Towards a Typology of Agreement Phenomena

Tania Avgustinova
Saarland University

Hans Uszkoreit
DFKI and Saarland University

Agreement phenomena are instances of co-variation of linguistic forms which is typically realised as feature congruity, i.e. compatibility of values of identical grammatical categories of syntactically combined linguistic items. Agreement is a relatively well-researched topic, especially in Slavic linguistics, cf. (Corbett, 2000a). However, the investigations have mainly concentrated on the linguistic items themselves (as agreement sources) and on the relevant properties of these items (in terms of agreement features and conditions). The nature of the relations holding between the "agreeing" items has not received proper attention yet.

1. Systematic relations

Our main hypothesis – articulated already in (Avgustinova and Uszkoreit, 2000) – is that systematic relations motivate shared patterns of variation cross-linguistically as well as across constructions. The ontology we develop[1] allows various degrees of abstraction as well as language-specific and construction-specific parameterisation. Thus, by design, it offers a novel typological perspective.

Being originally designed to systematise the inventory of syntactic relationships found across Slavic languages, the outlined approach allows us to specify more precisely the nature of the observable co-variation phenomena as well as to properly sub-classify them. In the presentation, a fairly pragmatic approach to terminology is adopted in order to ensure that all relevant distinctions are consistently made.

The dimensions of classification for (the arrays of) systematic relations discernible in syntactic constructions are sketched in (FIGURE 1).[2] The focus of our attention will be on segmental systematic relations in terms of *syntagmatics*, as they play a constitutive role in syntax. In accord with the traditional "form–function" perspective in theoretical linguistics, it is important to distinguish dimensions of *observable* syntagmatics (which is concerned with the overt linguistic form) and *structural* syntagmatics (which is concerned with the covert linguistic function). Structural syntagmatics is crucial in interpreting the observable syntagmatic relations which, in turn, can be classified as *combinatorial* (i.e. morphosyntactic) and *alignment* (i.e. configurational).

[1] Our use of the term ontology is fairly pragmatic namely, as representing a formal shared conceptualisation of a particular domain of interest. It describes concepts relevant for the domain, their relationships, as well as "axioms" about these concepts and relationships. Note that such a pragmatic approach does not presuppose any general all-encompassing ontology of language but rather "mini-ontologies" conceptualising the selected domain from various perspectives in a consistent way.

[2] The different shapes of edges connecting types in the graphical representation of hierarchies are significant. The 'square' edges indicate *conjunction* of types partitioning their super-type along various dimensions. The 'direct' edges indicate *disjunction* of types within the respective dimension. Cross-classifications encoding multiple inheritance are permitted with disjunctive but not with conjunctive types.

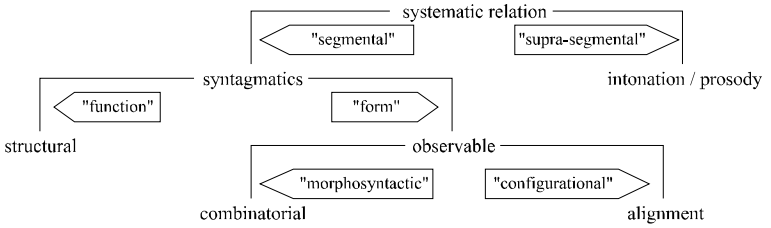

FIGURE 1. Systematic relations: dimensions of classification.

Syntagmatic regularities in morphosyntax reveal basic relations between properties of linguistic objects. Along with government and juxtaposition, co-variation belongs to what (Schmidt and Lehfeldt, 1995) regard as morphological signalling of direct syntactic relations, and (Avgustinova and Uszkoreit, 2000) call combinatorial syntagmatics. The latter is introduced as a separate dimension of classification within an HPSG-style multiple-inheritance type-hierarchy of systematic relations (FIGURE 2). The combinatorial syntagmatics encompasses observable relations of *assembling* (or "valence" in a broader sense) and *co-variation* (or "agreement" in a broader sense). Assembling includes what is traditionally considered government and juxtaposition. The former is understood as the determination by one element of the inflectional form of the other (i.e. form government; a classical instance thereof is case government), while the latter, in contrast, presupposes no overt morphological indication (its classical instance is case adjunction).

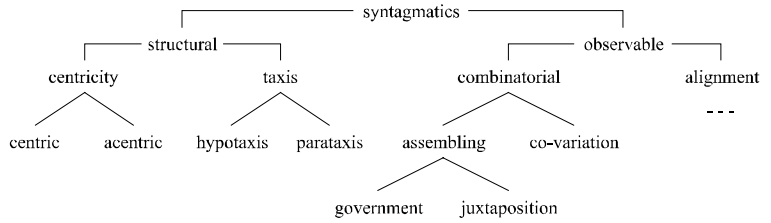

FIGURE 2. Syntagmatics.

Hypotaxis is a key notion in X-bar syntax.[3] Note that from the outlined perspective, the bar-level promoting relations are centric, while the bar-level preserving relations are acentric. Parataxis, in turn, is crucial for what can be called "mediation scheme".[4] The centricity dimension plays an important role in classifying (as well as predicting) paratactic phenomena. In particular, "restraint" mediation (e.g., control, resumption) is centric, while "coequal" mediation (e.g., co-dependence, coordination) is acentric (TABLE 1). The admissible cross-classifications in structural syntagmatics result in distinguishing four major types of relations. The *centric hypotaxis* is an 'endocentric' relational type representing the most structurally

[3] The X-bar scheme is a restrictive mechanism for delimiting possible syntactic (or morphological) structures. General assumptions: (i) every X^n is a projection of X; (ii) X^{max} is the maximal projection of X; (iii) every phrase has a head determining its specific properties; (iv) the head properties are preserved in all projections; (v) a head category X combines with a non-head category Y which can be a complement (bar-level promotion: $X^n \rightarrow Y^{max}\ X^{n-1}$), an adjunct (bar-level preservation: $X^n \rightarrow Y^{max}\ X^n$) or a specifier (special case of bar-level promotion: $X^{max} \rightarrow Y^{max}\ X^{max-1}$).

[4] The X-bar mechanism (modelling immediate relations) is irrelevant for parataxis which is generally not interpretable in terms of subordination.

marked option because there is a designated (central, or leading) element as well as a subordination relation between the items involved. The most structurally unmarked option, in turn, is the *acentric parataxis* which can be interpreted as an 'exocentric' relational type. The other possibilities include the *centric parataxis* which is an 'only-centric' relational type presupposing a designated element but no subordination, and the *acentric hypotaxis* which is an 'only-hypotactic' relational type involving subordination although none of the items is unambiguously interpretable as central.

	centric	*acentric*
hypotaxis "X-bar scheme"	*centric hypotaxis* "bar-level promotion"	*acentric hypotaxis* "bar-level preservation"
parataxis "mediation scheme"	*centric parataxis* "restraint"	*acentric parataxis* "coequal"

TABLE 1. Structural syntagmatics (cross-classification).

Looking at the ways structural syntagmatics is externalised by combinatorial syntagmatics helps us reveals various classes of phenomena. The admissible cross-classifications of the structural syntagmatic types with the assembling types gives us the result in (TABLE 2). For the sake of perspicuity, we mention here mainly phenomena that will come up in the examples later on. Note that not only well-known phenomena – like subcategorisation (with its more specific instance relational case), concordial case, control, marking, adjunction or coordination – can naturally be accommodated on such an approach but, crucially, new classes are systematically predicted (e.g., co-dependence).

	government	*juxtaposition*
centric hypotaxis "bar-level promotion"	subcategorisation *e.g., relational case, cross-referencing, object cliticisation*	marking
acentric hypotaxis "bar-level preservation"	governed modification *e.g., concordial case*	juxtaposed modification *e.g., (case) adjunction, secondary predication*
centric parataxis "restraint"	governed centric parataxis *e.g., control*	juxtaposed centric parataxis *e.g., relativising*
acentric parataxis "coequal"	governed acentric parataxis *e.g., co-dependence*	juxtaposed acentric parataxis *e.g., coordination*

TABLE 2. Predicted classes of assembling phenomena.

As our main topic of interest in this study is the typology of agreement phenomena, let us concentrate on how structural syntagmatics can be externalised via co-variation (TABLE 3). The centricity dimension of structural syntagmatics appears to be essential in classifying observable agreement phenomena. Taking into consideration how the sources of co-variation (i.e. the 'agreeing' items) are related to each other, we can distinguish two major types of co-variation: *asymmetric* and *balanced (distributed)*. The asymmetric co-variation is centric. It corresponds to the traditional directional concept, since one of the two co-variation sources is unambiguously interpretable as the trigger and the other one as the target of this relation. The trigger–target configuration can, more specifically, be *unidirectional*, if all co-varying grammatical categories are triggered at the same item, or *unstipulated*, if the items involved trigger different co-varying grammatical categories. The balanced (distributed) co-variation, in contrast, is acentric. Presupposing redundancy, it cannot be formulated in such directional terms. Intuitively, both co-variation sources are often interpretable as co-targets of an external trigger.

	co-variation		
	asymmetric unidirectional	*asymmetric unstipulated*	*balanced / distributed*
hypotaxis	agreement 1	agreement 2 (concord)	matching
parataxis	co-reference	agreement 3 (accord)	correlation
	centric		*acentric*

TABLE 3. Predicted classes of co-variation phenomena.

2. Morphosyntactic co-variation

The admissible cross-classifications with the structural taxis dimension result in six classes of co-variation phenomena. All known forms of agreement are obtained automatically and novel concepts of co-variation are predicted. The typology of morphosyntactic co-variation is sketched graphically in (FIGURE 3).

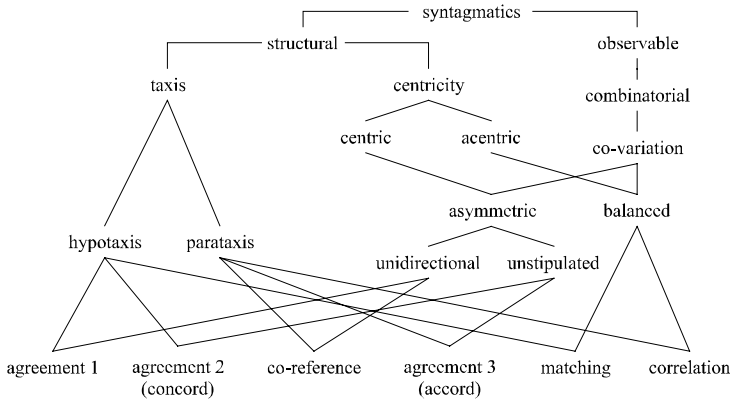

FIGURE 3. Types of co-variation.

Let us now look at examples from a Slavic language with a rich case system (Russian) and from another one with no cases in the nominal system but showing the phenomenon of "clitic doubling" (Bulgarian). The linguistic examples in (1)–(4) will be visualised as relational charts, a representation originally employed by (Avgustinova and Uszkoreit, 2000) which allows us to specify the array of systematic relations holding between any two items in the relevant "crossing" cell. A regular *affinity of assembling and co-variation* can be observed which is based on structural centricity, since either both relations involved are centric or at least one of them is. In addition, the actual co-occurrence of assembling and co-variation in a given array of systematic relations presupposes the same type of taxis.

(1)

Ona she.NOM.3SG.F	*rel-case [NOM]* *agr1 [SG.F]*	*co-dependence* *agr3 (accord) [SG]*
okazalas' turned.SG.F		*rel-case [INST]* *agr1 [SG]*
zdorovym healthy.INST.SG.M		*con-case [INST]* *agr2 (concord) [SG.M]*
rebënkom. child.INST.3SG.M		

'She turned out a healthy child.' (Russian)

(2)

Maria Mary.3SG.F	*cross-referencing* *agr1 [SG.F]*	*subcat*	*control* *co-reference [SG.F]*
	ja ACC.SG.F	*obj-cliticisation*	*control* *co-reference [SG.F]*
		vidjaxa saw.3PL	*secondary predication*
			maskirana. disguised.SG.F

'(They) saw Mary disguised.' (Bulgarian)

Agreement 1: this is hypotactic unidirectional co-variation. It holds, e.g., in number and gender between the verb (okazalas' 'turned out') and its subject (ona 'she'), or just in number between the same verb and its complement (rebënkom 'child') in (1). Co-variation in person, number and gender of the same type also holds between the verbal clitic pronoun (ja 'her') cliticized on the verb (vidjaxa 'saw') and the object (Maria 'Mary') cross-referenced by this clitic in (2). The trigger of the discussed co-variation is the nominal element, and the target is the verb or the clitic pronoun, respectively.

Agreement 2 (concord): this is a hypotactic unstipulated co-variation. Its prototypical instance can be found within nominal phrases, e.g., holding in number and gender between the adjective (zdorovym 'healthy') and the noun (rebënkom 'child') in (1). The trigger is the noun and the target is the adjective.

Co-reference: this is a paratactic unidirectional co-variation. In (2) it holds in number and gender between the object (Maria 'Mary') and the predicative adjective controlled by it (maskirana 'disguised'), but also between the verbal clitic (ja 'her') cross-referencing the object and the predicative adjective. The co-variation trigger here is the object noun or the verbal object clitic, respectively, while the target in both cases is the predicative adjective.

Agreement 3 (accord): this is a paratactic unstipulated co-variation. It holds in number between the subject (ona 'she') and the complement (rebënkom 'child') which are co-dependents of the same verb (okazalas' 'turned out') in (1). The trigger of co-variation is the subject, while the complement presents the co-variation target.

(3)

Ti you.2SG				*subcat* *agr1 [2SG]*
	si AUX.2SG	*marking* *matching [2SG.F]*		*marking* *matching [2SG.F]*
		štjala AUX.SG.F		
			da PRT	*marking*
				dojdeš. come.2SG

'You would come (reportedly).' (Bulgarian)

Matching: this is a hypotactic balanced co-variation. Its prototypical instance is the compatibility between the auxiliaries and the main verb in periphrastic forms). As discussed in (Avgustinova, 1997), the person–number–gender information in Bulgarian analytic (i.e. periphrastic) verb forms can be distributed among several components, namely, the main verb itself and a set of auxiliaries functioning as markers to it. The analytic verb form in (3) consists of two auxiliaries, a particle and a main verb (si štjala da dojdeš 'come.FUTURE.RENARRATIVE.2SG.F'). In fact, the balanced co-variation relation of matching

holds in all three grammatical categories between the 2nd person singular auxiliary (si) and the singular feminine auxiliary participle (štjala), as well as between this combination of auxiliaries (si štjala) and the 2nd person singular main verb (dojdeš 'come').

Correlation: this is a paratactic balanced co-variation. It is typically observed in relative clause constructions. So, in (4) it holds between the relative pronoun (kogoto 'whom') and the noun (studentyt 'the student') modified by the relative clause. The observed compatibility encompasses all three grammatical categories, i.e. person, number and gender.

(4)

Vliza enter.3SG	*subcat* *agr1 [3SG]*			
	studentyt, student.DEF.3SG.M		*relativising* *correlation [3SG.M]*	*adjunction*
		za about	*marking*	
			kogoto whom.SG.M	*subcat*
				govorixme. talked.1PL

'The student whom we talked about comes in.' (Bulgarian)

Our default assumption up to now was that co-variation could directly be encoded as structure sharing in terms of feature unification. For example, it is a common standard to use identically numbered boxes in the values of the relevant attributes. Although this appears to correspond to the most typical situation, such a view is definitely an oversimplification. It is well-known that the syntagmatic relation of coordination may affect co-variation in a non-monotonic way. In particular, (Corbett, 1998) observes: "An agreement controller consisting of conjoined noun phrases may well give rise to an agreement option. It may allow agreement with both or all the conjuncts, and it may allow agreement with just one conjunct."

Therefore we refine our ontology by partitioning the type *asymmetric* along two dimensions. The *arrangement* dimension covers the familiar distinction between *unidirectional* and *unstipulated* asymmetric co-variation – cf. (FIGURE 3), while the *compatibility* dimension introduces *monotonic* and *non-monotonic* co-variation as sub-types of asymmetric (FIGURE 4). The non-monotonic asymmetric co-variation is further specified with respect to the particular strategy employed. *Strategy A (resolved)* means that in establishing co-variation, conjoined noun phrases are treated as a semantically justified syntactic unit with a *resolved* index.[5] *Strategy B (partial)* means that the one of the conjuncts is favoured as decisive in establishing co-variation, mainly on *alignment* grounds. Both strategies are illustrated by the Czech example in (5). A resolved agreement 1 holds between the subject containing the conjunction of singular nominals (den i stát 'day and state') and the plural verb form (jsou opŚedeny 'are wrapped'). A partial agreement 2 (concord) holds within the subject itself between the singular demonstrative pronoun (tento 'this') and the conjunction of singular nominals (den i stát 'day and state').

(5) Tento den i stát jsou v našem podvədomí
this.SG day.SG and state.SG are.PL in our unconsciousness
opŚedeny mnoha mýty o ļeské jedineļnosti.
wrapped.PL many myths about Czech uniqueness
'This day and this state are surrounded in our unconsciousness by many myths about Czech uniqueness.' (Lidové noviny, ļ. 250/251 1998)

[5] The interested reader may consult (Corbett, 1998; Corbett, 2000b) for a detailed discussion of the so-called resolution rules, as well as for an extensive presentation of Slavic (and other) data and further references.

Let us consider how the observable Strategy B realisations fall out from our ontological representation as a result of admissible cross-classifications. In order to integrate the relevant alignment factors, we have to be more explicit about the configurational dimension of syntagmatics. So, the alignment relation is classified in (FIGURE 4) with respect to directionality (i.e. the mutual order of the trigger and the target) and periphery (left or right).

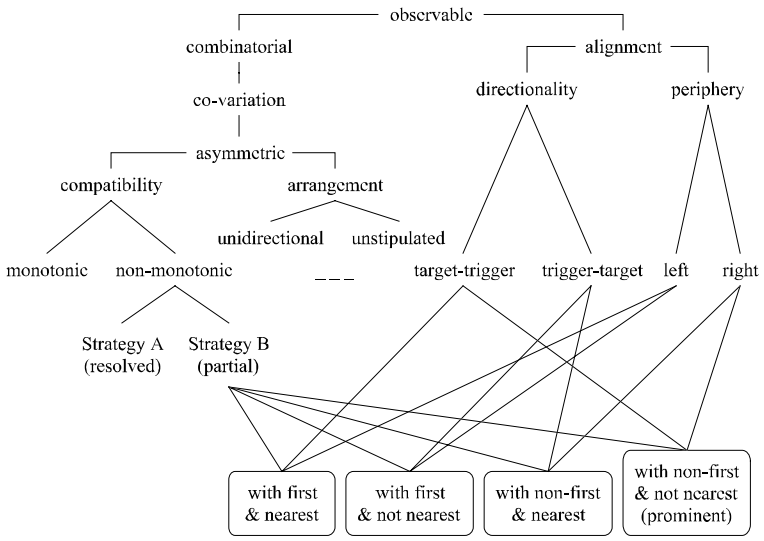

FIGURE 4. Partial co-variation with conjoined noun phrases.

The designated conjunct that determines the co-variation specifications at the target item can be both initial in the conjunction and the nearest to the target; initial in the conjunction but not the nearest to the target; non-initial in the conjunction but yet the nearest to the target; and, finally, neither initial in the conjunction nor the nearest to the target. While the first three variants of Strategy B are fairly common across languages, the fourth variant is also attested. In particular, (Corbett, 1998; Corbett, 2000b) mentions some interesting although limited evidence that in Serbo-Croatian Ļakavian dialects of 16^{th}-17^{th} centuries agreement has been attested "with the most important conjunct, even if this was not the nearest or the first".

3. Typology and Grammar Theory

In this section we will indicate how the proposed typology can be combined with formal grammatical descriptions. We will not attempt to propose any extensions to existing grammar models since the status of our generalisations with respect to each model still needs to be determined.

Grammatical theories contain complex descriptions of classes of grammatical objects, i.e., words, phrases and sentences. Grammar formalisms provide the means for formulating such complex descriptions. Generative grammars are the basis for producing or approving the correct representations with respect to a theory. We adopt here the constraint-based approach to generative grammar. Grammars are formulated and applied as complex constraints on permissible grammatical representations.

Grammatical representations encode both properties of individual objects and linguistically relevant relationships between two or more objects. We are concerned with the systematic relationships among the grammatical objects within a sentence. Depending on the theory, such relationships may be explicitly constrained by the grammar or they may be implicitly constrained through the interaction of several constraints. Some of these relationships are encoded in the lexicon, others are indirectly specified through the interaction of lexical information and syntactic rules or principles.

In today's more or less lexicalised theories of syntax, lexical representations of words contain explicit information about other objects the word can or must be combined with. Examples of such relationships are valence features of lexical heads such as HPSG's SUBCAT list or LFG's grammatical functions. Further examples are valence features of lexical adjuncts (HPSG's feature MOD) and even long-distance dependencies such as the reference to a missing NP in the infinitival VP-complement of "tough"-adjectives. Relationships among grammatical objects can also be encoded in rules or principles that combine such objects or license combinations. A phrase structure rule can be annotated by feature descriptions establishing such relationships. An example is the feature equation for testing the CASE of oblique objects in the English VP rule of LFG.

Relationships among grammatical objects are finally expressed in the grammatical representations of a larger unit to which the related objects belong. We can, for instance, read off all short or long distance dependencies within a sentence from the LFG f-structure or from the complete HPSG feature structure that the theories assign. The relationships that are realised in sentences and encoded in their representations can be the result of interacting statements. In (some variants of) HPSG, the relationship between the interrogative pronoun *who* and the embedded verb *see* in the (6) is the result of applying constraints from the lexical entries of *see, who* and a trace, the rule schema combining filler and sentential head and the non-local feature principle carrying the slash feature through the tree.

(6) Who_1 did you tell Mary to try to see $\__1$.

In order to arrive at a universal systematics of possible and actually realised relationships among grammatical objects, we will abstract away for the time being from the mechanisms that individual grammatical frameworks or grammar writers employ for specifying combinatory rules or constraints. We propose the following research strategy. Instead of trying to couch our ontology of relationships into an existing framework, we propose a formalisation that can be easily adapted to any cleanly defined constraint-based grammar model. A class of constraints called relational dependencies provides a universal means of introducing more abstract and modular specifications in grammar and lexicon (Dörre *et al.*, 1992). Relational dependencies are constraints that hold among typed feature structures. If we allow relational dependencies as part of our grammar specification language, they can be used within the specified types. They are constraints on permissible values of features with respect to other values. In HPSG, relational dependencies are employed at several places, i.e., to express complex linearization principles (Constituent Ordering Principle) or to express the fact that two lists are the concatenation of another list (Subcategorisation Principle).

Since we have based our notion of grammatical relationships on binary dependencies, we only need binary relational dependencies. Relational dependencies themselves can be expressed as feature structures with two attributes. These feature structures themselves can be typed. The types can be ordered in a multiple-inheritance hierarchy, preferably a semi-lattice. In this way we can construct a formal specification of the hierarchy of dependencies. The top element of the hierarchy is the type *rel-dep.* The values of the two attributes are of the least specific type assigned to grammatical objects. Borrowing terminology from HPSG we will assume that the type of these values is *sign*.

As we saw in (FIGURE 1), the most general type in our case study is *sys(tematic)-rel(ation)* for which the two attributes ARG1 and ARG2 are certainly appropriate (7). Its *syntagm(atics)* subtype specifies the values of the arguments as distinct linguistic entities. Borrowing terminology from HPSG we will assume that the type of these values is *sign*, which is ensured by the type *str(uctural)-syntagm(atics)*.

(7) $$\underset{sys-rel}{\begin{bmatrix}\text{ARG1}\\\text{ARG2}\end{bmatrix}} \quad \underset{syntagm}{\begin{bmatrix}\text{ARG1} & \boxed{\begin{smallmatrix}\text{EQ}\\\text{00}\end{smallmatrix}}\\\text{ARG2} & \boxed{\begin{smallmatrix}\text{EQ}\\\text{01}\end{smallmatrix}}\end{bmatrix}} \quad \underset{str-syntagm}{\begin{bmatrix}\text{ARG1} & \boxed{1}_{sign}\begin{bmatrix}\ \end{bmatrix}\\\text{ARG2} & \boxed{2}_{sign}\begin{bmatrix}\ \end{bmatrix}\end{bmatrix}}$$

We can now define a number of relationships among signs. The dimensions of centricity and taxis are encoded in (8) and (9), respectively. The *centric* type is associated with a disjunctive one-place predicate *center* identifying one of the related items as central. The *acentric* type, in contrast, is associated with a two-place predicate *x-center* establishing the unmarked case where neither of the items can unambiguously be identified as central.

(8) $$\underset{centric}{\begin{bmatrix}\text{ARG1}\ \boxed{1}\\\text{ARG2}\ \boxed{2}\\ center(\boxed{1}) \boxed{\begin{smallmatrix}\text{OR}\\\text{NOR}\end{smallmatrix}} center(\boxed{2})\end{bmatrix}} \quad \underset{acentric}{\begin{bmatrix}\text{ARG1}\ \boxed{1}\\\text{ARG2}\ \boxed{2}\\ x\text{–}center(\boxed{1},\boxed{2})\end{bmatrix}}$$

Similarly, the *hypotaxis* type is associated with a disjunctive one-place predicate *dominant* that identifies one of the related items as dominating the other. The *parataxis* type, in turn, is associated with a two-place predicate *para* establishing the unmarked case where neither of the items can unambiguously be identified as dominant.

(9) $$\underset{hypotaxis}{\begin{bmatrix}\text{ARG1}\ \boxed{1}\\\text{ARG2}\ \boxed{2}\\ dominant(\boxed{1}) \boxed{\begin{smallmatrix}\text{OR}\\\text{NOR}\end{smallmatrix}} dominant(\boxed{2})\end{bmatrix}} \quad \underset{parataxis}{\begin{bmatrix}\text{ARG1}\ \boxed{1}\\\text{ARG2}\ \boxed{2}\\ para(\boxed{1},\boxed{2})\end{bmatrix}}$$

The admissible cross-classifications from (TABLE 1) – i.e. within the structural syntagmatic dimension – are encoded in (10–13). The *centric-hypotaxis* type (10) states that the central item and the dominating item must coincide. The *centric-parataxis* type (11) specifies one of the items as central, while excluding any dominance between them. The *acentric-hypotaxis* type (12), in contrast, defines one of the items as dominant, while ensuring that none of them is central. Finally, the *acentric-parataxis* type (13) states that neither of the items can be unambiguously identified as central or dominant.

(10) $$\underset{centric-hypotaxis}{\begin{bmatrix}\text{ARG1}\ \boxed{1}\\\text{ARG2}\ \boxed{2}\\ \Big(center(\boxed{1}) \boxed{\begin{smallmatrix}\text{EQ}\\\text{00}\end{smallmatrix}} dominant(\boxed{1})\Big) \boxed{\begin{smallmatrix}\text{AND}\\\text{11}\end{smallmatrix}} \Big(center(\boxed{2}) \boxed{\begin{smallmatrix}\text{EQ}\\\text{01}\end{smallmatrix}} dominant(\boxed{2})\Big)\end{bmatrix}}$$

(11) $$\underset{centric-parataxis}{\begin{bmatrix}\text{ARG1}\ \boxed{1}\\\text{ARG2}\ \boxed{2}\\ \Big(center(\boxed{1}) \boxed{\begin{smallmatrix}\text{OR}\\\text{NOR}\end{smallmatrix}} center(\boxed{2})\Big) \boxed{\begin{smallmatrix}\text{EQ}\\\text{10}\end{smallmatrix}} para(\boxed{1},\boxed{2})\end{bmatrix}}$$

(12)
$$_{acentric\text{–}hypotaxis}\left[\begin{array}{l}\text{ARG1}\ \boxed{1}\\ \text{ARG2}\ \boxed{2}\\ x\text{–}cente\left(\boxed{1},\boxed{2}\right)\ \left(dominant\left(\boxed{1}\right)\ dominant\left(\boxed{2}\right)\right)\end{array}\right]$$

(13)
$$_{acentric\text{–}parataxis}\left[\begin{array}{l}\text{ARG1}\ \boxed{1}\\ \text{ARG2}\ \boxed{2}\\ x\text{–}center\left(\boxed{1},\boxed{2}\right)\ para\left(\boxed{1},\boxed{2}\right)\end{array}\right]$$

Turning now to the typology of agreement phenomena, let us consider how the classification from (FIGURE 3) will be encoded. The type *comb(inatorial)-syntagm(atics)* highlights certain key properties of the signs involved, and its subtype *covariation* specifies them as co-variation sources (14) by means of a two-place predicate *covar-sources*.

(14)
$$_{comb\text{–}syntagm}\left[\begin{array}{ll}\text{ARG1} & \boxed{1}_{sign}\left[\ldots\ \boxed{3}\right]\\ \text{ARG2} & \boxed{2}_{sign}\left[\ldots\ \boxed{4}\right]\end{array}\right]\quad _{covariation}\left[\begin{array}{ll}\text{ARG1} & \boxed{1}\left[\ldots\ \boxed{3}\right]\\ \text{ARG2} & \boxed{2}\left[\ldots\ \boxed{4}\right]\\ covar\text{–}sources\left(\boxed{3},\boxed{4}\right)\end{array}\right]$$

The *asym(metric)-covar(iation)* type (15) systematically differs from the *bal(anced)-covar(iation)* type (16) with respect to centricity.

(15)
$$_{asym\text{–}covar}\left[\begin{array}{ll}\text{ARG1} & \boxed{1}\left[\ldots\ \boxed{3}\right]\\ \text{ARG2} & \boxed{2}\left[\ldots\ \boxed{4}\right]\\ \left(center\left(\boxed{1}\right)\ center\left(\boxed{2}\right)\right)\ covar\text{–}sources\left(\boxed{3},\boxed{4}\right)\end{array}\right]$$

(16)
$$_{bal\text{–}covar}\left[\begin{array}{ll}\text{ARG1} & \boxed{1}\left[\ldots\ \boxed{3}\right]\\ \text{ARG2} & \boxed{2}\left[\ldots\ \boxed{4}\right]\\ x\text{–}center\left(\boxed{1},\boxed{2}\right)\ covar\text{–}sources\left(\boxed{3},\boxed{4}\right)\end{array}\right]$$

The more specific types *unidir(ectional)-asym(metric)-covar(iation)* and *unstip(ulated)-asym(metric)-covar(iation)* (17) actually differ with respect to the uniqueness of the co-variation trigger. It is identified in the former type by means of a disjunctive one-place predicate *trigger*, while the latter type is associated with a two-place predicate *trigger-target* which indicates that the trigger and the target cannot be unambiguously identified.

(17)
$$_{unidir\text{–}asym\text{–}covar}\left[\begin{array}{l}\text{ARG1}\left[\ldots\ \boxed{3}\right]\\ \text{ARG2}\left[\ldots\ \boxed{4}\right]\\ trigger\left(\boxed{3}\right)\ trigger\left(\boxed{4}\right)\end{array}\right]\quad _{unstip\text{–}asym\text{–}covar}\left[\begin{array}{l}\text{ARG1}\left[\ldots\ \boxed{3}\right]\\ \text{ARG2}\left[\ldots\ \boxed{4}\right]\\ trigger\text{–}target\left(\boxed{3},\boxed{4}\right)\end{array}\right]$$

Now, we are able to encode in terms of relational dependencies six distinct classes of co-variation phenomena, as predicted in (TABLE 3) and (FIGURE 3). The resulting types are sketched in (18–23).

(18)
$$\textit{agreement 1}\left[\begin{array}{l}\text{ARG1}\quad \boxed{1}\left[\ldots\ \boxed{3}\right]\\ \text{ARG2}\quad \boxed{2}\left[\ldots\ \boxed{4}\right]\\ \left(\textit{dominant}\ \left(\boxed{1}\right)\ \textit{trigger}\left(\boxed{3}\right)\right)\ \left(\textit{dominant}\ \left(\boxed{1}\right)\ \textit{trigger}\left(\boxed{4}\right)\right)\end{array}\right]$$

(19)
$$\textit{agreement 2 (concord)}\left[\begin{array}{l}\text{ARG1}\quad \boxed{1}\left[\ldots\ \boxed{3}\right]\\ \text{ARG2}\quad \boxed{2}\left[\ldots\ \boxed{4}\right]\\ \left(\textit{dominant}\ \left(\boxed{1}\right)\ \textit{dominant}\ \left(\boxed{2}\right)\right)\ \textit{trigger–target}\left(\boxed{3},\boxed{4}\right)\end{array}\right]$$

(20)
$$\textit{agreement 3 (accord)}\left[\begin{array}{l}\text{ARG1}\quad \boxed{1}\left[\ldots\ \boxed{3}\right]\\ \text{ARG2}\quad \boxed{2}\left[\ldots\ \boxed{4}\right]\\ \textit{para}\left(\boxed{1},\boxed{2}\right)\ \textit{trigger–target}\left(\boxed{3},\boxed{4}\right)\end{array}\right]$$

(21)
$$\textit{co–reference}\left[\begin{array}{l}\text{ARG1}\quad \boxed{1}\left[\ldots\ \boxed{3}\right]\\ \text{ARG2}\quad \boxed{2}\left[\ldots\ \boxed{4}\right]\\ \textit{para}\left(\boxed{1},\boxed{2}\right)\ \left(\textit{trigger}\left(\boxed{3}\right)\ \textit{trigger}\left(\boxed{4}\right)\right)\end{array}\right]$$

(22)
$$\textit{matching}\left[\begin{array}{l}\text{ARG1}\quad \boxed{1}\left[\ldots\ \boxed{3}\right]\\ \text{ARG2}\quad \boxed{2}\left[\ldots\ \boxed{4}\right]\\ \textit{acentric}\left(\boxed{1},\boxed{2}\right)\ \left(\textit{dominant}\ \left(\boxed{1}\right)\ \textit{dominant}\ \left(\boxed{2}\right)\right)\ \textit{covar–sources}\left(\boxed{3},\boxed{4}\right)\end{array}\right]$$

(23)
$$\textit{correlation}\left[\begin{array}{l}\text{ARG1}\quad \boxed{1}\left[\ldots\ \boxed{3}\right]\\ \text{ARG2}\quad \boxed{2}\left[\ldots\ \boxed{4}\right]\\ \textit{acentric}\left(\boxed{1},\boxed{2}\right)\ \textit{para}\left(\boxed{1},\boxed{2}\right)\ \textit{covar–sources}\left(\boxed{3},\boxed{4}\right)\end{array}\right]$$

Since we do attempt to propose a super-formalism or an interlingua of grammar formalisms, we will leave the formal interpretation of the predicates *center*, *x-center*, *dominant*, *para*, *covar-sources*, *trigger*, and *trigger-target* to the individual grammar models.

4. Conclusion and Outlook

The broad spectrum of agreement phenomena constitutes a challenge to any linguistic theory maintaining a universality claim and to any theoretically founded typological description. Because of the relational character of agreement, approaches to typology are needed that provide a fine-grained classification of possible relationships between grammatical units. We have proposed a multidimensional taxonomy that derives the space of possible relationships including agreement relations from a small of number of distinctions. We have demonstrated the descriptive power of such a taxonomy with a wide range of examples from several Slavic languages. We have finally shown how the descriptive device of relational dependency can be utilised to provide a formal framework for describing these relationships in such a way that the descriptions can be linked to constraint-based grammar formalisms.

We consider our proposal to be a first step towards a universal typology of relations that employs the power of multidimensional inheritance networks for a systematic and concise description.

The connections between our proposed dimensions and classes of thematic and semantic relations still remain to be investigated. The status of the typology needs to be determined with respect to constraint-based grammar models such as HPSG and LFG. The question is whether some or all of the dimensions can be derived from existing constraints of the theories. At this time, none of the theories provides a taxonomy of relations. Even in the sophisticated type hierarchy of HPSG, relations do not appear as types. We predict that the relational aspect of syntax will become more strongly reflected in the constraint system. Although we are not in a position at this point to submit a concrete proposal, we hope to contribute to this development.

References

Avgustinova, T. (1997). *Word order and clitics in Bulgarian*. Saarbrücken, Universität des Saarlandes / DFKI.

Avgustinova, T. and H. Uszkoreit (2000). *An ontology of systematic relations for a shared grammar of Slavic*. In Proceedings of the 18th International Conference on Computational Linguistics COLING'2000, Saarbrücken.

Corbett, G. G. (1998). *Agreement in Slavic (position paper). Workshop on Comparative Slavic Morphosyntax*. Indiana University. (http://www.indiana.edu/~slavconf).

Corbett, G. G. (2000a). *Agreement in the Slavonic Languages: A Provisional Bibliography*. (http://www.surrey.ac.uk/LIS/SMG/agr.html).

Corbett, G. G. (2000b). *Number*. Cambridge, UK, Cambridge University Press.

Dörre, J., A. Eisele and R. Seffert (1992). *Grammars as relational dependencies*. AIMS Report 7. Stuttgart, Institut für maschinelle Sprachverarbeitung.

Schmidt, P. and W. Lehfeldt (1995). Kongruenz - Rektion - Adjunktion. Systematische und historische Untersuchungen zur allgemeinen Morphosyntax und zu den Wortfügungen (slovosochetanija) im Russischen. München, Otto Sagner.

Department of Computational Linguistics
Saarland University
P. O. Box 151150, D-66043
Saarbrücken, Germany
avgustinova@coli.uni-sb.de / uszkoreit@dfki.de