessen was übrig war
 I have eaten what was left
 $\text{ACC} \in \text{OBJ CASE}$ $\{\text{NOM}, \text{ACC}\}$ $\text{NOM} \in \text{SUBJ CASE}$
 ‘I ate what was left.’



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 following assertions:

- (18)
- | | |
|-----------|--|
| was: | $(o \text{ CASE}) = \{\text{NOM, ACC}\}$ |
| gegessen: | $\text{ACC} \in (o \text{ CASE})$ |
| übrig: | $\text{NOM} \in (s \text{ CASE})$ |

As noted above, in a free relative construction the PRED and CASE values of the relative pronoun are shared between the matrix sentence and the relative clause, so for this example we also know that

- (19) $(o \text{ CASE}) = (s \text{ CASE})$

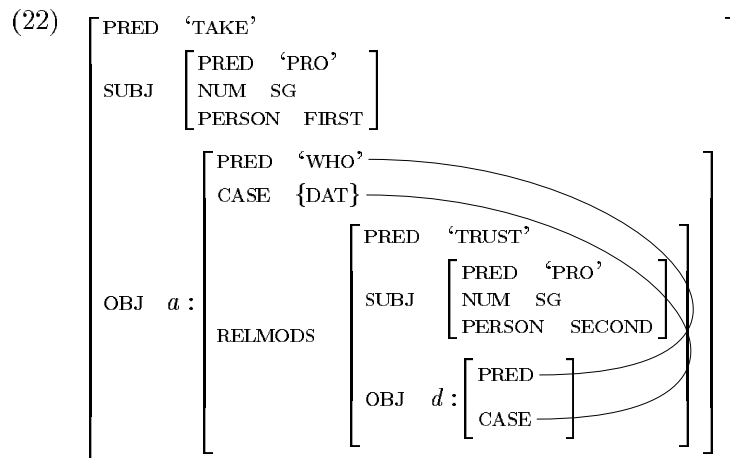
The conditions in (19) and (18) are mutually consistent: the case requirements for both verbs are satisfied and the sentence is marked as grammatical.

In contrast, the following example is not acceptable:

- (20)
- | | | | | |
|----------------------------------|---|------------------|-----|----------------------------------|
| *Ich nehme, | X | wem | du | vertraust. (German) |
| I take | | who | you | trust |
| $\text{ACC} \in \text{OBJ CASE}$ | | $\{\text{DAT}\}$ | | $\text{DAT} \in \text{OBJ CASE}$ |
- ‘I take who(ever) you trust.’

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 *nehme* ‘take’ requires an accusative object, and this constraint is not met. Letting *a* and *d* stand respectively for the f-structures of the accusative and dative objects, the relevant constraints for this example are:

- (21)
- | | |
|------------|---|
| | $(a \text{ CASE}) = (d \text{ CASE})$ |
| wem: | $(a \text{ CASE}) = \{\text{DAT}\}$ |
| nehme: | $\text{ACC} \in (a \text{ CASE})$ |
| vertraust: | $\text{DAT} \in (d \text{ CASE})$ [not satisfied] |



Our use of set representations as the formal encoding of feature values is a conservative solution to this descriptive problem. It is not an extension to LFG’s domain of f-structure entities, since sets have been part of the LFG f-structure ontology since the very beginning. Kaplan and Bresnan (1982) introduced sets of f-structures to represent collections of adjuncts and modifiers, and both Bresnan, Kaplan, and Peterson (1985) and Kaplan and Maxwell (1988) used sets to represent the conjoined elements of coordinate constructions. Thus, LFG’s functional description language has always included set membership \in 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

$$(23) \text{ was: } (\uparrow \text{ CASE}) = \{\text{NOM, ACC}\}$$

The expression $\{\text{NOM, ACC}\}$ in this equation is a *set designator* indicating that the feature value is a set and also providing an *exhaustive enumeration* of the set’s elements. Given only the set-membership relation of Kaplan and Bresnan (1982), the fact that NOM and ACC belong to the set would be encoded by the pair of assertions

$$(24) \text{ was: } \begin{array}{l} \text{NOM} \in (\uparrow \text{ CASE}) \\ \text{ACC} \in (\uparrow \text{ CASE}) \end{array}$$

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 (23), the assertions in (24) would be consistent with a further assertion that $\text{GEN} \in (\uparrow \text{ CASE})$.

The exhaustive enumeration implicit in the set designator notation also captures some of the disjunctive intuitions of feature indeterminacy: if x stands for some unknown value, we know that

$$(25) \quad x \in \{\text{NOM, ACC}\} \Rightarrow \text{NOM} = x \vee \text{ACC} = x \quad (\text{implicit disjunction}) \\ (\text{i.e., } \text{GEN} \neq x)$$

Whereas explicit disjunction as denoted by the Boolean operator \vee has the undesired wide-scope interpretation, elements in a designated set are ‘disjunctive’ only with respect to specific assertions of set membership.

Conversely, we also know that if y is defined by a set designator, then it does in fact contain all of the designated elements:

$$(26) \quad y = \{\text{NOM, ACC}\} \Rightarrow \text{NOM} \in y \wedge \text{ACC} \in y \quad (\text{implicit conjunction})$$

This is why the conjunction of differing verbal requirements is satisfiable with indeterminate pronoun specifications.

The original LFG formalism, without set designators, allows for a slightly different solution to German case resolution, expressed in terms of set membership statements like those in (24) together with ‘sub-c’ constraints for checking set membership:

(27) $ACC \in_c (\uparrow \text{ SUBJ CASE})$

As desired, a sub-c constraint for GEN would not be satisfied. However, without some other formal additions, this approach does not provide an analysis of the phenomenon of Xhosa noun class resolution discussed below.

3.2 Distributed values: Polish case

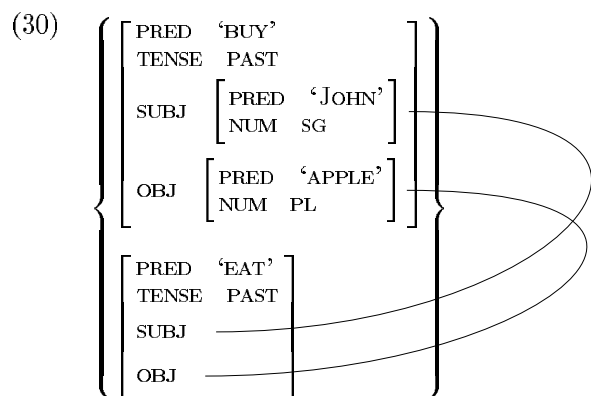
Case features in Polish coordinate constructions provide another instance of three-way indeterminate agreement. Dyla (1984, p. 701) discusses the following Polish example, in which a noun phrase is shared across a coordinate structure and different requirements are placed on its case feature in each conjunct:

(28) $\left[\begin{array}{ccc} \text{Kogo} & \text{Janek} & \text{lubi} \\ \text{who} & \text{Janek} & \text{likes} \\ \{ACC, GEN\} & & ACC \in OBJ \text{ CASE} \end{array} \right] \text{ and } \left[\begin{array}{ccc} \text{a Jerzy} & & \text{nienawidzi (Polish)} \\ \text{and Jerzy} & & \text{hates} \\ & & GEN \in OBJ \text{ CASE} \end{array} \right]$
 ‘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 example (29) in which *kogo* ‘who’ is replaced by *co* ‘what’:

(29) $\left[\begin{array}{ccc} *Co & \text{Janek} & \text{lubi} \\ \text{what} & \text{Janek} & \text{likes} \\ \{NOM, ACC\} & & ACC \in OBJ \text{ CASE} \end{array} \right] \text{ and } \left[\begin{array}{ccc} \text{a Jerzy} & & \text{nienawidzi (Polish)} \\ \text{and Jerzy} & & \text{hates} \\ & & GEN \in OBJ \text{ CASE} \end{array} \right]$
 ‘What 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, Kaplan, and Peterson (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. Thus the f-structure for a sentence like *John bought and ate apples* is:



This is an example of verb coordination described by the following phrase structure rule:

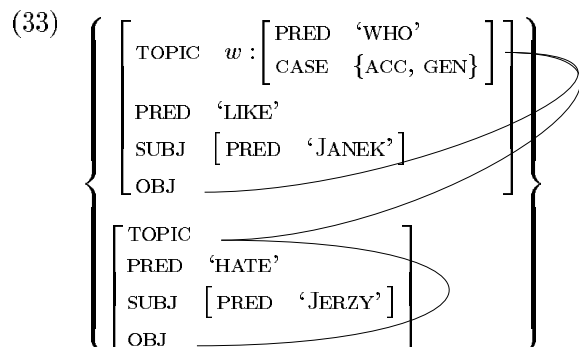
(31) $V \longrightarrow \begin{array}{c} V \\ \downarrow \in \uparrow \end{array} \text{ CONJ } \begin{array}{c} V \\ \downarrow \in \uparrow \end{array}$

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 $(\uparrow \text{ SUBJ}) = \downarrow$, and the NP *apples* provides its object. However, the original definition of function application (Kaplan and Bresnan, 1982) did not provide a value for such an application expression when \uparrow denotes a set of f-structures instead of a single f-structure, and thus it did not produce the desired effect of distributing the subject and object to each of the conjunct f-structures. Bresnan, Kaplan, and Peterson (1985) outlined a general theory of feature distribution in coordinate constructions by extending the definition of function application to sets:

(32) For any property P and set s : $P(s)$ iff $\forall f \in s.P(f)$

According to this definition, if the subject NP *John* is the subject of the coordinate structure, it is the subject of each of the conjuncts.¹

For our Polish examples, the effect of this definition is to distribute *kogo* ‘who’ so that it becomes the object in each of the conjunct f-structures:



The agreement pattern for these examples 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 $\{\text{ACC, GEN}\}$. 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.

(34)

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

3.3 Distributed checking: Xhosa noun classes

The Polish data 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 a noun phrase 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 language Xhosa, as discussed by Voeltz (1971).

In Xhosa, like other Bantu languages, nouns belong to one of a number of noun classes. We will 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 the second class, and so the noun class of a noun will be represented by a pair of numbers such as 1/2, indicating that the singular is in class 1 and the plural is 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:

(35)

umfana	nomfazi	bayagoduka
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.’

¹This definition of function application for sets differs from the later proposal of Kaplan and Maxwell (1988). They proposed that when f is a set, $(f a) = v$ iff v is the generalization of all of the elements of f applied to a . Among other things, this definition was intended to account for case agreement phenomena in coordination by allowing for disjunctively-specified case requirements to be resolved independently in each conjunct. However, their proposal does not extend to noncoordinate cases such as the German free relative, and also makes additional unwanted predictions.

$$(36) \left[\begin{array}{l} \text{PRED} \text{ 'GO.HOME'} \\ \text{SUBJ} \left\{ \begin{array}{l} \left[\begin{array}{l} \text{PRED} \text{ 'YOUNG.MAN'} \\ \text{CLASS} \text{ 1/2} \end{array} \right] \\ \left[\begin{array}{l} \text{PRED} \text{ 'YOUNG.WOMAN'} \\ \text{CLASS} \text{ 1/2} \end{array} \right] \end{array} \right\} \end{array} \right]$$

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:

$$(37) \text{ a. } \begin{array}{lll} *Igqira & nesanuse & ayagoiduka \\ \text{doctor} & \text{and-diviner} & \text{go home} \\ 5/6 & 7/8 & \text{SUBJ CLASS} = 5/6 \end{array}$$

$$\text{ b. } \begin{array}{lll} *Igqira & nesanuse & ziyagoiduka \\ \text{doctor} & \text{and-diviner} & \text{go home} \\ 5/6 & 7/8 & \text{SUBJ CLASS} = 7/8 \\ \text{'The doctor and diviner are going home.'} \end{array}$$

$$(38) \left[\begin{array}{l} \text{PRED} \text{ 'GO.HOME'} \\ \text{SUBJ} \left\{ \begin{array}{l} \left[\begin{array}{l} \text{PRED} \text{ 'DOCTOR'} \\ \text{CLASS} \text{ 5/6} \end{array} \right] \\ \left[\begin{array}{l} \text{PRED} \text{ 'DIVINER'} \\ \text{CLASS} \text{ 7/8} \end{array} \right] \end{array} \right\} \end{array} \right]$$

In cases where verb agreement is not involved, such coordinate structures are unproblematic; the problem with the examples in (37) is the lack of an appropriate agreeing form for the coordinate subject noun phrase.

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

$$(39) \begin{array}{lll} Izandla & neendlebe & zibomvu \\ \text{hands} & \text{and-ears} & \text{are-red} \\ 7/8 & 9/10 & \text{SUBJ CLASS} \in \{7/8, 9/10\} \\ \text{'The hands and the ears are red.'} \end{array}$$

We again use set representations to encode an indeterminate pattern of agreement. In this case, 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 example (39), the subject's noun class must be either 7/8 or 9/10, and a conjunction of nouns in these classes is permitted:

$$(40) f : \left[\begin{array}{l} \text{PRED} \text{ 'ARE.RED'} \\ \text{SUBJ} \left\{ \begin{array}{l} h : \left[\begin{array}{l} \text{PRED} \text{ 'HANDS'} \\ \text{CLASS} \text{ 7/8} \end{array} \right] \\ e : \left[\begin{array}{l} \text{PRED} \text{ 'EARS'} \\ \text{CLASS} \text{ 9/10} \end{array} \right] \end{array} \right\} \end{array} \right]$$

$$(41) \begin{array}{ll} \text{izandla} & (h \text{ CLASS}) = 7/8 \\ \text{neendlebe} & (e \text{ CLASS}) = 9/10 \\ \text{zibomvu} & (f \text{ SUBJ CLASS}) \in \{7/8, 9/10\} \end{array}$$

3.4 Previous proposals

Previous proposals for dealing with case agreement in German free relatives and related examples have involved more or less radical changes or augmentations to the architecture of constraint-based grammatical theories. 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. Ingria (1990) proposes to handle case agreement by means of a nondistinctness check rather than either a unification operation or an equality constraint. A relative pronoun such as German *was* is disjunctively marked for case, bearing the case $\text{NOM} \vee \text{ACC}$. Verbs check for consistency with this feature without instantiating it to one or the other 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.

This account builds into the formal apparatus a treatment of agreement features that incorrectly predicts that indeterminacy of agreement features is always tolerated in every language and for every speaker. As Zaenen and Karttunen (1984) point out, speakers sometimes vary in allowing feature resolution to take place; for example, only some speakers of Icelandic find the following example acceptable:

- (42) %hann stal og borðaði köku
 he stole and ate cookie
 OBJ=DAT OBJ=ACC {ACC,DAT}
- ‘I stole and ate the cookie.’
-

By expressing the properties of different features in different constructions in terms of simple mathematical relations and structures, we provide a coherent account that still has all the descriptively necessary flexibility. Further, unlike Ingria’s, the account we present makes use of structures and relations that are independently motivated in our theory.

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 cases of feature resolution and that a deductive approach is therefore to be preferred. Their primary argument against a constraint-based account is aimed at proposals for feature determination in coordination due to Sag et al. (1985) and Shieber (1987). 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, Kaplan, and Peterson (1985) approach, the account we present here does not suffer from such incorrect predictions.

Johnson and Bayer also give an account of German free relative case agreement which, like ours, properly classifies the grammatical and ungrammatical strings. They assign two different agreement features to each relative pronoun, one that must be compatible with the verb in the clause containing the free relative and one that agrees with the verb inside the relative. They point out that the essential elements of their deductive solution can be translated into any unification or equality-based theory to circumvent the German case difficulties. While this approach is technically correct, we regard it as having a major theoretical flaw: on this analysis it is a matter of pure stipulation that the internal and external agreement marks for most pronouns are identical.

We agree with Zaenen and Karttunen (1984), Ingria (1990) and Johnson and Bayer (1995) that simple equality constraints with atomic feature values such as NOM or ACC are inappropriate for characterizing case indeterminacy and case 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).

4 Feature determination in coordination

Another family of resolution problems concerns the determination of the features of coordinate noun phrases. A coordinate noun phrase has its own person and gender features just as a pronoun does. These are determined by the person and gender features of the conjuncts:

(43) Nosotros hablamos/*habláis/*hablan
We speak-1PL/*speak-2PL/*speak-3PL

(44) José y yo hablamos/*habláis/*hablan
José and I speak-1PL/*speak-2PL/*speak-3PL

Clearly, the person, number and gender features of a coordinate noun phrase are properties of a coordinate structure as a whole: that is, they are properties of the *set* of f-structures representing the coordinate phrase. We will refer to these features as *nondistributive* features, since they do not distribute to the individual conjuncts. Other features such as case or noun class, as we have seen, are *distributive* features, properties of each conjunct. Distributive features are properties of the individual conjunct noun phrases and not of the coordinate phrase, and are checked by inspecting the features of each conjunct, as in the Xhosa case described above. Thus we classify grammatical properties into two classes according to whether or not they apply to a set or distribute to its elements:

(45) SUBJ, OBJ, CASE, ... are distributive features.
PERSON, NUM, GEND are nondistributive features.

Following a proposal by John Maxwell (p.c.), we represent the distinction between distributive and nondistributive features by assuming that f-structures for coordinate structures are *hybrid objects*: they are structures with both elements and features. We now refine the Bresnan, Kaplan, and Peterson (1985) extended definition of function application given in (32) to take into account the distinction between distributive and nondistributive properties:

(46) For any *distributive* property P and set s , $P(s)$ iff $\forall f \in s.P(f)$
Any *nondistributive* property P holds of s itself.

To represent nondistributive attributes and values of the set itself, as opposed to the attributes and values of the individual conjuncts, we will use a notation like the following:

(47) ‘Bill and George’:

$$\left[\begin{array}{c} \left\{ \left[\begin{array}{cc} \text{PRED} & \text{BILL} \\ \text{NUM} & \text{SG} \end{array} \right] \right\} \\ \left\{ \left[\begin{array}{cc} \text{PRED} & \text{GEORGE} \\ \text{NUM} & \text{SG} \end{array} \right] \right\} \\ \text{NUM PL} \end{array} \right]$$

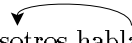
This represents the fact that while the number feature for each conjunct has the value SG, the coordinate phrase as a whole is plural, and its value for the feature NUM is PL.

Corbett (1983b) discusses the process whereby the person or gender features of a coordinate noun phrase are determined on the basis of the features of each conjunct. We will refer to this type of feature resolution as ‘feature determination in coordination’. In the following, we propose a method for representing person and gender features that permits the determination of the features of a coordinate noun phrase in a systematic way.²

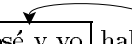
²It should be noted that in some cases, agreement with coordinate phrases is not statable in terms of features of the coordinate phrase as a whole, but instead with one of the conjuncts. For example, in some languages the verb may be permitted or required to agree with the closest conjunct of a coordinate phrase. Our theory describes how person and gender features can be systematically determined in coordinate structures, but does not specify when such a determination must take place. See Corbett (1991) for discussion of the latter issue.

4.1 Person determination

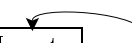
A first person plural verb form is used with first person plural pronominal subjects:

- (48)  Nosotros hablamos/*habláis/*hablan
We speak-1PL/*speak-2PL/*speak-3PL

Analogously, with a coordinate subject noun phrase containing a first person form, the first person plural form must in general be used:

- (49)  José y yo hablamos/*habláis/*hablan
José and I speak-1PL/*speak-2PL/*speak-3PL

Slovak (Corbett, 1983a, page 178) provides another illustration of the same dependency:

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

This phenomenon is generally 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 will bear first person features; if the coordinate structure contains a second person pronoun but no first person pronoun, the coordinate structure will bear second person features, and so on. This is represented in the following chart, where & represents coordination:

- (51) Spanish, Slovak, ...:
1st person & 2nd person = 1st person
1st person & 3rd person = 1st person
2nd person & 3rd person = 2nd person
3rd person & 3rd person = 3rd person

To permit person features to be represented so that the person feature of a coordinate structure can be computed on the basis of the person features of the conjuncts, we propose to make a different use of set values for agreement attributes than in previous sections. For person and gender features:

- (52) Marker sets encode complex values, not alternatives.

In particular, we refine the representation of a value such as FIRST for the PERSON feature, and propose that the value of the person feature is a set of marker values S and H:

- (53) Values of agreement features for Spanish, English, Slovak, ...:
1st (I, we): {S,H}
2nd (you): {H}
3rd (he, they): {}

S and H are mnemonic for Speaker and Hearer, but these values should be taken as an intuitive way of representing a complex syntactic value and not as part of a theory of reference. The value of the person feature for a coordinate structure is determined by taking the union of the person features of the conjuncts. This produces the correct result for English, Spanish, Slovak, and a range of other languages:

- (54) {S,H} (1st) \cup {H} (2nd) = {S,H} (1st)
{S,H} (1st) \cup {} (3rd) = {S,H} (1st)
{H} (2nd) \cup {} (3rd) = {H} (2nd)
{} (3rd) \cup {} (3rd) = {} (3rd)

For example, if a first person and third person noun phrase are coordinated, the person value for coordinate noun phrase will be first person, the union of the first and third person features:

$$(55) \left[\begin{array}{l} \text{PRED 'SPEAK'} \\ \\ \text{SUBJ } s : \left\{ \begin{array}{l} j : \left[\begin{array}{l} \text{PRED 'JOSE'} \\ \text{NUM SG} \\ \text{PERSON } \{\} \end{array} \right] \\ y : \left[\begin{array}{l} \text{PRED 'PRO'} \\ \text{NUM SG} \\ \text{PERSON } \{S,H\} \end{array} \right] \end{array} \right\} \\ \text{NUM PL} \\ \text{PERSON } \{S,H\} \end{array} \right]$$

$$(56) \quad \begin{array}{l} \text{José: } (j \text{ PERSON}) = \{\} \\ \text{yo: } (y \text{ PERSON}) = \{S,H\} \\ \text{José y yo: } (s \text{ PERSON}) = \{S,H\} \end{array}$$

4.2 Feature checking in coordination

We have seen that the desired value for the coordination's person feature are obtained if we represent the value of PERSON as a set and assume that the PERSON value for a coordinate structure is the union of the PERSON values of the conjuncts. This constraint is represented by the following annotation on a typical LFG phrase structure rule for coordinate noun phrases:

$$(57) \text{ NP} \longrightarrow \begin{array}{ccc} \text{NP} & \text{CONJ} & \text{NP} \\ \downarrow \in \uparrow & & \downarrow \in \uparrow \\ (\downarrow \text{PERSON}) \subseteq (\uparrow \text{PERSON}) & & (\downarrow \text{PERSON}) \subseteq (\uparrow \text{PERSON}) \end{array}$$

The annotations on this phrase structure rule require the value of the PERSON feature on each conjunct to be a subset of the value of the PERSON feature of the coordinate structure as a whole — in other words, the PERSON value of the coordinate structure must contain the PERSON values of the conjuncts. The smallest set that satisfies this requirement is the union of the PERSON values of each conjunct:

$$(58) \quad x \cup y \text{ is the smallest set } z \text{ such that } x \subseteq z \wedge y \subseteq z$$

The verb is annotated with a constraining equation which ensures that this smallest value is in fact the value of the PERSON feature of the coordinate phrase.

In the Spanish case discussed above, the conjuncts *José* and *yo* each have a PERSON value that is contained in the PERSON value of the coordinate phrase:

$$(59) \quad \boxed{\text{José y yo}} \text{ hablamos/*habláis/*hablan}$$

José and I speak-1PL/*speak-2PL/*speak-3PL

$$(60) \quad \begin{array}{l} \text{José: } (j \text{ PERSON}) = \{\} \\ \text{yo: } (y \text{ PERSON}) = \{S,H\} \\ \text{José y yo: } (s \text{ PERSON}) = \{S,H\} \end{array}$$

Here, $(s \text{ PERSON})$ includes the union of the person features of the conjuncts: the *smallest* value for $(s \text{ SUBJ PERSON})$ is $\{S,H\}$. The verb *hablamos* checks to ensure that this is the PERSON value for its subject:

$$(61) \quad \text{hablamos: } \begin{array}{l} (s \text{ SUBJ NUM}) = \text{PL} \\ (s \text{ SUBJ PERSON}) =_c \{S,H\} \end{array}$$

4.3 Intersection or union?

Sag et al. (1985) also propose the use of sets to encode the value for the person feature, but they provide a different means for determining the value of the person feature for the coordinate structure. They propose that the features on the coordinate NP are the intersection, not the union, of features on the conjuncts. Sag et al. (1985) propose the following set of markers as values for the person feature:

- (62) 1st $\{\}$
 2nd $\{+XSP\}$
 3rd $\{+XSP,+THP\}$

This also works for the case of English, Slovak, and Spanish:

- (63) $\{\}$ (1st) $\cap \{+XSP\}$ (2nd) = $\{\}$ (1st)
 $\{\}$ (1st) $\cap \{+XSP,+THP\}$ (3rd) = $\{\}$ (1st)
 $\{+XSP\}$ (2nd) $\cap \{+XSP,+THP\}$ (3rd) = $\{+XSP\}$ (2nd)
 $\{+XSP,+THP\}$ (3rd) $\cap \{+XSP,+THP\}$ (3rd) = $\{+XSP,+THP\}$ (3rd)

However, the particular set of features they propose does not work for languages with different verb agreement forms for inclusive and exclusive first person plurals. Fula (Forchheimer, 1953; Arnott, 1970) is such a language. 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. Intuitively, the reason that Sag et al.'s analysis fails is that their value for the first person number feature is the empty set, and intersecting the empty set with any other set gives the empty set. Under their analysis, it is not possible to distinguish different kinds of coordination involving a first person pronoun.

For Fula and other languages with an inclusive/exclusive distinction, the coordination facts are as follows:

- (64) Fula:
 1st singular & 2nd = 1st inclusive plural
 1st singular & 2nd & 3rd = 1st inclusive plural
 1st singular & 3rd = 1st exclusive plural
 1st exclusive plural & 3rd = 1st exclusive plural
 2nd & 3rd = 2nd
 3rd & 3rd = 3rd

We propose that Fula pronouns bear the following features:

- (65) Fula, analysis involving union:

	singular	plural
1st:	min: {S}	inclusive: enèn: {S,H} exclusive: mènèn: {S}
2nd:	an: {H}	onön: {H}
3rd:	ö: {}	kombè: {}

Assuming these features and assuming that the person feature of a coordinate noun phrase is obtained by computing the union of the person features of the conjuncts produces the correct result.

Of course, any analysis in which features are combined by union, such as the one above, has a dual corresponding analysis in which features are combined by intersection. This dual analysis is obtained simply by substituting the complement of the sets involved in each case:

- (66) Fula, alternative analysis involving intersection:

	singular	plural
1st:	min: {H}	inclusive: enèn: {} exclusive: mènèn: {H}
2nd:	an: {S}	onön: {S}
3rd:	ö: {S,H}	kombè: {S,H}

The values for the person feature are, we believe, much less intuitively easy to justify in this analysis than in the analysis involving union. Nevertheless, it is easy to verify that this analysis also adheres to the requirements if we assume that the value of the coordinate structure is obtained by intersecting the values of the conjuncts:

$$\begin{aligned}
 (67) \quad & \{H\} (1SG) \cap \{S\} (2nd) = \{\} (1INCPL) \\
 & \{\} (1INCPL) \cap \{S,H\} (3rd) = \{\} (1INCPL) \\
 & \{H\} (1SG) \cap \{S,H\} (3rd) = \{H\} (1EXCPL) \\
 & \{S\} (2nd) \cap \{S,H\} (3rd) = \{S\} (2nd) \\
 & \{S,H\} (3rd) \cap \{S,H\} (3rd) = \{S,H\} (3rd)
 \end{aligned}$$

The issue of whether union or intersection should be used to combine feature sets is, then, not an issue of whether an analysis is mathematically possible or impossible: exactly the same phenomena can be described no matter which way the sets are combined. The question, rather, is whether the marker sets required under one interpretation are more or less sensible than the sets required under the other.

5 Conclusion

Our analysis of feature indeterminacy and feature determination in coordination required extending LFG's f-description language with *set designators* and *subset constraints*, and adding *nondistributive features* to the structural domain. These changes enabled a simple and intuitive account of both these phenomena, and entailed minimal changes to LFG theory in comparison to other approaches. Although we have not discussed it here (but see Dalrymple and Kaplan 1997), our approach to feature determination accounts not only for determination of the person feature in coordination but also for gender determination. In sum, we have argued that set representations can provide a general theory of feature resolution, allowing for an account of complicated agreement phenomena in an intuitive way.

6 Acknowledgments

For helpful discussion of these and related issues, we are grateful to Amy Dahlstrom, Mark Johnson, John Maxwell, Paula Newman, María-Eugenia Niño, Annie Zaenen, and Arnold Zwicky.

References

- Arnott, D. W. 1970. *The Nominal and Verbal Systems of Fula*. Clarendon Press, Oxford.
- Bresnan, Joan, Ronald M. Kaplan, and Peter Peterson. 1985. Coordination and the flow of information through phrase structure. MS, Xerox PARC.
- Corbett, Greville. 1983a. *Hierarchies, targets, and controllers: Agreement patterns in Slavic*. Croom Helm, London.
- Corbett, Greville. 1983b. Resolution rules: Agreement in person, number, and gender. In Gerald Gazdar, Ewan Klein, and Geoffrey K. Pullum, editors, *Order, concord, and constituency*. Foris, Dordrecht, pages 175–206.
- Corbett, Greville. 1991. *Gender*. Cambridge University Press, Cambridge.
- Dalrymple, Mary and Ronald M. Kaplan. 1997. Resolution of syntactic features in description-based syntax. MS, Xerox PARC.
- Dyła, Stefan. 1984. Across-the-board dependencies and case in Polish. *Linguistic Inquiry*, 15(4):701–705.
- Forchheimer, Paul. 1953. *The Category of Person in Language*. de Gruyter, Berlin.
- Groos, Anneke and Henk van Reimsdijk. 1979. Matching effects in free relatives: A parameter of core grammar. In Adriana Belletti, Luciana Brandi, and Luigi Rizzi, editors, *Theory of Markedness in Generative Grammar: Proceedings of the 1979 GLOW Conference*. Scuola Normale Superiore de Pisa, Pisa.

- Ingria, Robert J. P. 1990. The limits of unification. In *Proceedings of the 28th Annual Meeting of the ACL*, pages 194–204, Pittsburgh. Association for Computational Linguistics.
- Johnson, Mark and Samuel Bayer. 1995. Features and agreement in lambek categorial grammar. In *Proceedings of the Formal Grammar Workshop*.
- Kaplan, Ronald M. and Joan Bresnan. 1982. Lexical-Functional Grammar: A formal system for grammatical representation. In Joan Bresnan, editor, *The Mental Representation of Grammatical Relations*. The MIT Press, Cambridge, MA, pages 173–281. Reprinted in Dalrymple, Kaplan, Maxwell, and Zaenen, eds., *Formal Issues in Lexical-Functional Grammar*, 29–130. Stanford: Center for the Study of Language and Information. 1995.
- Kaplan, Ronald M. and John T. Maxwell. 1988. Constituent coordination in Lexical-Functional Grammar. In *Proceedings of COLING-88*, volume 1, pages 303–305, Budapest. Reprinted in Dalrymple, Kaplan, Maxwell, and Zaenen, eds., *Formal Issues in Lexical-Functional Grammar*, 199–210. Stanford: Center for the Study of Language and Information. 1995.
- Pollard, Carl and Ivan A. Sag. 1994. *Head-Driven Phrase Structure Grammar*. The University of Chicago Press, Chicago.
- Pullum, Geoffrey K. and Arnold M. Zwicky. 1986. Phonological resolution of syntactic feature conflict. *Language*, 62(4):751–773.
- Sag, Ivan A., Gerald Gazdar, Thomas Wasow, and Steven Weisler. 1985. Coordination and how to distinguish categories. *Natural Language and Linguistic Theory*, 3:117–171.
- Shieber, Stuart M. 1987. Separating linguistic analyses from linguistic theories. In Peter Whitelock, Harold Somers, Paul Bennett, Rod Johnson, and Mary McGee Wood, editors, *Linguistic Theory and Computer Applications*. Academic Press, London, pages 1–36.
- Voeltz, Erhard. 1971. Surface constraints and agreement resolution: Some evidence from Xhosa. *Studies in African Linguistics*, 2(1):37–60.
- Zaenen, Annie and Lauri Karttunen. 1984. Morphological non-distinctiveness and coordination. In *Proceedings of ESCOL '84*, pages 309–320.

Mary Dalrymple <dalrymple@parc.xerox.com>
Ronald M. Kaplan <kaplan@parc.xerox.com>
Information Sciences and Technologies Laboratory
Xerox Palo Alto Research Center
3333 Coyote Hill Road
Palo Alto, CA 94304 USA