

A Theory of Agreement and Disagreement

Author(s): Stephen Wechsler and Larisa Zlatić

Proceedings of the Twenty-Fourth Annual Meeting of the Berkeley Linguistics Society: General Session and Parasession on Phonetics and Phonological Universals (1998), pp. 280-291

Please see “How to cite” in the online sidebar for full citation information.

Please contact BLS regarding any further use of this work. BLS retains copyright for both print and screen forms of the publication. BLS may be contacted via <http://linguistics.berkeley.edu/bls/>.

The Annual Proceedings of the Berkeley Linguistics Society is published online via [eLanguage](#), the Linguistic Society of America's digital publishing platform.

A Theory of Agreement and Disagreement

Stephen Wechsler and Larisa Zlatić
University of Texas at Austin

1. Introduction. A striking fact emerging from studies of agreement is that the expected agreement relations often break down, resulting in 'mixed' or 'hybrid' agreement, or, as we call it, *disagreement*. For example, certain common nouns trigger feminine agreement on modifiers but masculine on predicates and pronouns; or a noun may trigger masculine agreement, but is itself inflected in the pattern normally reserved for feminines. In this paper we propose a theory of agreement which explains this phenomenon of disagreement and predicts which classes of disagreement are possible, and which are impossible. We apply this theory to Serbo-Croatian (SC), and also show briefly that this theory predicts the main typological generalizations due to Greville Corbett (1983; 1991) and confirmed by further studies such as Barlow (1988). Our descriptive framework is Head-Driven Phrase Structure Grammar (HPSG, Pollard and Sag, 1994).

2. Two types of agreement: CONCORD and INDEX. Regular agreement in SC is illustrated in (1):

(1)	Ova	stara	knjiga	stalno	pada.
	this.F.SG	old.F.SG	book(fem.)	always	falls.3.SG
	Molim	vas	podignite	je.	
	please	you	pick.2pl	it.F.SG	

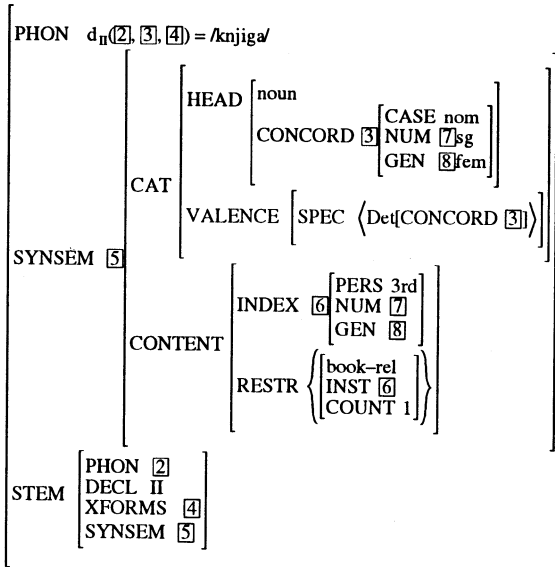
'This old book keeps falling. Please pick it up.'

The noun *knjiga* 'book' triggers feminine singular (f.sg) agreement on determiners (*ova*) and modifiers (*stara*), third person singular (3.sg) on finite predicates (*pada*), and third person feminine singular (3.f.sg) on coreferential pronouns (*je*). Participles also show gender and number agreement (see Section 7 below).

Following the basic lead of Kathol (to appear), Pollard and Sag (1994), and, in different variants, numerous other researchers, we distinguish two agreement features on the noun, called CONCORD (Kathol's AGR) and INDEX. CONCORD is a head feature, while INDEX is part of the noun's semantic content, as illustrated by the HPSG lexical sign for *knjiga* 'book.nom.sg' in (2). The CONCORD feature effects determiner and modifier agreement as a side-effect of functor-argument combination: the CONCORD value of the noun unifies with that of its specifier, as shown in (2); similarly, an attributive adjective specifies the features of the N' it modifies. INDEX agreement, on the other hand, arises through the sharing of referential indices between pronoun and antecedent. As shown in (2), the index contains person, number, and gender (PNG) features. Thus the matching condition on pronouns is a side-effect of the coreference ('coindexation') mechanism (see Pollard and Sag, 1994). The PNG features are grammaticalizations of referential anchoring conditions— PERSON: speaker → *1st*, addressee → *2nd*, other → *3rd*; NUMBER: aggregate → *pl*, other → *sg*; GENDER: male → *masc*, female → *fem*, other → *neut* — but cannot be reduced

to semantics (cp. grammatical gender, pluralia tantum, etc.). In addition, a finite predicate specifies the INDEX features of its nominative subject, enforcing agreement as a side-effect of functor-argument combination.

(2) Lexical sign for the feminine noun *knjiga* 'book.nom.sg':



We now motivate this split between concord and index, first theoretically and then empirically.

First, the feature sets differ: *person*, number, and gender for index but *case*, number, and gender for concord. Second, the syntactic domains differ. As explained above NP's and pronouns have referential indices, hence index agreement. Verb agreement, which is derived historically from incorporated pronouns, also involves index agreement. Concord agreement is typically (but not necessarily) NP-internal. Unlike index agreement, it includes non-referential elements such as determiners and adjectives.

Third, and most relevant here, the relation to other lexical properties differs for index and concord respectively. As we will show, INDEX features are closely related to semantics, while CONCORD features are closely related to phonology (declension class). In what follows, we will flesh out these index-semantics and concord-phonology relations. Then we turn to those exceptional noun classes where these relations break down.

3. SC gender and number constraints.

3.1. Phonology-concord constraints. The concord-phonology relation is explicated via the notion of declension class. We define a *declension* as a function d from stem phonology Φ_{stem} , CONCORD value C , and exceptional forms X (and animacy A , for d_1), to the inflected noun's phonology Φ_{noun} .

Following traditional descriptions, SC has three declensions, which we designate d_I , d_{II} , and d_{III} . Here is a partial statement of the functions:

$d_I(\psi, c, x, a) = \psi$ if $c = m.sg.nom$ or $[c=m.sg.acc \ \& \ a=-anim] \ \& \ x = \{ \}$
 $d_I(\psi, c, x, a) = \psi + /a/$ if $c = m.sg.gen$ or $[c=m.sg.acc \ \& \ a=+anim] \ \& \ x = \{ \}$
 $d_I(\psi, c, x) = \psi + /o-e/$ if $c = nt.sg.nom \vee acc \ \& \ x = \{ \}$
 ...etc.

Sample output forms are shown in this Table I at the end of the paper. There are other ways to set up the declension classes. The crucial distinction between declension and gender is that syntactic processes refer to gender but not declension; hence in (2) we make gender but not declension a head feature.

Concord and phonology are systematically related since concord *number* is an argument of all Declension functions, while concord gender is an argument of Declension class I. Moreover concord gender strongly influences assignment to declension class. As illustrated in Table I, declension class I nouns are either masculine or neuter, while Class II and III nouns are feminine. We capture this rule by positing a subsort of *noun-sign* called *dec=con* and formulating the following constraint:

$$(3) \text{ } dec=con: \begin{bmatrix} PHON & d_I \\ CONCD & m \vee n \end{bmatrix} \vee \begin{bmatrix} PHON & d_{II} \vee d_{III} \\ CONCD & f \end{bmatrix}$$

Most nouns belong to the sort *dec=con*; but some exceptional ones discussed below do not.

Evidence that (3) is part of SC competency grammar comes from *inquorate* nouns (Corbett, 1991:173), nouns with different gender in singular and plural. Interestingly, declension class and gender follow *dec=con* 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}). Note that the singular, being neuter, is declined with the Class I pattern, while the plural, being feminine, takes the Class III declension. Another example is *mač-e* ('kitten' nt., d_I) ~ *mačić-i* ('kittens', m., d_I). Certain collective nouns like *deca* 'children' (see Section 6 below) are also inquorates:

	<u><i>det-e</i></u> ('child')	<u><i>dec-a</i></u> ('children')
DECLENSION:	d_I	d_{II}
CONCORD:	nt.sg	f.sg
INDEX:	nt.sg	nt.pl

Once again, declension and concord covary in lockstep across the singular/plural distinction, following the constraint in (3): *deca* 'children' has feminine singular concord and is declined in class II. Moreover, the index features do not vary, as we demonstrate below (Section 6). This pattern, where concord but not index covaries with declension, is consistent across the SC language.

Note also the phonologically defined subclass of inquorates comprising /o/-final loans such as *auto*, *radio*, *kino*, and *torzo* (masculine in the singular, neuter in the plural). Summarizing, our basic observations are (i) the

Concord(gender)-Declension rule (3) is observed even in the face of irregularity across singular/ plural; and (ii) Concord (gender) irregularity is related to phonology.

3.2. Concord-index constraints. In regular nouns such as *knjiga* 'book' in (2) above, the number and gender values of the concord match those of the index. We posit that such nouns belong to a subsort of *noun-sign* called *con=ind* and formulate the following constraint:

$$(4) \text{ con=ind: } \left[\begin{array}{l} \text{CONCORD } \left[\begin{array}{l} \text{NUM } \boxed{1} \\ \text{GEN } \boxed{2} \\ \text{CASE} \end{array} \right] \\ \text{INDEX } \left[\begin{array}{l} \text{NUM } \boxed{1} \\ \text{GEN } \boxed{2} \\ \text{PERS} \end{array} \right] \end{array} \right]$$

Note that the boxed numerals ('tags') indicate structure sharing, each numeral representing a single node in a directed acyclic graph. Hence a noun belonging to this class has only one number value and one gender value (not two of each). Each of these values can be 'accessed' in two ways, via concord and index.

3.3. Index-semantics constraints. There are two constraints relating index to semantics, one for number and one for gender. Taking number first, in the normal case count nouns with aggregate reference are plural, while other nouns (mass and non-aggregate count) are singular. We formulate this with a feature COUNT declared only by count nouns, whose value encodes the cardinality of the reference. Here we need only consider two COUNT values, '1' (cardinality one) and '>1' (cardinality greater than one). The constraint is given in (5), where *ind=sem_N* is yet another subsort of *noun-sign*:

$$(5) \text{ ind=sem}_N: \left[\begin{array}{l} \text{INDEX } \boxed{2}\text{pl} \\ \text{RESTR } \left\{ \left[\begin{array}{l} \text{INST } \boxed{2} \\ \text{COUNT } >1 \end{array} \right] \dots \right\} \end{array} \right] \vee \left[\begin{array}{l} \text{INDEX } \boxed{1}\text{sg} \\ \text{RESTR } \neg \left\{ \left[\begin{array}{l} \text{INST } \boxed{1} \\ \text{COUNT } >1 \end{array} \right] \dots \right\} \end{array} \right]$$

A BLS audience member raised the problem of applying the notion of (non-) aggregate reference to nouns like *flock*, *committee*, etc. It is important to distinguish *reference* (a relation between language and an item in the world) from the *referent* itself (an item in the world). For us a noun with 'aggregate reference' is one such that its referent consists of more than one item of the sort described by that noun's stem: *apples* denotes more than one apple so it has aggregate reference. But *flock* does not denote more than one flock, so it has non-aggregate reference. Pluralia tantum nouns like *scissors* clearly also lack aggregate reference. It is this notion of aggregate reference that the COUNT feature encodes.

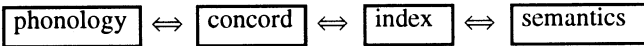
Turning to gender, SC nouns restricted to female (/male) referents have feminine (/masculine) index, a constraint apparently valid for Slavic generally. Hence nouns denoting 'man', 'husband', and 'bull' are masculine, while 'woman', 'wife', and 'hen' are feminine. This is formulated here:

(6) $ind = sem_G$:

$$\left(\left[\text{RESTR } \{ [female(\mathbb{I})], \dots \} \Rightarrow [\text{INDEX } \mathbb{I} fem] \right] \wedge \left(\left[\text{RESTR } \{ [male(\mathbb{I})], \dots \} \Rightarrow [\text{INDEX } \mathbb{I} masc] \right] \right) \right)$$

It is important to note that this constraint is ‘vacuously satisfied’ by any noun which is not sex-specific, either because it has no sex at all (‘book’, ‘table’, etc.) or because it is unspecific (‘child’, ‘human’, ‘president’, etc.)

Collecting the constraints in (3), (4), (5), and (6), we arrive at a ‘chain’ of dependencies between features of a common noun, where ‘ \Leftrightarrow ’ represents one or more constraints.



This chain can be broken in any of the three places depending on which constraint is violated. Technically the constraints are inviolable, but we capture such apparent violations through the system of sorts: for example, a noun violating the concord-index constraint simply does not belong to the subsort $con=ind$, and so on for the others. In the remainder of the paper we illustrate these three types of violations.

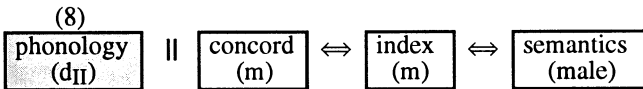
4. Phonology-concord mismatches. As our first example of a mismatch between phonology (i.e. declension) and concord, certain male names in /-a/ such as *Steva* and *Mika* trigger masculine agreement but nonetheless belong to declension II.

(7) Vratio mi je ovaj ludi
 returned-1SG me AUX.3SG this-NOM.M.SG crazy-NOM.M.SG

Steva violinu koju sam mu pozajmio.
 Steve(m.sg).NOM violin-ACC which AUX him-DAT.M.SG loaned

‘This(M) crazy(M) Steve returned to me the violin which I loaned him(M).’

Hence these names violate constraint (3). We show this situation schematically as follows:



Also certain nouns such as *sudija* ‘judge’, *sluga* ‘slave’, *gazda* ‘master’, *mušterija* ‘customer’, *kolega* ‘colleague’, etc. trigger feminine or masculine agreement depending on whether they refer to females or males:

- (9) a. Taj stari sudija_i je dobro sudio. On_i...
 that.M old.M judge AUX well judged.M he.M
 'That old (male) judge judged well.'
- b. Ta stara sudija_i je dobro sudila. Ona_i...
 that.F old.F judge AUX well judged.F she.F
 'That old (female) judge judged well.'

Since *sudija* belongs to declension class II, it violates (3) when referring to a male judge, as in (9)a.

How are we to capture the covariation between gender and sex shown in (9)a,b? The problem with assuming 'semantic agreement', where the sex (semantics) directly determines agreement, is that the syntactic mechanism of agreement would not be uniform: clearly sex does not determine agreement in (1) above, for example. Our constraints explain this covariation if we assume that *sudija* is sex-specific but disjunctive, as shown by the RESTRICTIONS value in (12)a below: it is either restricted to being male or restricted to being female. The agreement facts follow from a chain of deductions. Taking the male disjunct, by the constraint $ind=sem_G$ (see (6) above), $male(x)$ entails [INDEX masc]; hence, by $con=ind$ (see (4)), we also get [CONCORD masc]. (Constraint $dec=con$ is violated, as noted). Mutatis mutandis for the female disjunct. In contrast, a noun like *pas* 'dog' is grammatically masculine and has no sex specification at all:

- (10) taj verni pas
 that.M loyal.M dog
 'that loyal (male or female) dog'

The noun *mušterija* 'customer' has an optional disjunctive specification, hence it is midway between 'judge' and 'dog'. It exhibits dog-like behavior in (11)a (grammatically feminine but unspecified for sex) but is judge-like in (11)b (male and masculine). The female/feminine variant of (11)b overlaps with (11)a.

- (11) 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.'

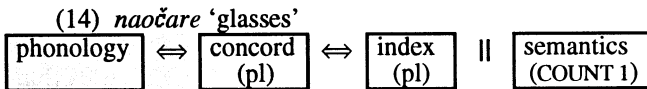
We summarize the three types as follows:

- (12) a. *sudij*-. {judge(x)} ∪ {male(x) ∨ female(x)}
 b. *pas*-. {dog(x)}
 c. *mušterij*-. {customer(x)} ∪ {male(x) ∨ female(x)}

5. Index-semantics mismatches. Pluralia tantum nouns, when used in the non-aggregate sense (hence specified [COUNT 1]), violate constraint (5) *ind=sem_N*. Examples include *kola* ‘car’, *vrata* ‘door’, *ledja* ‘back’, *makaze* ‘scissors’, *naočare* ‘glasses’, and *novine* ‘newspaper’.

- (13) Ove naočare_i su nove. One_i ...
 this.PL glasses are.3PL new-PL 3PL...
 ‘These glasses_i are new. They_i...’

This example is ambiguous, as in English, referring to one pair of glasses or more than one pair. On the former reading it violates (5), as shown here schematically:



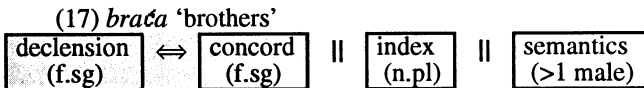
Violations of (6) *ind=sem_G* are rare, but *braća* ‘brothers’ is one case. While male-only, it has a nt. pl index:

- (15) ... braća_i ... Ona_i su došla
 ... brothers ... they.N.PL AUX.3.PL came.N.PL

Also, adjectives and determiners show f.sg:

- (16) ovu dobru braću
 this-ACC.F.SG good-ACC.F.SG brothers-ACC

Hence *braća* violates two constraints:



6. Index-concord mismatches. Collective nouns like *deca* ‘children’, *braća* ‘brothers’, and *gospoda* ‘gentlemen’ violate (4) *con=ind* (see Corbett, 1983). Such nouns trigger feminine singular on modifiers, plural on finite verbs and auxiliaries, and neuter plural on pronouns:

- (18) Posmatrali smo ovu dobru decu_i.
 watched.1.PL AUX this.F.SG good.F.SG children.ACC(f.sg[nt.pl])

Ona_i su spavala.
 they-N.PL AUX.3.PL slept-NT.PL

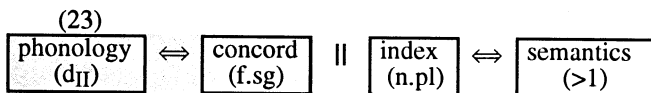
‘We watched those good children_i. They_i slept.’

- (19) ...deca_i... Mi smo ih_i videli.
 ...children... we Aux-PL them-ACC.PL saw
 '...children_i... We saw them_i.'
- (20) Ta dobra deca su došla.
 that.F.SG good.F.SG children(F.SG) Aux-PL came-PPRT.F.SG/N.PL
 'Those good children came.'
- (21) Ta dobra deca dolaze.
 that.F.SG good.F.SG children(F.SG) come.3.PL
 'Those good children came.'

We posit the following CONCORD and INDEX gender/number features for *deca*:

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

This situation is shown schematically as follows:



Note that the concord and phonology are still mutually consistent, as are index and semantics. Regarding the former, these nouns are inflected in the traditional second declension class, the norm for feminine (concord) nouns. Regarding the latter, *deca* 'children' is an aggregate and has a plural index, so it satisfies $ind=sem_N$, and it is not sex-specific so it vacuously satisfies $ind=sem_G$.

Relative pronouns show an interesting mixture of sensitivity to both concord and index features (Corbett, 1983). The nominative relative pronoun is *koja*, the nt.pl form.

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

But an accusative relative pronoun takes a f.sg form, and oblique cases can take either f.sg (*koje*) or plural (*kojih*) form.

- (25) deca koju sam video
 children who-ACC.F.SG Aux-1.SG saw
 'the children whom I saw'
- (26) deca koje/kojih se svi plače
 children who-GEN.F.SG./GEN.PL REFL all fear
 'children whom everyone fears'

We analyze these facts as follows. Relative pronouns, being heads of N'-modifiers, show CONCORD agreement (like other modifiers); being bound pronouns, they also show INDEX agreement. But the relative pronouns are not

specified for all features: most of the plurals do not mark (CONCORD) gender, and nominative and accusative relative pronouns are unmarked for (INDEX) PNG features. Details are omitted for lack of space (see Wechsler and Zlatic, 1997).

Relative clauses introduced by the complementizer *što* employ ordinary pronoun forms functioning as 'resumptive pronouns'. Since they are ordinary pronouns they show index agreement, irrespective of case.

(27) moja deca, što ih/*je svi hvale.
 my children, that 3.ACC.PL/*SG all praise
 'my children, whom everyone praises'

(28) moja deca, što ih/*je se svi plaće.
 my children, that 3.GEN.PL REFL all fear
 'my children, whom everyone fears'

7. Participles and predicate adjectives. Participles and predicate adjectives agree in gender and number with a nominative subject. But do they show CONCORD or INDEX agreement? That is, do they pattern with modifiers or with pronouns, when those differ? There is at least some evidence for INDEX agreement. The plural form *vladike* 'bishops' belongs to a group triggering feminine concord (although bishops are male). In this dialectal example the participle shows masculine rather than feminine agreement:

(29) Jekavski dialect (cited in Marković, 1954)
 ove dobre vladike su došli.
 these-F.PL good-F.PL bishops-F Aux came-M.PL
 'These good (male) bishops came.'

Coordination resolution also provides some evidence for index agreement. When items of different genders are coordinated, the coordinate structure triggers masculine plural agreement; when like gender is coordinated, the gender of the conjuncts is triggered. Coordinating *deca* with the female name *Jelena* in (30) yields masculine plural agreement, indicating mixed gender (the subscript in the gloss indicates index value; the non-subscripted item is concord value). Concord values on the conjuncts are both feminine singular, while index values are mixed (f. and nt.), so this means that the index value must be the relevant one. (This argument is due to Corbett 1983). Secondly, when *deca* is conjoined with the neuter plural NP *ta čudovišta* 'those monsters', neuter plural is favored, again suggesting index rather than concord agreement (31).

(30) Jelena i deca su došli/*došle.
 Yelena(F.SG[F.SG]) & children(F.SG[N.PL]) Aux.PL came.M.PL/*F.PL
 'Yelena and the children came.'

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

Appendix.

	Class I		Class II	Class III
	'window' (m)	'village' (nt)	'woman' (f)	'thing' (f)
Singular				
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 I. Serbo-Croatian Declension Classes

example	DECL	CONCD	INDEX	SEM
<i>knjiga</i> 'book'	II(f)	f	f	-
<i>rad</i> 'work'	I(m)	m	m	-
<i>žena</i> 'woman'	II(f)	f	f	F
<i>mužu</i> 'husband'	I(m)	m	m	M
<i>mušterija</i> 'customer' (when used for male)	II(f)	m	m	M
<i>sudija</i> 'judge' (when used for male)	II(f)	m	m	M
<i>Steva</i> 'Steve'	II(f)	m	m	M
<i>Jovanova</i> 'John's-f.sg'	II(f)	f	m	M
<i>Marijin:</i> 'Mary's-m.sg'	I(m)	m	f	F
<i>deca</i> 'children'	II(f.sg)	f.sg.	n.pl	>1
<i>gospoda</i> 'gentlemen'	II(f.sg)	f.sg	m.pl	M
<i>braća</i> 'brothers'	II(f.sg)	f.sg	n.pl	M, >1
<i>makaze</i> 'scissors' (COUNT 1)	pl	pl	pl	1

(m = masculine, f = feminine, M = male, F = female)

Table II. Declension, CONCORD, INDEX, and semantic features of sample Serbo-Croatian common nouns

References.

- Barlow, Michael. 1992. *A Situated Theory of Agreement*. (1988 Stanford PhD Dissertation). New York: Garland Publishing.
- Bresnan, Joan, and Sam A. Mchombo. 1987. Topic, pronoun, and agreement in Chichewa. *Language* 63 (4):741-782.
- Corbett, Greville. 1983. *Hierarchies, Targets and Controllers: Agreement Patterns in Slavic*. London: Croon Helm.
- Corbett, Greville. 1991. *Gender*. Cambridge: Cambridge University Press.
- Corbett, Greville G. 1998. Agreement in Slavic. *Position paper for the conference on Comparative Slavic Morphosyntax, Bloomington, Indiana, 5-7 June, 1998*.
- Kathol, Andreas. to appear. Agreement and the Syntax-Morphology Interface in HPSG. In *Readings in HPSG.*, edited by R. Levine and G. Green. Cambridge: Cambridge University Press.
- Marković, Svetozar. 1954. O kolebljivosti slaganja u rodu kod imenica čiji se prirodni i gramatički rod ne slažu (i o rodu ovih imenica). *Pitanja književnosti i jezika* 1:87-110.
- Pollard, Carl, and Ivan Sag. 1994. *Head Driven Phrase Structure Grammar*. Stanford and Chicago: CSLI Publications and University of Chicago Press.
- Wechsler, Stephen and Larisa Zlatic. 1997. (Dis)agreement in Serbo-Croatian. Manuscript. University of Texas at Austin.
- Wechsler, Stephen and Larisa Zlatic. to appear. Agreement in Discourse. In *Proceedings of the Conference on the Structure of Non-narrative Texts, University of Texas, Austin, February 13, 1998*.