A THEORY OF AGREEMENT AND ITS APPLICATION TO SERBO-CROATIAN

STEPHEN WECHSLER AND LARISA ZLATIĆ

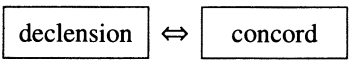
University of Texas at Austin

Four lexical features of a noun are relevant to agreement: (i) semantic conditions on reference, (ii) person, number, and gender features of the referential index, (iii) concord features, and (iv) declension class. These four features are correlated by a chain of binary constraints. When individual constraints are violated, the chain is broken, resulting in intricate patterns of mixed agreement. Three main types of mixed agreement are predicted, all of them attested in Serbo-Croatian. This theory helps explain Corbett's (1983) crosslinguistic AGREEMENT HIERARCHY.*

1. OVERVIEW OF A THEORY OF AGREEMENT. Any treatment of morphosyntactic agreement must face those aberrant cases of HYBRID or MIXED agreement. In certain situations for example a noun triggers one set of agreement features on adjectives but the NP headed by that noun triggers a different set of agreement features on verbs or coreferential pronouns; or a noun triggers masculine features on adjectives and pronouns but belongs to the declension class normally reserved for feminine nouns, and so on. These exceptional cases also shed light on regular agreement processes. We present a theory of agreement and apply it to Serbo-Croatian, which, like its Slavic relatives, is richly endowed with regular and hybrid agreement. The mixed agreement patterns that are predicted by the theory, and no others, are attested in that language.

Our theory may be summarized as follows. An inflected noun has two different feature sets that determine the agreement values it triggers: CONCORD features (case, number, and gender, for Serbo-Croatian); and the person, number, and gender features of the REFERENTIAL INDEX (Pollard & Sag 1994).[1] The concord feature of gender for a given noun is closely related to that noun's declension class, as schematized in 1.

(1)



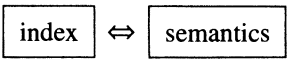
In contrast, the index features of gender and number are closely related to the noun's semantics, for example, whether the noun denotes a male or female, aggregate or nonaggregate entity.[2]

* For their many helpful comments on this work, we wish to thank Mark Aronoff, Wayles Browne, Greville Corbett, Manfred Krifka, Johanna Nichols, Gilbert Rappaport, Ivan Sag (and his seminar students), Carlota Smith, and two anonymous *Language* referees. We also thank the many native speakers of Serbo-Croatian who served as informants.

[1] In addition to these features, animacy plays an important role in Slavic agreement, but it is not treated in detail here. Briefly, the facts for Serbo-Croatian are as follows. Masculine accusative nouns alternate between two forms: for animates the accusative is homophonous with the genitive, while for inanimates the accusative is homophonous with the nominative. (That masculine gender rather than male sex is relevant here can be seen from nouns like *devojčurak*, a diminutive for a girl. This noun has masculine concord, and its accusative form is homophonous with the genitive.) Items showing concord with a masculine accusative noun exhibit the same alternation in form. This can be analyzed by treating animacy as a subgender of masculine. In our framework (see Appendix) this means the masculine value for the CONCORD:GENDER feature has two subsorts, for animate and inanimate masculine, respectively. See 68 and the discussion following it.

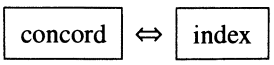
[2] For now the word *aggregate* may be interpreted in its colloquial sense: 'an assemblage or group of distinct particulars massed together' (from the *American Heritage Dictionary*). For example, *five apples* is an aggregate (of apples). Nonaggregates include singulars (*the apple*), mass terms (*water*), abstract terms (*democracy*), and so on. See § 4.3 for more discussion.

(2)



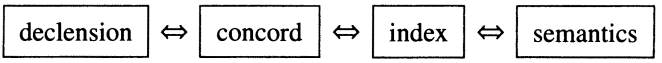
And finally, the concord and index features are systematically related to each other, as shown in 3.

(3)



Indeed, in the normal case, concord and index features are identical, although they differ in some exceptional cases discussed below. Putting these three binary correlation rules together produces a network of relationships.

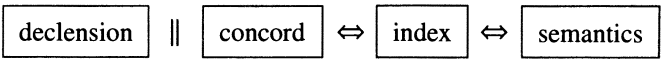
(4)

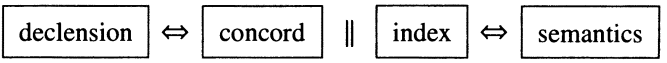


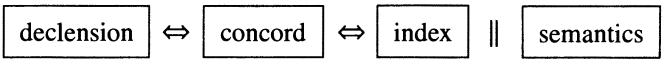
In a perfectly regular noun such as *žena* 'woman' or *knjiga* 'book', all types of information correlate: These nouns belong to the second declension class, which is normally for feminine nouns; they trigger feminine concord; their referential indices are feminine, so bound pronouns are feminine. For the index-semantics correlation, *žena* 'woman' is female-denoting, which satisfies the rule that female-denoting nouns are feminine, and *knjiga* 'book' is neither female- nor male-denoting, so it vacuously satisfies the rule. (In this article the terms MASCULINE, FEMININE, and GENDER refer to morphosyntactic classes, while MALE, FEMALE, and SEX refer to semantic classes.)

In other nouns however, this chain can be broken at any of the three points, as indicated by the double vertical bar.

(5)

a. 

b. 

c. 

For example, if the link between concord and index is severed, as in 5b, these will not correlate in the normal way, but declension and concord will continue to match, as will index and semantics. This theory predicts three types of mismatch, out of seven theoretically possible mismatching patterns.[3] Exactly these three types are attested in Serbo-Croatian, and we are unaware of any common or proper nouns exhibiting any of the nonpredicted patterns.[4]

[3] Abbreviating declension, concord, index, and semantics with d, c, i, and s respectively, the three mismatches shown in 5 can be represented as d/cis, dc/is, dci/s. The four unattested patterns are: ds/ci, di/cs, dcs/i, and c/dis. This makes seven logically possible patterns in all. In principle there could be nouns with more than one break, leading to additional mismatch types. One Serbo-Croatian noun of this kind is *braća* 'brothers' (see §6).

[4] This article deals only with common and proper nouns, and not pronouns, as agreement triggers. This approach is extended to the second person honorific pronoun *vi* in Wechsler 1999a. See also Corbett 1983 on honorific *vi*.

Taken together, the three constraints schematized in 1–3 predict an implicational hierarchy in which, for example, a noun's concord features cannot be semantically based unless the index features also are semantically based. This implicational hierarchy is broadly similar to Corbett's (1979, 1983, 1991, 1998) AGREEMENT HIERARCHY, which describes the major typological patterns governing agreement (see §8 for comparison).

2. BACKGROUND. Our analysis of agreement is formulated in the nonderivational, constraint-based theory of HEAD-DRIVEN PHRASE STRUCTURE GRAMMAR (Pollard & Sag 1994); see the appendix. HPSG assumes a lexicalist view of inflection, wherein the syntax deals only with fully inflected words. We use the standard terminology of AGREEMENT TRIGGER (also called CONTROLLER), typically a nominal, which determines agreement features on the AGREEMENT TARGET such as a verb or adjective. In *I am happy*, for example, the pronoun *I* is the trigger and *am* is the target, although agreement is not viewed as a directional process of copying or moving feature bundles, but rather as two elements specifying partial information about a single linguistic object. Agreement results from the fact that this information coming from two sources must be compatible. Hence **I is happy* fails because the third-person feature specified by the verb *is* for its subject clashes with the first-person feature specified by the pronoun *I*. For motivations of this approach see Barlow 1992 and Pollard & Sag 1994, ch. 2.

Following Pollard & Sag 1994 and Kathol 1999, structural agreement is further subdivided into INDEX AGREEMENT and CONCORD (regarding index agreement see also Farkas & Zec 1993).[5] Bresnan and Mchombo's (1987) distinction between anaphoric and grammatical agreement is similar. Although they are often conflated in studies of agreement, index agreement and concord result from different grammatical processes and consequently differ with respect to the domain of agreeing elements, the set of relevant agreement features, and the nature of the morphological realization (see §3). Further evidence for this distinction will be presented below. Given our theoretical framework, the lexical entry for a noun must have two distinct feature sets, INDEX and CONCORD, each accessed by a different set of syntactic processes. Consider the Serbo-Croatian example in 6.[6]

(6) Ov-a star-a knjig-a stalno pad-a.
this-NOM.F.SG old-NOM.F.SG book(F)-NOM.SG always fall-3SG
Molim vas, podignite je.
please you pick.2PL it.F.SG
'This old book keeps falling. Please pick it up.'

The noun *knjiga* 'book' has the three CONCORD values nominative, feminine, singular, and also has three values (third person, feminine, singular) on its INDEX. Normally, as in this case, the number and gender values match across CONCORD and INDEX. The determiner *ova* and adjective *stara* show concord with this head noun, while the finite verb *pada* and the pronoun *je* show index agreement with the NP *ova stara knjiga*. These types of agreement are compared further in the following section.

Data in the present study are from native Serbo-Croatian speaker informants, except where noted. Traditional scholarly sources on Serbo-Croatian and Slavic grammar include Bidwell 1966, Mrazović & Vukadinović 1990, and Stevanović 1974. Ivić 1960,

[5] Our CONCORD feature corresponds to Kathol's AGR feature. A third type discussed by Pollard and Sag (1994), honorific agreement (e.g. in Japanese and Korean), will not be treated here.

[6] Morpheme glosses: NOM nominative, GEN genitive, F feminine, SG singular, A(CC) accusative, INST instrumental, M masculine, PL plural, DAT dative, LOC locative, N(T) neuter, AUX auxiliary.

1963, and 1966 are specifically devoted to agreement mismatches in Serbo-Croatian. Beard 1995, an interesting generative study of the relation between agreement, declension and semantics in Slavic languages, influenced the present work, although there are substantial differences both in the theoretical approach and in the specific analyses proposed (a comparison with that work will not be undertaken here). Similar to this article in terms of descriptive goals are Corbett 1979, 1983, 1988, and 1998 (see §8). Corbett (1991:40–41) specifically attempts to predict the gender of nouns from other factors such as declension and semantics. Although our work could be used this way, our goal, more generally, is to formulate the relationships among these features as a set of constraints in order to predict patterns of exceptionality. See Corbett 1998 for a recent overview of the considerable literature on Slavic agreement; relevant generative works on Slavic morphology include Franks 1995 and Crockett 1976.

3. TYPES OF AGREEMENT. Agreement processes can be broadly divided into two types, called here GRAMMATICAL AGREEMENT and PRAGMATIC AGREEMENT. Generally speaking grammatical agreement results from structural properties of the grammar while pragmatic agreement results from structural properties of the world as described in an utterance. Grammatical agreement is further subdivided into INDEX AGREEMENT and CONCORD. This article deals primarily with index agreement and concord, but all three types must be introduced by way of background.

3.1. PRAGMATIC AGREEMENT. Pragmatic agreement results from the general condition that coreferential elements must have compatible referential properties. Consider the two-sentence discourse: *A cowboy approached the bar. She ordered a drink.* While the discourse is grammatical, it normally cannot be interpreted such that *she* and *a cowboy* are coreferential, because the two NPs conflict with respect to the sex of the referent: *she* must refer to a female, while *cowboy* must refer to a male. If the structure of the world is such that no object can satisfy both descriptions then it follows that these two NPs cannot be coreferential. Such mismatches can sometimes be accommodated, typically for rhetorical effect. By contrast, failure of grammatical agreement normally leads to ungrammaticality: **He are happy.*

3.2. PRAGMATIC AGREEMENT VERSUS INDEX AGREEMENT. Index agreement involves sharing of referential indices that are part of the semantic content of a nominal. These referential indices play an important role in semantic interpretation. The index is a variable which functions semantically as a RESTRICTED PARAMETER in the sense of situation semantics (Barwise & Perry 1983, Devlin 1991). The semantic CONTENT feature of a referential noun takes as its value a matrix with two features, the INDEX and a set of semantic RESTR(ictions) on that index. For example, a simplified version of the lexical entry for the noun *book* is given in 7.

(7) Simplified lexical sign for *book*

$$\left[\begin{array}{l} \text{PHON /book/} \\ \text{CONTENT} \left[\begin{array}{l} \text{INDEX } \boxed{1} \\ \text{RESTRICTIONS } \left\{\left[\begin{array}{l} \text{RELATION } \textit{book} \\ \text{INSTANCE } \boxed{1} \end{array}\right]\right\} \end{array}\right] \end{array}\right]$$

The CONTENT value of an NP is composed from its constituents according to a general principle of composition which will not be detailed here (the SEMANTICS PRINCIPLE, see

Pollard & Sag 1994). One consequence of that principle is that the INDEX value of an NP comes from its head noun. In the interpretation of an utterance in context, a special interpretation function called an ANCHOR maps, or 'anchors', each index onto some item in the model. This function is restricted by the utterance context, and also by the linguistic content: In ex. 7, the index must be anchored to an individual which is an INSTANCE of the 'book' RELATION, that is, it must be anchored to a book. The boxed numerals (called tags) in 7 indicate structure sharing, that is, multiple occurrences of a given tag represent a single node in a directed acyclic graph; see Shieber 1986. Pollard and Sag (1994) further propose that the index for a noun contains the grammatical features of person, number, and gender (hereafter PNG). Thus a slightly more complete representation of the value for the CONTENT feature in 7 is 8a. An abbreviation is shown in 8b.

(8) Semantic CONTENT value for *book*

a.

$$\left[\begin{array}{l} \text{INDEX } \boxed{1}\left[\begin{array}{l}\text{PERSON } \textit{3rd}\\ \text{NUMBER } \textit{sing}\\ \text{GENDER } \textit{neut}\end{array}\right] \\ \text{RESTRICTIONS } \left\{\left[\begin{array}{l}\text{RELATION } \textit{book}\\ \text{INSTANCE } \boxed{1}\end{array}\right]\right\}\end{array}\right]$$

b. abbreviation

$$\left[\begin{array}{l}\text{INDEX } \boxed{1}\left[\textit{3rd,sing,neut}\right] \\ \text{RESTR } \left\{\left[\textit{book } \left(\boxed{1}\right)\right]\right\}\end{array}\right]$$

The semantic index plays a deeper role in the morphosyntax than is sometimes assumed. For example, in local anaphoric binding the binder and anaphoric pronoun share a single index, so the matching condition on PNG features follows automatically (Pollard & Sag 1992).

There is a very close relation between the function of indices in semantic interpretation and the presence of PNG features, as the PNG features normally correlate directly with the referential anchoring conditions—this is the basis of so-called natural gender, person, and number. For example, any anchoring of the English noun *boy* must be to a male human and, in addition, its index has the feature [GENDER *masc*]. This forces the masculine form for any elements that share this index, such as bound pronouns, as in *The boy absented himself/*herself.* Note that in this example the 'fake reflexive' *himself* is not referential, so this cannot be an instance of pragmatic agreement.

The index's PNG features may be seen as grammaticalizations of anchoring conditions. The English personal pronoun *you* has the INDEX feature [PERSON *2nd*], and normally must be anchored to the addressee of the utterance. But the grammatical features cannot be reduced to referential anchoring conditions, as the correlation is not perfect: *you* can be used generically and not for the addressee in locutions such as 9a, on the interpretation paraphrased by 9b.

(9) a. You really have to watch yourself around here.
b. One really has to watch oneself around here.

Index agreement, unlike pragmatic agreement, is enforced within a syntactically defined domain. As noted above, English generics *one* and *you* are semantically similar, differ-

ing perhaps only in register, and one can switch between them in a discourse, as shown in 10a,b. But a bound reflexive must match its antecedent, as seen in 10c,d (contrast 9a,b).

(10) a. You really have to watch yourself around here. One can easily get in trouble.
b. One really has to watch oneself around here. You can easily get in trouble.
c. *You really have to watch oneself around here.
d. *One really has to watch yourself around here.

Certain structural syntactic mechanisms enforce identity of agreement features for phenomena such as subject-verb agreement and reflexive binding. But beyond the domain of those mechanisms, index agreement is not enforced and only pragmatic agreement can apply.[7]

To take a more complicated example, the Serbo-Croatian diminutive *devojče* 'girl' has neuter singular grammatical agreement features (both its index and concord features; see §3.3), so the prenominal elements and the participle must appear in neuter singular form. But a coreferential pronoun can be either neuter or feminine singular.

(11) Ovo malo devojče$_i$ je ušlo.
this.NT.SG little.NT.SG girl.(NT).SG AUX.3SG entered.NT.SG
a. Ono$_i$ je htelo da telefonira.
it.NT.SG AUX.SG wanted.NT.SG that telephone.3SG
b. Ona$_i$ je htela da telefonira.
she.F.SG AUX.SG wanted.F.SG that telephone.3SG
'This little girl$_i$ came in. She$_i$ wanted to use the telephone.'

The use of the neuter singular pronoun *ono* in 11a represents index agreement, while the use of the feminine singular pronoun *ona* in 11b is pragmatic agreement. The pronoun *ono* is lexically specified for a neuter plural index, while the pronoun *ona* in 11b is disjunctively specified: Either it has a feminine singular index, or it is restricted to a female referent. The NP *ovo malo devojče* 'this little girl' introduces a neuter singular index into the discourse representation (such an index within a discourse representation is called a DISCOURSE REFERENT in the discourse representation theory (DRT) literature; see Kamp & Reyle 1993). In this case index agreement is enforced not by syntactic structure but rather by discourse representation structure (DRS). Specifically, the DRS rule states that when two discourse referents are identified then their PNG features may match (Kamp & Reyle 1993:71), but this rule is optional. A pronoun can alternatively obey a set of pragmatic conditions, as in the use of the feminine singular pronoun *ona* in 11b. See Wechsler & Zlatić 1998 for a DRT analysis of pronoun agreement in discourse.

In addition to pronoun agreement, subject-verb agreement in the above examples is assumed, following Pollard & Sag 1994, to be index-based. This agreement is enforced as an automatic consequence of the argument-taking process. For example, a so-called third person singular verb in English is a verb that subcategorizes for a subject NP with third person singular index (see the appendix for details). An analysis of Serbo-Croatian subject-verb agreement as index agreement is defended below.

[7] This argument for distinguishing these two types of agreement is made in Bayer & Davis 1991 and Pollard & Sag 1994. See Dowty & Jacobson 1988 and Barlow 1992 for the contrary pan-pragmatic view of agreement.

3.3. CONCORD. In a Serbo-Croatian nominal, determiners and adjectives agree in gender, number and case with the head noun, as in 12.

(12) ov-a star-a knjig-a
this-NOM.F.SG old-NOM.F.SG book(F)-NOM.SG

This example is an instance of concord and not index agreement. Concord is the sharing of morphosyntactic head features between certain designated elements. As with the subject-verb agreement described above, this comes about automatically through the combination of head with dependents and adjuncts. For example, the subcategorization feature of every noun specifies that its determiner's concord features must match its own (see the appendix for details).

There are important theoretical and empirical differences between this type (concord) and the subject-verb (index) agreement. First, index agreement and concord involve different feature sets: PERSON, number and gender for the index; but CASE, number, and gender for concord. Inclusion of person as a feature of the index follows from the fact, noted above, that the index's features are grammaticalizations of the constraints on its anchoring in a discourse. In contrast, there are few if any languages in which concord includes person among its features, as noted by Lehmann (1982) and Kathol (1999).[8] Conversely, case is a feature of the NP and hence can be shared by its dependents (determiners, attributive adjectives, and so on), but obviously it could not be an index feature since a pronoun can stand in a different grammatical relation from its antecedent (nor is case involved with the referential anchoring conditions that give rise to index features). To put it more abstractly, case is a purely syntactic feature, dependent on local syntactic relations and thus inappropriate for the semantic index.

Unlike personal pronouns, which are marked for person, number, and gender, Serbo-Croatian relative pronouns agree only in number and gender—the two features in the intersection between the concord and index feature sets. The agreement behavior of relative pronouns reflects an interesting combination of concord and index (see §7.4).

In addition to the different feature sets, the two types of agreement also differ in their syntactic domains. Index agreement attaches to the referential index and hence applies to NP's and pronouns, which are referentially anchored. Agreement markers on verbs (especially finite verbs) typically derive historically from incorporated pronouns, and hence index agreement often (though not always) applies to verbs as well. But index agreement does not apply to NP-internal elements such as determiners and adjectives, for the simple reason that these elements are not referential and hence lack referential indices.[9] Hence any NP-internal agreement must be something else; here it is referred to as concord. But nothing forces concord to remain within the NP, and for NP-external agreement processes it is an empirical question whether they involve index or concord features. This issue arises with Serbo-Croatian participles and predicate adjectives, as for example, in 13.

(13) Ov-a star-a knjig-a je pa-l-a.
this-F.SG old-F.SG book(F)-NOM.SG AUX.3SG fall-PPRT-F.SG
'This old book fell.'

[8] Manfred Krifka pointed out to us one rather restricted counterexample: the Swahili form *-ote* 'all', shows person agreement in first and second person plural only: *sisi s-ote* 'all of us', but *ninyi ny-ote* 'all of you' (Ashton 1944). Other Swahili modifiers do not agree in person.

[9] A Slavic possessive adjective, like an English possessive pronoun, has a referential index, but that index is not shared with the noun it modifies (Zlatić 1997). Hence there is concord but no index agreement with the noun.

The participle *pala* agrees with the subject in gender and number (but not case or person). This is the intersection of the features involved in index and concord agreement, so it is not obvious whether the participle is agreeing with index or concord features of the subject. Participle agreement is taken up in §7.3.

To summarize, an inflected noun has a set of morphosyntactic (concord) features including number, gender and case specification, and in addition, a noun phrase is associated with a referential index with its own features of person, number, and gender. The inflected noun form *knjiga* in 13 has these features:

(14) *knjiga*:

$$\begin{bmatrix} \text{CONCORD} & \textit{f.sg.nom} \\ \text{INDEX} & \textit{3rd.f.sg} \end{bmatrix}$$

General principles of syntactic and semantic composition dictate that the NP headed by this noun share the same features. CONCORD is a HEAD feature, hence subject to the HEAD FEATURE PRINCIPLE (mother's and head daughter's HEAD values are identical). The INDEX is subject to the SEMANTICS PRINCIPLE (mother's and head daughter's INDEX values are identical; see Sag & Wasow 1999).

The concord feature is responsible for inflection on the NP-internal items (and also some predicate adjectives) while the index feature is responsible for pronouns and finite elements (verbs and auxiliaries). For a given noun the gender/number features of CONCORD and INDEX normally match, as in 14 where both are feminine singular. This gives the illusion (or so we claim) that a single feature bundle on the noun is responsible for all the agreeing items: determiner, adjective, verbs, and pronouns. However, with some nouns the values diverge, leading to mixed agreement patterns (see §7).

4. SERBO-CROATIAN GENDER AND NUMBER RULES. The three constraints diagrammed in 1, 2, and 3, respectively, specify (i) the relation between declension class and concord features, (ii) the relation between concord features and index features, and (iii) the relation between index features and semantic conditions. Details of the rules vary from language to language. This section gives the rules for Serbo-Croatian.

4.1. DECLENSION TO CONCORD. In traditional grammars, Serbo-Croatian common nouns are classified into the declension classes I, II, and III. Each class specifies a full morphological paradigm for the seven cases in singular and plural (with some syncretism). Declension class correlates with concord gender as follows: Class I nouns are either masculine or neuter; class II and III nouns are feminine. Examples appear in Table 1.

	CLASS I		CLASS II	CLASS III
SINGULAR	'window' (m)	'village' (nt)	'woman' (f)	'thing' (f)
Nominative	prozor	sel-o	žen-a	stvar
Accusative	prozor	sel-o	žen-u	stvar
Genitive	prozor-a	sel-a	žen-e	stvar-i
Dative/Locative	prozor-u	sel-u	žen-i	stvar-i
Instrumental	prozor-om	sel-om	žen-om	stvar-i
Vocative	prozor-e	sel-o	žen-o	stvar-i
PLURAL				
Nominative	prozor-i	sel-a	žen-e	stvar-i
Accusative	prozor-e	sel-a	žen-e	stvar-i
Genitive	prozor-a:	sel-a:	žen-a:	stvar-i:
Dative/Inst/Loc	prozor-ima	sel-ima	žen-ama	stvar-ima

TABLE 1. Serbo-Croatian declension classes

This relation between gender and declension class is expressed in the constraint on Serbo-Croatian nouns in 15.

(15) DECCON: $\begin{bmatrix}\text{DECL I} \\ \text{CONCORD} \mid \text{GENDER } m \vee n\end{bmatrix} \vee \begin{bmatrix}\text{DECL II} \vee \text{III} \\ \text{CONCORD} \mid \text{GENDER } f\end{bmatrix}$

The constraint DECCON (mnemonic for DEClension-CONcord) represents the normal case. Not all nouns abide by DECCON (see §5). For those regular nouns which do abide by this constraint, the disjunctive attribute-value matrix to the right of the colon is to be unified with the lexical sign for the noun stem.[10] The result is that any such regular noun must either (i) be in declension class I and have the CONCORD:GENDER value of *masculine* or *neuter*; or (ii) be in declension class II or III and have the CONCORD: GENDER value of *feminine*.

There are many alternative ways to set up the declensions, for example with four declension classes corresponding to the four columns in Table 1 above (for discussion see, for example, Beard 1997 and Carstairs 1983). Although the decision is not crucial, we adopt the traditional three-class system here because it is the most economical. (Also this system predicts an interesting asymmetry in declension-concord mismatches; see §5). Assuming that both declension class and gender determine the inflected form, then distinguishing the first two columns of the table by declension class would be redundant, because they are already distinguished by gender. The phonological form of an inflected noun is determined by the following factors: phonological form of the noun stem; declension class of the stem; concord features (stem gender; number of the noun; case of the noun); and animacy (for class I accusatives). This is formalized with a phonological function in the appendix.

Declension class and concord features (including concord gender) differ in an important respect: While the concord features are active in the morphosyntax, declension class is purely lexical, and invisible to syntax. This is captured within our formalism by making CONCORD a HEAD feature (hence subject to the head feature principle, as noted above), while the DECLENSION feature is entirely outside of the SYNSEM field, SYNSEM being the SYNtax-SEMantics feature bundle accessed by syntactic processes of subcategorization and modification. The DECLENSION feature is thus completely moribund, unaffected by feature percolation or other processes, and exists only on the noun stem and not on the noun word.[11]

Evidence that constraint 15 is part of the Serbo-Croatian competence grammar comes from INQUORATE nouns, that is, nouns with different gender in singular and plural (Corbett 1991:173, Ivić 1960:198 and 1966). Interestingly, declension class and gender observe the DECCON constraint in both singular and plural. Some inquorates have a different stem in singular and plural, such as *ok-o* ('eye' nt., d_I) ~ *oč-i* ('eyes' f., d_{III}). The singular, being neuter, is declined with the class I pattern, while the plural, being feminine, takes one of the two declensions normally reserved for feminines, in this case class III. Another example is *mač-e* ('kitten' nt., d_I) ~*mačić-i* ('kittens', m.,d_I). Certain collective nouns like *deca* 'children' (see Section 7 below) are also inquorates:

(16)

	det-e ('child')	dec-a ('children')
DECLENSION	I	II
CONCORD	nt.sg	f.sg
INDEX	nt.sg	nt.pl

[10] These constraints are implemented as subsorts of the *noun-stem* sort in the appendix.

[11] One exception is that undeclinable nouns (such as certain loans and female names) have special distributional properties related to oblique case (Wechsler & Zlatić 1999).

As seen in 16, declension and concord covary in lockstep across the singular/plural distinction, following the constraint in 15: singular *dete* 'child' triggers neuter concord (*ovo lepo dete* 'that.NT.SG beautiful.NT.SG child') and is declined in class I, while plural *deca* 'children' has feminine concord (*ova lepa deca* 'that.F.SG beautiful.F.SG children') and is declined in class II (it has singular concord despite its aggregate meaning; see §7.2).[12] Interestingly, unlike the concord gender features, the index gender features do not vary: Both singular and plural are neuter (this is demonstrated in §7.2). This pattern, where concord gender but not index gender covaries with declension, is consistent for Serbo-Croatian inquorates. In contrast, there are no nouns of the reverse sort, where the index gender covaries with declension but concord gender fails to. This supports the existence of the concord-declension constraint in 15 and the lack of any comparable index-declension constraint.

4.2. CONCORD TO INDEX. As noted above, concord features comprise number, gender, and case, while index features comprise number, gender, and person. In regular nouns such as *knjiga* 'book', the number and gender values of the concord match those of the index. Such regular nouns (the vast majority of the nouns) observe the constraint CONIND:

(17) CONIND

$$\left[\begin{array}{l}\text{CONCORD}\left[\begin{array}{l}\text{NUM } \boxed{1}\\ \text{GEN } \boxed{2}\\ \text{CASE}\end{array}\right]\\ \text{INDEX}\left[\begin{array}{l}\text{NUM } \boxed{1}\\ \text{GEN } \boxed{2}\\ \text{PERS}\end{array}\right]\end{array}\right]$$

The boxed tags indicate structure-sharing, hence, a noun subject to CONIND has only one number value and one gender value (not two of each). Each of these values, however, can be accessed in two ways, via concord and index.

4.3. INDEX TO SEMANTICS. The regular index-semantics relation has two components, for gender and number respectively. Taking gender first, Serbo-Croatian nouns lexically restricted to female (/male) referents have feminine (/masculine) index, a constraint apparently valid for Slavic generally (Corbett 1988). Hence *čovek* 'man', *muž* 'husband' and *bik* 'bull' are masculine, while *žena* 'woman/wife' and *kokoška* 'hen' are feminine. This is formulated here as the constraint $\text{INDSEM}_{\text{GEN}}$ (INDex-SEMantics$_{\text{GENder}}$) given in 18.

(18) $\text{INDSEM}_{\text{GEN}}$

$$\left[\left[\text{RESTR}\left\{\left[\textit{female}\left(\boxed{1}\right)\right],\ldots\right\}\right]\Rightarrow\left[\text{INDEX } \boxed{1}\textit{fem}\right]\right]\wedge$$

$$\left[\left[\text{RESTR}\left\{\left[\textit{male}\left(\boxed{1}\right)\right],\ldots\right\}\right]\Rightarrow\left[\text{INDEX } \boxed{1}\textit{masc}\right]\right]$$

[12] This covariation of declension and concord breaks down for nouns like *sudija* 'judge' (when referring to a male judge), for independent reasons discussed in §5.

This feature structure consists of a conjunction of two implications. The left conjunct of this rule states that an index lexically restricted to refer to females has feminine gender; the right conjunct states that an index lexically restricted to refer to males has masculine gender. It is important to note that this constraint is vacuously satisfied by any noun that is not sex-specific, either because it is inanimate (*knjiga* 'book', *sto* 'table', etc.) or because it is not specific to one sex (*kit* 'whale', *lekar* 'doctor', etc.)

Further evidence for the INDSEM$_{GEN}$ constraint comes from the agreement behavior of nouns such as *sudija* 'judge' (discussed more fully in §5). Nouns of this type trigger feminine or masculine agreement (both concord and index agreement) depending on whether they denote females or males.

(19) a. Taj stari sudija$_i$ je dobro sudio. On$_i$ . . .
that.M old.M judge AUX well judged.M 3M.SG
'That old (male) judge$_i$ judged well. He$_i$ . . .'

b. Ta stara sudija$_i$ je dobro sudila. Ona$_i$ . . .
that.F old.F judge AUX well judged.F 3F.SG
'That old (female) judge$_i$ judged well. She$_i$ . . .'

The covariation of gender across concord, index, and semantics follows on the assumption that *sudija* is disjunctively sex-specific.

(20) *sudija:*

$$\begin{bmatrix} \text{DECL II} \\ \text{CONTENT} \begin{bmatrix} \text{INDEX } x \\ \text{RESTR } \{judge(x),\ male(x) \vee female(x)\} \end{bmatrix} \end{bmatrix}$$

If the disjunction between male and female is resolved to male, then INDSEM$_{GEN}$ applies to give a masculine index gender and CONIND 17 gives masculine concord gender. If it is resolved to female, then the same constraints give feminine index and concord. Notice that the effect of this disjunctive sex specification ('male or female') differs from having no specification at all. For example, *kit* 'whale' has no sex specification at all but happens to be a masculine noun:

(21) a. taj kit
that.M whale
'that (male or female) whale'

b. Kit je podojio svoje mladunče.
whale AUX.SG nursed.3.M.SG self offspring
'The whale nursed her young.'

c. *kit:*

$$\begin{bmatrix} \text{DECL I} \\ \text{CONCORD} \mid \text{GEN } masc \\ \text{CONTENT} \begin{bmatrix} \text{INDEX } x \\ \text{RESTR } \{whale(x)\} \end{bmatrix} \end{bmatrix}$$

Hence *kit* is fully regular, satisfying all constraints. In particular, it vacuously satisfies the INDSEM$_{GEN}$ constraint, which applies only to sex-specific nouns. (In addition, it belongs to declension class I, and is masculine, thus abiding by DECCON.)

As for the index number feature, in the normal case count nouns with aggregate

reference are plural, while other nouns (mass, abstract, and nonaggregate count) are singular. This is formulated with a feature COUNT declared only by count nouns, whose value encodes the cardinality of the reference. Relevant here are two COUNT values: *one* (cardinality one) and *aggregate* (cardinality greater than one). The constraint INDSEM$_{NUM}$ (INDex-SEMantics$_{NUMber}$) is given in 22.

(22) INDSEM$_{NUM}$:

$$\begin{bmatrix} \text{INDEX } \boxed{2}\text{pl} \\ \text{RESTR } \left\{\left[\begin{array}{l}\text{INST } \boxed{2} \\ \text{COUNT } \textit{aggregate}\end{array}\right],\ldots\right\} \end{bmatrix} \vee \begin{bmatrix} \text{INDEX } \boxed{1}\text{sg} \\ \text{RESTR } \neg\left\{\left[\begin{array}{l}\text{INST } \boxed{1} \\ \text{COUNT } \textit{aggregate}\end{array}\right],\ldots\right\} \end{bmatrix}$$

The lexical entry for a count noun stem includes the feature COUNT with unspecified value. If that value is resolved to *aggregate* then the left side of the disjunction in 22 must be selected, resulting in plural index number; otherwise the right side is selected, resulting in singular number. The lexical entries for noncount noun stems lack the feature COUNT entirely, hence they can only match the right side of the disjunction, resulting in singular index number.

Serbo-Croatian collective nouns ending in *-ad*, such as *telad* '(a group of) calves', *unučad* 'grandchildren', *dugmad* 'buttons', and *burad* 'barrels', allow either singular or plural agreement with nondistributive predicates (23a), but require plural agreement with distributive predicates (23b).[13]

(23) a. Ova telad je/su pasla.
this.F.SG/NT.PL calves AUX.3SG/PL grazed.F.SG/NT.PL
'These calves were grazing.'

b. Ova telad imaju/*ima po dva roga.
this.NT.PL calves have.PL/*SG each two horns
'These calves have two horns (each).'

These agreement facts can be explained by assuming that *telad* is ambiguous between the aggregate interpretation (hence [COUNT *aggregate*]) and the nonaggregate one (perhaps one group, or a mass). The two respective CONTENT values for *telad* are shown in 24.

(24) a. CONTENT of singular *telad*

$$\begin{bmatrix} \text{INDEX } \boxed{1}\text{f.sg} \\ \text{RESTRICTION } \left\{\left[\begin{array}{l}\text{REL } \textit{group} \\ \text{INST } \boxed{1} \\ \text{MEMBERS [REL } \textit{calf}\text{]}\end{array}\right]\right\} \end{bmatrix}$$

b. CONTENT of plural *telad*

$$\begin{bmatrix} \text{INDEX } \boxed{1}\text{nt.pl} \\ \text{RESTRICTION } \left\{\left[\begin{array}{l}\text{REL } \textit{calf} \\ \text{INST } \boxed{1} \\ \text{COUNT } \textit{aggregate}\end{array}\right]\right\} \end{bmatrix}$$

[13] The predicate 'graze' is nondistributive in the sense that it does not force a distributive reading.

The IndSem$_{NUM}$ rule in 22 applies correctly to each variant of *telad*: the index in 24a is singular and its restrictions do not include the feature [COUNT *aggregate*]; while the index in 24b is plural and its restrictions do include the feature [COUNT *aggregate*]. (Hence the INDEX number values in 24 are derivable and need not be stipulated.) This ambiguity is clearly a lexical property of certain common nouns, since other collectives such as *deca* 'children' lack this ambiguity and must trigger plural INDEX agreement (although they trigger singular concord; see §7.1).

Certain English nouns like *faculty, staff*, and *nobility* show related behavior, in dialects which accept both 25a and 25b (examples from Pollard & Sag 1994:71):

(25) a. The faculty is voting itself a raise.
b. The faculty are voting themselves a raise.

As with *telad*, the noun *faculty* allows either aggregate reference ('faculty members') or nonaggregate reference (the institutional entity). The rule in 22 applies in the normal way to determine the appropriate index number value for each interpretation. What is slightly abnormal about a noun like *faculty* is that it lacks plural noun morphology in 25b despite the aggregate interpretation and plural agreement.

To summarize, the four rules given above relate syntax, morphology, and semantics in the following way: the DecCon constraint in 15 relates certain features of syntax (here called CONCORD) to morphology; the IndSem$_{GEN}$ and IndSem$_{NUM}$ constraints in 18 and 22 relate certain grammatical features (namely INDEX features) to semantics; and ConInd in 17 relates the morphology-related syntactic features (concord) to the semantics-related syntactic features (index). For example, a perfectly regular noun stem, such as *knjig-* 'book', observes all four constraints. The lexical sign for this noun is calculated by unifying the feature structures for the four respective constraints with the feature structure of the basic lexeme *knjig-* 'book'.

(26)

$$\left[\begin{array}{l} \text{PHON } /\text{knjig-}/ \\ \text{DECL II} \\ \text{SYNSEM} \left[\begin{array}{ll} \text{CATEGORY} & \left[\begin{array}{l} \text{HEAD } \textit{noun}[\text{CONCORD } \boxed{2}] \\ \text{VALENCE } [\text{SPR } \langle(\text{AP}[\textit{poss}; \text{CONCORD } \boxed{2}])\rangle] \end{array}\right] \\ \text{CONTENT} & \left[\begin{array}{l} \text{INDEX } \boxed{1} \\ \text{RESTR } \left\{\left[\begin{array}{l} \text{REL } \textit{book} \\ \text{INST } \boxed{1} \\ \text{COUNT } \textit{count} \end{array}\right]\right\} \end{array}\right] \end{array}\right] \end{array}\right]$$

$\wedge$ DecCon $\wedge$ ConInd $\wedge$ IndSem$_{GEN}$ $\wedge$ IndSem$_{NUM}$

The only idiosyncratic information stipulated in the above lexeme is declension class II. The agreement features are derived through application of the above constraints. The result of unification is given in 27.

(27)

$$\left[\begin{array}{l}\text{PHON /knjig-/}\\ \text{DECL II}\\ \text{SYNSEM}\left[\begin{array}{ll}\text{CAT} & \left[\begin{array}{l}\text{HEAD}\left[\begin{array}{ll}\textit{noun} & \\ \text{CONCORD}\ \boxed{2} & \left[\begin{array}{l}\text{CASE } \textit{case}\\ \text{NUM } \boxed{3}\\ \text{GEN } \boxed{4}\,\textit{fem}\end{array}\right]\end{array}\right]\\ \text{VALENCE}\left[\text{SPR}\left\langle(\text{AP}[\textit{poss};\ \text{CONCORD}\ \boxed{2}])\right\rangle\right]\end{array}\right]\\ \text{CONTENT}\left[\begin{array}{ll}\text{INDEX}\ \boxed{1} & \left[\begin{array}{l}\text{PERS } \textit{3rd}\\ \text{NUM } \boxed{3}\\ \text{GEN } \boxed{4}\,\textit{fem}\end{array}\right]\\ \text{RESTR} & \left\{\left[\begin{array}{l}\text{REL } \textit{book}\\ \text{INST } \boxed{1}\\ \text{COUNT } \textit{count}\end{array}\right]\right\}\end{array}\right]\end{array}\right]\end{array}\right]$$

Index and concord gender are now specified as feminine, but CASE, NUMber and COUNT values are unspecified. Not shown here is the fact that, due to the constraint INDSEM$_{\text{NUM}}$, the NUM and COUNT values covary, that is, they can be specified in two ways: [NUM sg, ¬ COUNT *aggregate*] or [NUM pl, COUNT *aggregate*]. Thus this noun STEM can be the basis for either singular or plural inflected nouns. The number and case values determine the phonological form of the inflected noun (see Appendix for details).

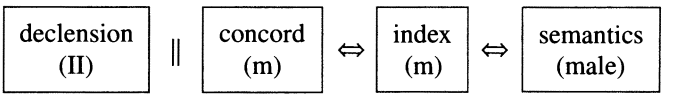
Each of the constraints is very simple, but they interact to predict a complex array of facts. The complexity arises because an individual constraint can be violated, while leaving the rest of the constraint system intact. The next three sections present such exceptions to the constraints.

5. DECLENSION-CONCORD MISMATCHES. Certain male names in *-a* such as *Steva* and *Mika* have all masculine grammatical properties (and refer to males) but nonetheless decline in class II, an otherwise feminine declension.

(28) Vratio mi je ovaj ludi Steva
returned.1SG me AUX.3SG this.NOM.M.SG crazy.NOM.M.SG Steve.NOM
violinu koju sam mu pozajmio.
violin-ACC which AUX.1SG 3DAT.M.SG loaned
'This (M) crazy (M) Steve returned to me the violin which I loaned him(M).'

This situation may be diagrammed as follows:

(29)



A similar pattern is found with a set of nouns that refer to humans and belong to the second declension class, having the nominative singular ending in *-a*, such as *sudija* 'judge', *sluga* 'slave', *gazda* 'master', *mušterija* 'customer', and *kolega* 'colleague'.

The *-a* nominative desinence is normally associated with feminine agreement. But as illustrated in 19 (repeated here), nouns of the *sudija* type trigger feminine or masculine agreement depending on whether they denote females or males (see also Ivić 1960).[14]

(19) a. Taj stari sudija$_i$ je dobro sudio. On$_i$. . .
that.M old.M judge AUX well judged.M 3M.SG
'That old (male) judge$_i$ judged well. He$_i$ . . .'

b. Ta stara sudija$_i$ je dobro sudila. Ona$_i$ . . .
that.F old.F judge AUX well judged.F 3F.SG
'That old (female) judge$_i$ judged well. She$_i$ . . .'

As noted above, the covariation of gender across concord, index, and semantics follows on the assumption that *sudija* is sex-specific but disjunctive (see 20 above). Thus when referring to a male judge, as in 19a, *sudija* resembles *Steva*: Concord and index agreement are masculine, the referent is male, but the noun declines as if it were feminine (declension class II).

As noted above, *kit* 'whale' has no sex specification at all but happens to be a masculine noun (declension class I), hence it triggers masculine concord regardless of the sex of the whale (see 21). Nouns like *mušterija* 'customer' and *pijanica* 'drunkard' represent a mix between the *sudija* and *kit* types (Ivić 1960:197). When in construction with feminine forms, *mušterija* is semantically unmarked for sex as in 30a; but with masculine forms, it is interpreted as male as in 30b.

(30) a. Ta mušterija je došla.
that.F customer AUX came.F
'That (male or female) customer came.'

b. Taj mušterija je došao.
that.M customer AUX came.M
'That (male) customer came.'

This pattern follows if the male sex specification is optional (31), and allows two possibilities for *mušterija* (32).

(31) *mušterija:*

$$\begin{bmatrix} \text{DECLENSION II} \\ \text{CONTENT} \begin{bmatrix} \text{INDEX } x \\ \text{RESTR } \{customer(x)\ (,\ male(x))\} \end{bmatrix} \end{bmatrix}$$

(32)

FEATURE	EXAMPLES ACCOUNTED FOR
a. [RESTR {*customer(x)*}]	30a (male or female customer)
b. [RESTR {*customer(x)*, *male(x)*}]	30b (male customer)

Variant 32a violates no constraints. It is unspecified for sex, so like *kit* 'whale' it vacuously satisfies IndSem$_{GEN}$, and the gender agreement feature is determined by the declension class, namely feminine. This accounts for the male and female variants of 30a (compare 21a). Variant 32b, with semantic restriction to males and declension class II (normally for feminines) necessarily violates some constraint. In particular, it violates DecCon. Masculine agreement is consistent with the semantics but inconsistent with the declension. Given the restriction to males, IndSem$_{GEN}$ forces masculine agreement. (Since this is a declarative framework, this is equivalent to saying that the masculine agreement targets force the noun variant with the restriction to males.)

[14] The nouns *sluga* 'slave', *gazda* 'master', and *kolega* 'colleague' pattern like *sudija*, but the feminine forms *služavka, gazdarica*, and *kolegi(ni)ca* can alternatively be used for female referents.

The plural forms of these nouns, such as *sudije* 'judges' and *gazde* 'masters', lose the disjunctive sex-specification of the singular. Instead all features are consistent with the declension class: Concord and index are feminine even when the noun refers to males.[15]

(33) Te sudije su došle.
those.F.PL judge(F).PL AUX.3PL came.F.PL
'Those (male and/or female) judges came.'

Since singular *sudija* and plural *sudije* differ, the noun stem underlying both must be disjunctively specified: The nonaggregate (COUNT *one*) variant has the sex restriction *male(x)* $\vee$ *female(x)* as shown in 20, while the aggregate (COUNT *aggregate*) variant lacks that restriction.

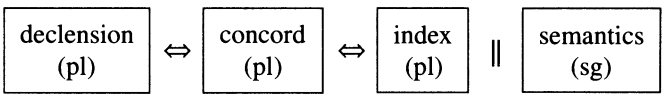
Notice that all of our examples of declension-concord mismatch involve masculine concord and declension class II. Interestingly, our declension system rules out most other hypothetical combinations. Mismatches involving declension class I are predicted to be impossible because a class I noun must have a concord gender value of masculine or neuter in order to select the proper subclass of class I, namely either 'class I masculine' or 'class I neuter' (shown in the first two columns of Table 1). A hypothetical irregular noun specified for class I declension and feminine concord would be of indeterminate form, since it would lie outside the range of the phonological spell-out function (see the discussion following 68 below). Assuming that the phonological form must be determined by the grammar, such hypothetical nouns are predicted to be impossible.

6. INDEX-SEMANTICS MISMATCHES. The index-semantics rule is split into two conditions, one for number (see 22) and one for gender (see 18). Pluralia tantum nouns, which are plentiful in Serbo-Croatian, violate the number condition, for example *kola* 'car', *vrata* 'door', *ledja* 'back', *makaze* 'scissors', *pantalone* 'pants', *naočare* 'glasses', and *novine* 'newspaper'. These have plural declension, concord, and index features, but can refer to nonaggregate entities (or else aggregates):

(34) Ove naočare$_i$ su nove. One$_i$ . . .
this.PL glasses be.3PL new.PL 3PL . . .
'These glasses$_i$ are new. They$_i$ . . .'

When the interpretation is nonaggregate ('this pair of glasses') these nouns illustrate this type of mismatch.

(35)



Serbo-Croatian gender mismatches between index and semantics are exemplified by diminutives like *devojče* 'girl', shown in 11 above: It denotes a female but nonetheless has neuter index. A similar but more unusual example is *braća* 'brothers'. The gender constraint INDSEM$_{GEN}$ in 18 applies to sex-specific nouns and states that male-only nouns must have masculine indices. Yet *braća* 'brothers', while male only, has a neuter plural index, as shown by the possibility of neuter plural coreferential personal pronouns, as in 36.

[15] Ivić 1960, citing Hraste 1953, notes that these plural nouns often trigger masculine agreement in Croatian dialects. In such dialects the sex restriction is disjunctive for both singular and plural.

(36) Sreo sam braću$_i$. Ona$_i$ su došla.
met.1SG AUX brothers.ACC they.N.PL AUX.3.PL came.N.PL
'I met the brothers. They came.'

The masculine plural pronoun *oni* can alternatively replace *ona* in 36. The masculine plural option is an instance of pragmatic agreement (see §3.2, especially ex. 11). In Serbo-Croatian, masculine plural is the form for pronouns referring to an aggregate of males (or an aggregate of mixed or unknown sex or gender). Pragmatic agreement is analyzed in Wechsler & Zlatić 1998, which describes the semantic and discourse correlates of the alternation between pragmatic agreement and grammatical agreement, the latter type being agreement determined by the grammatical features (concord and index) of the antecedent noun phrase. We focus here on grammatical agreement.[16]

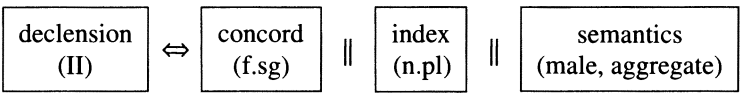
Relative pronouns also show neuter plural agreement with *braća* but adjectives and determiners show feminine singular agreement with *braća*. Compare 37 and 38.

(37) ov-u dobr-u brać-u
this-ACC.F.SG good-ACC.F.SG brothers-ACC

(38) ov-om dobr-om brać-om
this-INS.F.SG good-INS.F.SG brothers-INS

This noun is radically irregular, as it also violates the concord-index rule (CONIND, 17), as discussed in §7.1. As a consequence, *braća* actually has all three genders, and both number values! This is shown in 39.

(39) *braća* 'brothers'



In contrast, the male-specific noun *gospoda* 'gentlemen' abides by the index-semantics rule (although like *braća* it violates the concord-index rule). A pronoun bound by *gospoda* is masculine rather than neuter.[17]

(40) Sreo sam gospodu$_i$.
met.1SG AUX gentlemen.ACC
. . . *Ona$_i$ su bila jako ljubazna.
they.NT.PL AUX.PL were.NT.PL very kind.NT.PL
. . . Oni$_i$ su bili jako ljubazni.
they.M.PL AUX.PL were.M.PL very kind.M.PL
'I met the gentlemen. They were very kind.'

Nouns like *gospoda* 'gentlemen' are discussed further in the next section.

[16] Wechsler and Zlatić (1998) show that in strict/sloppy anaphora contexts, the coreferential ('strict') reading is possible with either structural or pragmatic agreement, while the sloppy reading requires structural agreement. For more on structural versus pragmatic agreement in Slavic, see Corbett 1983 and sources cited there. Corbett (1983, 1991) shows that in textual corpora pronouns that are closer to their antecedents tend to show structural agreement while those further away show pragmatic agreement. For more on the issue of grammatical versus pragmatic agreement generally, see Barlow 1992, Bosch 1987, 1988, 1989, Cornish 1987, 1994, 1999, Dowty & Jacobson 1988, Pollard & Sag 1994, Tasmowski & Verluyten 1985, and Tasmowski-DeRijk & Verluyten 1982.

[17] If the sentence containing the pronoun lacks any gender agreeing elements, then the neuter plural pronoun is slightly improved:

(i) . . . gospoda$_i$ . . . ? Ona$_i$ dolaze.
masters 3NT.PL come.3NT.PL
'. . . masters . . . They are coming.'

7. INDEX-CONCORD MISMATCHES. In our last type of mismatch the link between concord and index is severed. This type is considerably rarer than the others, but certain collective nouns of the second declension class have this property.

7.1. MIXED AGREEMENT WITH *deca*-TYPE COLLECTIVE NOUNS. The nouns in question have nominative ending in *-a* such as *deca* 'children', *braća* 'brothers', *gospoda* 'gentlemen', and *vlastela* 'landowners'. According to traditional Serbo-Croatian grammar books these nouns denote aggregate (plural) entities but bear a morphosyntactic (feminine) singular number marking. These nouns have singular counterparts *dete* 'child', *brat* 'brother', and *gospodin* 'gentleman', which behave normally with respect to agreement. The plurals will be referred to as nouns of the *deca* type.

As shown in 41, these nouns trigger mixed agreement: feminine singular on determiners, modifiers, and nonfinite predicate phrases, but neuter plural on coreferential pronouns.

(41) Posmatrali smo ovu dobru decu_i.
watched.1.PL AUX this.F.SG good.F.SG children.F.SG
Ona_i su spavala.
they.N.PL AUX.3PL slept.NT.PL
'We watched those good children_i. They_i slept.'

(42) Deca_i su spavala. Mi smo ih_i videli.
children AUX.3PL slept.NT.PL we AUX.PL them.ACC.PL saw
'The children_i slept. We saw them_i.'

As in the case of *braća* 'brothers' in 36, masculine plural pronouns can replace the neuter plural pronouns in 41 and 42. Again, the use of the masculine plural is pragmatic agreement, as Serbo-Croatian uses masculine plural for groups with mixed or unknown sex.

In addition, finite verbs and auxiliaries show plural agreement (they are unmarked for gender). In Serbo-Croatian only nominative NP's trigger finite agreement, as in 43.

(43) Ta dobra deca su došla.
that.F.SG good.F.SG children(F.SG) AUX.3PL come-PPRT.N.PL
'Those good children came.'

(44) Ta dobra deca dolaze.
that.F.SG good.F.SG children(F.SG) come.3.PL
'Those good children came.'

(The participle *došla* in 43, glossed here as neuter plural, is actually ambiguous between neuter plural and feminine singular; see §7.3.) To summarize, collective nouns of the *deca* type trigger feminine singular agreement on NP-internal items, neuter plural agreement on pronouns and finite verbs or auxiliaries, and an indeterminate form (feminine singular or neuter plural) on participles.

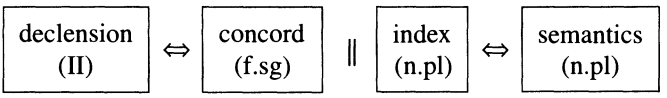
Accordingly, nouns of this type have a CONCORD value of feminine singular and INDEX value of neuter plural.

(45) *deca*:
$$\begin{bmatrix} \text{CONCORD} & \textit{fem.sg} \\ \text{INDEX} & \textit{nt.pl} \end{bmatrix}$$

With respect to number this mismatch of features captures the traditional notion of collective nouns, which have morphosyntactic singular number but refer to aggregate entities (e.g. Maretić 1899, Stevanović 1974, Mrazović & Vukadinović 1990). This follows since the CONCORD feature is associated with morphological concord within

the NP, while the INDEX feature is associated with the semantic anchoring conditions (46).

(46)



Nouns like *deca* belong to the second declension class, which contains feminine nouns; hence declension and concord match. Recall that the index-semantics rule has two components, for gender and number respectively. The number rule states that plural index correlates with aggregate reference, singular index elsewhere; clearly these nouns satisfy this condition.

The gender rule applies only to sex-specific nouns, and simply states that nouns restricted to female referents have feminine index, while nouns restricted to male referents have masculine index. This condition is vacuously satisfied by *deca* 'children' and *vlastela* 'landowners' but is violated by *braća* 'brothers'. With respect to gender *braća* actually has two breaks (see 39 above). Nouns of this kind are expected to be rare, since two constraints are violated, but nonetheless predicted to be possible according to our theory. Note that declension/concord are feminine while the semantics is male, hence these do not match in the usual way. Contrast a similar but apparently unattested hypothetical noun type in which the index value is the 'odd man out', while declension, concord and semantics are still mutually consistent (e.g. a noun just like *braća* but denoting females instead of males). On the present theory, such consistency between concord and semantics could only occur by coincidence, since there is no constraint directly relating concord to semantics. Instead they are related indirectly by two constraints, with index as an intermediary.

7.2. NOMINATIVE *deca*. Nominative forms of *deca*-type nouns pose a special problem, due to a quirk of Serbo-Croatian morphology. In nominative case only, feminine singular and neuter plural happen to be represented with homophonous endings on nouns and concord targets, as shown in Table 2 (see Corbett 1983:76).

	MASC	FEM	NEUT
singular	-ø	-a	-o/-e
plural	-i	-e	-a

TABLE 2. Serbo-Croatian nominative agreement endings (Det, Adj, Participle, etc.)

As a consequence, it is unclear a priori whether a sentence such as 47 is properly glossed as in (i) or (ii) or some combination of the two.

(47) Ta dobra deca su došla.
(i) that.F.SG good.F.SG children(F.SG) AUX.PL came.PPRT.F.SG
(ii) that.N.PL good.N.PL children(N.PL) AUX.PL came.PPRT.N.PL
'Those good children came.'

Our theory of agreement provides the following answer: via CONCORD agreement *deca* triggers feminine singular on the NP-internal elements (*ta*, *dobra*); via INDEX agreement it triggers third person plural on the finite auxiliary *su*; and the status of the participle *došla* is unclear. This is shown in 48.

(48) Ta dobra deca su došl-a.
that.F.SG good.F.SG children(F.SG$_{[N.PL]}$) AUX.3PL came.PPRT.??
'Those good children came.'

The auxiliary *su* is unambiguously plural, and so this clearly reflects the neuter plural index rather than the feminine singular concord value. As to the feminine singular on adjectives and determiners, recall that these modifiers show unambiguous feminine singular agreement with *deca* in all six cases besides nominative (feminine singular and neuter plural are homophonous in nominative case only).

7.3. PARTICIPLES AND PREDICATE ADJECTIVES. Turning now to the participle *došla* in 48, the situation is more complex, since Serbo-Croatian participles agree only with nominatives. Hence it is difficult to tell whether they show concord (here, f.sg) or index (here, n.pl) agreement. The same situation applies to predicate adjectives like *gladna* 'hungry (f.sg. or n.pl)'.

(49) Ta dobra deca su gladna.
that-F.SG good-F.SG children(F.SG$_{[N.PL]}$) AUX.3.PL hungry-??
'Those good children are hungry.'

Do predicate adjectives and participles show index or concord agreement? As a first clue, predicate adjectives show concord agreement when predicated of non-nominative NP's (such secondary predication is not possible with participles).

(50) Ja smatram decu gladnom/ *gladnim.
I consider children.ACC(F.SG$_{[N.PL]}$) hungry.INST.FEM.SG/*INST.PL
'I consider the children hungry.'

The secondary predicate appears in instrumental case, where plural and feminine singular are distinguished. As shown in 50, the adjective unambiguously shows feminine singular, hence concord and not index agreement. Assuming that primary and secondary predication are similar processes, this suggests that primary predicates also show concord agreement.

But data from coordination appear to point in the opposite direction, to index agreement. Normally in Serbo-Croatian two conjoined feminine singular subject NPs trigger feminine plural agreement on the predicate, while a mix of genders triggers masculine plural agreement. Corbett (1983) observes that when *deca* is conjoined with an unambiguously feminine singular noun, the result is masculine rather than feminine (plural) predicate agreement; compare 51 and 52.

(51) Jelena i deca su došli/ *došle.
Yelena(F.SG$_{[F.SG]}$) and children(F.SG$_{[N.PL]}$) AUX.3.PL came.M.PL/*F.PL
'Yelena and the children came.'

(52) Jelena i deca su gladni/ *gladne.
Yelena(F.SG$_{[F.SG]}$) and children(F.SG$_{[N.PL]}$) AUX.3.PL hungry.M.PL/*F.PL
'Yelena and the children are hungry.'

The features of *deca* relevant to triggering agreement on the participle are the neuter plural index features. Since the two conjuncts do not match (f.sg and n.pl), the rule dictates masculine plural agreement on the predicate adjective. This appears to contradict our earlier conclusion that participles show concord agreement.

Corbett's test is supported by a second coordination test (suggested to us by Wayles Browne). According to this test, two conjoined nominative neuter plural nouns should trigger neuter plural agreement on the predicate. And indeed, when *deca* is conjoined with an unambiguous neuter plural noun, such as *čudovišta* 'monsters' in 53, the result is neuter plural agreement on the predicate (or the dispreferred default masculine plural).

(53) Ta deca i ta čudovišta
that children(F.SG[N.PL]) and those monsters(N.PL[N.PL])
su se lepo igrala/ ?igrali.
AUX REFL well played.N.PL ?M.PL
'Those children and those monsters played well.'

Once again, neuter plural is the relevant value for *deca*—again supporting the view that predicates show index agreement, but contradicting the conclusion reached above on the basis of secondary predicates. One obvious solution to this conundrum is to posit that primary and secondary predication differ. Perhaps secondary predicate adjectives resemble attributive adjectives in showing concord agreement, which may in turn be related to the typical adjacency between the noun and the adjective, whether attributive or secondary predicate. But primary predicates resemble finite verbs and auxiliaries in showing index agreement, perhaps because they are more distant. This conclusion, however, depends on what theory of resolution rules for mixed coordinate structures is applied to 51–53. This issue will be left for future research.[18]

7.4. RELATIVE PRONOUNS. Relative pronouns such as *koji* show an interesting and seemingly anarchistic behavior with respect to agreement with *deca*. First of all, the nominative relative pronoun in a relative clause modifying *deca* takes the form *koja*. Its *-a* ending makes *koja* ambiguous between feminine singular and neuter plural, but it triggers plural on the predicate within the relative clause, showing that it is the neuter plural form. This fact was noted by Corbett (1983:78), from whom 54 is taken.

(54) deca koja su/ *je tada bila
children who.N.PL AUX.PL/*SG then were
'the children who were'

A relative pronoun in accusative case, however, takes a feminine singular form (*koju* in 55), and those in other oblique cases can take either feminine singular (*koje*) or plural (*kojih*) form, the latter unmarked for gender, as shown in 56.

(55) deca koju sam video
children who.ACC.F.SG AUX.1.SG saw
'the children whom I saw'

(56) deca koje/ kojih se svi plaše
children who.GEN.F.SG./GEN.PL REFL all fear
'children whom everyone fears'

This strange behavior is at least partially explained within our theory. Relative clauses, being N′-modifiers and hence NP-internal, are expected to show CONCORD agreement (if any), while the relative pronoun, being a bound pronoun, is expected to show INDEX agreement. Assuming that the relative clause's CONCORD agreement features are marked on the relative pronoun,[19] relative pronouns are expected to agree with the head noun in both concord and index features. (As noted above, Serbo-Croatian relative pronouns are unmarked for person, agreeing only in number and gender—the intersection of the feature sets involved in concord and index agreement.)

[18] On coordination resolution rules, see Corbett 1983, 1991, Farkas & Zec 1993, Moosally 1998, Dalrymple & Kaplan 2000, and Wechsler 1999b.

[19] This would follow if the relative pronoun heads its clause. If not then some device other than the head feature principle must force sharing of CONCORD values. One possible approach is the construction-based analysis of relative clauses (Sag 1997).

We posit (in 57) that *koja*, *koju*, and *koje* are all lexically specified for appearing in relative clauses that modify [CONCORD *f.sg*] nominals (like all forms in -a, *koja* is ambiguous between f.sg and n.pl, hence the disjunctive value *f.sg* $\vee$ *n.pl*). The HPSG feature 'MOD' (for modified) appears on adjuncts and unifies with information from the head, in this case the N′ the relative clause modifies (see the appendix for details).

(57)

$$koja: \begin{bmatrix} \text{CASE} & nom \\ \text{MOD} & \text{N}'\left[\text{CONCORD } f.sg \vee n.pl\right]_{\boxed{1}} \\ \text{INDEX} & \boxed{1} \end{bmatrix}$$

$$koju: \begin{bmatrix} \text{CASE} & acc \\ \text{MOD} & \text{N}'\left[\text{CONCORD } f.sg\right]_{\boxed{1}} \\ \text{INDEX} & \boxed{1} \end{bmatrix}$$

$$koje: \begin{bmatrix} \text{CASE} & gen \\ \text{MOD} & \text{N}'\left[\text{CONCORD } f.sg\right]_{\boxed{1}} \\ \text{INDEX} & \boxed{1} \end{bmatrix}$$

$$kojih: \begin{bmatrix} \text{CASE} & gen \\ \text{MOD} & \text{N}'_{\boxed{1}} \\ \text{INDEX} & \boxed{1}\, pl \end{bmatrix}$$

(The subscripted numbers abbreviate the feature [INDEX [1]].) But not all relative pronouns are fully specified for all features. In particular, PLURAL relative pronouns tend to be unspecified for morphological (CONCORD) gender—as is typical for Serbo-Croatian and crosslinguistically (Greenberg 1963, Dixon 1994). Thus the plural (INDEX) genitive pronoun *kojih* is unmarked for CONCORD. (At present, we have no explanation for the fact that, according to our analysis, the indices of nominative and accusative relative pronouns are unmarked for PNG features.) The result is precisely the complex agreement pattern seen above, illustrated in 58 (see 55–56).

(58)

a. Deca koju sam video

CONCORD & INDEX

$$\begin{bmatrix} \text{CASE} & nom \\ \text{CONCD} & f.sg \\ \text{INDEX} & n.pl \end{bmatrix} \begin{bmatrix} \text{CASE} & acc \\ \text{MOD N}' & [\text{CONCD } f.sg] \end{bmatrix}$$

'the children whom I saw'

b. Deca koje / kojih se svi plaše

CONCORD & INDEX

$$\begin{bmatrix} \text{CASE} & nom \\ \text{CONCD} & f.sg \\ \text{INDEX} & n.pl \end{bmatrix} \begin{bmatrix} \text{CASE} & gen \\ \text{MOD N}' & [\text{CONCD } f.sg] \end{bmatrix} \begin{bmatrix} \text{CASE} & gen \\ \text{MOD} & \text{N}'[\text{INDEX } [1]] \\ \text{INDEX} & [1]\, pl. \end{bmatrix}$$

CONCORD & INDEX

'the children whom everyone fears'

In short, all relative pronouns agree with their head nouns in both CONCORD and INDEX, but not all pronouns are specified for all features. The 'morphologically' (CONCORD) singular forms *koju* and *koje* are unmarked for INDEX features, while the plural INDEX form *kojih* is unmarked for CONCORD features.

7.5. RELATIVE CLAUSES INTRODUCED BY *što*. Another type of relative clause, introduced by the complementizer *što*, employs ordinary pronouns functioning like resumptive pronouns. Following Browne 1986, these will be called *što*-relatives, as distinct from the *koji*-relatives described in the previous section. Since the pronouns in *što*-relatives are ordinary pronouns they show index agreement, irrespective of case. Hence when the antecedent is *deca*, that pronoun must be plural rather than singular.

(59) moja deca, što ih/ *je nije briga nizašta.
my children, that 3.ACC.PL/*SG neg.AUX care nothing
'my children, whom nothing worries' (i.e. 'who don't worry about anything')

Thus *što*-relatives directly contrast with *koji*-relatives.

(60) a. Accusative *koji*-relative (= 55)
deca koju sam video
children who-ACC.F.SG AUX-1.SG saw
'the children whom I saw'
b. Accusative *što*-relative
moja deca, što ih/ *je svi hvale
my children, that 3.ACC.PL/*SG all praise
'my children, whom everyone praises'

(61) a. Genitive *koji*-relative (= 56)
deca koje/ kojih se svi plaše
children who-GEN.F.SG./GEN.PL REFL all fear
'children whom everyone fears'
b. Genitive *što*-relative
moja deca, što ih se svi plaše
my children, that 3.GEN.PL REFL all fear
'my children, whom everyone fears'

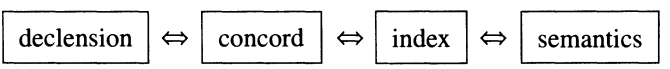
As these examples show, the *što*-relative pronoun must be plural, indicating index agreement. The singular pronoun counterpart to 61b is ungrammatical.

(62) *moja deca, što je se svi plaše
my children, that 3.GEN.F.SG REFL all fear
'my children, whom everyone fears'

In sum, the resumptive pronouns in *što*-relative clauses are ordinary pronouns and thus show index and not concord agreement.

8. THEORETICAL AND CROSSLINGUISTIC IMPLICATIONS. In essence our universal theory holds that agreement features are subject to three types of constraints, arranged as in 4 above, repeated here.

(4)



Of the two morphosyntactic features, concord is a head feature, hence accessible to complement saturation and adjunction processes, while the index is involved in bound

anaphora, since bound anaphora is modeled on coindexation (identity of indices), and thus has access to index features but not head features. As shown in 4, the grammar of a language determines a systematic relation between concord and declension on the one hand, and between index and semantics on the other. We propose that this aspect of the architecture of agreement is universal. But languages vary in the particular constraints, the number and nature of exceptions to those constraints, and the inventory of agreement features.

How well does our theory fare crosslinguistically? Corbett (1983, 1991) surveyed agreement processes across many languages to determine whether a given agreement type is more likely to involve SYNTACTIC AGREEMENT (agreement *ad formam*, 'grammatical' agreement) or SEMANTIC AGREEMENT (agreement *ad sensum*). Corbett summarized the striking crosslinguistic generalizations he discovered in the following agreement hierarchy (Corbett 1983:10, 1991:226, 1998).

(63) The Agreement Hierarchy
attributive < predicate < relative pronoun < personal pronoun
←syntactic agreement— —semantic agreement→
As we move rightwards along the hierarchy, the likelihood of semantic agreement will increase monotonically (that is, with no intervening decrease).

Corbett supported this hierarchy with evidence from hybrid agreement. If, for example, a given noun shows a split between the features determined on personal pronouns and relative pronouns, then the relative pronouns will show syntactic and the personal pronouns semantic agreement, rather than the other way around. If there is a split between all pronouns and predicates, then the predicate shows syntactic and the pronouns semantic agreement, and so on.

This fascinating typological observation is largely explained by our theory of agreement. All pronouns have indices and hence show index agreement with any coindexed items (normally the antecedent). According to our theory the index is directly connected to the semantics and furthermore mediates any connection between the semantics and concord (see example 4 above). Hence concord elements should never show semantic agreement unless the pronouns also do.

The split between personal and relative pronouns really subsumes two different observations, both explained by the present account. First, relative pronouns are always anaphorically bound,[20] while personal pronouns can be deictic or exophoric (in the sense of Halliday 1985). In other words, personal pronouns can refer directly to a class of referents rather than receiving their interpretation syntactically via a linguistic antecedent (this is deixis if the referent is present in the speech context, exophora if it is not). Deictic and exophoric pronouns will thus undergo pragmatic agreement (see §3). Hence if there is a mismatch between personal and relative pronouns, then any personal pronouns being used deictically or exophorically will go the semantics way while the relative pronouns will agree in index features. Second, as discussed above, relative pronouns can in principle also show some concord properties because they are NP-internal.

Now consider the attributive versus predicative distinction. Attributive modifiers are unlikely to show index agreement since they lack referential indices; any agreement they do show is what we have called concord. Nouns like *sudija* 'judge' appear to

[20] With the possible exception of pronouns in free relative clauses like *what* in *John ate what I ordered.*

trigger semantic agreement on attributive adjectives (as in 19), but this is really concord agreement which is forced to correlate with index and semantics by the constraints. Hence attributive adjectives show semantic agreement only if pronouns also do. What is therefore predicted to be impossible is a noun triggering this apparent semantic agreement on attributive adjectives, but grammatical agreement (concord) on pronouns.

Turning to predicates, our theory makes no prediction as to whether they should show concord or index agreement. Indeed, in Serbo-Croatian they are mixed. Interestingly, the split is closely related to the attributive modifiers: predicate forms that can alternatively function attributively, namely adjectives, show concord in secondary predication. Finite verbs and auxiliaries, which never serve as attributive modifiers, show index agreement. For Serbo-Croatian participles and primary predicate adjectives the situation is unclear, as discussed in §7.3.

To take a single non-Serbo-Croatian example as illustration, Spanish titles like *Majestad* '(his) majesty' show concord-index disagreement (like *deca*). This noun, which refers to a male, triggers feminine on attributive modifiers, but masculine on predicates and pronouns (ex. 64 adapted from Corbett 1991:230).

(64) Su Majestad suprema$_i$ esta contento. Él$_i$. . .
his majesty supreme.F is happy.M he . . .
'His supreme majesty is happy. He . . . '

(65) *Majestad*: $\begin{bmatrix} \text{CONCORD} & \textit{fem} \\ \text{INDEX} & \textit{masc} \end{bmatrix}$

As 65 shows, we can assign feminine concord and masculine index features to *Majestad*, on the assumption that predicate adjectives in Spanish show index agreement.

While our chain of constraints in 4 has clear parallels with Corbett's agreement hierarchy (63), the two constructs differ in theoretical status. Our diagram 4 represents a general schema for rules in the competence grammar of human languages, for which the specific Serbo-Croatian instantiation is presented in detail above. In contrast, Corbett's agreement hierarchy represents statistical patterns governing corpora, as Corbett (1998:36) makes clear.[21] The chain of constraints thus represents a particular explanation for many of the facts described in terms of the agreement hierarchy. According to that explanation, a noun can have up to two sets of grammatical agreement features, one associated with the noun's form (CONCORD features) and the other with the noun's meaning (INDEX features). The syntax puts those feature sets to different uses. Aspects of this explanation are in the spirit of informal explanations offered by Corbett and others, but now they have been incorporated into a formalized competence grammar. As a result, a set of grammatical constraints which are local (each constraint governs only two features) and categorical (the constraints are all or nothing) explains textual data that are describable in terms of an implicational hierarchy. It would be interesting to see whether any other observed implicational universals, typically stated in terms of hierarchies, are amenable to constraint-based analyses of the sort proposed here.

[21] 'The important claim is that at the level of the corpus the constraints of the Agreement Hierarchy will hold. That is to say, the overall likelihood of semantic agreement will be greater with targets to the right on the hierarchy . . . but this constraint need not apply at the level of the individual sentence' (Corbett 1998: 36).

9. CONCLUSION. We hope to have contributed to the understanding of agreement in two ways. First, morphosyntactic features, which we call concord features, were distinguished from features of the referential index (essentially following Kathol 1999). Second, we have provided a theory that predicts the range of possible mixed agreement patterns for nouns. Table 3 summarizes the Serbo-Croatian data.

		LEXICAL FEATURES OF THE NOUN			
NOUN TYPE	EXAMPLE	DECL	CONCD	INDEX	SEM
regular (unsexed)	*knjiga* 'book'	II	f	f	–
regular (unsexed)	*rad* 'work'	I	m	m	–
regular (female)	*žena* 'woman'	II	f	f	F
regular (male)	*muž* 'husband'	I	m	m	M
sex-differentiated	*mušterija* (m) 'customer' (when applied to males)	II	m	m	M
sex-differentiated	*sudija* (m) 'judge' (when applied to males)	II	m	m	M
male names in-*a*	*Steva* 'Steve'	II	m	m	M
collective (unsexed)	*deca* 'children'	II	f.sg.	nt.pl	PL
collective (male)	*gospoda* 'gentlemen'	II	f.sg	m.pl	M.PL
collective (male)	*braća* 'brothers'	II	f.sg	nt.pl	M.PL
female diminutives	*devojče* 'girl'	I	nt	nt	F
plur.tantum	*makaze* 'scissors' (when nonaggregate)	pl	pl	pl	sg

TABLE 3. Summary of Serbo-Croatian data
(nt = neuter, m = masculine, f = feminine, pl = plural, M = male, F = female, PL = aggregate reference)

Only contiguous cells in a row are connected by constraints, so we predict exactly the pattern observed. This is not insignificant. Taking just those minimally heterogenous cases with two types represented, there are seven logical possibilities a priori, that is, assuming no theory at all (see n. 3). Of those, our theory predicts three to be possible, and exactly those three and none of the others are found. Taking those highly irregular forms with three types represented in a single noun, we naturally expect those to be rare. We have found only one example, *braća* 'brothers'; it is one of the three patterns predicted to be possible, out of six logical possibilities.

Looking beyond Serbo-Croatian, we showed that our theory explains some of the typological generalizations noted by Corbett 1983 and 1991. Testing our theory on other languages is a task for future research.

APPENDIX: FORMALIZING THE ACCOUNT

1. THE SYNTAX OF CONCORD. In HPSG theory, concord and subject-verb agreement do not require a special module, but rather follow automatically from the normal combinatorial processes that combine a head with its dependents. We illustrate subject-verb agreement and NP-internal agreement in 66 by giving partial lexical specifications of the relevant words.

(66) Moja stara knjiga je pala.
my-F.NOM.SG old-F.NOM.SG book-NOM.SG AUX.3.SG fall-PPRT.F.SG
'My old book fell.'

Following Zlatić 1997, we assume that the possessive *moja* is categorically an adjective and functions as a specifier selected by the head noun, while other prenominal elements such as adjectives and determiners are adjuncts. Let us begin with the sign for the noun stem *knjig-* 'book'.

(67)
$$\left[\begin{array}{l} \textit{noun-stem} \\ \text{PHON /knjig-/} \\ \text{DECL II} \\ \text{XFORMS } \{\,\} \\ \text{SYNSEM} \left[\begin{array}{l} \text{CAT} \left[\begin{array}{l} \text{HEAD} \left[\begin{array}{l} \textit{noun} \\ \text{CONCORD } \boxed{3} \left[\begin{array}{l} \text{CASE } \textit{case} \\ \text{NUM } \boxed{1} \\ \text{GEN } \boxed{2}\textit{fem} \end{array}\right] \end{array}\right] \\ \text{VALENCE } \left[\text{SPR } \langle \text{AP}[\textit{poss}; \text{CONCORD } \boxed{3}]\rangle\right] \end{array}\right] \\ \text{CONTENT} \left[\begin{array}{l} \text{INDEX } \boxed{4} \left[\begin{array}{l} \text{PERS } \textit{3rd} \\ \text{NUM } \boxed{1} \\ \text{GEN } \boxed{2} \end{array}\right] \\ \text{RESTR } \left\{\left[\begin{array}{l} \text{REL } \textit{book} \\ \text{INST } \boxed{4} \\ \text{COUNT } \textit{count} \end{array}\right]\right\} \end{array}\right] \end{array}\right] \end{array}\right]$$

In HPSG a linguistic sign is an attribute-value matrix (AVM) comprising three main features, which encode phonological (PHON), syntactic (CATEGORY), and semantic (CONTENT) information.[22] Syntactic and semantic information are brought together in the SYNSEM field. The boxed numbers are tags indicating structure-sharing (token identity). Within the value for CATEGORY we have two features: the HEAD feature, whose value is automatically shared with all projections of this word (via the head feature principle); and the VALENCE feature, which specifies any dependents subcategorized by this word.

Inflected nouns belong to the sort noun-word, with the following constraint in 68.[23]

(68)
$$\textit{noun-word} \Rightarrow \left[\begin{array}{l} \text{PHON } \phi(\boxed{1}, \boxed{2}, \boxed{3}, \boxed{4}) \\ \text{SYNSEM } \boxed{5} \left[\text{CAT} \left[\text{HEAD} \left[\begin{array}{l} \textit{noun} \\ \text{CONCORD } \boxed{3} \end{array}\right]\right]\right] \\ \text{STEM} \left[\begin{array}{l} \textit{noun-stem} \\ \text{PHON } \boxed{1} \\ \text{DECL } \boxed{2} \\ \text{XFORMS } \boxed{4} \\ \text{SYNSEM } \boxed{5} \end{array}\right] \end{array}\right]$$

As indicated by the tag $\boxed{5}$, the SYNSEM value of an inflected noun is token-identical with that of its stem. Inflectional form is determined by the function ϕ, which takes four arguments: the stem phonology p, declension class d, CONCORD value c (which includes case, number, and gender; see n. 1 concerning animacy) and exceptional forms x. The output of this function is the inflected noun's phonological form. Here is a partial statement of the function:

$\phi\ (p, d, c, x) = p$ if $c =$ [*masc.sg.nom* or *masc-inan.sg.acc*] & $x = \{\,\}$
$= p$+/a/ if $d =$ I & [c $=$ *gen* or *masc-anim.sg.acc*] & $x = \{\,\}$
$= p$+/o~e/ if $c =$ *nt.sg.nom*/*acc* & $x = \{\,\}$
. . . etc.

See Table 1 for the full inflectional paradigm. As explained in n. 1, masculine accusatives are homophonous

[22] The values for features can be any of four types: atomic symbols (e.g. 3rd is the value for the attribute PER(son)); complex values (e.g. the value for CATEGORY is an AVM with two features, with attributes HEAD and VALENCE); a list of values, indicated by angle brackets (e.g. VALENCE features such as SPR (specifier) take list values; for SPR the list is of length one); or a set of values (e.g. the value for RESTRICTION is a unary set).

[23] Note that the tags do not carry over from one AVM to the next. For example, $\boxed{4}$ in 67 has no relation to $\boxed{4}$ in 68.

with genitive if animate, or nominative if inanimate. Concord elements such as determiners and adjectives share this split with the head noun they modify, so the animacy distinction must be represented in the CONCORD:GENDER feature. The gender values *masc-anim* and *masc-inan*, for animate and inanimate respectively, are subsorts of masculine.

The nominative singular form *knjiga* 'book' is obtained by substituting the sign for the noun stem /knjig-/ in 67 (with number resolved to singular) for the value of the feature STEM in 68, resulting in 69.

(69)
$$\left[\begin{array}{l}\text{PHON } \Phi(\boxed{1}\text{/knjig-/}, \boxed{2}\text{II}, \boxed{3}\textit{f.sg.nom}, \boxed{4}\{\ \}) = \text{/knjiga/} \\ \text{SYNSEM } \boxed{6}\left[\begin{array}{ll}\text{CAT} & \left[\begin{array}{ll}\text{HEAD} & \left[\begin{array}{ll}\textit{noun} & \\ \text{CONCORD } \boxed{3} & \left[\begin{array}{ll}\text{CASE } \textit{nom} \\ \text{NUM } \boxed{7}\textit{sing} \\ \text{GEN } \boxed{8}\textit{fem}\end{array}\right]\end{array}\right] \\ \text{VALENCE} & \left[\text{SPR } \langle \text{AP}[\textit{poss}; \text{CONCORD } \boxed{3}]\rangle\right]\end{array}\right] \\ \text{CONTENT} & \left[\begin{array}{ll}\text{INDEX } \boxed{6} & \left[\begin{array}{ll}\text{PERS } \textit{3rd} \\ \text{NUM } \boxed{7} \\ \text{GEN } \boxed{8}\end{array}\right] \\ \text{RESTR} & \left\{\left[\begin{array}{ll}\text{REL} & \textit{book} \\ \text{INST} & \boxed{6} \\ \text{COUNT} & \textit{one}\end{array}\right]\right\}\end{array}\right]\end{array}\right] \\ \text{STEM } \left[\begin{array}{ll}\text{PHON} & \boxed{1}\text{/knjig-/} \\ \text{DECL} & \boxed{2}\text{II} \\ \text{XFORMS} & \boxed{4}\{\ \} \\ \text{SYNSEM} & \boxed{5}\end{array}\right]\end{array}\right]$$

The agreement processes illustrated here are side-effects of normal valence saturation as enforced by the valence principle (Pollard & Sag 1994:348). Along with its specifier's category (namely AP[poss], that is, possessive adjective phrase), this noun indicates the CONCORD features of that AP. When the noun combines with its possessive AP these features unify with the corresponding features on the AP itself, and any clash in values leads to failure, hence ungrammaticality. In this example, the tags [7] and [8] indicate that the CONCORD and INDEX share GENder and NUMber values (namely feminine singular). The tag [3] indicates that specifier's CONCORD value is identical to the CONCORD value of the noun itself (nominative, feminine, singular).

An important (and often misunderstood) aspect of the HPSG lexicon is that all predictable information in a complex lexical sign such as 67—which turns out to be most of the information in it—is factored out and stated only once in the grammar, for all lexical signs to which it applies.[24] For example, probably every noun's CONCORD value is token-identical with the CONCORD value of its specifier, so this can be stated in a single place for all nouns. The lexical item is given in its long form here for illustrative purposes only. See §2 of this appendix.

As shown in the CONTENT field, the referential INDEX introduced by the noun *knjiga* is semantically constrained to be anchored to one book. As noted above, positing person, number, and gender features on the INDEX automatically explains agreement with bound pronouns.

Continuing our illustration of 66, the lexical entry of the possessive *moja* 'my' indicates three agreement features: case, gender, and number. These features will automatically be structure-shared with the head noun when this argument of the head noun saturates the head noun's SPR (specifier) subcategorization feature.

(70) *moja:*
$$\left[\begin{array}{ll}\text{PHON /moja/} \\ \text{SYNSEM } \left[\text{CAT } \left[\text{HEAD } \left[\begin{array}{ll}\textit{adj} & \\ \text{CONCORD} & \left[\begin{array}{ll}\text{CASE } \textit{nom} \\ \text{NUM } \textit{sg} \\ \text{GEN } \textit{fem}\end{array}\right]\end{array}\right]\right]\right]\end{array}\right]$$

[24] This is accomplished with an inheritance hierarchy. See Flickinger 1987.

The auxiliary *je* in 66 has no specification for its own person and number features, but rather, this information is encoded on the index of its SUBJect specification, ensuring that its subject will be third person singular.[25]

(71)

$$\textit{je:}\left[\begin{array}{l}\text{PHON } /\text{je}/\\ \text{SYNSEM}\left[\begin{array}{l}\text{CAT}\left[\begin{array}{l}\text{HEAD}\left[\begin{array}{l}\textit{verb}\\ \text{VFORM } \textit{finite}\\ \text{AUX } +\end{array}\right]\\ \text{VALENCE}\left[\begin{array}{l}\text{SUBJ } \langle \boxed{2}\,\text{NP}[\textit{nom}]_{3rd,sg}\rangle\\ \text{COMPS } \langle \text{VP}[\text{SUBJ } \langle \boxed{2}\rangle]:\boxed{1}\rangle\end{array}\right]\end{array}\right]\\ \text{CONTENT } \boxed{1}\end{array}\right]\end{array}\right]$$

Assuming that participles exhibit index agreement with their subjects, 72 is the sign for *pala* 'fall.PPRT.F.SG'

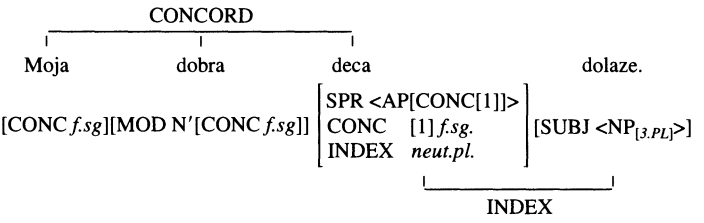
(72)

$$\textit{pala:}\left[\begin{array}{l}\text{PHON } /\text{pala}/\\ \text{SYNSEM}\left[\begin{array}{l}\text{CAT}\left[\begin{array}{l}\text{HEAD}\left[\begin{array}{l}\textit{verb}\\ \text{VFORM } \textit{past.part}\end{array}\right]\\ \text{VALENCE}\left[\text{SUBJ } \langle \text{NP}[\textit{nom}]_{\boxed{1}f.sg}\rangle\right]\end{array}\right]\\ \text{CONTENT}\left[\begin{array}{l}\text{REL } \textit{fall}\\ \text{FALLER } \boxed{1}\end{array}\right]\end{array}\right]\end{array}\right]$$

Once again, agreement between the verb and the subject is encoded directly in the verb's subcategorization frame.

We show below how agreement proceeds with nominative *deca*.

(73) Moja dobra deca dolaze.
my.F.SG good.F.SG children come.PAST.3.PL
'My good children came.'



Through the mechanisms explained above, feminine singular CONCORD values are determined on the determiner and adjective, while neuter plural INDEX value is determined on the finite verb.

2. THE STRUCTURE OF THE LEXICON OF NOUN STEMS. Noun stems are characterized by the constraint in 74.

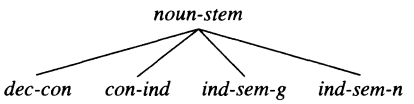
(74)

$$\textit{noun-stem} \Rightarrow \left[\begin{array}{l}\text{SYNSEM} \mid \text{CAT}\left[\begin{array}{l}\text{HEAD}\left[\begin{array}{l}\textit{noun}\\ \text{CONCORD } \boxed{1}\end{array}\right]\\ \text{VALENCE}\left[\text{SPR } \langle(\text{AP}[\textit{poss};\ \text{CONCORD } \boxed{1}])\rangle\right]\end{array}\right]\end{array}\right]$$

[25] In this framework, the auxiliary is treated on a par with raising verbs, i.e. verbs that do not assign a semantic role to their subjects. More precisely, the subject of an auxiliary has the entire SYNSEM value structure-shared with that of the subject of its 'unsaturated' VP complement. For simplicity the lexical sign for *je* covers only those uses with overt subjects.

This declarative constraint specifies that any object satisfying the description to the left of the double arrow must also satisfy the description to the right of the arrow. Hence in this instance any item belonging to the sort noun-stem must unify with the feature structure to the right of the arrow. The noun-stem sort has four subsorts, one for each of the constraints detailed in 15, 17, 18, and 22 above.

(75)



The four constraints discussed above are encoded (in 76–79) as constraints on each of the four subsorts.

(76) $\textit{dec-con} \Rightarrow \begin{bmatrix}\text{DECL I} \\ \text{CONCD} \mid \text{GEN } m \vee n\end{bmatrix} \vee \begin{bmatrix}\text{DECL II} \vee \text{III} \\ \text{CONCD} \mid \text{GEN } f\end{bmatrix}$

(77)

$$\textit{con-ind} \Rightarrow \begin{bmatrix} \text{CONCORD} & \begin{bmatrix}\text{NUM } \boxed{1} \\ \text{GEN } \boxed{2} \\ \text{CASE}\end{bmatrix} \\ \text{INDEX} & \begin{bmatrix}\text{NUM } \boxed{1} \\ \text{GEN } \boxed{2} \\ \text{PERS}\end{bmatrix} \end{bmatrix}$$

(78) *ind-sem-g* $\Rightarrow$

$$\left[\left[\text{RESTR } \left\{\left[\textit{female}\left(\boxed{1}\right)\right],\ldots\right\}\right] \Rightarrow \left[\text{INDEX } \boxed{1}\textit{fem}\right]\right] \wedge$$

$$\left[\left[\text{RESTR } \left\{\left[\textit{male}\left(\boxed{1}\right)\right],\ldots\right\}\right] \Rightarrow \left[\text{INDEX } \boxed{1}\textit{masc}\right]\right]$$

(79) *ind-sem-n* $\Rightarrow$

$$\begin{bmatrix}\text{INDEX } \boxed{2}\textit{pl} \\ \text{RESTR } \left\langle\left[\begin{matrix}\text{INST } \boxed{2} \\ \text{COUNT } \textit{aggr}\end{matrix}\right],\ldots\right\rangle\end{bmatrix} \vee \begin{bmatrix}\text{INDEX } \boxed{1}\textit{sg} \\ \text{RESTR} \neg \left\langle\left[\begin{matrix}\text{INST } \boxed{1} \\ \text{COUNT } \textit{aggr}\end{matrix}\right],\ldots\right\rangle\end{bmatrix}$$

Hence to specify that a particular noun stem follows the declension-concord constraint we stipulate that it belongs to sort dec-con; and so on for the other constraints. (In addition, the noun stem will inherit the specification of *noun-stem* given above, and that of any other supersorts.) Following Malouf 1998, subsorts form a cover of their supersort but do not partition it, so a given noun stem can belong to more than one of the four subsorts given above. A fully regular noun stem belongs to all four.

To give our account of Serbo-Croatian predictive power, we need to disallow nouns violating more than one constraint. (The one case we have found of a noun violating two constraints, *braća* 'brothers', will be relegated to a special class of pathologically irregular nouns.) A hypothetical noun violating more than one constraint could have number or gender values that accidentally correlate in a way that we wish to rule out in principle. For example, a hypothetical noun violating both *con-ind* and *ind-sem-g* could be specified for feminine concord and female sex (semantics), but masculine index. So we will stipulate that all nouns (except the radically irregular *braća*) must observe at least three of the four constraints. For convenience combinations of subsorts are abbreviated as follows:

$$\begin{aligned}
\textit{REG} &= \textit{dec-con} \wedge \textit{con-ind} \wedge \textit{ind-sem-n} \wedge \textit{ind-sem-g} \\
\textit{NoDC} &= \textit{con-ind} \wedge \textit{ind-sem-n} \wedge \textit{ind-sem-g} \\
\textit{NoCI} &= \textit{dec-con} \wedge \textit{ind-sem-n} \wedge \textit{ind-sem-g} \\
\textit{NoISn} &= \textit{dec-con} \wedge \textit{con-ind} \wedge \textit{ind-sem-g} \\
\textit{NoISg} &= \textit{dec-con} \wedge \textit{con-ind} \wedge \textit{ind-sem-n}
\end{aligned}$$

(*REG* = regular; *NoDC* = no Declension-Concord constraint; etc.) Each common noun's lexical sign specifies *REG*, *NoDC*, *NoCI*, *NoISn*, or *NoISg*. In 80a we see the regular noun stem /muž-/ 'husband'. The sort name *REG* means that it abides by all four constraints; it is a count noun, so it has the COUNT feature. Since both singular and plural nouns are formed on this stem, it is underspecified for the COUNT feature. Spelling out the effects of the *REG* sort, we arrive at the representations in 80b and c, corresponding to the stems for singular and plural nouns, respectively. In this case all agreement and declension features are deduced from the semantics of the noun.

(80) a.
$$\begin{bmatrix} \textit{REG} \\ \text{PHON /muž-/} \\ \text{CONTENT|RESTR } \{\textit{husband(x)}, \textit{male(x)}, [\text{COUNT } \textit{count}]\} \end{bmatrix}$$

b.
$$\begin{bmatrix} \textit{REG} \\ \text{PHON} \quad \text{/muž/} \\ \text{DECL} \quad \text{I} \\ \text{CONCORD } \textit{m.sg} \\ \text{INDEX} \quad \textit{3rd.m.sg} \\ \text{RESTR } \{\textit{husband(x)}, \textit{male(x)}, [\text{COUNT } \textit{one}]\} \end{bmatrix}$$

c.
$$\begin{bmatrix} \textit{REG} \\ \text{PHON} \quad \text{/muž/} \\ \text{DECL} \quad \text{I} \\ \text{CONCORD } \textit{m.pl} \\ \text{INDEX} \quad \textit{3rd.m.pl} \\ \text{RESTR } \{\textit{husband(x)}, \textit{male(x)}, [\text{COUNT } \textit{aggregate}]\} \end{bmatrix}$$

For the regular noun stem /knjig-/ 'book' in 67, declension needs to be lexically stipulated, as in 81.

(81)
$$\begin{bmatrix} \textit{REG} \\ \text{PHON /knjig-/} \\ \text{DECL II} \\ \text{CONTENT | RESTR } \{\textit{book(x)}, [\text{COUNT } \textit{count}]\} \end{bmatrix}$$

The rest of the information in 67 is derived. Turning to irregular noun stems, *dec-* 'children' in 82a is specified for NoCI, singular CONCORD and neuter INDEX. The other agreement and declension features in 82b are deduced.

(82) a.
$$\begin{bmatrix} \textit{NoCI} \\ \text{PHON} \quad \text{/dec-/} \\ \text{DECL} \quad \text{II} \\ \text{CONCORD } \textit{sg} \\ \text{INDEX} \quad \textit{nt} \\ \text{CONTENT|RESTR } \{\textit{child(x)}[\text{COUNT } \textit{aggregate}]\} \end{bmatrix}$$

b.
$$\begin{bmatrix} \textit{NoCI} \\ \text{PHON} \quad \text{/dec-/} \\ \text{DECL} \quad \text{II} \\ \text{CONCORD } \textit{f.sg} \\ \text{INDEX} \quad \textit{nt.pl} \\ \text{CONTENT|RESTR } \{\textit{child(x)}[\text{COUNT } \textit{aggregate}]\} \end{bmatrix}$$

REFERENCES

ASHTON, E. O. 1944. Swahili grammar (including intonation). London: Longman.
BARLOW, MICHAEL. 1992. A situated theory of agreement. New York: Garland Publishing.
BARWISE, JON, and JOHN PERRY. 1983. Situations and attitudes. Cambridge, MA: MIT Press.

BAYER, SAMUEL, and ANTHONY DAVIS. 1991. Agreeing with scissors and silverware; Disagreeing with Dowty and Jacobson, MS.
BEARD, ROBERT. 1995. The gender-animacy hypothesis. Journal of Slavic Linguistics 3.59–96.
——. 1997. How many declensions? Proceedings of the sixth conference on formal approaches to Slavic linguistics. Bloomington, IN: Indiana University Linguistics Club.
BIDWELL, CHARLES E. 1966. An outline of Serbo-Croatian nominal morphology. Washington, DC: U.S. Department of Health, Education, and Welfare Educational Resources Information Center.
BOSCH, PETER. 1987. Pronouns under control? A reply to Liliane Tasmowski and Paul Verluyten. Journal of Semantics 5.65–78.
——. 1988. Representing and accessing focussed referents. Language and Cognitive Processes 3.207–31.
——. 1989. Coherence and cohesion: Comments on Roger G. van de Velde's paper "Man, verbal text, inferencing, and coherence". Connexity and coherence—Analysis of text and discourse, ed. by W. Heydrich, F. Neubauer, J. S. Petofi and E. Sozer, 218–27. Berlin: Walter de Gruyter.
BRESNAN, JOAN, and SAM A. MCHOMBO. 1987. Topic, pronoun, and agreement in Chicheŵa. Language 63.741–82.
BROWNE, WAYLES. 1986. Relative clauses in Serbo-Croatian in comparison with English. The Yugoslav SerboCroatian-English contrastive project, new study series, vol. 4, ed. by Rudolf Filipović. Zagreb, Yugoslavia: Institute of Linguistics, Faculty of Philosophy, University of Zagreb.
CARSTAIRS, ANDREW. 1983. Paradigm economy. Journal of Linguistics 19.115–28.
CORBETT, GREVILLE. 1979. The agreement hierarchy. Journal of Linguistics 15.203–24.
——. 1983. Hierarchies, targets and controllers: Agreement patterns in Slavic. London: Croom Helm.
——. 1988. Agreement: A partial specification based on Slavonic data. Agreement in natural languages, ed. by C. Ferguson and M. Barlow, 23–43. Stanford: CSLI Publications.
——. 1991. Gender. Cambridge: Cambridge University Press.
——. 1998. Agreement in Slavic. Position paper for the conference on comparative Slavic morphosyntax, Bloomington, Indiana, 5–7 June, 1998.
CORNISH, FRANCIS. 1987. Anaphoric pronouns: under linguistic control or signalling particular discourse representations? A contribution to the debate between Peter Bosch, and Liliane Tasmowski and Paul Verluyten. Journal of Semantics 5.233–60.
——. 1994. Agreement and discourse: From cohesion to coherence. French Language Studies 4.191–213.
——. 1999. Anaphora, discourse, and understanding—evidence from English and French. Oxford: Oxford University Press.
CROCKETT, DINA. 1976. Agreement in contemporary standard Russian. Columbus, OH: Slavica.
DALRYMPLE, MARY, and RONALD M. KAPLAN. 2000. Feature indeterminacy and feature resolution in description-based syntax. Language 76.759–98.
DEVLIN, KEITH. 1991. Logic and information. Cambridge: Cambridge University Press.
DIXON, R.M.W. 1994. Ergativity. Cambridge: Cambridge University Press.
DOWTY, DAVID, and PAULINE JACOBSON. 1988. Agreement as a semantic phenomenon. Paper presented at the Eastern States Conference on Linguistics, Ohio State University, Columbus, Ohio.
FARKAS, DONKA F., and DRAGA ZEC. 1993. Agreement and pronominal reference. Santa Cruz, CA: Linguistic Research Center, University of California.
FLICKINGER, DANIEL. 1987. Lexical rules in the hierarchical lexicon. Stanford, CA: Stanford University dissertation.
FRANKS, STEVEN. 1995. Parameters of Slavic morphosyntax. New York: Oxford University Press.
GREENBERG, JOSEPH. 1963. Some universals of grammar with particular reference to the order of meaningful elements. Universals of language, ed. by J. Greenberg, 73–113. Cambridge, MA: MIT Press.

HALLIDAY, M.A.K. 1985. An introduction to functional grammar. London: E. Arnold.
HRASTE, M. 1953. Moje kolege ili moji kolege? [My (feminine) colleague or my (masculine) colleague?]. Jezik 2.61–62.
IVIĆ, MILKA. 1960. Obeležavanje imeničkog roda u srpskohrvatskom jeziku [Marking gender of nouns in the Serbo-Croatian language]. Naš jezik (Beograd) n.s. 10.7–10, 192–211.
——. 1963. Relationship of gender and number in Serbocroatian substantives. International Journal of Slavic Linguistics and Poetics 6.51–57.
——. 1966. Tipovi imeničkog roda u srpskohrvatskom književnom jeziku [Types of gender of nouns in Serbo-Croatian literary language]. Zbornik za filologiju i lingvistiku (Novi Sad) 9.41–47.
KAMP, HANS, and UWE REYLE. 1993. From discourse to logic—introduction to modeltheoretic semantics of natural language, formal logic, and discourse representation theory. Dordrecht: Kluwer.
KATHOL, ANDREAS. 1999. Agreement and the syntax-morphology interface in HPSG. Studies in contemporary phrase structure grammar, ed. by R. Levine and G. Green, 223–74. New York: Cambridge University Press.
LEHMANN, CHRISTIAN. 1982. Universal and typological aspects of agreement. Das sprachliche Erfassen von Gegenständen II: die Techniken und ihre Zusammenhang in Einzelsprachen, ed. by H. Seiler and F. J. Stachowiak, 201–67. Tübingen: Narr.
MALOUF, ROBERT. 1998. Mixed categories in the hierarchical lexicon. Stanford, CA: Stanford University dissertation.
MARETIĆ, TOMA. 1899. Gramatika i stilistika hrvatskoga ili srpskoga književnog jezika [The grammar and stylistics of Croatian or Serbian literary language]. Zagreb: Hartman.
MOOSALLY, MICHELLE. 1998. Noun phrase coordination: Ndebele agreement patterns and cross-linguistic variation. Austin, TX: University of Texas dissertation.
MRAZOVIĆ, PAVICA, and ZORA VUKADINOVIĆ. 1990. Gramatika srpsko-hrvatskog jezika za strance [The grammar of Serbo-Croatian language for foreigners]. Novi Sad: Dobra Vest.
POLLARD, CARL, and IVAN SAG. 1992. Anaphors in English and the scope of binding theory. Linguistic Inquiry 23.261–303.
——, ——. 1994. Head-driven phrase structure grammar. Stanford and Chicago: CSLI Publications and University of Chicago Press.
SAG, IVAN A. 1997. English relative clause constructions. Journal of Linguistics 33.431–83.
——, and THOMAS WASOW. 1999. Syntactic theory: A formal introduction. Stanford: CSLI Publications.
SHIEBER, STUART. 1986. An introduction to unification-based approaches to grammar. CSLI lecture notes no. 4. Stanford: CSLI Publications.
STEVANOVIĆ, MIHAILO. 1974. Savremeni srpskohrvatski jezik, 2, sintaksa [The contemporary Serbo-Croatian language, 2, syntax]. Beograd: Naučna Knjiga.
TASMOWSKI, LILIANE, and PAUL VERLUYTEN. 1985. Control mechanisms of anaphora. Journal of Semantics 4.341–70.
TASMOWSKI-DERIJK, LILIANE, and PAUL VERLUYTEN. 1982. Linguistic control of pronouns. Journal of Semantics 1.323–46.
WECHSLER, STEPHEN. 1999a. Four agreement puzzles. Paper read at the sixth international conference on head-driven phrase structure grammar, Edinburgh, Scotland.
——. 1999b. Gender resolution in coordinate structures. Proceedings of the workshop on the structure of spoken and written texts, ed. by Carlota Smith, 1–22. Austin, TX: University of Texas Linguistics Department.
——, and LARISA ZLATIĆ. 1998. Agreement in discourse. Proceedings of the conference on the structure of non-narrative texts, University of Texas, Austin, February 13, 1998.
——, ——. 1999. Syntax and morphological realization in Serbo-Croatian. Slavic in HPSG, ed. by R. Borsley and A. Przepiorkowski, 283–309. Stanford: CSLI Publications.

ZLATIĆ, LARISA. 1997. The structure of the Serbian noun phrase. Austin, TX: University of Texas dissertation.

Wechsler
Dept. of Linguistics
University of Texas
Austin, TX 78712-1196
[wechsler@mail.utexas.edu]
[lzlatic@alumni.utexas.net]

[Received 4 March 1997;
revision received 15 July 1999;
accepted 20 August 1999]